

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 611 977**

51 Int. Cl.:

C12N 9/90	(2006.01)
C12N 9/92	(2006.01)
C12N 9/12	(2006.01)
C12P 7/06	(2006.01)
C12P 7/10	(2006.01)
C12N 1/21	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.12.2012 PCT/US2012/070456**

87 Fecha y número de publicación internacional: **18.07.2013 WO13106172**

96 Fecha de presentación y número de la solicitud europea: **19.12.2012 E 12808652 (7)**

97 Fecha y número de publicación de la concesión europea: **02.11.2016 EP 2794877**

54 Título: **Modificación del gen pnp para utilización mejorada de xilosa en zymomonas**

30 Prioridad:

20.12.2011 US 201161577871 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
11.05.2017

73 Titular/es:

**E. I. DU PONT DE NEMOURS AND COMPANY
(100.0%)
Chestnut Run Plaza, 974 Center Road
P.O. Box 2915, Wilmington, DE 19805, US**

72 Inventor/es:

**CAIMI, PERRY, G.;
QI, MIN;
TAO, LUAN;
VIITANEN, PAUL, V. y
YANG, JIANJUN**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 611 977 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

Modificación del gen *pnp* para utilización mejorada de xilosa en *Zymomonas*

Esta invención se realizó con el apoyo del gobierno de los Estados Unidos bajo el contrato nº DE-FC36-07G017056, concedido por el Departamento de Energía. El gobierno tiene ciertos derechos en esta invención.

5 Campo de la invención

La invención se refiere a los campos de la microbiología y la ingeniería genética. Más específicamente, el gen *pnp* en el genoma de *Zymomonas* se identificó como una diana para la modificación para proporcionar una utilización mejorada de la xilosa y la producción de etanol.

Antecedentes de la invención

10 La producción de etanol por microorganismos proporciona una fuente de energía alternativa a los combustibles fósiles y por tanto es un campo importante de la investigación actual. Es deseable que los microorganismos que producen etanol, así como otros productos útiles, puedan usar xilosa como fuente de carbono, ya que la xilosa es la pentosa principal en la biomasa lignocelulósica hidrolizada. La biomasa puede proporcionar un sustrato de carbono disponible en abundancia y de bajo coste. La *Zymomonas mobilis* y otros etanológenos bacterianos que no utilizan xilosa de forma natural, han sido modificados genéticamente para la utilización de la xilosa mediante la introducción de genes que codifican 1) xilosa isomerasa, que cataliza la conversión de xilosa en xilulosa; 2) xilulocinasa, que fosforila la xilulosa para formar 5-fosfato de xilulosa; 3) transcetolasa; y 4) transaldolasa (documentos US 5514583, US 5712133, US 6566107, WO 95/28476, Feldmann et al. (1992) Appl. Microbiol. Biotechnol. 38: 354 - 361, Zhang y otros (1995) Science 267: 240 - 243; Yanase et al. (2007) Appl. Environ. Microbiol., 73: 2592 - 2599). Típicamente, las regiones de codificación utilizadas eran de genes de *E. coli*.

Incluso con la expresión de esta ruta de utilización de xilosa, típicamente las cepas modificadas genéticamente no crecen ni producen etanol igual de bien en xilosa como en glucosa. Las cepas modificadas genéticamente para la utilización de xilosa han sido adaptadas por pasos en serie sobre medio de xilosa, dando como resultado cepas con una utilización mejorada de xilosa como se describe en la Patente de EE.UU. 7.223.575 y en la Patente de EE.UU. 7.741.119. Este último documento también describe la inactivación del locus GFOR que codifica glucosa-fructosa oxidorreductasa para mejorar la utilización de xilosa. En la Patente de EE.UU. 7.998.722 se describe la ingeniería genética para la utilización mejorada de xilosa mediante la expresión de xilosa isomerasa de *E. coli* a partir de un promotor de gen de gliceraldehído-3-fosfato deshidrogenasa (Pgap) de *Zymomonas mobilis* mutado altamente activo.

30 Sigue existiendo la necesidad de cepas manipuladas por ingeniería genética de *Zymomonas* y otros etanológenos bacterianos con una utilización mejorada de xilosa y producción de etanol en medio que contenga xilosa, y procedimientos para usar estas cepas para producir etanol.

Sumario de la invención

35 La invención proporciona células de *Zymomonas* o *Zymobacter* recombinantes que utilizan xilosa en las que se modifica un gen *pnp* endógeno que codifica polinucleótido fosforilasa. Además las células tienen actividad aumentada de ribosa-5-fosfato isomerasa (RPI) y actividad de xilosa isomerasa no limitante. La utilización de xilosa y la producción de etanol en medios que contienen xilosa están mejoradas en estas células.

En consecuencia, la invención proporciona una célula huésped bacteriana recombinante que comprende:

- a) una ruta metabólica de xilosa que comprende al menos un polipéptido que tiene actividad de xilosa isomerasa;
- 40 b) al menos una modificación genética que incrementa la actividad de ribosa-5-fosfato isomerasa en la célula huésped en comparación con la actividad de ribosa-5-fosfato isomerasa en la célula huésped que carece de dicha modificación genética, en donde dicha al menos una modificación genética que incrementa la actividad de ribosa-5-fosfato isomerasa se elige entre el grupo que consiste en:
 - 45 i) sobreexpresión de un gen endógeno que codifica un polipéptido que tiene actividad de ribosa-5-fosfato isomerasa; y/o
 - ii) expresión de al menos un gen no endógeno que codifica un polipéptido que tiene actividad de ribosa-5-fosfato isomerasa y
- c) al menos una modificación genética en la secuencia de un gen endógeno que codifica polinucleótido fosforilasa que acorta la región de codificación lo que da como resultado la expresión de una proteína C-terminal truncada;
- 50 en donde la célula huésped bacteriana utiliza xilosa para producir etanol y la actividad de xilosa isomerasa no es limitante en la célula huésped bacteriana, en donde la actividad de xilosa isomerasa no es limitante cuando un

aumento en la actividad de xilosa isomerasa no mejora la utilización de xilosa y la producción de etanol en dicha célula.

En otra forma de realización, la invención proporciona un procedimiento para producir etanol que comprende:

- a) proporcionar la célula huésped recombinante descrita anteriormente; y
- 5 b) cultivar la célula huésped de (a) en un medio que comprende xilosa, con lo que la xilosa se convierte en etanol.

Breve descripción de los dibujos y descripciones de las secuencias

Los solicitantes han realizado los siguientes depósitos biológicos bajo los términos del Tratado de Budapest sobre Reconocimiento del Depósito de Microorganismos para los fines de Procedimiento de Patentes:

Información sobre las cepas depositadas

Referencia de identificación del depositante	Designación del depósito internacional	Fecha del depósito
<i>Zymomonas</i> ZW658	ATCC No PTA-7858	12 sept. 2006

- 10 La Figura 1 muestra un diagrama de rutas para el metabolismo de la xilosa y la producción de etanol.
- La Figura 2 muestra un mapa de plásmido de pMODlinker-Spec-GapRpi.
- La Figura 3 muestra para las cepas B9, B11, I, cada una con sobreexpresión de RPI y la cepa progenitora ZW801-4: crecimiento en mRM3-X10 (A); consumo de xilosa y producción de ribulosa a las 54 horas (B), y crecimiento en mRM3-G10 (C).
- 15 La Figura 4 muestra para siete transformantes de la cepa I con xilosa isomerasa adicional (I (cm1-5,8,9)) así como testigos de la cepa I (I-1,2) y ZW801 (801-1,2): crecimiento en mRM3-X10 (A), y crecimiento en mRM3-G10 (B).
- La Figura 5 muestra el crecimiento en mRM3-X10 bajo condiciones de pH controlado de dos transformantes de cepa I con xilosa isomerasa (I (cm1,9)), así como testigos de cepa I y ZW801-4:
- 20 La Figura 6 muestra el crecimiento en mRM3-X10 para diez transformantes de cepa B11 con xilosa isomerasa adicional (B11 (cm1-10)), junto con los testigos B11 (B11-1,2) (A); o diez transformantes de la cepa B9 con xilosa isomerasa adicional (B9 (cm1-10)), como junto con los controles B9 (B9-1,2) (B).
- La Figura 7 muestra una alineación de secuencias de aminoácidos de la polinucleótido fosforilasa (SEQ ID NO: 2) codificada por pnp de la cepa ZM4 de *Z. mobilis* de tipo silvestre y la proteína de fusión codificada por el gen pnp modificado de la cepa I (SEQ ID NO: 9).
- 25 La Figura 8 muestra los mapas de plásmidos de pZX21 (A), pZX52 (B) y pZX6 (C).
- La Figura 9 muestra gráficos de crecimiento, glucosa utilizada y etanol producido para cultivos desarrollados en mRM3-G10 de ZW1-X109 (A), ZW1-X210 (B) y testigo ZW1 (C).
- La Figura 10 muestra gráficas de crecimiento, glucosa utilizada y etanol producido para cultivos desarrollados en mRM3-X10 de ZW1-109 (A), ZW1 - 210 (B) y el testigo ZW1 (C).
- 30 La Figura 11 muestra los mapas de plásmidos de pPNP-1 (A), pPNP-IN (B), pPNP-C (C) y pPNP-M (D).
- La Figura 12 muestra gráficas de crecimiento, glucosa utilizada y etanol producido para cultivos desarrollados en mRM3-G10 de ZW1-X109-PNPi (A), ZW1-X109-PNPc (B), ZW1-X109-PNPm (C), ZW1-X109-PNPin (D) y testigo ZW1-X109 (E).
- 35 La Figura 13 muestra gráficas de crecimiento, glucosa utilizada y etanol producido para cultivos desarrollados en mRM3-X10 de ZW1- X109-PNPi (A), ZW1-X109-PNPc (B), ZW1-X109-PNPm (C), ZW1-X109-PNPin (D) y testigo ZW1-X109 (E).
- La Figura 14 muestra gráficas de crecimiento, xilosa usada y etanol producido para cultivos desarrollados en mRM3-X10 de ZW801-PNPi (A), ZW801-PNPc (B) y testigo ZW801 (C).
- 40 Las secuencias que siguen son conformes con 37 C.F.R. 1.821-1.825 ("Requisitos para las solicitudes de patente que contienen secuencias de nucleótidos y/o descripciones de secuencias de aminoácidos - las reglas de secuencia") y son consistentes con la norma ST.25 (2009) de la World Intellectual Property Organization (WIPO) y los requisitos del listado de secuencias de la EPO y PCT (Reglas 5.2 y 49.5 (a-bis), y la Sección 208 y el Anexo C de las Instrucciones Administrativas). Los símbolos y el formato utilizados para los datos de secuencias de nucleótidos y aminoácidos cumplen con las reglas expuestas en 37 C.F.R. art. 1.822.

ES 2 611 977 T3

- La SEQ ID NO: 1 es la secuencia de nucleótidos de la región de codificación de pnp de la cepa ZM4 de *Zymomonas mobilis*.
- La SEQ ID NO: 2 es la secuencia de aminoácidos de la polinucleótido fosforilasa codificada por pnp de *Zymomonas mobilis*, cepa ZM4.
- 5 La SEQ ID NO: 3 es la secuencia de nucleótidos de la región codificadora de pnp de la cepa NCIMB 11163 de *Zymomonas mobilis*.
- La SEQ ID NO: 4 es la secuencia de aminoácidos de la polinucleótido fosforilasa codificada por pnp de *Zymomonas mobilis* cepa NCIMB 11163.
- 10 La SEQ ID NO: 5 es la secuencia de nucleótidos de la región codificadora de pnp de *Zymomonas mobilis* cepa ATCC 10988.
- La SEQ ID NO: 6 es la secuencia de aminoácidos de la polinucleótido fosforilasa codificada por pnp de la cepa ATCC 10988 de *Zymomonas mobilis*.
- La SEQ ID NO: 7 es la secuencia de nucleótidos de la región codificadora de pnp de *Zymomonas mobilis pomaceae* ATCC 29192.
- 15 La SEQ ID NO: 8 es la secuencia de aminoácidos de la polinucleótido fosforilasa codificada por pnp de *Zymomonas mobilis* cepa ATCC 29192.
- La SEQ ID NO: 9 es la secuencia de aminoácidos de la proteína de fusión codificada por pnp modificada de la cepa I, que tiene 709 aminoácidos N-terminales nativos y 14 aminoácidos C-terminales adicionales.
- 20 La SEQ ID NO: 10 es la secuencia de aminoácidos de una proteína de fusión codificada por pnp modificada que tiene 695 aminoácidos N-terminales nativos y 2 aminoácidos C-terminales adicionales.
- La SEQ ID NO: 11 es la secuencia de aminoácidos de una proteína de fusión codificada por pnp modificada que tiene 368 aminoácidos N-terminales nativos y 10 aminoácidos C-terminales adicionales.
- La SEQ ID NO: 12 es la secuencia de aminoácidos de una proteína de fusión codificada por pnp modificada que tiene 32 aminoácidos N-terminales nativos y 17 aminoácidos C-terminales adicionales.
- 25 La SEQ ID NO: 13 es la secuencia de nucleótidos de la región de codificación de la proteína RPI de *Z. mobilis*.
- La SEQ ID NO: 14 es la secuencia de aminoácidos de la proteína RPI de *Z. mobilis*.
- La SEQ ID NO: 15 es la secuencia de nucleótidos de la región codificadora de la proteína RPI de *E. coli*.
- La SEQ ID NO: 16 es la secuencia de aminoácidos de la proteína RPI de *E. coli*.
- 30 La SEQ ID NO: 17 es la secuencia de nucleótidos de la región de codificación de la xilosa isomerasa de *Actinoplanes missouriensis* que es de codón optimizado para la expresión en *Zymomonas mobilis*.
- La SEQ ID NO: 18 es la secuencia de aminoácidos de la xilosa isomerasa de *Actinoplanes missouriensis*.
- La SEQ ID NO: 19 es la secuencia de nucleótidos de la región de codificación de ribulosa-fosfato 3-epimerasa de *Zymomonas mobilis*.
- La SEQ ID NO: 20 es la secuencia de aminoácidos de la ribulosa-fosfato 3-epimerasa de *Zymomonas mobilis*.
- 35 La SEQ ID NO: 21 es la secuencia de nucleótidos del promotor mutante P_{gapS} (también denominado gap 801).
- La SEQ ID NO: 22 es la secuencia de nucleótidos de la casete de expresión de RPI que está en el plásmido pZB188aadA/Gap/Zymo RPI/EcoliSL, que está situada entre los sitios NcoI y NotI únicos.
- Las SEQ ID NO: 23 y 24 son los cebadores PPI-F y PPI-R-Sbfl.
- La SEC ID N°: 25 es la secuencia de nucleótidos del plásmido pZX21.
- 40 La SEQ ID NO: 26 es la secuencia de nucleótidos del fragmento GFO-L.
- La SEQ ID NO: 27 es la secuencia de nucleótidos de la secuencia de codificación de gfor.
- La SEQ ID NO: 28 es la secuencia de nucleótidos del fragmento GFO-R.

ES 2 611 977 T3

- La SEQ ID NO: 29 es la secuencia de nucleótidos de un gen *xylA* quimérico de 1.661 pb que contiene el promotor Super GAP de *Z. mobilis* de 304 pb, una secuencia de codificación *xylA* de *A. missouriensis* de 1.185 pb y un *araD* 3'UTR de *E. coli* de 166 pb con un sitio 5'XbaI.
- 5 La SEQ ID NO: 30 es la secuencia de nucleótidos de un gen *xylB* quimérico de 1.960 pb que contiene un P_{eno} de 191 pb, una secuencia de codificación *xylB* de pb de *E. coli* de 1.455 pb y un *xylB* 3'UTR de *E. coli* de 314 pb.
- La SEQ ID NO: 31 es la secuencia de nucleótidos de un marcador *aadA* de 1.014 pb (para resistencia a la espectinomicina, Spec-R) ligada por sitios *lox*.
- La SEQ ID NO: 32 es la secuencia de nucleótidos del vector lanzadera pZX52.
- La SEQ ID NO: 33 es la secuencia de nucleótidos del fragmento LDH-L.
- 10 La SEQ ID NO: 34 es la secuencia de nucleótidos del fragmento LDH-R.
- La SEQ ID NO: 35 es la secuencia de nucleótidos de la secuencia de codificación *ldhA*.
- La SEQ ID NO: 36 es la secuencia de nucleótidos de un operón P_{gapT}-Tal-Tkt de 3.339 pb que contiene un mutante T de 304 pb del promotor GAP de *Z. mobilis* (P_{gapT}), una región de codificación Tal de *E. coli* de 954 pb, una región de codificación Tkt de *E. coli* de 1.992 pb y una Tkt 3'UTR de *E. coli* de 68 pb.
- 15 La SEQ ID NO: 37 es la secuencia de nucleótidos del promotor P_{gapT}.
- La SEQ ID NO: 38 es la secuencia de nucleótidos de un operón P_{eno}-Rpi-Rpe de 1.443 pb que contiene un P_{eno} de 191 pb, una secuencia de codificación Rpi de *Z. mobilis* de 471 pb Rpi, una secuencia de codificación de Rpe de *Z. mobilis* de 663 pb, y un *xyl A* 3'UTR de *E. coli* de 35 pb.
- La SEQ ID NO: 39 es la secuencia de nucleótidos del vector lanzadera pZX6 de DCO.
- 20 La SEQ ID NO: 40 es la secuencia de nucleótidos del fragmento PNP-L.
- La SEQ ID NO: 41 es la secuencia de nucleótidos del fragmento PNP-R.
- Las SEQ ID NO: 42 a 54 son cebadores de PCR.
- La SEQ ID NO: 55 es la secuencia de nucleótidos del vector suicida de DCO pPNP - 1.
- 25 La SEQ ID NO: 56 es la secuencia de nucleótidos de la secuencia flanqueante aguas arriba para integración en el gen *pnp*.
- La SEQ ID NO: 57 es la secuencia de nucleótidos de la secuencia flanqueante aguas abajo para la integración en el gen *pnp*.
- La SEQ ID NO: 58 es la secuencia de nucleótidos del vector suicida de DCO pPNP - IN.
- La SEQ ID NO: 59 es la secuencia de nucleótidos del fragmento PNP-U.
- 30 La SEQ ID NO: 60 es la secuencia de nucleótidos del fragmento PNP-D.
- La SEQ ID NO: 61 es la secuencia de nucleótidos del vector suicida de DCO pPNP-C.
- La SEQ ID NO: 62 es la secuencia de nucleótidos del vector suicida de DCO pPNP-M.
- La SEQ ID NO: 63 es la secuencia de nucleótidos del fragmento PNPm-L.
- La SEQ ID NO: 64 es la secuencia de nucleótidos del fragmento PNPm-R.
- 35 La SEQ ID NO: 65 es la secuencia de nucleótidos de la región de codificación para la proteína RPI de *Z. mobilis* con el codón de inicio Mutado a ATG.

Descripción detallada

- Se describen en el presente texto células bacterianas que utilizan xilosa, y en particular células *Zymomonas* o *Zymobacter*, que tienen modificaciones genéticas de un gen *pnp* endógeno (modificación *pnp*). Las células están además modificadas genéticamente para que tengan una expresión incrementada de la actividad de ribosa-5-fosfato isomerasa (RPI), en comparación con las células sin esta modificación genética. Además, las células no están limitadas en la actividad de xilosa isomerasa en ausencia de la modificación de *pnp*. Las células con estas propiedades tienen una utilización de xilosa mejorada, que se desea para que crezcan en medios que contienen xilosa incluyendo la biomasa sacarificada, lo que conduce al aumento de la producción de etanol. El etanol es un
- 40

compuesto importante para uso en la sustitución de los combustibles fósiles, y la biomasa sacarificada proporciona una fuente de carbono renovable abundantemente disponible para la producción de etanol por fermentación.

Las definiciones que siguen se pueden utilizar para la interpretación de las reivindicaciones y de la memoria descriptiva:

5 Como se usa en la presente memoria, se pretende que los términos "comprende", "que comprende", "incluye", "que incluye", "tiene", "que tiene", "contiene" o "que contiene" o cualquier otra variación de los mismos, cubran una inclusión no exclusiva. Por ejemplo, una composición, una mezcla, proceso, método, artículo o aparato que comprenden una lista de elementos no están necesariamente limitados a sólo esos elementos, sino que pueden incluir otros elementos no enumerados expresamente o inherentes a dicha composición, mezcla, proceso, método, artículo o aparato. Además, a menos que se establezca expresamente lo contrario, "o" se refiere a una o inclusiva y no a una o exclusiva. Por ejemplo, una condición A o B se satisface por cualquiera de los siguientes: A es verdadero (o presente) y B es falso (o no presente), A es falso (o no presente) y B es verdadero (o presente), y A y B son ambos verdaderos (o presentes).

15 Además, se entiende que los artículos indefinidos "un" y "uno" que preceden a un elemento o componente de la invención no son restrictivos con respecto al número de casos (es decir, apariciones) del elemento o componente. Por tanto, "un" o "uno" deben leerse para incluir uno o al menos uno, y la forma de la palabra singular del elemento o componente también incluye el plural a menos que el número indique obviamente que es singular.

20 Las expresiones "invención" o "presente invención" tal como se utilizan en el presente texto son expresiones no limitantes y no pretenden referirse a cualquier realización individual de la invención particular sino que comprende todas las realizaciones posibles como se describen en la memoria descriptiva y en las reivindicaciones.

25 Como se usa en el presente texto, el término "aproximadamente" que modifica la cantidad empleada de un ingrediente o sustancia reaccionante de la invención, se refiere a la variación en la cantidad numérica que puede ocurrir, por ejemplo, a través de procedimientos de mediciones típicas y manejo de líquidos usados para hacer concentrados o usar soluciones en el mundo real; por error inadvertido en estos procedimientos; a través de diferencias en la fabricación, fuente o pureza de los ingredientes empleados para hacer las composiciones o llevar a cabo los métodos; y similares. El término "aproximadamente" abarca también cantidades que difieren debido a diferentes condiciones de equilibrio para una composición que resulta de una mezcla inicial particular. Modificado o no por el término "aproximadamente", las reivindicaciones incluyen equivalentes a las cantidades. En una forma de realización, el término "aproximadamente" significa dentro del 10% del valor numérico indicado, preferiblemente dentro del 5% del valor numérico indicado.

La expresión "sustrato de carbono" o "sustrato de carbono fermentable" se refiere a una fuente de carbono capaz de ser metabolizada por organismos huéspedes de la presente invención y en particular fuentes de carbono elegidas entre el grupo que consiste en monosacáridos, oligosacáridos y polisacáridos.

35 "Gen" se refiere a un fragmento de ácido nucleico que expresa una proteína específica o una molécula de ARN funcional, que puede incluir opcionalmente secuencias reguladoras que preceden (secuencias 5' no codificadoras) y que siguen (secuencias 3' no codificadoras) a la secuencia de codificación. "Gen nativo" o "gen de tipo silvestre" se refiere a un gen tal como se encuentra en la naturaleza con sus propias secuencias reguladoras. "Gen quimérico" se refiere a cualquier gen que no es un gen nativo, que comprende secuencias reguladoras y codificadoras que no se encuentran juntas en la naturaleza. En consecuencia, un gen quimérico puede comprender secuencias reguladoras y secuencias codificadoras que se derivan de fuentes diferentes, o secuencias reguladoras y secuencias codificadoras derivadas de la misma fuente, pero dispuestas de una manera distinta a la que se encuentra en la naturaleza. "Gen endógeno" se refiere a un gen nativo en su localización natural en el genoma de un organismo. Un gen "extraño" se refiere a un gen que no se encuentra normalmente en el organismo huésped, pero que se introduce en el organismo huésped por transferencia génica. Los genes extraños pueden comprender genes nativos insertados en un organismo no nativo, o genes quiméricos. Un "locus" es una región de un genoma que contiene un gen.

50 La expresión "construcción genética" o "constructo genético" se refiere a un fragmento de ácido nucleico que codifica la expresión de una o más proteínas específicas o moléculas de ARN funcionales. En un constructo genética el gen puede ser nativo, quimérico o de naturaleza extraña. Típicamente, una construcción genética comprenderá una "secuencia de codificación". Una "secuencia de codificación" se refiere a una secuencia de ADN que codifica una secuencia de aminoácidos específica.

55 "Promotor" o "regiones de control de la iniciación" se refiere a una secuencia de ADN capaz de controlar la expresión de una secuencia de codificación o ARN funcional. En general, una secuencia de codificación se localiza 3' respecto a una secuencia promotora. Los promotores pueden derivarse en su totalidad de un gen nativo, o estar compuestos de diferentes elementos derivados de diferentes promotores que se encuentran en la naturaleza, o incluso comprenden segmentos de ADN sintéticos. Los expertos en la técnica entienden que diferentes promotores pueden dirigir la expresión de un gen en diferentes tejidos o tipos de células, o en diferentes etapas de desarrollo, o en

respuesta a diferentes condiciones ambientales. Los promotores que hacen que un gen se exprese en la mayoría de los tipos celulares en la mayoría de las veces se denominan por lo general "promotores constitutivos".

5 La expresión "modificación genética" se refiere, no inclusivamente, a cualquier modificación, mutación, delección de bases, adición de bases, modificación de codones, sobreexpresión génica, supresión génica, modificación o sustitución de promotores, adición de genes (bien sea simple o multicopia), expresión o supresión antisentido, o cualquier otro cambio en los elementos genéticos de una célula huésped o cepa bacteriana, tanto si producen un cambio en el fenotipo como si no.

La expresión "célula huésped bacteriana recombinante" se refiere a una célula bacteriana que comprende al menos un gen heterólogo, o constructo genético o fragmento de ácido nucleico.

10 El término "expresión", como se usa en el presente texto, se refiere a la transcripción y a la acumulación estable de ARN de codificación (ARNm) o funcional derivado de un gen. La expresión también puede referirse a la traducción de ARNm en una proteína. "Inhibición antisentido" se refiere a la producción de transcritos de ARN antisentido capaces de suprimir la expresión de la proteína diana. "Sobreexpresión" se refiere a la producción de un producto génico en organismos transgénicos que excede los niveles de producción en organismos normales o no transformados. "Cosupresión" se refiere a la producción de transcritos de ARN sentido o fragmentos capaces de suprimir la expresión de genes extraños idénticos o sustancialmente similares o endógenos (documento US 5.231.020). El término "transformación", como se utiliza en el presente texto, se refiere a la transferencia de un fragmento de ácido nucleico a un organismo huésped, con el resultado de una herencia genéticamente estable. El ácido nucleico transferido puede estar en forma de un plásmido mantenido en la célula huésped, o algo de ácido nucleico transferido puede ser integrado en el genoma de la célula huésped. Los organismos huésped que contienen los fragmentos de ácido nucleico transformados se denominan organismos "transgénicos" o "recombinantes" o "transformados".

25 Los términos "plásmido" y "vector", como se usan en el presente texto, se refieren a un elemento cromosómico extra que a menudo lleva genes que no forman parte del metabolismo central de la célula, y usualmente en forma de moléculas de ADN bicatenarias circulares. Tales elementos pueden ser secuencias de replicación autónoma, secuencias integradoras del genoma, secuencias de fago o nucleótidos, lineales o circulares, de un ADN o ARN de cadena sencilla o doble, derivado de cualquier fuente, en donde cierto número de secuencias de nucleótidos han sido unidas o recombinadas en una construcción única que es capaz de introducir un fragmento promotor y secuencia de ADN para un producto génico seleccionado junto con la secuencia 3' no traducida apropiada, en una célula.

30 La expresión "unido operativamente" se refiere a la asociación de secuencias de ácido nucleico en un único fragmento de ácido nucleico de tal modo que la función de uno es afectada por el otro. Por ejemplo, un promotor está unido operativamente con una secuencia de codificación cuando es capaz de afectar a la expresión de esa secuencia de codificación (es decir, que la secuencia de codificación está bajo el control transcripcional del promotor). Las secuencias de codificación pueden estar operativamente unidas a secuencias reguladoras en orientación de sentido o antisentido.

35 La expresión "marcador seleccionable" significa un factor de identificación, habitualmente un antibiótico o un gen de resistencia química, que es capaz de ser seleccionado basándose en el efecto del gen marcador, es decir, resistencia a un antibiótico, en el que el efecto se utiliza para rastrear la herencia de un ácido nucleico de interés y/o para identificar una célula u organismo que ha heredado el ácido nucleico de interés.

40 Como se usa en el presente texto, la expresión "degeneración de codones" se refiere a la naturaleza en el código genético que permite la variación de la frecuencia de nucleótidos sin afectar la secuencia de aminoácidos de una proteína codificada. El experto en la técnica está al tanto del "sesgo de codones" exhibido por una célula huésped específica en el uso de codones de nucleótidos para especificar un aminoácido dado. Por tanto, cuando se sintetiza un gen para una expresión mejorada en una célula huésped, es deseable diseñar el gen de forma tal que su frecuencia del uso de codones se acerque a la frecuencia del uso preferido del codón de la célula huésped.

45 La expresión "codón optimizado", ya que se refiere a genes o regiones de codificación de moléculas de ácido nucleico para la transformación de varios huéspedes, se refiere a la alteración de codones en el gen o regiones de codificación de las moléculas de ácido nucleico para reflejar el uso típico de codones del organismo huésped sin alterar la proteína codificada por el ADN.

50 El término "lignocelulósico" se refiere a una composición que comprende tanto lignina como celulosa. El material lignocelulósico puede comprender también hemicelulosa.

El término "celulósico" se refiere a una composición que comprende celulosa y componentes adicionales, incluyendo hemicelulosa.

55 El término "sacarificación" se refiere a la producción de azúcares fermentables a partir de polisacáridos.

El término "biomasa pretratada" significa biomasa que ha sido sometida a un pretratamiento físico, químico y/o térmico para aumentar la accesibilidad de los polisacáridos en la biomasa antes de la sacarificación.

5 "Biomasa" se refiere a cualquier material celulósico o lignocelulósico e incluye materiales que comprenden celulosa, y opcionalmente que comprenden además hemicelulosa, lignina, almidón, oligosacáridos y/o monosacáridos. La biomasa también puede comprender componentes adicionales tales como proteínas y/o lípidos. La biomasa puede derivarse de una sola fuente, o la biomasa puede comprender una mezcla derivada de más de una fuente; por ejemplo, la biomasa podría comprender una mezcla de mazorcas de maíz y rastrojo de maíz, o una mezcla de hierba y hojas. La biomasa incluye, pero no se limita a ellos, cultivos de bioenergía, residuos agrícolas, residuos sólidos urbanos, residuos sólidos industriales, lodos de la fabricación de papel, residuos de jardinería, madera y 10 residuos forestales. Los ejemplos de biomasa incluyen, pero no se limitan a ellos, mazorcas de maíz, residuos de cosecha tales como cáscaras de maíz, rastrojos de maíz, hierbas, paja de trigo, paja de cebada, heno, paja de arroz, pasto varilla, papel residual, bagazo de caña de azúcar, sorgo, soja, componentes obtenidos de la molienda de granos, árboles, ramas, raíces, hojas, astillas de madera, serrín, malezas y arbustos, verduras, frutas, flores y estiércol animal.

15 "Hidrolizado de biomasa" se refiere al producto resultante de la sacarificación de la biomasa. La biomasa también puede ser pretratada o preprocesada antes de la sacarificación.

La expresión "ruta metabólica de la xilosa" o "ruta de utilización de xilosa" se refiere a una serie de enzimas (codificadas por genes) que metabolizan la xilosa a través de fructosa-6-fosfato y/o gliceraldehído-6-fosfato e incluyen 1) xilosa isomerasa, que cataliza la conversión de xilosa en xilulosa; 2) xilulocinasa, que fosforila la xilulosa para formar xilulosa 5-fosfato; 3) transcetolasa; y 4) transaldolasa. 20

La expresión "xilosa isomerasa" se refiere a una enzima que cataliza la interconversión de D-xilosa y D-xilulosa. Las xilosa isomerasas (XI) pertenecen al grupo de enzimas clasificadas como EC 5.3.1.5.

La expresión "valor E", como se conoce en la técnica de la bioinformática, es "valor esperado" que proporciona la probabilidad de que ocurra por casualidad una coincidencia. Proporciona la significación estadística de la 25 coincidencia con una secuencia. Cuanto más bajo sea el valor E, tanto más significativo es el acierto.

La expresión "ribosa-5-fosfato isomerasa" o "RPI" se refiere a una enzima que cataliza la interconversión de ribulosa-5-fosfato y ribosa-5-fosfato. Las ribosas-5-fosfato isomerasas pertenecen al grupo de enzimas clasificadas como EC 5.3.1.6.

La expresión "ribulosa fosfato 3-epimerasa" o "RPE" se refiere a una enzima que cataliza la interconversión de D-ribulosa 5-fosfato y D-xilulosa 5-fosfato y se clasifica como EC 5.1.3.1. 30

La expresión "Z. mobilis RPI-A" se refiere a RPI de Z. mobilis que ha sido etiquetada en la técnica como RPI-A. Sin embargo, la proteína RPI de Z. mobilis tiene una identidad de secuencia más próxima a la proteína RPI-B de E. coli (36%) que a la proteína RPI-A de E. coli (20%) y el análisis adicional de RPIs descritos en la solicitud de patente de EE. UU. 13/161734 de propiedad común y en trámite, publicada como WO2012/006061, sitúa RPI de Z. mobilis en el grupo RPI-B. Sin embargo, en el presente texto el RPI de Z. mobilis se denomina RPI-A por coherencia con su nombre conocido públicamente. 35

La expresión "gen pnp" se refiere a un gen que codifica polinucleótido fosforilasa, también denominado polirribonucleótido nucleotidiltransferasa. Esta enzima es bifuncional con una actividad fosforilítica 3' a 5' exorribonucleasa y una actividad terminal 3' de oligonucleótido polimerasa. Está involucrada en el procesamiento y 40 degradación del ARNm y se clasifica como EC 2.7.7.8.

La expresión "aminoácidos nativos" se refiere a aminoácidos cuando se presentan en posiciones de la secuencia peptídica que está codificada por un gen endógeno.

La expresión "aminoácidos no nativos" se refiere a aminoácidos en posiciones que no están codificadas por un gen endógeno.

45 La expresión "secuencia de aminoácidos N-terminal" se refiere a la secuencia de aminoácidos que comienza en el término N de un polipéptido. El primer aminoácido N-terminal se cuenta como "1".

El término "heterólogo" significa que no se encuentra de forma natural en la posición de interés. Por ejemplo, un gen heterólogo se refiere a un gen que no se encuentra de modo natural en el organismo huésped, sino que se introduce en el organismo huésped por transferencia génica. Por ejemplo, una molécula heteróloga de ácido nucleico que está presente en un gen quimérico es una molécula de ácido nucleico que no se encuentra naturalmente asociada con los otros segmentos del gen quimérico, tal como las moléculas de ácido nucleico que tienen la región de codificación y segmentos promotores que no están asociados de forma natural entre sí. 50

Como se usa en el presente texto, una "molécula aislada de ácido nucleico" es un polímero de ARN o ADN que es de una o de dos cadenas, que contiene opcionalmente bases nucleotídicas sintéticas, no naturales o alteradas. Una

molécula de ácido nucleico aislada en forma de un polímero de ADN puede comprender uno o más segmentos de ADNc, ADN genómico o ADN sintético.

Un fragmento de ácido nucleico es "hibridable" con otro fragmento de ácido nucleico, tal como un ADNc, ADN genómico, o ARN, cuando una forma monocatenaria del fragmento de ácido nucleico puede hibridarse con el otro fragmento de ácido nucleico bajo las condiciones apropiadas de temperatura y fuerza iónica de la disolución. Las condiciones de hibridación y lavado son bien conocidas y se ejemplifican en Sambrook, J., Fritsch, E. F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2ª ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989), en particular el Capítulo 11 y la Tabla 11.1 en ese texto. Las condiciones de temperatura y fuerza iónica determinan las condiciones de "rigor" de la hibridación. Las condiciones de rigor pueden ajustarse a cribar fragmentos moderadamente similares (tales como secuencias homólogas de organismos relacionados de lejos), a fragmentos altamente similares (tales como genes que duplican enzimas funcionales de organismos estrechamente relacionados). Los lavados post-hibridación determinan las condiciones de rigor. Un conjunto de condiciones preferidas utiliza una serie de lavados comenzando con 6X SSC, SDS al 0,5% a temperatura ambiente durante 15 min, después se repite con 2X SSC, SDS al 0,5% a 45°C durante 30 min, y a continuación se repite dos veces con 0,2X SSC, SDS al 0,5% a 50°C durante 30 min. Un conjunto de condiciones rigurosas más preferido utiliza temperaturas más altas en las que los lavados son idénticos a los anteriores excepto que la temperatura de los dos lavados finales de 30 minutos en SSC 0,2X, SDS al 0,5% se aumentó a 60°C. Otro conjunto preferido de condiciones altamente rigurosas utiliza dos lavados finales en 0,1X SSC, SDS al 0,1% a 65°C. Un conjunto adicional de condiciones de rigor incluye hibridación a 0,1X SSC, SDS al 0,1%, 65°C y lavados con 2X SSC, SDS al 0,1% seguido por 0,1X SSC, SDS al 0,1%, por ejemplo.

La hibridación requiere que los dos ácidos nucleicos contengan secuencias complementarias, aunque dependiendo del rigor de la hibridación son posibles desajustes entre bases. El rigor apropiado para hibridar ácidos nucleicos depende de la longitud de los ácidos nucleicos y del grado de complementación, variables bien conocidas en la técnica. Cuanto mayor es el grado de similitud u homología entre dos secuencias de nucleótidos, tanto mayor será el valor de Tm para híbridos de ácidos nucleicos que tengan esas secuencias. La estabilidad relativa (correspondiente a Tm más alta) de las hibridaciones de ácidos nucleicos disminuye en el siguiente orden: ARN: ARN, ADN: ARN, ADN: ADN. Para híbridos de más de 100 nucleótidos de longitud, se han deducido ecuaciones para el cálculo de Tm (véase Sambrook et al., supra, 9.50 - 9.51). Para hibridaciones con ácidos nucleicos más cortos, es decir, oligonucleótidos, la posición de desajustes se hace de mayor importancia, y la longitud del oligonucleótido determina su especificidad (véase Sambrook et al., Supra, 11.7 - 11.8). En una forma de realización, la longitud para un ácido nucleico hibridable es al menos aproximadamente 10 nucleótidos. Preferiblemente, una longitud mínima para un ácido nucleico hibridable es al menos aproximadamente 15 nucleótidos; más preferiblemente al menos aproximadamente 20 nucleótidos; y lo más preferiblemente la longitud es al menos aproximadamente 30 nucleótidos. Además, un profesional experto en la técnica reconocerá que la temperatura y la concentración de sal de la solución de lavado pueden ajustarse según sea necesario de acuerdo con factores tales como la longitud de la sonda.

El término "complementario" se usa para describir la relación entre las bases de nucleótidos que son capaces de hibridarse entre sí. Por ejemplo, con respecto al ADN, la adenosina es complementaria a la timina y la citosina es complementaria a la guanina.

La expresión "identidad porcentual", como se conoce en la técnica, es una relación entre dos o más secuencias de polipéptidos o dos o más secuencias de polinucleótidos, como se determina comparando las secuencias. En la técnica, "identidad" también significa el grado de relación de secuencias entre las secuencias de polipéptidos o polinucleótidos, según sea el caso, tal como se determina por la coincidencia entre cadenas de tales secuencias. La "identidad" y la "similitud" se pueden calcular fácilmente por métodos conocidos, incluyendo, pero sin limitarse a ellos, los descritos en: 1.) *Computational Molecular Biology* (Lesk, A. M., Ed.) Oxford University: NY (1988); 2.) *Biocomputing: Informatics and Genome Projects* (Smith, D. W., Ed.) Academic: NY (1993); 3.) *Computer Analysis of Sequence Data*, parte I (Griffin, A. M., y Griffin, H. G., Eds.) Humana: NJ (1994); 4.) *Sequence Analysis in Molecular Biology* (von Heinje, G., Ed.) Academic (1987); y 5.) *Sequence Analysis Primer* (Gribskov, M. y Devereux, J., Eds.) Stockton: NY (1991).

Los métodos preferidos para determinar la identidad se diseñan para dar la mejor coincidencia entre las secuencias ensayadas. Los métodos para determinar la identidad y la similitud se codifican en programas informáticos disponibles públicamente. Las alineaciones de secuencias y los cálculos de porcentaje de identidad pueden realizarse utilizando el programa MegAlign™ de la colección de bioinformática LASERGENE (DNASTAR Inc., Madison, WI). Se lleva a cabo la alineación múltiple de las secuencias utilizando el "Método de alineación Clustal" que abarca diversas variedades del algoritmo incluyendo el "método de alineación Clustal V" correspondiente al método de alineación marcado Clustal V (descrito por Higgins y Sharp, CABIOS, 5: 151 - 153 (1989); Higgins, D.G. et al., *Comput. Appl. Biosci.*, 8: 189 - 191 (1992)) y se encuentran en el programa MegAlign™ de la colección de computación de bioinformática LASERGENE (DNASTAR Inc.). Para alineaciones múltiples, los valores por defecto corresponden a GAP PENALTY (penalización de espacio) = 10 y GAP LENGTH PENALTY (penalización de longitud de espacio) = 10. Los parámetros por defecto para alineaciones por parejas y cálculo de la identidad porcentual de las secuencias de proteínas utilizando el método Clustal son KTUPLE = 1, GAP PENALTY = 3, WINDOW = 5 y DIAGONALS SAVED = 5. Para los ácidos nucleicos estos parámetros son KTUPLE = 2, GAP PENALTY = 5,

WINDOW = 4 y DIAGONALS SAVED = 4. Después de la alineación de las secuencias usando el programa Clustal V, es posible obtener una "identidad porcentual" visualizando la tabla "distancias de secuencia" en el mismo programa. Adicionalmente el "método de alineación Clustal W" está disponible y corresponde al método de alineación marcado Clustal W (descrito por Higgins y Sharp, CABIOS 5: 151 - 153 (1989); Higgins, D. G. et al., Comput. Appl. Biosci. 8: 189 - 191 (1992)) y encontrado en el programa MegAlign™ v6.1 de la colección de computación de bioinformática LASERGENE (DNASTAR Inc.). Los parámetros por defecto para la alineación múltiple (GAP PENALTY = 10, GAP LENGHT PENALTY = 0,2, Delay Divergen Seqs (%) = 30, peso de transición de ADN = 0,5, Matriz de Peso de Proteína = Serie Gonnet, Matriz de Peso de ADN = IUB). Después de la alineación de las secuencias usando el programa Clustal W, es posible obtener una "identidad porcentual" visualizando la tabla de "distancias de secuencia" en el mismo programa.

Un experto en la técnica entiende bien que hay muchos niveles de identidad de secuencia que son útiles en la identificación de polipéptidos, de otras especies, en donde tales polipéptidos tienen la misma o similar función o actividad. Entre los ejemplos útiles del porcentaje de identidades se pueden incluir, pero sin limitarse a ellos: 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, o 95%, o cualquier porcentaje entero de 25% a 100% puede ser útil en la descripción de la presente invención, tal como 25%, 26%, 27%, 28%, 29%, 30%, 31%, 32%, 33%, 34%, 35%, 36%, 37%, 38%, 39%, 40%, 41%, 42%, 43%, 44%, 45%, 46%, 47%, 48%, 49%, 50%, 51%, 52%, 53%, 54%, 55%, 56%, 57%, 58%, 59%, 60%, 61%, 62%, 63%, 64%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% o 99%. Los fragmentos de ácido nucleico adecuados no sólo tienen las homologías anteriores, sino que típicamente codifican un polipéptido que tiene al menos 50 aminoácidos, preferiblemente al menos 100 aminoácidos, y más preferiblemente al menos 150 aminoácidos.

La expresión "software de análisis de secuencias" se refiere a cualquier algoritmo de ordenador o programa de software que sea útil para el análisis de secuencias de nucleótidos o de aminoácidos. El "software de análisis de secuencias" puede estar disponible comercialmente o ser desarrollado independientemente. El software de análisis de secuencias típico incluirá, pero no sin limitarse a ello: 1.) la colección GCG de programas (Wisconsin Package Version 9.0, Genetics Computer Group (GCG), Madison, WI); 2.) BLASTP, BLASTN, BLASTX (Altschul et al., J. Mol. Biol., 215: 403 - 410 (1990)); 3.) DNASTAR (DNASTAR, Inc. Madison, WI); 4.) Sequencher (Gene Codes Corporation, Ann Arbor, MI); y 5.) el programa FASTA que incorpora el algoritmo Smith-Waterman (W.R. Pearson, Comput. Methods Genome Res., [Proc. Int. Symp.] (1994), Meeting Date 1992, 111-20 Editor(es): Suhai, Sandor. Plenum: Nueva York, NY). Dentro del contexto de esta solicitud se entenderá que cuando se utiliza el software de análisis de secuencia para el análisis, que los resultados del análisis se basarán en los "valores por defecto" del programa referenciado, a menos que se especifique lo contrario. Como se usa aquí, "valores por defecto" significará cualquier conjunto de valores o parámetros que originalmente se cargan con el software cuando se inicializa por primera vez.

Las técnicas estándar de ADN recombinante y de clonación molecular usadas en la presente invención son bien conocidas en la técnica y son descritas por Sambrook, J. y Russell, D., Molecular Cloning: A Laboratory Manual, tercera edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (2001); y por Silhavy, T. J., Bannan, M. L. y Enquist, L. W., Experiments with Gene Fusions, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1984); y por Ausubel, F. M. et. al., Short Protocols in Molecular Biology, 5ª Ed. Current Protocols, John Wiley and Sons, Inc., N.Y., 2002.

La presente invención se refiere a células manipuladas genéticamente de bacterias que utilizan xilosa, y en particular *Zymomonas* o *Zymobacter* que han mejorado la utilización de xilosa cuando se fermentan en medios que contienen xilosa. Un reto para mejorar la producción de etanol por fermentación de un biocatalizador en medios que incluyen hidrolizado de biomasa, producido típicamente por pretratamiento y sacarificación de biomasa, es la obtención de la utilización óptima de la xilosa. La xilosa es uno de los azúcares de pentosa predominantes en la biomasa lignocelulósica hidrolizada, siendo el otro la arabinosa. Los solicitantes han descubierto que la modificación del gen pnp endógeno, en combinación con el aumento de la expresión de ribosa-5-fosfato isomerasa y en presencia de actividad de xilosa isomerasa no limitante, en células que utilizan xilosa conduce a un aumento en la eficiencia de la utilización de la xilosa y, por tanto, a mayores rendimientos de etanol cuando la fermentación es en un medio que contiene xilosa.

Modificación del gen pnp endógeno

La presente invención se dirige a células de *Zymomonas* o *Zymobacter* que utilizan xilosa modificada genéticamente que tienen un gen pnp endógeno modificado. La región de codificación del gen pnp endógeno codifica una proteína con actividad de polinucleótido fosforilasa. La proteína codificada también se llama polirribonucleótido nucleotidiltransferasa. Se encontró aquí que modificaciones en la región de codificación del gen pnp endógeno mejoran la utilización de xilosa en células de *Zymomonas* manipuladas por ingeniería genética que se diseñan adicionalmente como se describe más adelante.

Cualquier gen de *Zymomonas* o *Zymobacter* que se identifica como codificador de una proteína con actividad de polinucleótido fosforilasa o de polirribonucleótido nucleotidiltransferasa, puede proporcionar el gen endógeno pnp diana para la modificación como se describe en el presente texto. La región de codificación de pnp de la cepa ZM4

de *Zymomonas mobilis* tiene la secuencia de SEQ ID NO: 1. Las regiones de codificación de pnp endógeno conocidas procedentes de otras cepas de *Zymomonas* tienen secuencias con identidades con la SEQ ID NO: 1 del 99% (*Z. mobilis* NCIMB 11163; SEQ ID NO: 3), 98% (*Z. mobilis* ATCC 10988; SEQ ID NO: 5) y 83% (*Z. mobilis pomaceae* ATCC 29192; SEQ ID NO: 7). Cualquiera de estas secuencias, o cualquier secuencia con al menos

5 aproximadamente 95%, 96%, 97%, 98% o 99% de identidad con una cualquiera de estas secuencias e identificada como codificadora de un polinucleótido fosforilasa o polirribonucleótido nucleotidiltransferasa se puede usar como diana para la modificación como se describe más adelante. Las secuencias del gen pnp endógeno diana adicionales se pueden identificar usando el análisis BLAST u otros análisis de comparación de secuencias que son bien conocidos por los expertos en la técnica.

10 En las células presentes, la región de codificación de pnp se modifica para acortar la región de codificación en el extremo 3', dando como resultado la expresión de una proteína truncada C-terminal, en comparación con la proteína codificada de forma natural. La polinucleótido fosforilasa codificada nativa de *Zymomonas mobilis* es una proteína de aproximadamente 748 aminoácidos, que es cualquiera de las SEQ ID NO: 2, 4, 6, 8 o cualquier secuencia con al menos 95%, 96%, 97%, 98%, o 99% de identidad respecto a cualquiera de estas secuencias e identificada como

15 polinucleótido fosforilasa o polirribonucleótido nucleotidiltransferasa. En una realización, la proteína truncada expresada a partir de la región de codificación de pnp modificada conserva al menos aproximadamente 350 aminoácidos de la secuencia de aminoácidos N terminal codificada por el gen endógeno que codifica la polinucleótido fosforilasa, que son aminoácidos N terminales nativos. La proteína truncada conserva al menos

20 540, 550, 560, 570, 580, 590, 600, 610, 620, 630, 640, 650, 660, 670, 680, 690, 700 o 710 de los aminoácidos N-terminales nativos.

En una realización, la modificación genética a la región de codificación de pnp endógeno añade una secuencia de codificación para aminoácidos no nativos adyacentes a, y en marco con, la región de codificación nativa truncada de forma que se produce una proteína de fusión. Por ejemplo, la región de codificación para adicionalmente entre

25 aproximadamente 1 y aproximadamente 20 aminoácidos, que no están codificados por el gen endógeno, se puede añadir adyacente a, y en marco con, la región de codificación nativa truncada produciendo una proteína de fusión que tiene hasta aproximadamente 20 aminoácidos no nativos en el término C. La proteína truncada C-terminal es entonces parte de la proteína de fusión. En diversas realizaciones pueden haber 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, o más aminoácidos no nativos en el extremo C-terminal. Algunos ejemplos no

30 limitantes de proteínas de fusión que pueden resultar de la modificación genética de un gen endógeno que codifica polinucleótido fosforilasa incluyen proteínas que: 1) conservan 709 aminoácidos N-terminales nativos e incluyen 14 aminoácidos C-terminales adicionales (SEQ ID NO: 9); 2) conservan 695 aminoácidos N-terminales nativos e incluyen 2 aminoácidos C-terminales adicionales (SEQ ID NO: 10); 3) conservan 368 aminoácidos N-terminales nativos e incluyen 10 aminoácidos C-terminales adicionales (SEQ ID NO: 11); y 4) conservan 32 aminoácidos N-

35 terminal nativos e incluyen 17 aminoácidos C-terminales adicionales (SEQ ID NO: 12).

La región de codificación de pnp endógeno puede modificarse como se ha descrito antes por cualquier método conocido por un experto en la técnica. Típicamente, la región de codificación es señalada por secuencias de ADN de dirección de recombinación que son porciones de la región de codificación de pnp y puede incluir ADN genómico adyacente circundante. Las secuencias de ADN de dirección de recombinación dirigen la inserción directa de

40 secuencias de ADN unidas por ellas en el gen pnp endógeno por recombinación homóloga. En una realización la secuencia de ADN unida incluye una secuencia de codificación para hasta aproximadamente 20 aminoácidos que está diseñada para estar en marco de lectura con una posición en la región de codificación de pnp nativo, como se ha descrito anteriormente, después de la integración por recombinación homóloga. Alternativamente, toda la región de codificación de pnp nativa puede ser reemplazada, usando recombinación homóloga, con una región de

45 codificación diseñada para producir una proteína con un truncamiento C-terminal como se ha descrito anteriormente. Además, la región de codificación de sustitución puede codificar aminoácidos no nativos adicionales en el término C de la proteína como se ha descrito anteriormente, dando como resultado la producción de una proteína de fusión.

Zymomonas o *Zymobacter* que utilizan xilosa de manipulación genética

El gen pnp endógeno se modifica en una célula de *Zymomonas* o *Zymobacter* que contiene una ruta metabólica de

50 xilosa que confiere la capacidad de utilizar xilosa para la producción de etanol. *Zymobacter palmae* es una bacteria productora de etanol que ha sido manipulada genéticamente para la utilización de xilosa expresando genes para la utilización de xilosa como se describe a continuación para *Zymomonas*, usando gliceraldehído-3-fosfato deshidrogenasa de *Z. mobilis* y promotores de enolasa (Yanase et al. Applied and Environmental Microbiology (2007) 73: 2592 - 2599).

Las cepas de *Zymomonas*, tales como *Z. mobilis*, han sido manipuladas para la fermentación de xilosa para dar etanol. Típicamente se han introducido cuatro genes en *Z. mobilis* para la expresión de cuatro enzimas implicadas en el metabolismo de la xilosa formando una ruta metabólica de xilosa en la célula (Figura 1 en negrita) como se describe en la patente de EE.UU. n.º 5.514.583, patente de EE. UU. n.º 5.712.133, patente de EE. UU. n.º 6.566.107, documento WO 95/28476, Feldmann et al. ((1992) Appl. Microbiol. Biotechnol. 38: 354 - 361), y Zhang et al. ((1995) Science 267: 240 - 243). Estos incluyen genes que codifican xilosa isomerasa que cataliza la conversión de xilosa en xilulosa, y xilulocinasa que fosforila la xilulosa para formar xilulosa 5-fosfato. Adicionalmente se expresan

55 Las cepas de *Zymomonas*, tales como *Z. mobilis*, han sido manipuladas para la fermentación de xilosa para dar etanol. Típicamente se han introducido cuatro genes en *Z. mobilis* para la expresión de cuatro enzimas implicadas en el metabolismo de la xilosa formando una ruta metabólica de xilosa en la célula (Figura 1 en negrita) como se describe en la patente de EE.UU. n.º 5.514.583, patente de EE. UU. n.º 5.712.133, patente de EE. UU. n.º 6.566.107, documento WO 95/28476, Feldmann et al. ((1992) Appl. Microbiol. Biotechnol. 38: 354 - 361), y Zhang et al. ((1995) Science 267: 240 - 243). Estos incluyen genes que codifican xilosa isomerasa que cataliza la conversión de xilosa en xilulosa, y xilulocinasa que fosforila la xilulosa para formar xilulosa 5-fosfato. Adicionalmente se expresan

60 Science 267: 240 - 243). Estos incluyen genes que codifican xilosa isomerasa que cataliza la conversión de xilosa en xilulosa, y xilulocinasa que fosforila la xilulosa para formar xilulosa 5-fosfato. Adicionalmente se expresan

transcetolasa y transaldolasa, dos enzimas de la ruta de pentosa fosfato que convierten xilulosa 5-fosfato en productos intermedios que acoplan el metabolismo de la pentosa a la ruta glicolítica de Entner-Doudoroff permitiendo el metabolismo de xilosa a etanol (véase la Figura 1). Las secuencias de ADN que codifican estas enzimas pueden obtenerse a partir de cualquiera de los numerosos microorganismos que pueden metabolizar la xilosa, tales como bacterias entéricas, y algunas levaduras y hongos. Las fuentes para las regiones de codificación pueden incluir *Xanthomonas*, *Klebsiella*, *Escherichia*, *Rhodobacter*, *Flavobacterium*, *Acetobacter*, *Gluconobacter*, *Rhizobium*, *Agrobacterium*, *Salmonella*, *Pseudomonads* y *Zymomonas*. Se utilizan normalmente las regiones de codificación de *E. coli*.

Los genes endógenos pueden proporcionar parte de una ruta de fermentación de xilosa, o pueden alterarse por cualquier técnica de manipulación genética conocida para proporcionar una proteína con actividad enzimática útil para el metabolismo de la xilosa. Por ejemplo, la transcetolasa endógena puede complementar otras actividades enzimáticas introducidas en la creación de una ruta de utilización de xilosa.

Las cepas de *Zymomonas* o *Zymobacter* que son diseñadas adicionalmente para utilizar otros azúcares que, como la xilosa, no son sustratos naturales, también se pueden usar en el presente procedimiento. Un ejemplo es una cepa de *Z. mobilis* diseñada para la utilización de arabinosa y se describe en la patente de EE.UU. n° 5.843.760. Las cepas pueden modificarse de otras formas adicionales para mejorar la utilización de xilosa y la producción de etanol.

Aumento de la expresión de RPI

Las presentes células se diseñan adicionalmente para el aumento de la expresión de la actividad de ribosa-5-fosfato isomerasa (RPI). La RPI cataliza la interconversión de ribulosa-5-fosfato y ribosa-5-fosfato (véase la Figura 1). El aumento de la expresión de RPI se describe en la solicitud de patente de EE. UU. 13/161734, de titularidad común y co-pendiente, publicada como WO2012/006061, que describe el aumento de expresión de RPI que confiere una mayor eficacia de la utilización de xilosa asociada con la reducción de producción de ribulosa.

Se puede lograr una mayor expresión de RPI usando cualquier proteína o polipéptido con actividad de ribosa-5-fosfato isomerasa en *Zymomonas* como se describe en la solicitud de patente de EE. UU. 13/161734, publicada como WO2012/006061. Los polipéptidos que tienen actividad de ribosa-5-fosfato isomerasa tienen la clasificación EC 5.3.1.6. Hay dos grupos de enzimas ribosa-5-fosfato isomerasa que se denominan RPI-A y RPI-B. Las enzimas RPI-B pertenecen a la familia RpiB/LacA/LAcB de azúcar-fosfato isomerasas. *E. coli* tiene ambos tipos de proteínas RPI que se anota como RPI-A. Sin embargo, la proteína RPI de *Z. mobilis* tiene una identidad de secuencia más cercana a la de la proteína RPI-B de *E. coli* (36%) que la proteína RPI-A de *E. coli* (20%). Otro análisis de RPIs descritos en la solicitud de patente de EE.UU. 13/161734, publicada como WO2012/006061, puso la RPI de *Z. mobilis* en el grupo RPI-B.

Las secuencias de proteínas RPI que pueden usarse en las presentes células son muy diversas como se ejemplifica por las proteínas RPI de *Z. mobilis* y *E. coli* (SEQ ID NO: 14 y 16, respectivamente; regiones de codificación SEQ ID NO: 13 y 15, respectivamente). Las proteínas RPI que pueden usarse en los presentes microorganismos se pueden identificar usando análisis bioinformático. La identificación utilizando un análisis bioinformático de estructura/función basado en Modelo de Markov de Perfil Oculto (usando el algoritmo hmmsearch del paquete de software HMMER; Janelia Farm Research Campus, Ashburn, VA), la identificación de residuos de sitio activo y el cribado de aminoácidos de identificación adicional se describen en el Ejemplo 8 de la solicitud de patente de EE.UU. USSN 13/161734, publicada como WO2012/006061.

Se describen ejemplos de proteínas RPI-A y RPI-B que se adaptan a estos criterios y que pueden usarse en los presentes microorganismos, en la solicitud de patente de EE.UU. USSN 13/161734, publicada como US2012 0156746A1.

Los RPI adicionales pueden identificarse fácilmente en la bibliografía y en bases de datos de bioinformática, como es bien conocido por un profesional experto en la materia y como se ha descrito anteriormente. La identificación de proteínas y/o secuencias de codificación utilizando bioinformática es típicamente a través de BLAST (descrito anteriormente) buscando bases de datos disponibles para el público con secuencias de aminoácidos de RPI o secuencias de codificación, tales como las que se proporcionan en el presente texto. Las identidades se basan en el método Clustal W de alineación usando los parámetros por defecto de GAP PENALTY = 10, GAP LENGTH PENALTY = 0,1 y la serie Gonnet 250 de matriz de peso de proteína.

En la presente célula, se hace una modificación genética que incrementa la actividad de ribosa-5-fosfato isomerasa en comparación con la actividad de ribosa-5-fosfato isomerasa de la célula que carece de la modificación genética. El aumento de la expresión de la actividad de RPI puede obtenerse expresando una molécula de ADN aislada que codifica una proteína que tiene actividad de ribosa-5-fosfato isomerasa que es activa en la célula huésped. Las proteínas útiles con actividad de ribosa-5-fosfato isomerasa pertenecen a la clasificación EC 5.3.1.6 e incluyen las proteínas ribosa-5-fosfato isomerasa A y ribosa-5-fosfato isomerasa B que se describen anteriormente.

Cualquier método para aumentar la actividad de una enzima en una célula puede usarse para aumentar la actividad de RPI. Tales métodos son bien conocidos por los profesionales expertos en la técnica e incluyen el aumento del número de copias del gen de codificación y/o expresión por un gen que contiene un promotor de expresión elevada.

Las cepas presentes pueden diseñarse para el aumento de la expresión de un región de codificación de RPI endógena, y/o expresión de una región de codificación de RPI heteróloga introducida para dar un aumento de actividad enzimática. Además, la actividad de RPI puede aumentarse por mutación y cribado de genes mutados expresados para identificar las células que tienen mayor actividad.

- 5 Típicamente, el aumento de expresión de RPI se logra por transformación con una molécula de ADN que codifica RPI que está operativamente unida a un promotor en un gen u operón quimérico. Las secuencias de codificación para RPIs que se pueden usar incluyen cualquier secuencia que codifican las proteínas RPI-A y RPI-B descritas anteriormente.

10 Cuando se usa una región de codificación heteróloga, la secuencia puede ser de codón optimizado para una máxima expresión en la célula huésped diana, como es bien conocido por los expertos en la técnica. Si el codón de iniciación nativo es GTG, puede cambiarse a ATG para una mayor expresión de proteína. Los métodos para la expresión génica en bacterias son bien conocidos en la técnica. La expresión de genes en bacterias requiere típicamente un promotor, unido operativamente a una región de codificación de interés, y un terminador transcripcional. Los promotores que pueden usarse son promotores que se expresan en células de *Zymomonas* o *Zymobacter* tales como los promotores del gen de gliceraldehído-3-fosfato deshidrogenasa (promotor GAP, P_{gap}) de *Z. mobilis*, del gen de enolasa de *Z. mobilis* (promotor ENO; P_{eno}), y del gen que codifica la xilosa isomerasa de *Actinoplanes missouriensis* (promotor GI, P_{gi}). Los promotores de expresión particularmente elevada que pueden usarse son los promotores de P_{gap} con mutaciones que causan expresión elevada como se describe en la patente de EE.UU. 7.998.722.

20 Un gen quimérico o un operón para la expresión de RPI se construye típicamente en, o se transfiere a, un vector para otras manipulaciones. Los vectores son bien conocidos en la técnica. Ciertos vectores son capaces de replicarse en una amplia gama de bacterias huésped y se pueden transferir por conjugación. La secuencia completa y anotada de pRK404 y tres vectores relacionados: pRK437, pRK442, y pRK442 (H) están disponibles. Estos derivados han demostrado ser herramientas valiosas para manipulación genética en bacterias gram-negativas (Scott et al., Plasmid 50 (1): 74 - 79 (2003)).

25 Son particularmente útiles para la expresión en *Zymomonas* los vectores que pueden replicarse tanto en *E. coli* como en *Zymomonas*, tal como pZB188 que se describe en la patente de EE.UU. 5.514.583.

30 Los vectores pueden incluir plásmidos para la replicación autónoma en una célula, y plásmidos para llevar los constructos que se han de integrar en genomas bacterianos. Los plásmidos para la integración de ADN pueden incluir transposones, regiones de secuencia de ácido nucleico homóloga respecto al genoma bacteriano diana, u otras secuencias que soportan la integración. Un tipo de vector adicional puede ser un transposoma producido usando, por ejemplo, un sistema que está disponible comercialmente de EPICENTRE®. Es bien sabido cómo elegir un vector apropiado para el huésped diana deseado y la función deseada.

35 Las células bacterianas pueden ser manipuladas genéticamente introduciendo un vector que tiene un gen quimérico que comprende una región de codificación de RPI por métodos bien conocidos, tales como usar la transformación congelación-descongelación, la transformación mediada por calcio, la electroporación, o la conjugación. Cualquier célula bacteriana a manipular para mejorar la utilización de xilosa aumentando la expresión de una enzima RPI es una célula huésped diana para transformación para diseñar una cepa tal como se describe en la presente memoria. Son células huésped particularmente adecuadas las células de *Zymomonas* y *Zymobacter*. El gen quimérico introducido puede mantenerse en la célula en un plásmido de replicación estable, o integrarse en el genoma después de la introducción.

40 Para construir una cepa con un gen quimérico o un operón de RPI integrado en el genoma de la célula bacteriana, pueden ser utilizados métodos que son bien conocidos en la técnica tales como recombinación homóloga, inserción de transposones o inserción de transposoma. En la recombinación homóloga, las secuencias de ADN que flanquean un sitio de integración diana se ponen uniendo un gen de resistencia a la espectinomicina, u otro marcador seleccionable, y el gen quimérico RPI que da lugar a la inserción del marcador seleccionable y el gen quimérico RPI en el sitio genómico diana. Además, el marcador seleccionable puede limitarse por sitios de recombinación específicos del sitio, de modo que después de la expresión de la correspondiente recombinasa específica del sitio, el gen de resistencia es escindido del genoma.

45 Además, el promotor del gen que expresa RPI endógeno puede ser reemplazado por un gen expresado de forma más elevada para aumentar la actividad de RPI en la célula. Esto puede conseguirse mediante recombinación homóloga usando vectores y métodos que se han descrito anteriormente.

Actividad de xilosa isomerasa

55 Las presentes células tienen un nivel de actividad de xilosa isomerasa que no es limitante. La actividad de xilosa isomerasa (XI) es no limitante cuando no es la etapa limitante de la velocidad para la función de las rutas de utilización de xilosa y de producción de etanol como se muestra en la Figura 1. La actividad de xilosa isomerasa no es limitante cuando un aumento en la actividad de xilosa isomerasa no mejora la utilización de xilosa y la producción de etanol. Esta situación indica que uno o más pasos de la ruta son limitantes. Como se encuentra aquí, la actividad

de xilosa isomerasa en células ZW658, medida usando un ensayo que incluye xilosa, NADH, MgSO₄, trietanolamina y sorbitol deshidrogenasa que se describe en el Ejemplo 11 del presente texto, es de aproximadamente 0,25 µmoles de producto/mg de proteína/minuto. Este nivel de actividad de xilosa isomerasa no es limitante en la cepa ZW658 como se describe en Ejemplo 3 del presente texto. La cepa ZW658 es un precursor de la cepa ZW801-4 utilizada en el presente texto como se describe la sección de construcción de la cepa de *Zymomonas mobilis*, en el presente texto. En una realización, la actividad de XI no limitante es mayor que aproximadamente 0,25 µmoles de producto/mg de proteína/minuto, medida en un extracto libre de células usando el ensayo descrito en el Ejemplo 11 en el presente texto. La actividad de XI no limitante puede ser mayor que aproximadamente 0,25, 0,3, 0,35, 0,4, 0,45, o 0,5 µmoles de producto/mg de proteína/minuto.

10 Cuando la actividad de XI es limitante en una célula, la célula puede ser manipulada genéticamente para tener actividad de XI no limitante aumentando el nivel de expresión de una enzima XI o introduciendo la expresión de una enzima XI más altamente activa. El aumento del nivel de expresión puede ser por cualquier método conocido por un experto en la técnica tal como el aumento del número de copias de un gen que codifica XI o utilizando un promotor más altamente activo para expresar una enzima XI.

15 Por ejemplo, se puede expresar una región de codificación de XI utilizando un promotor mutado que tiene un nivel de expresión aumentado en comparación con el promotor no mutado. Un ejemplo de un promotor de alta expresión mutado es el promotor mutado del gen de gliceraldehído-3-fosfato deshidrogenasa de *Z. mobilis* que se describe en el documento US 7.989.206, denominado en el presente texto promotor Super GAP (P_{gapS}).

20 Las enzimas xilosa isomerasa que pueden proporcionar alta actividad por expresión en múltiples copias o a partir de un promotor de alta expresión mutado son cualquiera de las descritas en el documento US 7.998.722. Como se describe allí, las enzimas xilosa isomerasas pertenecen a EC 5.3.1.5 y pueden identificarse utilizando un perfil HMM (descrito antes para RPI) y además cuatro aminoácido del sitio catalítico que se ha encontrado que son característicos de las xilosa isomerasas.

25 Alternativamente, o además, se puede obtener una actividad de XI elevada expresando una enzima XI que tiene actividad más alta en las células de *Zymomonas* que en la XI de *E. coli* utilizada comúnmente. La publicación de solicitud de patente de los EE. UU. US20110318801 describe que las enzimas de xilosa isomerasa identificadas por EC 5.3.1.5 y que pertenecen al Grupo I tienen mayor actividad que la XI de *E. coli*, que pertenece al Grupo II. Además, una XI del Grupo I puede expresarse a partir de un promotor de la expresión elevada mutado, como se describió anteriormente, para obtener una alta actividad de XI en una célula huésped.

30 Las xilosa isomerasas del Grupo I, como se describe en la publicación de solicitud de patente de EE.UU. US20110318801, se refieren a una xilosa Isomerasa que pertenece al Grupo I como se define por al menos uno de los criterios siguientes: a) cae dentro de un agrupamiento de identidad de secuencia de umbral 50% que incluye la XI de *A. missouriensis* que se prepara utilizando análisis de bioinformática filogenético molecular como en el Ejemplo 4 de la publicación de solicitud de patente de EE.UU. US20110318801; b) se ajusta sustancialmente a los aminoácidos para el Grupo I en las posiciones que determinan la especificidad (SDP) identificados usando el análisis GroupSim de los conjuntos del Grupo I y II determinados a partir del análisis filogenético molecular en el Ejemplo 4 de la publicación de solicitud de patente de EE. UU. US20110318801; y/o c) tiene un valor E de 1 E-15 o menos cuando se inquiriere usando un modelo oculto de Markov de perfil preparado como se describe en el Ejemplo 4 de la publicación de solicitud de patente de EE.UU. US20110318801. Por ejemplo la XI de *Actinoplanes missouriensis* (por ejemplo, SEQ ID NO: 18) fue identificada como perteneciente al Grupo I y cuando se expresa en células de *Zymomonas mobilis* proporciona una mayor actividad que la que expresó similarmente XI de *E. coli*.

Los genes quiméricos, vectores, transformación, integración, optimización de codones y expresión de XI son como se describen para RPI anteriormente y son bien conocidos por un experto en la técnica.

Aumento de la expresión de RPE

45 En una forma de realización, las células presentes se diseñan adicionalmente para incrementar la expresión de la actividad de la ribulosa-fosfato 3-epimerasa (RPE). La RPE cataliza la interconversión de D-ribulosa 5-fosfato y D-xilulosa 5-fosfato (véase la Figura 1) y se clasifica como EC 5.1.3.1. Se realiza al menos una modificación genética que aumenta la actividad de RPE en la célula en comparación con la actividad de RPE en la célula que carece de la modificación genética. Las modificaciones para aumentar la expresión son las descritas anteriormente para RPI, y pueden usar cualquier enzima que pertenezca a EC 5.1.3.1 con actividad de RPE. Por ejemplo, pueden expresarse en la célula copias adicionales de la RPE de *Z. mobilis* (SEQ ID NO: 20; secuencia de codificación SEQ ID NO: 19), o se puede expresar una región codificadora de RPE a partir de un promotor de expresión elevada.

Fermentación de una cepa mejorada que utiliza xilosa

55 Las presentes células *Zymomonas* o *Zymobacter* que utilizan xilosa manipulada genéticamente pueden usarse en la fermentación para producir etanol. Como ejemplo, se describe la producción de etanol por una cepa de *Z. mobilis* de la invención.

Para la producción de etanol, la *Z. mobilis* que utiliza xilosa recombinante que tiene actividad de RPI incrementada, actividad de XI no limitante y modificación del gen *pnp* endógeno, se pone en contacto con medio que contiene xilosa. La xilosa puede ser el único azúcar, pero típicamente el medio contiene una mezcla de azúcares incluyendo xilosa y glucosa. El medio puede contener hidrolizado de biomasa que incluye estos azúcares que se derivan de la biomasa celulósica o lignocelulósica tratada.

Cuando la concentración de azúcares mixtos es tan alta que se inhibe el crecimiento, el medio incluye sorbitol, manitol, o una mezcla de los mismos como se describe en el documento US 7.629.156 de propiedad común. El galactitol o el ribitol pueden reemplazar o combinarse con sorbitol o manitol. Las células *Z. mobilis* crecen en el medio en el que tiene lugar la fermentación y se produce etanol. La fermentación se conduce sin aire suplementario, oxígeno u otros gases (que pueden incluir condiciones tales como fermentación anaerobia, microaerobia o microaerofílica), durante al menos aproximadamente 24 horas, y puede funcionar durante 30 o más horas. El tiempo para alcanzar la máxima producción de etanol es variable, dependiendo de las condiciones de fermentación. Típicamente, si están presentes inhibidores en el medio, se requiere un periodo de fermentación más largo. Las fermentaciones pueden llevarse a cabo a temperaturas que están entre aproximadamente 30°C y aproximadamente 37°C, a un pH de aproximadamente 4,5 a aproximadamente 7,5.

Las presentes células de *Z. mobilis* pueden crecer en medio que contiene azúcares mixtos incluyendo xilosa en fermentadores a escala de laboratorio, y en fermentación a mayor escala en donde se producen cantidades comerciales de etanol. Cuando se desea una producción comercial de etanol, se pueden aplicar diversas metodologías de cultivo. Por ejemplo, la producción a gran escala a partir de las presentes células de *Z. mobilis* puede producirse mediante metodologías de cultivo por lotes o tandas y cultivo continuo. Un método clásico de cultivo por tandas es un sistema cerrado en el que la composición del medio se ajusta al principio del cultivo y no se somete a alteraciones artificiales durante el proceso de cultivo. Así pues, al comienzo del proceso de cultivo el medio se inocula con el organismo deseado y se permite que se produzca el crecimiento o la actividad metabólica sin añadir nada al sistema. Típicamente, sin embargo, un cultivo "discontinuo o por tandas" es por tandas con respecto a la adición de fuente de carbono y a menudo se hacen intentos en factores de control tales como el pH y la concentración de oxígeno. En sistemas por tandas, las composiciones de metabolito y biomasa del sistema cambian constantemente hasta el momento en que se termina el cultivo. Dentro de los cultivos por tandas, las células se moderan a través de una fase estática de retardo hasta una fase logarítmica de alto crecimiento y finalmente a una fase estacionaria en la que la velocidad de crecimiento disminuye o se detiene. Si no se tratan, las células en la fase estacionaria morirán finalmente. Las células en fase logarítmica son a menudo responsables de la mayor parte de la producción de producto final o intermedio en algunos sistemas. Se puede obtener en otros sistemas la producción en fase estacionaria o post-exponencial.

Una variación del sistema de lotes estándar es el sistema *Fed-Batch* o de lote alimentado. Los procedimientos de cultivo de lote alimentado son también adecuados para el crecimiento de las presentes células *Z. mobilis* y comprenden un sistema de lotes típico con la excepción de que el sustrato se añade en incrementos a medida que el cultivo progresa. Los sistemas de lote alimentado son útiles cuando la represión de los catabolitos es apta para inhibir el metabolismo de las células y donde es deseable tener cantidades limitadas de sustrato en el medio. La medida de la concentración de sustrato real en sistemas de lote alimentado es difícil y por lo tanto se estima basándose en los cambios de factores medibles tales como el pH y la presión parcial de gases residuales tales como el CO₂. Los métodos de cultivo por lotes y de lote alimentado son habituales y bien conocidos en la técnica y se pueden encontrar ejemplos en *Biotechnology: A Textbook of Industrial Microbiology*, Crueger, Crueger y Brock, segunda edición (1989) Sinauer Associates, Inc., Sunderland, MA, o Deshpande, Mukund V., *Appl. Biochem. Biotechnol.*, 36, 227, (1992).

La producción comercial de etanol puede realizarse también con un cultivo continuo. Los cultivos continuos son sistemas abiertos en los que continuamente se añade a un biorreactor un medio de cultivo definido, y simultáneamente se elimina una cantidad igual de medio acondicionado para su procesamiento. Los cultivos continuos generalmente mantienen las células a una densidad de fase líquida constante alta, en la que las células están principalmente en crecimiento en fase logarítmica. Alternativamente, se puede practicar el cultivo continuo con células inmovilizadas en las que se añaden continuamente carbono y nutrientes, y los productos valiosos, los subproductos o los productos residuales se retiran continuamente de la masa celular. La inmovilización de las células puede realizarse utilizando una amplia gama de soportes sólidos compuestos de materiales naturales y/o sintéticos, como conoce un experto en la técnica.

El cultivo continuo o semicontinuo permite la modulación de un factor o cualquier número de factores que afectan al crecimiento celular o a la concentración final del producto. Por ejemplo, un método mantendrá un nutriente limitante tal como la fuente de carbono o el nivel de nitrógeno a una tasa fija y permitirá moderar todos los demás parámetros. En otros sistemas, una serie de factores que afectan al crecimiento pueden ser alterados continuamente mientras que la concentración celular, medida por la turbidez del medio, se mantiene constante. Los sistemas continuos se esfuerzan por mantener las condiciones de crecimiento en estado estacionario y por ello la pérdida de células debida al medio que se saca debe equilibrarse con la velocidad de crecimiento celular en el cultivo. Los métodos de modulación de nutrientes y factores de crecimiento para procesos de cultivo continuos así como técnicas para maximizar la velocidad de formación de producto son bien conocidos en la técnica de la microbiología industrial y se detalla una amplia variedad de métodos en Brock, supra.

Particularmente adecuado para la producción de etanol es un régimen de fermentación como el que sigue. Las células deseadas de *Z. mobilis* de la presente invención se hacen crecer en matraces de agitación en medio semi-complejo a una temperatura de aproximadamente 30°C a aproximadamente 37°C, con agitación, a aproximadamente 150 rpm, en agitadores orbitales, y después se transfieren a un fermentador de siembra de 10 L que contiene un medio similar. El cultivo de siembra se hace crecer en el fermentador de siembra de forma anaerobia hasta que el valor de la DO₆₀₀ está entre 3 y 6, cuando se transfiere al fermentador de producción en el que se optimizan los parámetros de fermentación para la producción de etanol. Los volúmenes de inóculo típicos transferidos desde el tanque de siembra al tanque de producción están entre aproximadamente el 2% y aproximadamente el 20% v/v. El medio de fermentación típico contiene componentes del medio mínimos tales como fosfato de potasio (1,0 - 10,0 g/L), sulfato de amonio (0 - 2,0 g/L), sulfato de magnesio (0 - 5,0 g/L), una fuente de nitrógeno compleja tal como productos basados en extracto de levadura o soja (0 - 10 g/L). Una concentración final de aproximadamente 5 mM de sorbitol o manitol está presente en el medio. Azúcares mixtos que incluyen xilosa y al menos un azúcar adicional tal como glucosa (o sacarosa), proporcionando una fuente de carbono, se añaden continuamente al recipiente de fermentación al agotarse la fuente de carbono añadida inicialmente (50-200 g/L) para maximizar la velocidad del etanol y el título. Las velocidades de alimentación de la fuente de carbono se ajustan dinámicamente para asegurar que el cultivo no está acumulando glucosa en exceso, lo que podría conducir a la acumulación de subproductos tóxicos tales como el ácido acético. Con el fin de maximizar el rendimiento de etanol producido a partir del sustrato utilizado, el crecimiento de biomasa se restringe por la cantidad de fosfato que se añade inicialmente al lote o bien que se alimenta en el curso de la fermentación. La fermentación se controla a pH 5,0 - 6,0 usando una solución cáustica (tal como hidróxido amónico, hidróxido potásico o hidróxido sódico) y ácido sulfúrico o o bien ácido fosfórico. La temperatura del fermentador se controla entre 30°C y 35°C. Con el fin de minimizar la formación de espuma, se añaden al recipiente según sea preciso agentes antiespumantes (cualquier clase basada en silicona, base orgánica, etc.). Un antibiótico, para el cual hay un marcador resistente a los antibióticos en la cepa, tal como kanamicina, puede usarse opcionalmente para reducir la contaminación al mínimo.

Cualquier conjunto de condiciones descritas anteriormente, y adicionalmente variaciones en estas condiciones que son bien conocidas en la técnica, son condiciones adecuadas para la producción de etanol por células *Zymomonas* recombinantes que utilizan xilosa.

Ejemplos

La presente invención se define con más detalle en los Ejemplos que siguen. Debe entenderse que estos Ejemplos, aunque indican realizaciones preferidas de la invención, se dan solamente a título de ilustración.

Métodos generales

Las técnicas estándar de ADN recombinante y clonación molecular usadas aquí son bien conocidas en la técnica y son descritas por Sambrook, J., Fritsch, E. F. y Maniatis, T., *Molecular Cloning: A Laboratory Manual*, 2ª edición, Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989) (en lo sucesivo "Maniatis"); y por Silhavy, T. J., Bennan, M. L. y Enquist, L. W., *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1984); y por Ausubel, F. M. et al., *Current Protocols in Molecular Biology*, publicado por Greene Publishing Assoc. y Wiley - Interscience, Hoboken, NJ (1987).

El significado de las abreviaturas es como sigue: "kb" significa kilobase(s), "bp" significa pares de bases, "nt" significa nucleótido(s), "hr" significa hora(s), "min" significa minuto(s), "L" significa litro(s), "ml" significa mililitro(s), "μl" significa microlitro(s), "μg" significa microgramo(s), "ng" significa nanogramo(s), "mM" significa milimolar, "μM" significa micromolar, "nm" significa nanómetro(s), "μmol" significa micromol(es), "pmol" significa picomol(es), "Cm" significa cloranfenicol, "Cms" significa sensible al cloranfenicol, "Specr" o "Spec-R" significa resistencia a la espectinomomicina, "Sp^s" significa espectinomomicina, "XI" es xilosa isomerasa, "XK" es xilulocinasa, "TAL" es transaldolasa, "TKT" es transcetolasa, "DO₆₀₀" significa densidad óptica medida a una longitud de onda de 600 nm, PCR significa reacción en cadena de la polimerasa, "kDa" significa kilodalton(s), "g" significa la constante de gravitación, "bp" significa par(es) de bases, "kbp" significa kilopar(es) de bases, "HPLC" significa cromatografía líquida de alto rendimiento y "GC" significa cromatografía de gases, "RM" significa medio rico que contiene 10 g/L de extracto de levadura más 2 g/L de KH₂PO₄, "MM" significa medio de acoplamiento que contiene 10 g/L de extracto de levadura, 5 g/L de triptona, 2,5 g/L de (NH₄)₂SO₄ y 0,2 g/L de KH₂ PO₄.

50 Transformación de *Z. mobilis*

Se introdujo ADN de plásmido replicante y no replicante en *Z. mobilis* usando electroporación, esencialmente como se describe en el documento US 5.514.583.

De forma resumida, las reacciones de transformación de 50 μl contenían ~ 10¹⁰ células/ml en glicerol al 10% (v/v) y ~ 0,5 - 2,0 mg de ADN de plásmido no metilado que se aisló de *E. coli* SCS110. Las reacciones de control se trataron de forma idéntica, pero no recibieron ningún ADN plasmídico. Los ajustes para el electroporador eran 16 kv/cm, 200 Ω y 25 μF, y el ancho mínimo de la cubeta era 0,1 cm. Después de la electroporación, las reacciones de transformación se diluyeron con medio MMG (50 g/L de glucosa, 10 g/L de extracto de levadura, 5 g/L de triptona, 2,5 g/L de (NH₄)₂SO₄, 0,2 g/L de K₂HPO₄ y 1 mM MgSO₄) y las células se dejaron recuperar a 30°C antes de

ponerlas en placas sobre medio MMG que contenía 1,5% de agar (placa de agar MMG) con o sin antibióticos como se indica. Las placas se incubaron en una cámara anaeróbica a 30°C hasta la aparición de colonias. En los Ejemplos que siguen se describen detalles adicionales.

Experimentos en matraces de agitación

5 A menos que se indique otra cosa, todos los experimentos descritos a continuación se llevaron a cabo a 30°C en matraces de agitación (tubos de ensayo de 15 ml de forma cónica con tapa suelta sin apretar) usando medios sintéticos que contenían glucosa o xilosa como única fuente de carbono. El medio mRM3-G10 contiene 10 g/L de extracto de levadura, 2 g/L de K₂HPO₄, 1 g/L de MgSO₄ (7H₂O) y 100 g/L de glucosa. El medio mRM3-X10 es idéntico pero contiene 100 g/L de xilosa en vez de glucosa. El crecimiento celular se monitorizó mediante espectrofotometría siguiendo los cambios en la densidad óptica a 600 nm en función del tiempo. En las leyendas del texto y de las figuras, "OD", "DO" o "DO600" significa densidad óptica a 600 nm. A los intervalos indicados durante los estudios de crecimiento en matraz de agitación, se retiraron partes alícuotas de los cultivos de 1,0 ml para análisis por HPLC usando un aparato Agilent 1100 equipado con un detector de índice de refracción (Hewlett-Packard, Palo Alto, CA) para determinar las concentraciones de glucosa, xilosa, ribulosa y etanol que estaban presentes en el caldo de fermentación. Antes del análisis por HPLC, las células fueron eliminadas por centrifugación y el sobrenadante se filtró a través de un filtro de tubo de centrifuga Spin-X de acetato de celulosa de 0,22 µm (Costar, número de catálogo 8160) para eliminar partículas pequeñas. Los compuestos se separaron en una columna Aminex HPX-87H (Bio-Rad) que funcionó a 55°C bajo condiciones isocráticas usando una velocidad de flujo de 0,6 ml/min y H₂SO₄ 0,01 N como fase móvil. Se usaron patrones auténticos de concentración conocida para cuantificar los picos de interés y todos los resultados se expresan en g/L.

Construcción de la cepa de *Zymomonas mobilis*

En el documento US 7.741.119 se proporciona descripción detallada de la construcción de la cepa recombinante que utiliza xilosa, ZW801-4, a partir de la cepa madre de tipo silvestre, ZW1. La cepa ZW801-4 se derivó a partir de la cepa ZW800, que se derivó de la cepa ZW658, todo ello como se describe en el documento US 7.741.084.

25 Se construyó ZW658 integrando dos operones, P_{gap}xylAB y P_{gap}tal/tkt, que contienen cuatro genes que utilizan xilosa que codifican xilosa isomerasa (xylA), xilulocinasa (xylB), transaldolasa (tal) y transcetolasa (tkt), con regiones de codificación de genes de *E. coli*, en el genoma de ZW1 (cambio de denominación de la cepa ZM4; ATCC 31821) a través de eventos de transposición secuencial seguido por la adaptación en medio selectivo que contiene xilosa para producir la cepa X13L3, que fue renombrada ZW641. Una posterior adaptación de ZW641 sobre medios de crecimiento que contienen xilosa dio lugar a ZW658, que crece mucho mejor en xilosa y se depositó al amparo del Tratado de Budapest como ATCC PTA-7858. Como se describe en el documento US 7.989.206 de propiedad común, ZW658 tiene mucha más actividad de xilosa isomerasa debido a una mutación puntual en el promotor (P_{gap}) que expresa la región de codificación xylA. Este promotor (SEQ ID NO: 21), denominado en el presente texto promotor 801 GAP o bien promotor Super GAP o P_{gaps}, tiene una "T" en vez de una "G" en la posición 116 en la SEQ ID NO: 21, cuando se compara con el nativo P_{gap} en ZW641 (el promotor 641 GAP). El P_{gaps} tiene una fuerza de expresión de 3 a 4 veces más alta que la P_{gap} en *Z. mobilis*.

En ZW658, el gen que codifica glucosa-fructosa oxidorreductasa fue inactivado por inserción usando recombinación homóloga de doble entrecruzamiento mediada por el huésped, recombinación homóloga y resistencia a espectinomicina como marcador seleccionable para crear la cepa ZW800. El marcador de resistencia a la espectinomicina, que estaba limitado por sitios loxP, se eliminó por recombinación específica del sitio usando Cre recombinasa para crear la cepa ZW801-4.

Preparación de extractos libres de células de *Zymomonas* para ensayo enzimático

Las células se hicieron crecer en 50 ml de RM + 2% de glucosa a 30°C durante la noche hasta una DO₆₀₀ de 1,0 - 1,2. Las células se recolectaron por centrifugación a 4500 rpm durante 10 min a 4°C. El sobrenadante se desechó y el sedimento celular se lavó con 25 ml de tampón de sonicación enfriado con hielo (Tris 10 mM, pH 7,6, MgCl₂ 10 mM), seguido de centrifugación a 4500 rpm durante 10 min. El sedimento se resuspendió en 2,0 - 2,5 ml de tampón de sonicación más ditiotreitól 1 mM. Se centrifugó una alícuota de 500 µl durante 1 min en una centrifuga eppendorf a 4°C. La mayor parte del sobrenadante se desechó, dejando unos 10 - 20 µL para evitar que el sedimento se secase. Las células se congelaron y se almacenaron a aproximadamente 80°C hasta que se ensayaron. Antes del ensayo, las células se descongelaron y se resuspendieron con 500 µL de tampón de sonicación más ditiotreitól 1 mM. La mezcla se sometió a sonicación 2x durante 45 segundos a un ciclo de trabajo del 62% y un control de salida de 2 usando un sonicador Branson 450, dejando enfriar las muestras aproximadamente entre 3 y 5 minutos entre las sonicaciones. Las muestras se centrifugaron a 14.000 rpm durante 60 minutos en una microcentrifuga Beckman a 4°C. El sobrenadante se pasó a un nuevo tubo y se mantuvo a 4°C. Se usó el ensayo Pierce BCA para determinar las concentraciones de proteína.

Ejemplo 1

Construcción de pMODlinker-Spec-GapRpi

El PMODLinker-Spec-GapRpi (Figura 2) es un plásmido que puede usarse para generar un transposón que contiene un casete de expresión de RPI de *Z. mobilis* y un casete de resistencia a Spec flanqueado lox. El casete de expresión de RPI de *Z. mobilis* se generó mediante PCR usando el plásmido pZB188aad/Gap/Zymo RPI/EcoliSL como molde. Este último plásmido se describe en el documento US 7.989.206. Contiene un casete de expresión para la ribosa 5-fosfato isomerasa (RPI) de *Z. mobilis*, que consiste en (de 5' a 3') una secuencia de promotor GAP 641 de longitud completa, el marco de lectura abierto entero de la RPI de *Z. mobilis* (SEQ ID NO: 13), y la región de tallo-bucle XylA que está presente en la región intergénica del operón XylA/B de *E. coli*. La construcción de expresión de RPI entera descrita anteriormente (SEQ ID NO: 22) está situada entre los sitios NcoI y NotI únicos del plásmido pZB188aad/Gap/Zymo RPI/EcoliSL.

Se insertó un fragmento de ADN generado por PCR que contenía esta casete de expresión en pMOD-Linker-Spec, que también se describe en el documento US 7.989.206. El pMOD-Linker-Spec fue derivado del vector comercialmente disponible pMOD™-2 <MCS> Transposon Construction Vector (Nº de catálogo MOD0602; EPICENTRE® Biotechnologies, Madison, WI). El sitio de multi-clonación original se reemplazó por sitios de restricción únicos para AsiSI, FseI y SbfI. Un fragmento de ADN que confiere resistencia a espectinomomicina (SpeCr) y tiene un sitio lox^p de tipo silvestre en los dos extremos fue insertado entre los sitios AsiSI y FseI creando pMOD-Linker-Spec.

El PMOD-Linker-Spec se digirió secuencialmente con SbfI y FseI, y el fragmento de vector grande de 3,6 kb fue purificado a partir de un gel de agarosa al 1%. A continuación, el gen quimérico de RPI de *Z. mobilis* con su promotor 641 GAP asociado y terminador Xyl A de *E. coli* fue amplificado por PCR a partir del plásmido pZB188aad/Gap/Zymo RPI/EcoliSL usando el cebador PPI-F (SEQ ID NO: 23) y el cebador PPI-R-SbfI (SEQ ID NO: 24). El producto de PCR de 0,96 kpb resultante se cortó también con FseI y SbfI, y el fragmento de ADN purificado se ligó después con el fragmento de vector pMOD-Linker-Spec digerido con FseI/SbfI para dar pMODlinker-Spec-GapRpi, que se muestra en la Figura 2.

Ejemplo 2

Sobreexpresión de RPI en ZW801-4: Generación y caracterización de la cepa I

Un transposón generado a partir de pMODlinker-Spec-GapRpi (Ejemplo 1) fue introducido en la cepa ZW801-4 (véase Métodos Generales) para aumentar la expresión de RPI. El elemento transponible en este plásmido que se integra aleatoriamente en el ADN después de la conversión en un transposoma es el fragmento de ADN entero que está situado entre dos extremos de mosaico (ME) en el vector, que incluye tanto la casete de expresión de RPI de *Z. mobilis* como la casete Spec^r. El transposoma se generó *in vitro* esencialmente como se describe en el documento US 7.989.206 usando el protocolo general que se describe en el manual de instrucciones EPICENTRE® para el vector de construcción de transposón EZ :: TN™ pMOD™ -2 <MCS> (Nº de cat. MOD0602). El transposoma resultante fue electroporado en células ZW801-4 y los transformantes se recuperaron en placas de agar MMG y MMX (el mismo medio con 50 g/L de xilosa en vez de 50 g/L de glucosa) que contenían espectinomomicina (200 mg/ml). Dado que el elemento transponible se inserta aleatoriamente en el ADN, lo que puede tener como resultado eventos perjudiciales de rotura génica y/o efectos de posición que alteran el nivel de expresión del gen RPI integrado, se llevó a cabo un experimento preliminar para determinar qué transformantes era mejor perseguir. Así pues, se cribaron 30 colonias de las placas mRM3-G10/Spec y 6 de las placas mRM3-X10/Spec, en relación con el crecimiento en experimentos en matraz de agitación (véanse los métodos generales) usando mRM3-X10 como medio de ensayo, y las tres cepas que crecieron a los valores finales más altos de DO (cepas B9, B11 e I) se seleccionaron para caracterización más profunda. Las cepas B9 y B11 se recuperaron de las placas MMG/Spec y la cepa I se recuperó de una placa MMX/Spec.

En la solicitud de patente de Estados Unidos USSN 13/161734, de propiedad común y copendiente, publicada como US 20120156746 A1, se describió la RPI como la enzima limitante de la velocidad para el metabolismo de la xilosa en ZW801-4, y expresar RPI a partir de un plásmido multicopia en esta cepa tuvo como resultado un mejor crecimiento y una más rápida utilización de la xilosa en experimentos en matraces de agitación con mRM3-X10. Aunque las células crecieron hasta un valor DO final más alto con RPI adicional, hubo poco o ningún cambio en la velocidad de crecimiento exponencial. La sobreexpresión plasmídica de RPI en ZW801-4 también redujo la cantidad de ribulosa que se produjo, y esto tuvo como resultado un rendimiento metabólico más alto para la producción de etanol a partir de xilosa.

Como se muestra en la Figura 3, se obtuvieron resultados similares para las cepas ensayadas que tienen una única copia del transposón de expresión RPI integrado aleatoriamente en el genoma de ZW801-4. Cultivos de siembra de las cepas B9, B11 y I, junto con el control ZW801-4, se cultivaron en mRM3-X10 y se usaron para inocular mRM3-X10 en matraces de agitación a un valor inicial de DO₆₀₀ de ~ 0,1 para cada cultivo. Los cultivos se desarrollaron a 30°C y se ensayaron a lo largo del tiempo para ver el crecimiento por DO₆₀₀, y para xilosa, ribulosa y etanol como se describe en los Métodos Generales. Como se muestra en la Figura 3A, las tres cepas que contienen el transposón

de expresión de RPI crecieron a una DO casi 50% más alta que la cepa madre ZW801-4. Como se muestra en la Figura 3B, también consumieron xilosa a una velocidad mucho mayor y produjeron menos ribulosa que la cepa de control: para un tiempo de 54 h, casi toda la xilosa había desaparecido para B9, B11 y la cepa I, mientras que la cepa madre ZW801-4 sólo había utilizado ~ 75% de la xilosa. Para un valor del tiempo de 71 horas, había todavía ~ 8 g/L de xilosa residual para ZW801-4. Las tres cepas con el transposón de expresión de RPI produjeron también menos ribulosa que ZW801-4 a pesar del hecho de que consumieron más xilosa; la concentración final de ribulosa en el medio de crecimiento a las 72 horas para ZW801-4 fue de 5,3 g/L frente a 3,2 - 3,6 g/L para las otras tres cepas.

Las mismas 4 cepas fueron cultivadas en medio mRM3-G10 a 30°C y se ensayaron a lo largo del tiempo para ver el crecimiento mediante DO₆₀₀ (Figura 3C). Al contrario que los resultados obtenidos con xilosa, las cuatro cepas mostraron tasas de crecimiento similares cuando crecieron en glucosa, alcanzando una DO₆₀₀ de aproximadamente 1 a las 8 h y de 5,2 - 5,5 a las 20 h. Este resultado sugiere que el efecto estimulador de la sobreexpresión de RPI sobre el crecimiento en los experimentos en frascos de agitación de xilosa con la cepa B9, B11 y I depende de la fuente de carbono.

Ejemplo 3

Sobreexpresión de xilosa isomerasa de *E. coli* en cepas B9, B11 y I

Aunque las cepas B9, B11 y I se comportaron de forma muy similar en los experimentos en frasco de agitación con glucosa y xilosa como se describió anteriormente, experimentos posteriores revelaron una diferencia importante entre la cepa I y las otras dos cepas que también tienen una copia integrada del transposón de expresión de RPI. Esta diferencia se encontró evaluando si una mayor expresión de xilosa isomerasa en las 3 cepas tendría como resultado un aumento adicional en la velocidad de flujo de carbono a través de la ruta de xilosa manipulada genéticamente.

La xilosa isomerasa fue la enzima limitante de la velocidad para el metabolismo de la xilosa en la cepa ZW641, que se superó en la cepa ZW658 por una mutación puntual en el promotor P_{gap} que conduce el operón XylA/B de *E. coli*, lo que tuvo como resultado un aumento de la expresión de xilosa isomerasa (véase la construcción de la cepa en Métodos Generales). Como se describe en la solicitud de patente de Estados Unidos USSN 13/161734 de propiedad común y copendiente publicada US20120156746A1, experimentos subsiguientes establecieron que la nueva etapa limitadora de la velocidad para el metabolismo de xilosa tanto en ZW658 como en ZW801-4 (un derivado de ZW658 que tiene inactivación del gen de glucosa-fructosa-oxidoreductasa) era RPI.

Basándose en estas observaciones, fue de interés determinar si la eliminación del cuello de botella de RPI, como se proporciona mediante el aumento de la expresión de RPI en las cepas B9, B11 y I, permitiría un nivel más alto de expresión de la xilosa isomerasa para dar como resultado otro aumento en la tasa de flujo de carbono a través de la ruta de xilosa. Para probar esta hipótesis los autores de la presente invención utilizaron un transposón que se generó a partir del plásmido pMODlinker-Cm-801GapXylA. Este plásmido es idéntico al pMOD-Linker-Spec-801 GapXylA que se describe en el documento US 7.989.206, excepto que tiene un casete de resistencia a Cm entre los dos sitios lox en vez de un casete de resistencia a Spec. El pMOD-Linker-Spec-801 apXylA es pMOD-Linker-Spec con un fragmento de ADN añadido que se obtuvo de ZW801-4 que contiene el P_{gap}, la región de codificación de XylA y la región de tallo-bucle que está entre los marcos de lectura abiertos XylA y XylB. Por tanto, el promotor que impulsa XylA es el promotor 801gap (SEQ ID NO: 21), que tiene la mutación puntual G a T en la posición 116 que la hace más activa, y es el mismo promotor mutante que impulsa el operón XylA/B de *E. coli* en ZW658.

Además del casete de resistencia a Cm, el transposón que se generó a partir de pMODlinker-Cm-801GapXylA (denominado más adelante como "transposón 801GapXylA-Cm") contiene el casete de expresión de xilosa isomerasa anterior que consiste en el promotor 801 GAP, el marco de lectura abierto XylA de *E. coli* y la región tallo-bucle XylA estabilizadora que sigue inmediatamente al codón de parada como se describe en detalle en el documento US 7.989.206. El transposón 801GapXylA-Cm se convirtió en un transposoma como se describe en el Ejemplo 2 y este último fue electroporado en la cepa I. Los transformantes se seleccionaron en placas de agar MMG que contenían cloranfenicol (120 µg/mL).

En un estudio preliminar, diez colonias de Cm^r elegidas aleatoriamente que fueron recuperadas de la reacción de transformación con el transposón 801GapXylA-Cm fueron ensayadas para ver el crecimiento en mRM3-X10 a 30°C en un experimento en matraz de agitación en el que se midió periódicamente la DO₆₀₀. Aunque las diez cepas crecieron mejor en xilosa que la cepa madre (basándose en las velocidades iniciales de crecimiento y los valores finales de DO), sólo siete de ellas fueron seleccionadas para una caracterización adicional: I (cm1), I (cm2), I (cm3), I (cm 4), I (cm 5), I (cm 8) y I (cm 9). Para examinar estas cepas con más detalle, se repitió el experimento en matraz de agitación usando cultivos de siembra que se desarrollaron en mRM3-X10 y ZW801-4 y la cepa I como controles. La DO inicial fue de 0,075 en todos los casos y las dos cepas de control se realizaron por duplicado.

De forma consistente con los experimentos anteriores, la cepa I, que tiene un nivel de expresión de RPI más alto, creció mucho mejor en xilosa que ZW801-4, como se muestra en la Figura 4A. Obsérvese que las curvas de crecimiento para los duplicados de cepa I eran superponibles y no se pueden distinguir. La concentración final de

ribulosa en el medio de cultivo, a un tiempo de 56 horas, fue más bajo para la cepa I que para ZW801-4: 1,8 y 1,9 g/L frente a 3,4 y 4,2 g/L. La sobreexpresión de xilosa isomerasa en la cepa I mejoró grandemente el crecimiento en xilosa. Las siete cepas que contenían el transposón 801GapXylA-Cm crecieron con mucha más rapidez que ZW801-4 y la cepa I, y sus curvas de crecimiento fueron virtualmente idénticas a pesar del hecho de que tienen sitios de inserción del transposón diferentes. Esto fue confirmado realmente mediante la secuenciación de ADN de los sitios de inserción para I (cm1) - I (cm6). Aunque los sitios de inserción del transposón no se determinaron para las otras cuatro cepas, están probablemente también en localizaciones genómicas diferentes debido a la frecuencia de transformación de *Z. mobilis* extraordinariamente baja y al procedimiento EPICENTRE® que se utilizó.

El aumento de la expresión de la actividad de la enzima xilosa isomerasa en la cepa I también incrementó la velocidad de crecimiento exponencial, y esto tuvo como resultado una acumulación más rápida de biomasa y una utilización más rápida de la xilosa. Los valores de punto final para la ribulosa para las siete cepas que contenían el transposón 801 GapXylA-Cm estaban en el intervalo entre 2,9 - 3,7 g/L. Aunque estos valores son similares a la cantidad de ribulosa que produjo ZW801-4, son en realidad más pequeños sobre la base del azúcar utilizado ya que a las 56 h las cepas que sobreexpresaron xilosa isomerasa habían consumido toda la xilosa en el medio mientras que ZW801-4 había utilizado menos del 80%.

Estos resultados demuestran claramente que la xilosa isomerasa es la enzima limitante de la velocidad para el metabolismo de la xilosa en la cepa I. También revelan que la cepa I tiene suficiente actividad enzimática RPI para soportar un nivel de expresión de xilosa isomerasa mucho más elevado. El hecho de que los siete derivados de cepa I que contenían el transposón 801GapXylA-Cm en diferentes localizaciones en el cromosoma crecieran con la misma cinética en el medio que contiene xilosa, sugiere claramente que los eventos de ruptura de genes y/o efectos de posición causados por el transposón no contribuyeron significativamente al fenotipo observado (es decir, mejor crecimiento en xilosa). Esta conclusión viene también apoyada por el experimento que se muestra en la Figura 4B, en la que las siete cepas que contenían el transposón crecieron todas así como ZW801-4 y la cepa I en experimentos en matraz de agitación con mRM3-G10 a 30°C.

Las observaciones anteriores se confirmaron usando condiciones controladas por el pH de la forma que sigue. Durante los experimentos en matraces de agitación, el pH del medio de crecimiento puede caer en más de 1 unidad de pH, y esto tiene un efecto inhibitorio sobre el crecimiento de *Z. mobilis* especialmente cuando la xilosa es la única fuente de carbono. Para obviar este problema se llevó a cabo el experimento de biorreactor controlado por el pH que se muestra en la Figura 5. Como todos los derivados de la cepa I que sobreexpresan la xilosa isomerasa se comportaron de la misma manera en matraces de agitación, sólo dos de ellos se usaron para este experimento, I (cm1) y I (cm9). Los cultivos de siembra se desarrollaron en mRM3-X10 a 30°C en matraces de agitación hasta una DO de ~0,5, y la DO de partida en los biorreactores fue 0,035. Los biorreactores contenían también mRM3-X10, y la temperatura y el pH se mantuvieron constantes en 30°C y 5,8, respectivamente, utilizando KOH 2 N para el control del pH.

Las curvas de crecimiento para la cepa I y ZW801-4 fueron muy similares bajo condiciones controladas por el pH (Figura 5), aunque estaba claro que la cepa I creció ligeramente mejor hacia el final del experimento. Es curioso, sin embargo, el hecho de que ninguna de las dos cepas produjo mucha ribulosa bajo estas condiciones; los valores de punto final para ZW801-4 y la cepa I fueron 1,92 y 1,16 g/L, respectivamente. Estas observaciones sugieren que RPI no es un cuello de botella tan grande para el metabolismo de xilosa en ZW801-4 cuando el pH del crecimiento se mantiene en 5,8. También indican que el gen RPI nativo en ZW801-4 es casi capaz de mantener el ritmo con el nivel existente de actividad de xilosa isomerasa bajo tales condiciones. Sin embargo, estaba claro a partir de las curvas de crecimiento que se muestran en la Figura 5 que la xilosa isomerasa que sobreexpresaba la cepa I mejoraba en gran medida el metabolismo de la xilosa incluso cuando el pH del medio de crecimiento se mantenía constante a pH 5,8.

El transposón 801GapXylA-Cm fue también sometido a electroporación en las cepas B9 y B11 usando el mismo procedimiento que se usó para la cepa I, y diez transformantes primarios para cada una de estas cepas fueron ensayados para ver el crecimiento en xilosa en experimentos en frasco de agitación con mRM3-X10 a 30°C (Figura 6). Aunque algunos de los transformantes sí que crecieron mejor en xilosa que la cepa madre, en ambos casos los resultados no fueron casi tan drásticos como los que se obtuvieron con la cepa I. Hubo también un alto grado de variación entre los diferentes transformantes, lo que sugiere que los eventos de ruptura de genes y/o efectos de posición insercional estaban contribuyendo a los fenotipos observados. Estas observaciones sugirieron que podría haber algo significativo sobre la ubicación cromosómica del transposón de expresión de RPI en la cepa I, que es beneficioso para el crecimiento y metabolismo de la xilosa, además de un nivel elevado de expresión de RPI.

Se determinó que el sitio de inserción para el transposón de expresión de RPI en la cepa I mediante secuenciación de ADN estaba entre los nts 543506 y 543507 del genoma *Z. mobilis* (número de orden GenBank AE008692). La secuenciación de la región de inserción indicó que la integración del transposón provocó un desplazamiento de marco en el extremo 3' del marco de lectura abierto del gen pnp que codifica la polirribonucleótido nucleotidil transferasa. Una alineación del producto génico pnp de tipo silvestre y el producto génico pnp predicho para la cepa I se muestra en la Figura 7. Obsérvese que a la proteína mutante le faltan los 39 últimos restos de aminoácidos de la proteína nativa (conserva 709 aminoácidos a partir del término N) y tiene 14 nuevos aminoácidos en su término C (SEQ ID NO: 9). No se sabe si esta proteína hipotética es funcional y, de serlo, hasta qué punto podría contribuir al

fenotipo de la cepa I además de la sobreexpresión de RPI. Al margen de la respuesta, los resultados anteriores indican que la sobreexpresión de RPI en la misma localización que en la cepa I, en la región de codificación de pnp, dio como resultado un mejor crecimiento en xilosa y permitió que fuese eficaz un mayor nivel de expresión de xilosa isomerasa.

5 Ejemplo 4

Constructos vectoriales para la construcción de cepas de *Z. mobilis* que utilizan xilosa usando integración dirigida

Se construyó una cepa nueva de *Z. mobilis* que utiliza xilosa, introduciendo genes quiméricos xylA, xylB, tal y tkt en la cepa ZW1. Las regiones de codificación de xylB, tal y tkt procedían de genes de *E. coli* como en la cepa ZW658 descrita en Métodos Generales. La región de codificación de xylA era de *Actinoplanes missouriensis* (AMxylA) que se describe en la publicación de solicitud de patente de EE. UU. US20110318801 de propiedad común y copendiente, como codificadora de una enzima que tiene mayor actividad que la xilosa isomerasa de *E. coli* en *Z. mobilis*. La región de codificación para AMxylA fue optimizada en codón para la expresión en *Z. mobilis* (SEQ ID NO: 17). También se introdujeron copias adicionales de genes rpi y rpe de *Z. mobilis* con el fin de aumentar las actividades de ribosa-5-fosfato isomerasa (RPI) y ribulosa-fosfato 3-epimerasa (RPE). Se diseñaron vectores de transformación de doble entrecruzamiento (DCO) para integrar específicamente los genes quiméricos en regiones diana en el genoma de *Z. mobilis*.

Se usaron métodos estándar de recombinación molecular para construir vectores suicidas de integración con DCO (doble entrecruzamiento). Para expresar xilosa isomerasa y xilulosa quinasa en *Z. mobilis* se construyó un vector suicida de DCO pZX21 de 10.250 pb (SEQ ID NO: 25, figura 8A). Este vector tiene una estructura pBluescript que contiene un origen de replicación para *E. coli* pero no origen de replicación de *Z. mobilis*, y así no puede propagarse en *Z. mobilis* convirtiéndose en un vector suicida. Contiene secuencias de ADN del gen de *Z. mobilis* que codifica glucosa-fructosa oxidoreductasa, GFO-L y GFO-R, flanqueando las secuencias que se han de integrar. Ambos fragmentos se sintetizaron mediante PCR, utilizando ADN genómico de *Z. mobilis* como molde. El fragmento GFO-L de 1.186 pb (SEQ ID NO: 26) incluye los primeros 654 pb (de nt-1 a nt-653) de la secuencia de codificación de gfor (SEQ ID NO: 27) y 533 pb de la secuencia genómica aguas arriba. El fragmento GFO-R de 1.446 pb (SEQ ID NO: 28) incluye los últimos 480 pb (de nt-823 a nt-1302) de la secuencia de codificación de GFOR y 966 pb de la secuencia genómica aguas abajo. La integración directa de las secuencias de GFO-L y GFO-R en el locus gfor, reemplazando un segmento de la secuencia de codificación de gfor (de nt-655 a nt-822) en el genoma de *Z. mobilis*. Esto interrumpe la expresión de glucosa-fructosa oxidoreductasa, lo que reduce la producción de xilitol y aumenta la producción de etanol como se describe en el documento US 7.741.119.

La región en pZX21 entre GFO-L y GFO-R incluye tres genes quiméricos. Uno es un gen xylA quimérico de 1.661 pb (SEQ ID NO: 29) que contiene el promotor Super GAP de *Z. mobilis* de 304 pb (P_{gapS} , descrito en el documento US 7.989.206), una secuencia de codificación de xylA de *A. missouriensis* de 1.185 pb (AMxylA) y un araD 3'UTR de *E. coli* de 166 pb con un sitio 5' XbaI (ECaraD 3'UTR). La región codificadora de AMxylA se optimizó para la expresión en *Z. mobilis* de acuerdo con el sesgo de codón de *Z. mobilis* ZM4 (SEQ ID NO: 17). El 3'UTR de ECaraD era del operón de araBAD de *E. coli*. El segundo gen es un gen xylB quimérico de 1.960 pb (SEQ ID NO: 30) que contiene un P_{eno} de 191 pb, una secuencia codificadora de xylB de *E. coli* de 1.455 pb (ECxylB) y un xylB 3'UTR de 314 pb de *E. coli* (ECxylB 3'UTR). P_{eno} es un promotor fuertemente constitutivo a partir del ADN genómico de *Z. mobilis* que tiene aproximadamente 28% de actividad de P_{gap} . El tercer gen es un marcador aadA de 1.014 pb (para la resistencia a la espectinomicina, Spec-R) limitado por sitios lox (SEQ ID NO: 31). El marcador se puede eliminar después de la integración expresando Cre recombinasa.

Para expresar transaldolasa, transcetolasa, ribosa-5-P-isomerasa y D-ribulosa-P-3-epimerasa en *Z. mobilis*, fue construido un vector de lanzadera de DCO de 12.198 pb pZX52 (SEQ ID NO: 32, Figura 8B). Este vector es un vector de lanzadera *Zymomonas-E. coli* que se basa en el vector pZB188 (Zhang et al. (1995) Science 267: 240 - 243; documento US 5.514.583), que incluye un fragmento de ADN genómico de *Z. mobilis* de 2.582 pb que contiene un origen de replicación que permite al vector replicarse en células de *Zymomonas* y un origen de replicación de *E. coli* de 909 pb (Ori). Tiene un marcador de resistencia al cloranfenicol de 911 pb (Cm-R) para la selección de transformantes de *E. coli* o bien de *Z. mobilis*. pZX52 contiene secuencias de ADN procedentes del gen ldhA de *Z. mobilis* que codifica lactato deshidrogenasa, LDH-L (875 pb; SEQ ID NO: 33) y LDH-R (1.149 pb, SEQ ID NO: 34), flanqueando las secuencias que se han de integrar. Estas secuencias dirigen la integración en la secuencia de codificación de ldhA (SEQ ID NO: 35) en el genoma *Z. mobilis* entre los nucleótidos 493 y 494, interrumpiendo así la expresión de la lactato deshidrogenasa.

La región en pZX52 entre LDH-L y LDH-R incluye dos operones quiméricos. El primero es un operón P_{gapT} -Tal-Tkt de 3.339 pb (SEQ ID NO: 36) que contiene un mutante T de 304 pb del promotor GAP de *Z. mobilis* (P_{gapT}), una región de codificación Tal de *E. coli* de 954 pb (ECT_{al}), una región de codificación Tkt de *E. coli* de 1.992 pb, y un Tkt 3'UTR de *E. coli* de 68 pb (ECTkt 3'UTR). Este operón es idéntico al operón Tal-Tkt de *E. coli* existente en la naturaleza excepto para el promotor P_{gapT} (SEQ ID NO: 37), que es un P_{gap} con un cambio de "G" a "A" en la posición 83 en SEQ ID NO : 21 y una "T" que falta en la posición 285. El otro operón quimérico es un operón P_{eno} -Rpi-Rpe de 1.443 pb (SEQ ID NO: 38), que contiene P_{eno} de 191 pb, una secuencia de codificación Rpi de *Z. mobilis*

de 471 bp con el primer codón cambiado a ATG (SEQ ID NO: 65) (ZMRpi), una secuencia de codificación de Rpe de *Z. mobilis* de 663 pb (ZMRpe) y un xylA 3'UTR de *E. coli* de 35 pb (ECxylA 3'UTR).

Se construyó otro vector lanzadera de DCO denominado pZX6 (SEQ ID NO: 39; Figura 8C). Este vector de 12.704 pb es una modificación de pZX52 que tiene secuencias de LDH-L y LDH-R reemplazadas con secuencias del gen pnp de *Z. mobilis* que codifica la polinucleótido fosforilasa. El fragmento PNP-L de 1.318 pb (SEQ ID NO: 40) es un segmento de la secuencia codificadora de pnp (SEQ ID NO: 1) de nt-767 a nt-2.084, mientras que el fragmento PNP-R de 1.225 pb (SEQ ID NO: 41) incluye los últimos 59 pb (de nt-2189 a nt-2247) de la secuencia de codificación de pnp y 1.166 pb de la secuencia genómica aguas abajo. Por consiguiente, pZX6 es capaz de dirigir la integración del operón P_{gapT} - Tal - Tkt y el operón P_{eno} - Rpi - Rpe en el gen pnp endógeno cerca del extremo de la secuencia de codificación de pnp y reemplazar un segmento de la secuencia de codificación de pnp (de nt - 2.084 a nt - 2.188) en el genoma de *Z. mobilis*.

Ejemplo 5

Desarrollo de cepas de *Z. mobilis* que utilizan xilosa

La cepa ZW1 se transformó con dos plásmidos en dos etapas. Se prepararon células competentes de ZW1 cultivando células de siembra durante la noche en mRM3-G5 (extracto de levadura al 1%, KH_2PO_4 15 mM, MgSO_4 4mM y 50 g/L de glucosa) a 30°C con una agitación a 150 rpm hasta un valor DO_{600} próximo a 5. Se recolectaron las células y se resuspendieron en medio fresco a un valor de DO_{600} de 0,05. Las células se desarrollaron bajo las mismas condiciones hasta la mitad de la fase logarítmica (DO_{600} próximo a 0,5). Las células se recolectaron y se lavaron dos veces con agua helada y después una vez con 10% de glicerol enfriado con hielo. Las células competentes resultantes se recogieron y resuspendieron en glicerol al 10% enfriado con hielo hasta un valor DO_{600} cercano a 100. Dado que la transformación de *Z. mobilis* requiere ADN no metilado, los plásmidos DCO pZX21, pZX52 y pZX6 se transformaron cada uno en células competentes de *E. coli* SCS110 (Stratagene, La Jolla, CA). Para cada transformación, se cultivó una colonia de células transformadas en 10 mL de LB-Amp100 (caldo LB que contiene 100 mg/L de ampicilina) durante la noche a 37°C. Se preparó ADN a partir del cultivo de 10 ml, utilizando el kit QIAprep Spin DNA Miniprep (Qiagen).

Se mezcló aproximadamente 1 μg de ADN de pZX21 no metilado, con 50 μL de células competentes ZW1 en una cubeta de electroporación de 1 MM (VWR, West Chester, PA). El ADN plasmídico fue electroporado en las células a 2,0 KV usando un aparato Transporater Plus BT720 (BTX-Genetronics, San Diego, CA). Las células transformadas se recuperaron en 1 mL de medio MMG5 (10 g/L de glucosa, 10 g/L de extracto de levadura, 5 g/L de triptona, 2,5 g/L de $(\text{NH}_4)_2\text{SO}_4$, 2 g/L de K_2HPO_4 y 1 mM de MgSO_4) 4 horas a 30°C, y se cultivaron en placas MMG5-Spec250 (MMG5 con 250 mg/L de espectinomomicina y 15 g/L de agar) durante 3 días a 30°C, dentro de un tarro anaeróbico con un AnaeroPack (Mitsubishi Gas Chemical, New York, NY).

Dado que pZX21 es un vector suicida DCO, las colonias Spec^R supervivientes tenían integrado el segmento P_{gapS} -AMxylA :: P_{eno} -ECxylB :: Spec-R en el locus gfor. Las colonias fueron raspadas y cultivadas en una placa MMG5-Spec250 fresca, y luego se sometieron a PCR para observar la integración del gen quimérico. La primera PCR utilizó cebador directo ara285 (SEQ ID NO: 42) y el cebador inverso ara120 (SEQ ID NO: 43) para inspeccionar la recombinación de doble entrecruzamiento mediada por el fragmento GFO-L en pZX21. El cebador ara285 coincide con un segmento de la secuencia genómica de *Z. mobilis* que está 494 pb aguas arriba del fragmento GFO-L en el genoma, mientras que ara120 complementa los últimos 18 pb de P_{gapS} y los primeros 17 pb de AMxylA en pZX21. Si la integración hubiera ocurrido como se había diseñado, la PCR amplificaría un fragmento de 1.903 pb a partir de los transformantes. La 2ª PCR utilizó el cebador directo ara46 (SEQ ID NO: 44) y el cebador inverso ara274 (SEQ ID NO: 45) para inspeccionar la recombinación de doble entrecruzamiento mediada por el fragmento GFO-R en pZX21. El cebador ara46 es una secuencia cerca del extremo del gen Spec^R en pZX21, mientras que ara274 complementa un segmento de ADN genómico de *Z. mobilis* que está 83 pb aguas abajo del fragmento GFO-R. Esta PCR amplificaría un fragmento de 1.864 pb de las colonias que tuviesen una integración con éxito. Ambas inspecciones dieron los productos de PCR esperados y confirmaron así la integración transgénica certera. La cepa resultante se denominó ZW1-pZX21.

En la segunda etapa, se transformó ZW1-pZX21 con pZX52 y se seleccionó en una placa MMG5-Spec250-CM120 (MMG5-Spec250 con 120 mg/L de cloranfenicol). Debido a que pZX52 es un vector lanzadera de DCO que tiene el marcador Cm^R para la selección del plásmido y un segmento de integración sin marcador (P_{gapT} -ECTal-ECTkt :: P_{eno} -ZMRpi-ZMRpe), las colonias recuperadas deben contener no sólo el constructo integrado previamente P_{gapS} -AMxylA :: P_{eno} -ECxylB :: Spec-R en el genoma de *Z. mobilis*, sino también la construcción no integrada P_{gapT} -ECTal-ECTkt :: P_{eno} -ZMRpi-ZMRpe en el plásmido pZX52 propagado. Estos transformantes deben tener todos los genes necesarios para la ruta de utilización de xilosa. Para demostrar que todos los transgenes eran funcionales en *Z. mobilis*, diez colonias seleccionadas se sometieron a un ensayo de crecimiento de 48 horas en xilosa. En el ensayo se inocularon 2 mL de mRM3-G5-Spec200-CM120 (mRM3-G5 con 200 mg/L de espectinomomicina y 120 mg/L de cloranfenicol) en un tubo de fondo redondo de polipropileno Falcon de 14 mL, con una colonia seleccionada y se cultivaron durante la noche a 30°C con una agitación de 150 rpm. Los tubos se taparon fuertemente, pero se perforó un orificio en la parte superior de la tapa usando una aguja 23G1 para el alivio de la presión durante el crecimiento de las células y la fermentación. Las células se recolectaron, se lavaron con MRM3X10 (MRM3 con 100 g/L de

xilosa) y se resuspendieron en mRM3-X10-Spec200-CM120 (mRM3-X10 que contenía 200 mg/L de espectinomicina y 120 mg/L de cloranfenicol) para tener una DO_{600} de partida de 0,1. Se pusieron 5 mL de la suspensión en un nuevo tubo de fondo redondo de polipropileno Falcon de 14 mL. Los tubos fueron tapados con un orificio en la parte superior. Las células se cultivaron durante 48 horas a 30°C con una agitación de 150 rpm y se midió la DO_{600} en un espectrofotómetro Shimadzu UV-1201. Después se centrifugó 1 mL de cultivo a 10.000 x g para eliminar las células. El sobrenadante se filtró a través de un filtro de tubo de centrifuga Costar Spin-X de 0,22 μ m (Corning Inc, Corning, NY) y se analizó xilosa y etanol corriendo a través de una columna de exclusión de iones BioRad Aminex HPX-A7H (BioRad, Hercules, CA) con H_2SO_4 0,01 N a una velocidad de 0,6 mL/min a 55°C en un sistema Agilent 1100 HPLC (Agilent Technologies, Santa Clara, CA). Los resultados indicaron que los 10 transformantes habían adquirido la ruta de utilización de xilosa para la producción de etanol. La nueva cepa se denominó ZW1-pZX21-pZX52 y uno de los cultivos se utilizó en experimentos adicionales.

ZW1-pZX21-pZX52 pasó después por tres procesos de post-transformación secuencialmente para la integración del constructo P_{gapT} -ECTal-ECTkt :: P_{eno} -ZMRpi-ZMRpe.

(1) La cepa se adaptó en xilosa. En este procedimiento, se suspendió ZW1-pZX21-pZX52 en un medio mRM3-G1X9-Spec200-CM120 de 5 mL (MRM3 con 10 g/L de glucosa, 90 g/L de xilosa, 200 mg/L de espectinomicina y 120 mg/L de cloranfenicol) con un valor de DO_{600} inicial de 0,2 y se cultivó durante 3 a 4 duplicaciones a 30°C (valor de DO_{600} de 0,2 a 2; un pasaje). El cultivo se diluyó después hasta el valor inicial de DO_{600} y se cultivó para otro paso. En total se completaron 4 pasajes (aproximadamente 15 duplicaciones).

(2) El curado del plásmido y la integración del constructo P_{gapT} -ECTal-ECTkt :: P_{eno} -ZMRpi-ZMRpe se llevaron a cabo cultivando 10 mL del conjunto de células de adaptación en 2 mL de medio mRM3-G5-Spec200 a una temperatura más alta (37°C) durante la noche. El cultivo de 10 mL se diluyó después en 2 mL de medio mRM3-G5-Spec200 y se cultivó para otro paso. En total se realizaron 5 pasajes a 37°C en medio de glucosa. Como resultado del crecimiento a alta temperatura, la mayoría de la población no debía alojar más el plásmido pZP52, pero el constructo P_{gapT} -ECTal-ECTkt :: P_{eno} -ZMRpi-ZMRpe (carente de un marcador selectivo) debía haber sido integrado en el gen *ldhA* del genoma de *Z. mobilis*. Una minoría de la población puede mantener pZX52, sin integración.

(3) La población se enriqueció cultivando 50 mL del conjunto de células en 2 mL de mRM3-X10-Spec200 a 30°C durante la noche. La población enriquecida se desarrolló en una placa MMG5-Spec250 a 30°C durante la noche. Se seleccionaron colonias individuales y se rasparon en placas MMG5 y placas MMGS-CM120 en réplica. Después de incubar a 30°C durante una noche, las colonias que crecieron en MMG5 pero no en MMGS-CM120 se seleccionaron para una inspección por PCR adicional. La primera PCR utilizó el cebador directo ara45 (SEQ ID NO: 46) y el cebador inverso ara356 (SEQ ID NO: 47) para inspeccionar la recombinación de doble entrecruzamiento mediada por el fragmento LDH-L en pZX52. El cebador ara45 coincide con un segmento de ADN genómico de *Z. mobilis* que está 86 pb aguas arriba del fragmento LDH-L en el genoma, y ara356 complementa el fragmento (de nt-91 a nt-112) de la región de codificación de ECTal en pZX52. La PCR amplificaría un fragmento de 1.383 pb de las colonias si la integración hubiera ocurrido tal como se había diseñado. La 2ª PCR utilizó el cebador directo ara354 (SEQ ID NO: 48) y el cebador inverso ara43 (SEQ ID NO: 49) para inspeccionar la recombinación de doble entrecruzamiento mediada por el fragmento LDH-R en pZX52. El cebador ara354 es una secuencia cerca del extremo 3' de ZMRpe en pZX52. El cebador ara43 complementa un segmento de ADN genómico de *Z. mobilis* que está 122 pb aguas abajo del fragmento LDH-R. Esta PCR amplificaría un fragmento de 1.468 pb de las colonias cuando la recombinación fuese como se esperaba. Ambas PCR produjeron fragmentos de ADN con los tamaños esperados, lo que confirmó que el constructo P_{gapT} -ECTal-ECTkt :: P_{eno} -ZMRpi-ZMRpe había sido integrado certeramente como se diseñó en todas las colonias inspeccionadas. Las colonias resultantes se denominaron ZW1 - X109.

En un segundo planteamiento, la cepa ZW1-pZX21 fue transformada con el vector lanzadera pZX6 de DCO y se realizaron los tres procedimientos post-transformación como se describió anteriormente para ZW1-X109, excepto que la adaptación fue para 10 pasajes en vez de para 4 pasajes. Por consiguiente, el constructo P_{gapT} -ECTal-ECTkt:: P_{eno} -ZMRpi-ZMRpe fue dirigido al gen *pnp* endógeno. Como se describió para la construcción de ZW1-X109, el ensayo de crecimiento de 48 horas se preformó antes de los tres procedimientos de post-transformación para asegurarse de que todos los transgenes estaban funcionando como se esperaba. Después de los tres procedimientos de post-transformación, la integración también fue inspeccionada por PCR. La primera PCR utilizó cebador directo ara340 (SEQ ID NO: 50) y cebador inverso ara356 (SEQ ID NO: 47) para inspeccionar la recombinación de doble entrecruzamiento mediada por el fragmento PNP-L en pZX6. El cebador ara340 coincide con el ADN genómico de *Z. mobilis* que está 75 pb aguas arriba del fragmento PNP-L. El cebador ara356 usado aquí complementa un fragmento (de nt-91 a nt-112) de ECTal en pZX6. La PCR produjo un fragmento de 1.815 pb a partir de los transformantes, como se esperaba para un evento de integración preciso. La segunda PCR usó cebador directo ara354 (SEQ ID NO: 48) y cebador inverso ara339 (SEQ ID NO: 51) para inspeccionar la recombinación de doble entrecruzamiento mediada por el fragmento PNP-R en pZX6. En este caso, el cebador ara354 coincide con una secuencia próxima al extremo 3' de ZMRpe en pZX6, y el cebador ara339 complementa un segmento de ADN genómico de *Z. mobilis* que está 59 pb después de la secuencia del fragmento PNP-R. Esta PCR amplificó un fragmento de 1.549 pb de los transformantes, un tamaño que se esperaba para una integración con éxito. Por consiguiente, la inspección de PCR confirmó que la construcción P_{gapT} -ECTal-ECTkt :: P_{eno} -ZMRpi-ZMRpe se había integrado certeramente en todas las colonias inspeccionadas. Esta nueva cepa se denominó ZW1-X210.

En resumen, dos cepas de *Z. mobilis* que utilizan xilosa fueron reconstruidas *de novo* a partir de ZW1 de tipo silvestre. Ambos tenían un constructo $P_{\text{gapS-AMxylA}} :: P_{\text{eno-ECxylB}} :: \text{Spec-R}$ integrado en el locus *gfor*. La cepa ZW1-X109 tenía un constructo $P_{\text{gapT-ECTal-ECTkt}} :: P_{\text{eno-ZMRpi-ZMRpe}}$ integrado en el locus *ldhA*, mientras que la cepa ZW1-X210 tenía el mismo constructo integrado en el gen *pnp* endógeno. Ambas cepas tenían un gen marcador en el constructo integrado $P_{\text{gapS-AMxylA}} :: P_{\text{eno-ECxylB}} :: \text{Spec-R}$, que se pudo eliminar mediante la introducción de Cre recombinasa.

Ejemplo 6

Caracterización de nuevas cepas de *Z. mobilis* que utilizan xilosa

La capacidad de las cepas ZW1-X109 y ZW1-X210 para fermentar xilosa fue demostrada mediante un ensayo estándar de crecimiento como se describe en el Ejemplo 5. Para determinar cuantitativamente el crecimiento y los perfiles metabólicos de estas nuevas cepas y compararlas con ZW1, estas cepas fueron caracterizadas en ensayos de fermentación en matraz de agitación. Primero se sometieron a la fermentación en matraz agitado usando MRM3G10 con el fin de determinar su metabolismo basal de glucosa. Las cepas fueron desarrolladas durante la noche en 2 mL de mRM3-G5-Spec250 a 30°C, con una agitación de 150 rpm. Las células se recolectaron, se lavaron con mRM3-G10 y se resuspendieron en mRM3G10 para tener una DO_{600} de partida de 0,1. Se pusieron 20 mL de la suspensión en un tubo de centrifuga VWR con tapón roscado de 45 ml y se cultivaron a 30°C con una agitación de 150 rpm. Para evitar la acumulación de presión debido a la evaporación del etanol durante la fermentación, las tapas se cerraron herméticamente y luego se aflojaron una vuelta. En el curso del tiempo, se midió la DO_{600} en un espectrofotómetro Shimadzu UV-1201 a las 0, 10 y 24 horas. A cada valor del tiempo se centrifugó 1 mL de cultivo a 10.000 x g para eliminar las células. El sobrenadante se filtró a través de un filtro de tubo de centrifuga Costar Spin-X de 0,22 μm y se analizó haciéndole correr a través de una columna de exclusión de iones BioRad Aminex HPX-A7H con H_2SO_4 0,01 N a una velocidad de 0,6 mL/min a 55°C en un sistema Agilent 1100 HPLC. La ZW1 de tipo silvestre se desarrolló sin antibióticos y se analizó como testigo. Los resultados que se dan en la Figura 9 muestran que cada cepa agotó rápidamente la glucosa disponible en 24 horas; tanto la cepa ZW1-X109 como la cepa ZW1-X210 (Figura 8A, 8B, respectivamente) utilizaron glucosa de manera similar a ZW1 (Figura 8C). Por ejemplo, después de 10 horas de fermentación, ZW1-X109 había utilizado aproximadamente el 34,8% de glucosa (una reducción de 102,7 g/L a 66,9 g/L) para apoyar un título de etanol de 16,4 g/L y un crecimiento de biomasa hasta un valor de DO_{600} de 4,88; ZW1-X210 había utilizado aproximadamente el 32,1% de glucosa (una reducción de 102,7 g/L a 69,7 g/L) para apoyar un título de etanol de 15,2 g/L y un crecimiento de biomasa hasta un valor de DO_{600} de 4,68; ZW1 había utilizado aproximadamente el 33,9% de glucosa (una reducción de 103,1 g/L a 68,2 g/L) para apoyar un título de etanol de 16,4 g/L y un crecimiento de biomasa hasta un valor de DO_{600} de 4,6. Por lo tanto, las dos cepas nuevas tienen un metabolismo basal de glucosa robusto.

La fermentación en matraz de agitación se llevó a cabo en 20 mL de mRM3-X10 con el fin de determinar la capacidad de cada cepa para usar la xilosa. El valor de OD_{600} y de las concentraciones tanto de xilosa como de etanol se midieron a las 0, 24, 48 y 72 horas. La Figura 10 es un resumen de los resultados para ZW1-X109 (A), ZW1-210 (B) y ZW1 (C). Los resultados también confirman que las dos nuevas cepas fueron capaces de fermentar la xilosa. Al cabo de 72 horas de fermentación, ZW1-X109 había utilizado aproximadamente el 64,2% de xilosa (una reducción de 105,6 g/L a 37,8 g/L) para apoyar un título de etanol de 31,5 g/L y el crecimiento de biomasa hasta un valor DO_{600} de 3,51; ZW1-X210 había utilizado casi toda la xilosa disponible (una reducción de 105,6 g/L a 1,6 g/L) para apoyar un título de etanol de 48,5 g/L y un crecimiento de biomasa hasta un valor de DO_{600} de 5,22. Sin embargo, ZW1 no pudo crecer en mRM3-X10 debido a la falta de la ruta metabólica de la xilosa. Por tanto, entre las nuevas cepas, la ZW1-X210 pudo fermentar xilosa más rápido que la ZW1-X109, en el medio de azúcar que solo contiene xilosa. La diferencia principal entre ZW1-X109 y ZW1-X210 es que el constructo $P_{\text{gapT-ECTal-ECTkt}} :: P_{\text{eno-ZMRpi-ZMRpe}}$ se insertó en el locus *ldhA* en ZW1-X109 y en el gen *pnp* endógeno en ZW1-X210. Este resultado indica que la interrupción del gen *pnp* puede beneficiar el metabolismo de la xilosa en *Z. mobilis*.

Ejemplo 7

Constructos vectoriales para la interrupción del gen *pnp* endógeno en cepas de *Z. mobilis*

Para ensayar directamente si la interrupción del gen *pnp* endógeno beneficia al metabolismo de la xilosa, se construyeron cuatro vectores suicidas DCO que se muestran en la Figura 4 mediante métodos estándar de recombinación molecular.

pPNP-1 (SEQ ID NO: 55, Figura 11A) es un vector DCO basado en pUC18 de 5.548 pb. Su andamiaje contiene un origen de la replicación de 653 pb (pUC Ori) y un marcador de resistencia a la ampicilina de 1.144 pb (Amp-R) que permiten que el vector se propague y se seleccione para *E. coli*. El pPNP-1 contiene un gen de resistencia al cloranfenicol (Cm-R) de 911 pb y secuencias flanqueantes que dirigen la integración de este gen en el gen *pnp* endógeno. La secuencia flanqueante aguas arriba de 891 pb (SEQ ID NO: 56) consiste en una secuencia PNPI-L de 775 pb (región de codificación *pnp* (SEQ ID NO: 1) de nt-1.345 a nt-2.119), un segmento de 9 pb directamente invertido (DR: Direct Reversed), un elemento ME (extremo de mosaico) de 19 pb, y un elemento Lox de 34 pb. La secuencia flanqueante de 1.030 pb aguas abajo (SEQ ID NO: 57) consiste en un elemento ME de 19 pb, un segmento de 9 pb directamente invertido (DR) y una secuencia PNPI-R de 916 pb. La secuencia PNPI-R incluye las

últimas 116 pb de la región codificadora de pnp (de nt-2.129 a nt-2.244) y una secuencia de ADN de 800 pb aguas abajo de la secuencia codificadora de pnp. Tanto PNPI-L como PNPI-R fueron amplificados a partir de ADN genómico de *Z. mobilis* ZW1. En pPNP-I, las secuencias PNPI-L y PNPI-R dirigen la integración de la secuencia de ADN entre ellas en el genoma de *Z. mobilis* entre nt-2.119 y nt-2.129 de la región de codificación de pnp a través de una recombinación homóloga de doble entrecruzamiento. La integración interrumpe el gen pnp endógeno y es similar al genotipo de la cepa I descrita en el Ejemplo 3, con el transgén integrado como Cm-R en vez de Rpi :: Spec-R. Tiene por resultado un producto de proteína de fusión 723-aa pnp truncada como se describe en el Ejemplo 3, que es 25 aminoácidos más corto que el producto de proteína pnp de tipo silvestre (una polinucleótido fosforilasa de 748 aminoácidos; SEQ ID NO: 2). La proteína truncada comparte los primeros 709 restos de aminoácidos con el tipo silvestre, pero tiene una nueva secuencia de 14 aa unida al término C (SEQ ID NO: 9).

El pPNP-IN (figura 11B; SEQ ID NO: 58) es también un vector DCO basado en pUC18 con un tamaño de 6.471 pb. Se derivó directamente de pPNP-1 reemplazando PNPI-L por PNP-U y PNPI-R por PNP-D. PNP-U (SEQ ID NO: 59) es un fragmento de ADN genómico de 1.369 pb de *Z. mobilis* que consiste en los primeros 96 pb de la región codificadora de pnp y una secuencia de 1.273 pb aguas arriba de la región de codificación de pnp. PNP-D (SEQ ID NO: 60) es un fragmento de ADN genómico de 1.251 pb de *Z. mobilis* que incluye una parte de la secuencia de codificación de pnp, desde nt-97 a nt-1.347. En el vector pPNP-IN, PNP-U y PNP-D son fragmentos de recombinación homóloga que dirigen la integración de la secuencia entre dos elementos DR de 9 pb en el genoma de *Z. mobilis* entre nt-96 y nt-97 de la región codificadora de pnp. La integración interrumpe el gen pnp endógeno y da como resultado un producto de proteína pnp truncada con sólo 49 restos de aa (SEQ ID NO: 12), que es 699 aa más corto que el producto de proteína pnp de tipo silvestre. Esta proteína corta comparte los primeros 32 restos de aa con el tipo silvestre, y por tanto tiene 17 nuevos restos de aminoácidos unidos al término C.

El pPNP-C (SEQ ID NO: 61, Figura 11C) es un vector basado en pBluescript de 6.342 pb. Su andamiaje consiste en un sitio de replicación f1 (+) (f1 (+) Ori), un marcador de resistencia a la ampicilina (Amp-R) y un sitio de replicación pUC (pUC Ori) para propagación y selección en *E. coli*. En el vector, un gen de resistencia al cloranfenicol de 911 pb (Cm-R) está flanqueado por un fragmento PNP-L de 1.318 pb y un fragmento PNP-R de 1.225 pb. Los fragmentos PNP-L y PNP-R se amplificaron a partir de ADN genómico de *Z. mobilis* ZW1 y son idénticos a los de pZX6 descritos en el Ejemplo 4. El fragmento PNP-L es un segmento de la secuencia de codificación de pnp, de nt-767 a nt-2.084, mientras que el fragmento PNP-R incluye los últimos 56 pb (de nt-2.189 a nt-2.244) de la secuencia codificadora de pnp, su codón de parada y la secuencia aguas abajo de 1.166 pb. Por lo tanto, pPNP-C dirige la integración del marcador Cm-R en la región de codificación de pnp entre nt-2.085 y nt-2.188. Este sitio de integración está cerca del extremo 3' de la secuencia de codificación de pnp y 34 pb aguas arriba del sitio de integración diana de pPNP-1. La región de codificación de pnp truncada produce una proteína de 697 aa (SEQ ID NO: 10), que es 51 restos de aminoácidos más corta que el tipo silvestre y comparte 695 restos de aminoácidos con el tipo silvestre, con 2 nuevos aminoácidos en el término C.

Al igual que pPNP-C, el pPNP-M (SEQ ID NO: 62), Figura 11 D) es un vector basado en pBluescript de 6.322 pb, que también tiene una secuencia de cadena principal consistente en un f1 (+) Ori, un Amp-R, y un pUC Ori. Sin embargo, en pPNP-M, el transgén de resistencia a cloranfenicol de 911 (Cm-R) está flanqueado por un fragmento PNPm-L de 1.200 pb y un fragmento PNPm-R de 1.324 pb. Ambos fragmentos flanqueantes se amplificaron a partir del ADN genómico de *Z. mobilis* ZW1. PNPm-L (SEQ ID NO: 63) incluye una secuencia genómica de 96 pb aguas arriba de la secuencia codificadora de pnp y los primeros 1.104 pb de la secuencia codificadora de pnp (de nt-1 a nt-1.104), mientras que PNPm-R (SEQ ID NO: 64) incluye los últimos 1140 pb (de nt-1,105 a nt-2244) de la secuencia codificadora de pnp, su codón de parada y la secuencia aguas abajo de 181 pb. Por consiguiente, pPNP-M es capaz de dirigir la integración del marcador Cm-R en el gen pnp endógeno entre nt-1.104 y nt-1.105 en el genoma de *Z. mobilis*. El sitio de integración está cerca de la mitad de la secuencia de codificación de pnp y 1.015 pb aguas arriba del sitio de integración para pPNP-I. Tiene como resultado un producto de proteína pnp de 378-aa truncada (SEQ ID NO: 11), 370 aa más corto que el producto de proteína pnp de tipo silvestre. La proteína truncada comparte los primeros 368 restos de aa con el tipo silvestre, pero tiene una nueva secuencia de 10 aa unida en el término C.

Ejemplo 8

Interrupción del gen pnp endógeno en la cepa ZW1-X109

Para determinar si la interrupción del gen pnp endógeno beneficia la utilización de xilosa en *Z. mobilis*, se transformó ZW1-X109 por separado con pPNP-I, pPNP-C, pPNP-M y pPNP-IN como se describe en el Ejemplo 5. Dado que todos los vectores son vectores suicidas, los transformantes fueron seleccionados directamente en una placa MMG5-CM120 (MMG5 con 120 mg/L de cloranfenicol y 15 g/L de agar). Las cepas resultantes se denominaron ZW1-X109-PNPI, ZW1-X109-PNPc, ZW1-X109-PNPm y ZW1-X109-PNPIn. Unas pocas colonias de cada cepa se rasparon y se desarrollaron en placas MMG5-CM120 frescas. La integración de los transgenes se confirmó por inspección mediante PCR.

En la inspección mediante PCR se emplearon cinco cebadores. El cebador directo ara448 (SEQ ID NO: 52) es una secuencia situada en el comienzo del marcador Cm-R en todos los plásmidos pPNP. El cebador inverso ara339 (SEQ ID NO: 51) complementa un segmento de secuencia genómica de *Z. mobilis* aguas abajo del gen pnp. Se usaron para inspeccionar la recombinación por doble entrecruzamiento mediada por los siguientes fragmentos de

recombinación: PNPI-R en pPNP-I, PNP-R en pPNP-C, PNPm-R en pPNP-M y PNP-D en pPNP-IN. Si la integración era exitosa, estos dos cebadores amplificarían un producto de PCR de 2.393 pb de la cepa ZW1-X109-PNPI, 2.256 pb de la cepa ZW1-X109-PNPc, 3.361 de la cepa ZW1-X109-PNPm y 4.565 pb de la cepa ZW1-X109-PNPIn. El cebador directo ara340 o 4R0 (SEQ ID NO: 50 y 53, respectivamente) y el cebador inverso ara449 (SEQ ID NO: 54) se usaron para inspeccionar la recombinación de doble entrecruzamiento mediada por los siguientes fragmentos de recombinación: PNPI-L en pPNP-I, PNP-L en pPNP-C, PNPm-L en pPNP-M, y PNP-U en pPNP-IN. El cebador ara340 coincide con un segmento (de nt-702 a nt-724) de la secuencia de codificación de pnp situada aguas arriba de las secuencias de PNPI-L y PNP-L en el ADN genómico de *Z. mobilis*. El cebador 4R0 es un segmento de la secuencia genómica de *Z. mobilis* fuera del gen pnp y aguas arriba de las secuencias PNPm-L y PNP-U. El cebador ara449 complementa una secuencia en el extremo del marcador Cm-R en plásmidos pPNP. Por tanto, cuando la integración era exitosa, ara340 y ara449 podían amplificar un producto de PCR de 2.551 pb de la cepa ZW1-X109-PNPI y un producto de PCR de 2.306 pb de la cepa ZW1-X109-PNPc, mientras que 4R0 y ara449 podían amplificar un producto de PCR de 3.424 pb a partir de la cepa ZW1-X109-PNPm, y un producto de PCR de 2.430 pb de la cepa ZW1-X109-PNPIn. Las reacciones PCR estándar usando PCR Supermix de Invitrogen se realizaron directamente en cepas recién cultivadas. Los resultados demostraron una integración precisa.

Ejemplo 9

Caracterización de las cepas de integración de pnp

Las cepas ZW1-X109-PNPI, ZW1-X109-PNPc, ZW1-X109-PNPm y ZW1-X109-PNPIn fueron caracterizadas adicionalmente en fermentaciones en matraz de agitación para determinar su crecimiento y sus perfiles metabólicos. La cepa madre ZW1-X109, que contiene un gen pnp endógeno ininterrumpido, fue utilizada como control. La fermentación siguió un protocolo estándar como se describe en el Ejemplo 6, excepto que el volumen de cultivo de células se redujo de 20 mL a 10 mL, por lo cual se usaron tubos de fondo redondo Falcon de 14 mL tapados, con un orificio perforado en la parte superior, en vez de los tubos de centrifuga VWR de 45 mL. La primera fermentación en matraz de agitación se llevó a cabo en mRM3-G10. A las 0, 10 y 24 horas de fermentación se midió la DO₆₀₀ en un espectrofotómetro Shimadzu UV-1201, mientras que las concentraciones de glucosa y etanol se determinaron por un sistema de HPLC Agilent 1100 usando una columna de exclusión de iones Aminex HPX-A7H de BioRad. Los resultados se representan gráficamente en la Figura 12, indicando que las cuatro cepas de pnp interrumpido tienen un metabolismo basal de glucosa similar a la cepa madre ZW1-X109. Por ejemplo, después de 10 horas de fermentación, ZW1-X109-PNPI (A) utilizó aproximadamente 37,7% de glucosa (una reducción de 120,8 g/L a 75,3 g/L) para apoyar un título de etanol de 22,9 g/L y un crecimiento de biomasa a un valor de DO₆₀₀ de 5,36; ZW1-X109-PNPc (B) utilizó aproximadamente 37,7% de glucosa (una reducción de 120,8 g/L a 75,3 g/L) para apoyar un título de etanol de 23,0 g/L y un crecimiento de biomasa hasta un valor de DO₆₀₀ de 4,98; ZW1-X109-PNPm (C) utilizó aproximadamente 39,1% de glucosa (una reducción de 120,8 g/L a 73,6 g/L) para apoyar un título de etanol de 23,5 g/L y un crecimiento de biomasa hasta un valor de DO₆₀₀ de 5,14; ZW1-X109-PNPIn (D) utilizó aproximadamente 36,8% de glucosa (una reducción de 120,8 g/L a 76,4 g/L) para apoyar un título de etanol de 22,6 g/L y un crecimiento de biomasa hasta un valor DO₆₀₀ de 5,32. La cepa madre ZW1-X109 (E) utilizó aproximadamente 39,5% de glucosa (una reducción de 120,8 g/L a 73,4 g/L) para apoyar un título de etanol de 24,0 g/L y un crecimiento de biomasa hasta un valor DO₆₀₀ de 5,62. Después de 24 horas de fermentación, todas las cepas de suspensión con pnp interrumpido agotan la glucosa para apoyar un título de etanol próximo a 58,1 g/L y un crecimiento de biomasa hasta un valor de DO₆₀₀ de aproximadamente 7,90, mientras que ZW1-X109 también agotó la glucosa para apoyar un título de etanol de 57,3 g/L y un crecimiento de biomasa hasta un valor de DO₆₀₀ de 7,64.

La fermentación en matraz de agitación se llevó a cabo después en 10 mL de mRM3-X10 como se describió anteriormente. El valor de la DO₆₀₀ y las concentraciones de xilosa y etanol se midieron a las 0, 24, 48 y 72 horas. La Figura 13 muestra gráficas de los resultados. Muestra que las cuatro cepas de pnp interrumpido utilizaron xilosa en la fermentación mejor que la cepa madre ZW1-X109. Después de 48 horas de fermentación, ZW1-X109-PNPI (A) había utilizado ya aproximadamente el 95,5% de la xilosa (una reducción de 117,9 g/L a 5,3 g/L) para apoyar un título de etanol de 54,7 g/L y un crecimiento de la biomasa hasta un valor de DO₆₀₀ de 5,34. Después de otras 24 horas de fermentación, casi había consumido la xilosa dejando sólo 1,4 g/L para apoyar un título de etanol de 56,0 g/L y un crecimiento de la biomasa hasta un valor de DO₆₀₀ de 5,62. En el mismo período de 72 horas, ZW1-X109-PNPc (B) utilizó aproximadamente 94,2% de xilosa (una reducción de 117,9 g/L a 6,8 g/L) para apoyar un título de etanol de 53,6 g/L y un crecimiento de la biomasa hasta un valor de DO₆₀₀ de 4,64; ZW1-X109-PNPm (C) utilizó aproximadamente 83,3% de xilosa (una reducción de 117,9 g/L a 19,7 g/L) para apoyar un título de etanol de 47,5 g/L y un crecimiento de biomasa hasta un valor de DO₆₀₀ de 4,24; ZW1-X109-PNPIn (D) utilizó aproximadamente 79,8% de glucosa (una reducción de 117,9 g/L a 23,8 g/L) para apoyar un título de etanol de 45,5 g/L y un crecimiento de biomasa hasta un valor de DO₆₀₀ de 4,04. La cepa madre ZW1-X109 (E) utilizó aproximadamente 61,7% de xilosa (una reducción de 117,9 g/L a 45,1 g/L) para apoyar un título de etanol de 34,8 g/L y un crecimiento de biomasa de hasta un valor de DO₆₀₀ de 3,26. Los resultados también indican que las cepas con menos truncamiento de la proteína codificada pnp, tales como ZW1-X109-PNPI y ZW1-X109-PNPc, usan xilosa más eficazmente que las cepas con truncamientos más grandes, tales como ZW1-X109-PNPm y ZW1-X109-PNPIn. La clasificación de estas cepas por orden de su capacidad para utilizar xilosa, siendo la mejor la primera, fue: ZW1-X109-PNPI, ZW1-X109-PNPc, ZW1-X109-PNPm, ZW1-X109-PNPIn y finalmente ZW1-X109.

En resumen, este ejemplo demuestra que la integración en el gen *pnp* endógeno de *Z. mobilis* no afecta al metabolismo de la glucosa sino que mejora la utilización de xilosa en la fermentación.

Ejemplo 10 (comparativo)

Interrupción del gen *pnp* endógeno en ZW801-4 en ausencia de sobreexpresión de RPI

- 5 Los efectos de las modificaciones del gen *pnp* endógeno se ensayaron en ZW801-4 (véanse los Métodos Generales) usando los dos constructos suicidas que tenían el mayor efecto sobre el crecimiento y el metabolismo de la xilosa en ZW1-X109 (p. ej. pPNP-1 y pPNP-C, descritos en Ejemplo 7). ZW801-4 se transformó con cada uno de estos vectores suicidas por separado, como se describe en el Ejemplo 8. Dos transformantes por cada constructo se evaluaron en experimentos en frascos de agitación con xilosa, y se incluyó ZW801-4 como control o testigo (también hecho por duplicado). El medio de crecimiento era mRM3-X10 y la temperatura 33°C. Las muestras de cada cultivo se ensayaron en relación con DO₆₀₀, xilosa y etanol como se describió anteriormente. Los resultados de este experimento se muestran en la Figura 14. Solamente el transformante preparado con el constructo de pPNP-1 mostró una ligera mejora en comparación con el control. Obsérvese que la recombinación homóloga de doble entrecruzamiento del constructo suicida pPNP-1 con el cromosoma de ZW801-4 da como resultado exactamente la misma modificación del gen *pnp* que está presente en la cepa I.
- 10
- 15

Ejemplo 11

Ensayo de la actividad de xilosa isomerasa

- Se midió la actividad de xilosa isomerasa en ZW658 (véanse los Métodos Generales) en una reacción que contenía 20 µL de extracto libre de células (véanse los Métodos Generales), NADH 0,256 mM, xilosa 50 mM, MgSO₄ 10 mM, trietanolamina 10 mM y 1 U/mL de SDH (sorbitol deshidrogenasa) a 30°C. El valor de A₃₄₀ se obtuvo de un lector de placas durante 3 - 5 min. La actividad de XI se calculó de la forma siguiente:
- 20

1 unidad de XI corresponde a la formación de 1 µmol de D-xilulosa por minuto a 30°C

- U (µmoles/min) = pendiente (dA₃₄₀/min) * volumen de reacción (µL) / 6220/0,55 cm (moles de NADHP → NAD es 6220 A₃₄₀ por mol por L en una cubeta de 1 cm) (longitud de recorrido de 200 µL por pocillo en microplaca = 0,55 cm).
- 25

Actividad específica (µmoles/min - mg) = µmoles/min/concentración de proteína (mg)

La actividad medida para ZW658 fue 0,25 ± 0,033 µmoles de producto/mg de proteína/minuto.

Listado de secuencias

<110> E. I. du Pont de Nemours and Company
 <120> Modificación del gen pnp para utilización mejorada de xilosa en *zymomonas*
 <130> CL5440PCT

5 <160> 65
 <170> PatentIn versión 3.5

<210> 1
 <211> 2247
 <212> ADN

10 <213> *Zymomonas mobilis*

<400> 1

```

atgttcgata ttaaacgcca ggaaatcgat tggggcggaa aaaaactgac actggaaacc      60
ggacaggttg cccgtcaggc agatggcgcc gtcattgcga ccttaggtga aacggtcgta      120
ttatgcgcgg taacggcagc aaaaacggta aaagaaggtc aggatttctt tcctttgacc      180
gtccattatc aggaaaaata ttcagcagca ggccgtattc cgggtggctt tttcaagcgt      240
gaacgtggcg caaccgaacg ggaaacgctg atttcacggt taatcgaccg tccaatccgt      300
cctctgtttc cggaaggttt ctataacgaa accttgggtca ttgcgcaggt catgtcctat      360
gacggcgaga atgaaccgga tatcttggcg atgatcgctg cttctgcggc tcttgctctt      420
tccggtgtgc ctttcttggg ccccatcggg gctgcccggt tgggttatca agatggcgag      480
ttcattctta acccgacctt ggaacagctt gaaaaaagtg atcttgatct ggttgtcggg      540
gctacccgtg atgccgtgat gatggttgaa tcggaagcga atgagcttcc cgaagaagtc      600
atgtcfaatg ccgtttcttt tgcgcatgaa tctttacagc cggttatcaa agctatcacc      660
aatctggcag aacaggccgc taaagagcct tgggaactgg tcagctatga tgacagcgca      720
ttggctgcca aagtogaaga actctgctac gacaatttcg ataaggccta tcgtctgact      780
cgcaaggctg aacgtggtga cgccttgagc aaggccaaag cggttcttga cgaagccttc      840
ccagaagctg atccgacaga aaagctgcgc atccagaagc ttgcgaagaa gctggaagca      900
aaaatcgtcc gcaccgccat tctgaaagaa ggccggagaa ttgacggacg cgatctgaaa      960
acagttcgcc cgatccgctc tcaggttggg ttcttgcccc gcacgcatgg ttctgccctg     1020
tttacgcgtg gtgaaacaca ggctttggtt tcaaccaccc ttggaacggc ggatgctgaa     1080
cagatgatcg acggtttaac cggccttcat tatgaacgct tcatgctgca ttacaacttc     1140
ccccatatt cggtcgggtg agttggtcgt tttggtgctc cgggtcgtcg tgaatcggc     1200
    
```

ES 2 611 977 T3

catggtaaac tggcatggcg tgcgcttcat ccggttttgc cgagcaaggc tgatttcccg 1260
tataccatcc gtgttttgtc ggatatacacc gaatctaatag gttcctcttc catggcaacc 1320
gtttgcggtg gctgccttgc attgatggat gccgggtgtc ccttaacgcg tccggtttcc 1380
ggtatcgcca tgggtcttat tctggaaaaa gacggcttcg ctattttgtc cgatatcatg 1440
ggtgatgaag atcacttggg tgatatggac ttttaagtcg ccggtaccga aaaaggtatc 1500
accagcctcc agatggacat caaggttgct ggcattaccg aagaaatcat gcagaaagct 1560
ttggaacagg ctaaaggtgg ccgtgctcat atcttgggtg aaatgtccaa agcgctgggt 1620
gaagtccgct ccgaaatttc taatttggca ccgcgcattg aaacaatgag cgtacaaaaa 1680
gacaaaatcc gtgatgttat cggaacgggc ggaaaagtta tccgtgaaat cgtggcgacc 1740
acaggtgcca aggtcgatat cgaagatgac ggcacggttc gtctgtcttc ttctgatccg 1800
gccaatattg aagcagcccg tgaatggatc aatggatttg ttgaagaacc ggaagtaggc 1860
aaaatctata acggtaaagt cgtcaatata gttgatttcg gtgccttcgt aaacttcatg 1920
ggtggccgtg acggcttggg acatgtttcg gaaatcaaga acgaacgtgt caacaaggtc 1980
agcgatgtcc tgtccgaagg tcaggaagtc aaagtcaagg ttcttgaaat tgacaaccgt 2040
ggcaaggttc gcctgtctat gcgtgttgc gatcaggaaa ccggcgcaga gctggatgat 2100
aaccgtccgc cacgtgagaa cgcagaacgt ccgggtgggt agcgtcctcg tcgtgatcgg 2160
ggcctcgtc gggaatctgg cgatcgtccg gcaagacgtg atatggaacc ggaatttgc 2220
ccggcattcc tgcgcaaaga tagctaa 2247

<210> 2

<211> 748

<212> PRT

5 <213> Zymomonas mobilis

<400> 2

Met Phe Asp Ile Lys Arg Gln Glu Ile Asp Trp Gly Gly Lys Lys Leu
1 5 10 15

Thr Leu Glu Thr Gly Gln Val Ala Arg Gln Ala Asp Gly Ala Val Ile
20 25 30

Ala Thr Leu Gly Glu Thr Val Val Leu Cys Ala Val Thr Ala Ala Lys
35 40 45

Thr Val Lys Glu Gly Gln Asp Phe Phe Pro Leu Thr Val His Tyr Gln
50 55 60

Glu Lys Tyr Ser Ala Ala Gly Arg Ile Pro Gly Gly Phe Phe Lys Arg
65 70 75 80

ES 2 611 977 T3

Glu Arg Gly Ala Thr Glu Arg Glu Thr Leu Ile Ser Arg Leu Ile Asp
 85 90 95

Arg Pro Ile Arg Pro Leu Phe Pro Glu Gly Phe Tyr Asn Glu Thr Leu
 100 105 110

Val Ile Ala Gln Val Met Ser Tyr Asp Gly Glu Asn Glu Pro Asp Ile
 115 120 125

Leu Ala Met Ile Ala Ala Ser Ala Ala Leu Ala Leu Ser Gly Val Pro
 130 135 140

Phe Leu Gly Pro Ile Gly Ala Ala Arg Val Gly Tyr Gln Asp Gly Glu
 145 150 155 160

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

Leu Val Val Gly Ala Thr Arg Asp Ala Val Met Met Val Glu Ser Glu
 180 185 190

Ala Asn Glu Leu Pro Glu Glu Val Met Leu Asn Ala Val Ser Phe Ala
 195 200 205

His Glu Ser Leu Gln Pro Val Ile Lys Ala Ile Ile Asn Leu Ala Glu
 210 215 220

Gln Ala Ala Lys Glu Pro Trp Glu Leu Val Ser Tyr Asp Asp Ser Ala
 225 230 235 240

Leu Ala Ala Lys Val Glu Glu Leu Cys Tyr Asp Asn Phe Asp Lys Ala
 245 250 255

Tyr Arg Leu Thr Arg Lys Ala Glu Arg Val Asp Ala Leu Ser Lys Ala
 260 265 270

Lys Ala Val Leu Asp Glu Ala Phe Pro Glu Ala Asp Pro Thr Glu Lys
 275 280 285

Leu Arg Ile Gln Lys Leu Ala Lys Lys Leu Glu Ala Lys Ile Val Arg
 290 295 300

Thr Ala Ile Leu Lys Glu Gly Arg Arg Ile Asp Gly Arg Asp Leu Lys
 305 310 315 320

Thr Val Arg Pro Ile Arg Ser Gln Val Gly Phe Leu Pro Arg Thr His
 325 330 335

ES 2 611 977 T3

Gly Ser Ala Leu Phe Thr Arg Gly Glu Thr Gln Ala Leu Val Ser Thr
 340 345 350
 Thr Leu Gly Thr Ala Asp Ala Glu Gln Met Ile Asp Gly Leu Thr Gly
 355 360 365
 Leu His Tyr Glu Arg Phe Met Leu His Tyr Asn Phe Pro Pro Tyr Ser
 370 375 380
 Val Gly Glu Val Gly Arg Phe Gly Ala Pro Gly Arg Arg Glu Ile Gly
 385 390 395 400
 His Gly Lys Leu Ala Trp Arg Ala Leu His Pro Val Leu Pro Ser Lys
 405 410 415
 Ala Asp Phe Pro Tyr Thr Ile Arg Val Leu Ser Asp Ile Thr Glu Ser
 420 425 430
 Asn Gly Ser Ser Ser Met Ala Thr Val Cys Gly Gly Cys Leu Ala Leu
 435 440 445
 Met Asp Ala Gly Val Pro Leu Thr Arg Pro Val Ser Gly Ile Ala Met
 450 455 460
 Gly Leu Ile Leu Glu Lys Asp Gly Phe Ala Ile Leu Ser Asp Ile Met
 465 470 475 480
 Gly Asp Glu Asp His Leu Gly Asp Met Asp Phe Lys Val Ala Gly Thr
 485 490 495
 Glu Lys Gly Ile Thr Ser Leu Gln Met Asp Ile Lys Val Ala Gly Ile
 500 505 510
 Thr Glu Glu Ile Met Gln Lys Ala Leu Glu Gln Ala Lys Gly Gly Arg
 515 520 525
 Ala His Ile Leu Gly Glu Met Ser Lys Ala Leu Gly Glu Val Arg Ser
 530 535 540
 Glu Ile Ser Asn Leu Ala Pro Arg Ile Glu Thr Met Ser Val Pro Lys
 545 550 555 560
 Asp Lys Ile Arg Asp Val Ile Gly Thr Gly Gly Lys Val Ile Arg Glu
 565 570 575
 Ile Val Ala Thr Thr Gly Ala Lys Val Asp Ile Glu Asp Asp Gly Thr

ES 2 611 977 T3

580	585	590
Val Arg Leu Ser Ser Ser Asp Pro Ala Asn Ile Glu Ala Ala Arg Glu 595 600 605		
Trp Ile Asn Gly Ile Val Glu Glu Pro Glu Val Gly Lys Ile Tyr Asn 610 615 620		
Gly Lys Val Val Asn Ile Val Asp Phe Gly Ala Phe Val Asn Phe Met 625 630 635 640		
Gly Gly Arg Asp Gly Leu Val His Val Ser Glu Ile Lys Asn Glu Arg 645 650 655		
Val Asn Lys Val Ser Asp Val Leu Ser Glu Gly Gln Glu Val Lys Val 660 665 670		
Lys Val Leu Glu Ile Asp Asn Arg Gly Lys Val Arg Leu Ser Met Arg 675 680 685		
Val Val Asp Gln Glu Thr Gly Ala Glu Leu Asp Asp Asn Arg Pro Pro 690 695 700		
Arg Glu Asn Ala Glu Arg Arg Gly Gly Glu Arg Pro Arg Arg Asp Arg 705 710 715 720		
Gly Pro Arg Arg Glu Ser Gly Asp Arg Pro Ala Arg Arg Asp Met Glu 725 730 735		
Pro Glu Phe Ala Pro Ala Phe Leu Arg Lys Asp Ser 740 745		

<210> 3

<211> 2247

<212> ADN

5 <213> Zymomonas mobilis

<400> 3

atgttcgata ttaaacgcca ggaaatcgat tggggcgcaa aaaaactgac actggaaacc	60
ggacaggttg cccgtcaggc agatggcgcc gtcattgcga ccttaggtga aacggtcgta	120
ttatgcgcg taacggcagc aaaaacggta aaagaaggtc aggatttctt tcctttgacc	180
gtccattatc aggaaaaata ttcagcagca ggccgtattc ccggtggctt tttcaagcgt	240
gaacgtggcg caaccgaacg ggaaacgctg atttcacggt taatcgaccg tccaatccgt	300
cctctgtttc cggaaggttt ctataacgaa accttggtca ttgcgcaggt catgtcctat	360
gacggcgaga atgaaccgga tatcttgggc atgatcgccg cttctgcggc ccttgccttt	420

ES 2 611 977 T3

tccgggtgtgc ctttccttgg tcctatcggg gctgcccgtg tggggttatca agatggcgag 480
 ttcattctta acccgacctt ggaacagctt gaaaaaagtg atcttgatct ggttgtcggg 540
 gctacccgtg atgccgtgat gatggttgaa tcggaagcga atgagcttcc cgaagaagtc 600
 atgctcaatg ctgtttcttt tgcgcatgaa tctttacagc cggttatcaa agctatcatc 660
 aatctggcag aacaggccgc taaagagcct tgggaactgg tcagctatga tgacagcgca 720
 ttggctgcc aagtogaaga actctgctac gacaatttcg ataaggccta tcgtctgact 780
 cgcaaggctg aacgtgttga cgccttgagc aaggccaaag cggttcttga cgaagccttc 840
 ccagaagctg atccgacaga aaagctgccc atccagaagc ttgcgaagaa gctggaagca 900
 aaaattgtcc gcaccgccat tctgaaagaa ggccggagaa ttgacggacg cgatctgaaa 960
 acagttcgcc cgatccgctc tcaggttggg ttcttgcccc gcacgcatgg ttctgccctg 1020
 tttacgcgtg gtgaaacaca ggctttgggt tcaaccacco ttggaacggc ggatgctgaa 1080
 cagatgatcg acggtttaac cggccttcat tatgaacgct tcatgctgca ttataacttc 1140
 cctccttatt cggtcgggtg agttggctgt tttggctgctc cgggtcgtcg tgaaatcggc 1200
 catggtaaac tggcatggcg tgcgcttcat ccggttttgc cgagcaaggc tgatttcccg 1260
 tataccatcc gcgttttgtc ggatatacacc gaatctaag gttcctcttc tatggcaacc 1320
 gtttgccgtg gctgccttgc attgatggat gccggtgttc ccttaacgcg tccggtttcc 1380
 ggtatcgcca tgggtcttat tctagaaaaa gacggcttcg ctattttgtc cgatatcatg 1440
 ggtgatgaag atcacttggg tgatatggac tttaaggctg ccggtaccga aaaaggatc 1500
 accagcctcc agatggacat caaggttgct ggcatcacc aagaaatcat gcagaaagct 1560
 ttggaacagg ctaaagggtg ccgtgctcat atcttgggtg aaatgtccaa agcgtgggt 1620
 gaagtcgct ccgaaatttc taatttggca ccgcgcattg aaacgatgag cgtacaaaaa 1680
 gacaaaatcc gtgatgttat cggaacgggc ggaaaagtta tccgtgaaat cgtggcaacc 1740
 acaggtgcc aagtcgatat cgaagatgac ggaacggctc gtctgtcttc ttccgatcct 1800
 gccaatattg aagcagcccg tgaatggatc aatggtattg ttgaagaacc ggaagtaggc 1860
 aaaatctata acggtaaagt cgtcaatata gttgatctcg gtgccttcgt aaacttcag 1920
 ggccggccgtg acggcttggg acatgtttcg gaaatcaaga acgaacgtgt caacaaggctc 1980
 agcgatgtcc tgtctgaagg tcaggaagtc aaagtcaagg ttcttgaaat tgacaaccgt 2040
 ggcaaggctc gcctgtctat gcgtgttgc gatcaggaaa ccggcgcaga gctggatgat 2100
 aaccgtccgc cacgtgagaa cgcagaacgt ccgggtggtg agcgtcctcg tcgtgatcgg 2160
 ggccctcgtc gggaaatctgg cgatcgtccg gcaagacgtg acatggaacc ggaatttgc 2220
 ccggcattcc tgcgcaaaga tagctaa 2247

<210> 4
 <211> 748
 <212> PRT
 <213> *Zymomonas mobilis*
 <400> 4

5

ES 2 611 977 T3

Met Phe Asp Ile Lys Arg Gln Glu Ile Asp Trp Gly Gly Lys Lys Leu
 1 5 10 15

Thr Leu Glu Thr Gly Gln Val Ala Arg Gln Ala Asp Gly Ala Val Ile
 20 25 30

Ala Thr Leu Gly Glu Thr Val Val Leu Cys Ala Val Thr Ala Ala Lys
 35 40 45

Thr Val Lys Glu Gly Gln Asp Phe Phe Pro Leu Thr Val His Tyr Gln
 50 55 60

Glu Lys Tyr Ser Ala Ala Gly Arg Ile Pro Gly Gly Phe Phe Lys Arg
 65 70 75 80

Glu Arg Gly Ala Thr Glu Arg Glu Thr Leu Ile Ser Arg Leu Ile Asp
 85 90 95

Arg Pro Ile Arg Pro Leu Phe Pro Glu Gly Phe Tyr Asn Glu Thr Leu
 100 105 110

Val Ile Ala Gln Val Met Ser Tyr Asp Gly Glu Asn Glu Pro Asp Ile
 115 120 125

Leu Ala Met Ile Ala Ala Ser Ala Ala Leu Ala Leu Ser Gly Val Pro
 130 135 140

Phe Leu Gly Pro Ile Gly Ala Ala Arg Val Gly Tyr Gln Asp Gly Glu
 145 150 155 160

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

Leu Val Val Gly Ala Thr Arg Asp Ala Val Met Met Val Glu Ser Glu
 180 185 190

Ala Asn Glu Leu Pro Glu Glu Val Met Leu Asn Ala Val Ser Phe Ala
 195 200 205

His Glu Ser Leu Gln Pro Val Ile Lys Ala Ile Ile Asn Leu Ala Glu
 210 215 220

ES 2 611 977 T3

Gln Ala Ala Lys Glu Pro Trp Glu Leu Val Ser Tyr Asp Asp Ser Ala
 225 230 235 240

Leu Ala Ala Lys Val Glu Glu Leu Cys Tyr Asp Asn Phe Asp Lys Ala
 245 250 255

Tyr Arg Leu Thr Arg Lys Ala Glu Arg Val Asp Ala Leu Ser Lys Ala
 260 265 270

Lys Ala Val Leu Asp Glu Ala Phe Pro Glu Ala Asp Pro Thr Glu Lys
 275 280 285

Leu Arg Ile Gln Lys Leu Ala Lys Lys Leu Glu Ala Lys Ile Val Arg
 290 295 300

Thr Ala Ile Leu Lys Glu Gly Arg Arg Ile Asp Gly Arg Asp Leu Lys
 305 310 315 320

Thr Val Arg Pro Ile Arg Ser Gln Val Gly Phe Leu Pro Arg Thr His
 325 330 335

Gly Ser Ala Leu Phe Thr Arg Gly Glu Thr Gln Ala Leu Val Ser Thr
 340 345 350

Thr Leu Gly Thr Ala Asp Ala Glu Gln Met Ile Asp Gly Leu Thr Gly
 355 360 365

Leu His Tyr Glu Arg Phe Met Leu His Tyr Asn Phe Pro Pro Tyr Ser
 370 375 380

Val Gly Glu Val Gly Arg Phe Gly Ala Pro Gly Arg Arg Glu Ile Gly
 385 390 395 400

His Gly Lys Leu Ala Trp Arg Ala Leu His Pro Val Leu Pro Ser Lys
 405 410 415

Ala Asp Phe Pro Tyr Thr Ile Arg Val Leu Ser Asp Ile Thr Glu Ser
 420 425 430

Asn Gly Ser Ser Ser Met Ala Thr Val Cys Gly Gly Cys Leu Ala Leu
 435 440 445

Met Asp Ala Gly Val Pro Leu Thr Arg Pro Val Ser Gly Ile Ala Met
 450 455 460

Gly Leu Ile Leu Glu Lys Asp Gly Phe Ala Ile Leu Ser Asp Ile Met
 465 470 475 480

ES 2 611 977 T3

Gly Asp Glu Asp His Leu Gly Asp Met Asp Phe Lys Val Ala Gly Thr
 485 490 495

Glu Lys Gly Ile Thr Ser Leu Gln Met Asp Ile Lys Val Ala Gly Ile
 500 505 510

Thr Glu Glu Ile Met Gln Lys Ala Leu Glu Gln Ala Lys Gly Gly Arg
 515 520 525

Ala His Ile Leu Gly Glu Met Ser Lys Ala Leu Gly Glu Val Arg Ser
 530 535 540

Glu Ile Ser Asn Leu Ala Pro Arg Ile Glu Thr Met Ser Val Pro Lys
 545 550 555 560

Asp Lys Ile Arg Asp Val Ile Gly Thr Gly Gly Lys Val Ile Arg Glu
 565 570 575

Ile Val Ala Thr Thr Gly Ala Lys Val Asp Ile Glu Asp Asp Gly Thr
 580 585 590

Val Arg Leu Ser Ser Ser Asp Pro Ala Asn Ile Glu Ala Ala Arg Glu
 595 600 605

Trp Ile Asn Gly Ile Val Glu Glu Pro Glu Val Gly Lys Ile Tyr Asn
 610 615 620

Gly Lys Val Val Asn Ile Val Asp Phe Gly Ala Phe Val Asn Phe Met
 625 630 635 640

Gly Gly Arg Asp Gly Leu Val His Val Ser Glu Ile Lys Asn Glu Arg
 645 650 655

Val Asn Lys Val Ser Asp Val Leu Ser Glu Gly Gln Glu Val Lys Val
 660 665 670

Lys Val Leu Glu Ile Asp Asn Arg Gly Lys Val Arg Leu Ser Met Arg
 675 680 685

Val Val Asp Gln Glu Thr Gly Ala Glu Leu Asp Asp Asn Arg Pro Pro
 690 695 700

Arg Glu Asn Ala Glu Arg Arg Gly Gly Glu Arg Pro Arg Arg Asp Arg
 705 710 715 720

Gly Pro Arg Arg Glu Ser Gly Asp Arg Pro Ala Arg Arg Asp Met Glu
 725 730 735

Pro Glu Phe Ala Pro Ala Phe Leu Arg Lys Asp Ser
 740 745

<210> 5
 <211> 2247

ES 2 611 977 T3

<212> ADN

<213> Zymomonas mobilis

<400> 5

atgttcgata ttaaacgccca ggaaatcgat tggggcggga aaaaactgac actggaaacc	60
ggacaggttg cccgtcaggc agatggcgcc gtcattgcga ccttaggtga aacggtcgta	120
ttatgcgcgg taacggcagc aaaaacggta aaagaaggtc aggatttctt tcctttgacc	180
gtccattatc aggaaaaata ttcagcagca ggccgtatc cgggtggctt tttcaagcgt	240
gaacgtggcg caaccgaacg ggaaacgctg atttcacggt taatcgaccg tccaatccgt	300
cctctgtttc cggaaggttt ctataacgaa accttggtca ttgcgcaggt catgtcctat	360
gacggcgaga atgaaccgga tatcttggcg atgatcgccg cttctgcggc ccttgctctt	420
tccgggtgtc ctttccttgg tcccatcggg gctgcccgtg tgggttatca agatggcgag	480
ttcattctta acccgacctt ggaacagctt gaaaaaagtg atcttgatct ggttgtcggg	540
gctacccgtg atgccgtgat gatggttgaa tcggaagcga atgagcttcc ogaagaagtc	600
atgctcaatg ccgtttcttt tgcgcatgaa tctttacagc cggttatcaa agctatcatc	660
aatctggcag aacaggccgc taaagagcct tgggaactgg tcagctatga tgacagcgca	720
ttggctgcca aagtccaaga actctgctac gacaatttcc ataaggccta tcgtctgact	780
cgtaaggctg agcgtgttga agccttgagc aaggccaaag cggttcttga cgaagccttc	840
ccagaagctg atccgacaga aaagctgcgt atccagaagc tcgcaagaa gctggaagca	900
aaaatcgtcc gcaccgccat tctgaaagaa ggccggagaa ttgacggacg cgatctgaaa	960
acagttcgcc cgatccgctc tcaggttgga ttcttgcccc gcacgcatgg ttctgctctg	1020
tttacgcgcg gtgaaacaca ggctttggtt tcaaccaccc ttggaacggc ggatgctgaa	1080
cagatgatcg acggtttaac cggccttcat tatgaacgct tcatgctgca ttataacttc	1140
cctccttatt cggtcgggtga agttggtcgt tttggggctc cgggtcgtcg tgaaatcggc	1200
catggtaaac tggcatggcg tgcgcttcat ccggttttgc cgagcaaggc tgatttcccg	1260
tataccatcc gcgttttgtc ggatatcacc gaatctaata gttcctcttc catggcaacc	1320
gtttgcggtg gctgccttgc cttgatggat gccggtgttc ccttaacgcg tccggtttcc	1380
ggtatcgcca tgggtcttat tctggaaaaa gacggcttcg ctattttgtc ggatatcatg	1440
ggtgatgaag atcacttggg tgatatggac ttaaggctc cgggtaccga aaaaggatc	1500
accagcctcc agatggacat caaggttgct ggcattaccg aagaaatcat gcagaaagct	1560

ES 2 611 977 T3

ttggaacagg ctaaagggtg ccgtgctcat atcttgggtg aaatgtccaa agcgctgggt 1620
 gaagtccgct ccgaaatttc taatttggca ccgcgcattg aaacaatgag cgtaccaaaa 1680
 gacaaaatcc gtgatgttat cggaacgggc ggaaaagtta tccgtgaaat cgtggcgacc 1740
 acaggtgcca aggtcgatat cgaagatgac ggacacggttc gtctgtcttc ttccgatccg 1800
 gccaatattg aagcagcccg tgaatggatc aatggtattg ttgaagaacc ggaagtaggc 1860
 aaaatctata acggtaaagt cgtcaatata gttgatttcg gtgccttcgt aaacttcatg 1920
 ggtggccgtg acggcttggg acatgtttcg gaaatcaaga acgaacgtgt caacaaggtc 1980
 agcgatgtcc tgtctgaagg tcaggaagtc aaagtcaagg ttcttgaaat tgacaaccgt 2040
 ggcaaggttc gcctgtctat gcgtgttgc gatcaggaaa ccggcgcaga gctggatgat 2100
 aaccgtccgc cacgtgagaa cgcagaacgt cgcggtggtg agcgtcctcg tcgtgatcgg 2160
 ggccctcgtc gggaatctgg cgatcgccg gcaagacgtg acatggaacc ggaatttgct 2220
 ccggcattcc tgcgcaaaga tagctaa 2247

<210> 6

<211> 748

<212> PRT

5 <213> Zymomonas mobilis

<400> 6

Met Phe Asp Ile Lys Arg Gln Glu Ile Asp Trp Gly Gly Lys Lys Leu
 1 5 10 15

Thr Leu Glu Thr Gly Gln Val Ala Arg Gln Ala Asp Gly Ala Val Ile
 20 25 30

Ala Thr Leu Gly Glu Thr Val Val Leu Cys Ala Val Thr Ala Ala Lys
 35 40 45

Thr Val Lys Glu Gly Gln Asp Phe Phe Pro Leu Thr Val His Tyr Gln
 50 55 60

Glu Lys Tyr Ser Ala Ala Gly Arg Ile Pro Gly Gly Phe Phe Lys Arg
 65 70 75 80

Glu Arg Gly Ala Thr Glu Arg Glu Thr Leu Ile Ser Arg Leu Ile Asp
 85 90 95

Arg Pro Ile Arg Pro Leu Phe Pro Glu Gly Phe Tyr Asn Glu Thr Leu
 100 105 110

Val Ile Ala Gln Val Met Ser Tyr Asp Gly Glu Asn Glu Pro Asp Ile
 115 120 125

ES 2 611 977 T3

Leu Ala Met Ile Ala Ala Ser Ala Ala Leu Ala Leu Ser Gly Val Pro
 130 135 140

Phe Leu Gly Pro Ile Gly Ala Ala Arg Val Gly Tyr Gln Asp Gly Glu
 145 150 155 160

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

Leu Val Val Gly Ala Thr Arg Asp Ala Val Met Met Val Glu Ser Glu
 180 185 190

Ala Asn Glu Leu Pro Glu Glu Val Met Leu Asn Ala Val Ser Phe Ala
 195 200 205

His Glu Ser Leu Gln Pro Val Ile Lys Ala Ile Ile Asn Leu Ala Glu
 210 215 220

Gln Ala Ala Lys Glu Pro Trp Glu Leu Val Ser Tyr Asp Asp Ser Ala
 225 230 235 240

Leu Ala Ala Lys Val Glu Glu Leu Cys Tyr Asp Asn Phe Asp Lys Ala
 245 250 255

Tyr Arg Leu Thr Arg Lys Ala Glu Arg Val Glu Ala Leu Ser Lys Ala
 260 265 270

Lys Ala Val Leu Asp Glu Ala Phe Pro Glu Ala Asp Pro Thr Glu Lys
 275 280 285

Leu Arg Ile Gln Lys Leu Ala Lys Lys Leu Glu Ala Lys Ile Val Arg
 290 295 300

Thr Ala Ile Leu Lys Glu Gly Arg Arg Ile Asp Gly Arg Asp Leu Lys
 305 310 315 320

Thr Val Arg Pro Ile Arg Ser Gln Val Gly Phe Leu Pro Arg Thr His
 325 330 335

Gly Ser Ala Leu Phe Thr Arg Gly Glu Thr Gln Ala Leu Val Ser Thr
 340 345 350

Thr Leu Gly Thr Ala Asp Ala Glu Gln Met Ile Asp Gly Leu Thr Gly
 355 360 365

Leu His Tyr Glu Arg Phe Met Leu His Tyr Asn Phe Pro Pro Tyr Ser

ES 2 611 977 T3

Gly Lys Val Val Asn Ile Val Asp Phe Gly Ala Phe Val Asn Phe Met
625 630 635 640

Gly Gly Arg Asp Gly Leu Val His Val Ser Glu Ile Lys Asn Glu Arg
645 650 655

Val Asn Lys Val Ser Asp Val Leu Ser Glu Gly Gln Glu Val Lys Val
660 665 670

Lys Val Leu Glu Ile Asp Asn Arg Gly Lys Val Arg Leu Ser Met Arg
675 680 685

Val Val Asp Gln Glu Thr Gly Ala Glu Leu Asp Asp Asn Arg Pro Pro
690 695 700

Arg Glu Asn Ala Glu Arg Arg Gly Gly Glu Arg Pro Arg Arg Asp Arg
705 710 715 720

Gly Pro Arg Arg Glu Ser Gly Asp Arg Pro Ala Arg Arg Asp Met Glu
725 730 735

Pro Glu Phe Ala Pro Ala Phe Leu Arg Lys Asp Ser
740 745

<210> 7

<211> 2265

<212> ADN

5 <213> Zymomonas mobilis

<400> 7

```

atgttcgata ttaaacgcca ggaaatogat tggggcggaa agaagctgac gctggaaacc      60
ggacaggttg cccgtcaggc agatggcgcc gtcacgcgca ccttaggtga aacggtcgtg      120
ctttgcgcgg taacagcggc caaaacgggt aaagagggac aggatTTTTT ccctttaacg      180
gttcattatc aagaaaaata ttcagcagcc ggtcgtatcc ccggtggctt tttcaagcgt      240
gaacgtggtg ctagcgaacg cgaaactttg gtttcacgct tgattgatcg tccaattcgc      300
ccccttttcc cggacggttt ttataacgaa accttactta tcgctcaggt catgtcttat      360
gacggcgaaa atgaacctga catcttagcc atgattgcgg cctcggctgc tcttgcgctt      420
tccggtgtgc ctttcttggg cccaattggt gctgcgcgtg ttggctatca agatggcgaa      480
tatattttaa atccgacctt ggctcagctc gaaaacagcg atcttgatct ggtagtcggt      540
gcaacgcgcg atgcogtgat gatggttgaa tcggaagcaa aagagctatc cgaagaaatc      600
atgcttgatg cggtttcctt tgcgcataaa tctttacagc ctgttatcaa ggogatcatc      660
aatcttgccg agcaagccgc gaaagaaccg tgggagctct caagctatga tgacacagct      720
ttggctgcaa aagttgaaga actttgcaaa gatagccttg ataaggccta tcgtctgacc      780

```

ES 2 611 977 T3

aaaaaaagtg aacgtgtaga ggctatcttct aaggccaaag ccgttttgga tgaagctttc 840
 cccgatgctg atgcctcgga aaaactacgc attcagaaat tggcgaaaaa acttgaagcc 900
 aaaattgttc gcaactgcat cttaaagaa ggctcgtcga ttgatggtcg tgatctaaaa 960
 acggttcgct ctatccgttc acaggttggg ttcttacctc ggacccatgg gtctgcgctc 1020
 tttacgcggg gtgaaaccca agccttgggt tccacaacct ttggaaccgc agatgctgag 1080
 caaatgattg atggcctgaa tggccttcat tacgaacgct ttatgctgca ttataacttc 1140
 ccaccttatt ccggttggga agtgggtcgt tttggcgctc ctggccgctg tgaaatcggg 1200
 catggtaaac tggcatggcg tgctttacat cctgtgcttc ctagcaaggc tgacttcctc 1260
 tatacgatcc gcgttctatc cgatattacg gaatcaaacg gttcttcctc gatggcaacg 1320
 gtctgcggtg gctgtcttgc tttgatggat gcgggcgttc ccttgaagcg tccggctctc 1380
 ggcattgcga tgggccttat tcttgaaaaa gatggttttg ccattcttcc cgatattatg 1440
 ggtgatgaag atcacttagg ggatatggac tttaaggtag ccggtacaga agaaggcatt 1500
 accagccttc agatggacat taaggttgct ggtatcactg aagaaatcat gggtaaggct 1560
 ttggaacagg caaaagccgg ccgtgcccat attttgggtg aaatgtccaa agctttgggt 1620
 gaagtctgct cggaactttc gaatttagcg cctcgtattg aaacaatgag cgttcctaaa 1680
 gacaaaattc gtgatgttat tggaaactggc ggtaaagtca ttcgtgagat tgttgcgaca 1740
 accggcgcga aagttgacat tgaagatgac ggcaccgtac gcttgtcttc ttctgatccg 1800
 gctcagatcg aagctgcccg taattggatt accggtatca tcgaagaacc ggaagtcggc 1860
 aaaatttata acggttaagg tgtaaacatt gttgatttcg gtgcctttgt gaatttcag 1920
 ggtggccgtg atggtctggt tcacgtctct gaaattaaaa acgagcgcgt gaacaaggctc 1980
 agtgacgttc tggccgaagg ccaggaagtt aaggttaagg tgcttgaat tgacaatcgc 2040
 ggtaaagtcc gcttgtcaat gcgtgttgc gatcaggaaa ctggcgcgga actggaagac 2100
 aatcgtccgc ctagggaagc tcgtgaagtc agtgatcgcg gtccacgggg tgatcggcct 2160
 cgtcgcgatc gtggcccacg tcgcgaaccg cagaatggtt caaaccttc aggccgtgat 2220
 atggaacccg aatttgctcc ggctttttta cgaaaagatg attaa 2265

<210> 8

<211> 754

<212> PRT

5 <213> Zymomonas mobilis

<400> 8

Met Phe Asp Ile Lys Arg Gln Glu Ile Asp Trp Gly Gly Lys Lys Leu
 1 5 10 15

ES 2 611 977 T3

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

ES 2 611 977 T3

Lys Ala Val Leu Asp Glu Ala Phe Pro Asp Ala Asp Ala Ser Glu Lys
 275 280 285

Leu Arg Ile Gln Lys Leu Ala Lys Lys Leu Glu Ala Lys Ile Val Arg
 290 295 300

Thr Ala Ile Leu Lys Glu Gly Arg Arg Ile Asp Gly Arg Asp Leu Lys
 305 310 315 320

Thr Val Arg Pro Ile Arg Ser Gln Val Gly Phe Leu Pro Arg Thr His
 325 330 335

Gly Ser Ala Leu Phe Thr Arg Gly Glu Thr Gln Ala Leu Val Ser Thr
 340 345 350

Thr Leu Gly Thr Ala Asp Ala Glu Gln Met Ile Asp Gly Leu Asn Gly
 355 360 365

Leu His Tyr Glu Arg Phe Met Leu His Tyr Asn Phe Pro Pro Tyr Ser
 370 375 380

Val Gly Glu Val Gly Arg Phe Gly Ala Pro Gly Arg Arg Glu Ile Gly
 385 390 395 400

His Gly Lys Leu Ala Trp Arg Ala Leu His Pro Val Leu Pro Ser Lys
 405 410 415

Ala Asp Phe Pro Tyr Thr Ile Arg Val Leu Ser Asp Ile Thr Glu Ser
 420 425 430

Asn Gly Ser Ser Ser Met Ala Thr Val Cys Gly Gly Cys Leu Ala Leu
 435 440 445

Met Asp Ala Gly Val Pro Leu Lys Arg Pro Val Ser Gly Ile Ala Met
 450 455 460

Gly Leu Ile Leu Glu Lys Asp Gly Phe Ala Ile Leu Ser Asp Ile Met
 465 470 475 480

Gly Asp Glu Asp His Leu Gly Asp Met Asp Phe Lys Val Ala Gly Thr
 485 490 495

Glu Glu Gly Ile Thr Ser Leu Gln Met Asp Ile Lys Val Ala Gly Ile
 500 505 510

Thr Glu Glu Ile Met Gly Lys Ala Leu Glu Gln Ala Lys Ala Gly Arg
 515 520 525

ES 2 611 977 T3

Ala His Ile Leu Gly Glu Met Ser Lys Ala Leu Gly Glu Val Arg Ser
530 535 540

Glu Leu Ser Asn Leu Ala Pro Arg Ile Glu Thr Met Ser Val Pro Lys
545 550 555 560

Asp Lys Ile Arg Asp Val Ile Gly Thr Gly Gly Lys Val Ile Arg Glu
565 570 575

Ile Val Ala Thr Thr Gly Ala Lys Val Asp Ile Glu Asp Asp Gly Thr
580 585 590

Val Arg Leu Ser Ser Ser Asp Pro Ala Gln Ile Glu Ala Ala Arg Asn
595 600 605

Trp Ile Thr Gly Ile Ile Glu Glu Pro Glu Val Gly Lys Ile Tyr Asn
610 615 620

Gly Lys Val Val Asn Ile Val Asp Phe Gly Ala Phe Val Asn Phe Met
625 630 635 640

Gly Gly Arg Asp Gly Leu Val His Val Ser Glu Ile Lys Asn Glu Arg
645 650 655

Val Asn Lys Val Ser Asp Val Leu Ala Glu Gly Gln Glu Val Lys Val
660 665 670

Lys Val Leu Glu Ile Asp Asn Arg Gly Lys Val Arg Leu Ser Met Arg
675 680 685

Val Val Asp Gln Glu Thr Gly Ala Glu Leu Glu Asp Asn Arg Pro Pro
690 695 700

Arg Glu Ala Arg Glu Val Ser Asp Arg Gly Pro Arg Gly Asp Arg Pro
705 710 715 720

Arg Arg Asp Arg Gly Pro Arg Arg Glu Pro Gln Asn Gly Ser Asn His
725 730 735

Ser Gly Arg Asp Met Glu Pro Glu Phe Ala Pro Ala Phe Leu Arg Lys
740 745 750

Asp Asp

<210> 9

<211> 723

<212> PRT

5 <213> secuencia artificial

<220>

<223> proteína de fusión manipulada genéticamente

<400> 9

ES 2 611 977 T3

Met Phe Asp Ile Lys Arg Gln Glu Ile Asp Trp Gly Gly Lys Lys Leu
 1 5 10 15

Thr Leu Glu Thr Gly Gln Val Ala Arg Gln Ala Asp Gly Ala Val Ile
 20 25 30

Ala Thr Leu Gly Glu Thr Val Val Leu Cys Ala Val Thr Ala Ala Lys
 35 40 45

Thr Val Lys Glu Gly Gln Asp Phe Phe Pro Leu Thr Val His Tyr Gln
 50 55 60

Glu Lys Tyr Ser Ala Ala Gly Arg Ile Pro Gly Gly Phe Phe Lys Arg
 65 70 75 80

Glu Arg Gly Ala Thr Glu Arg Glu Thr Leu Ile Ser Arg Leu Ile Asp
 85 90 95

Arg Pro Ile Arg Pro Leu Phe Pro Glu Gly Phe Tyr Asn Glu Thr Leu
 100 105 110

Val Ile Ala Gln Val Met Ser Tyr Asp Gly Glu Asn Glu Pro Asp Ile
 115 120 125

Leu Ala Met Ile Ala Ala Ser Ala Ala Leu Ala Leu Ser Gly Val Pro
 130 135 140

Phe Leu Gly Pro Ile Gly Ala Ala Arg Val Gly Tyr Gln Asp Gly Glu
 145 150 155 160

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

Leu Val Val Gly Ala Thr Arg Asp Ala Val Met Met Val Glu Ser Glu
 180 185 190

Ala Asn Glu Leu Pro Glu Glu Val Met Leu Asn Ala Val Ser Phe Ala
 195 200 205

His Glu Ser Leu Gln Pro Val Ile Lys Ala Ile Ile Asn Leu Ala Glu
 210 215 220

ES 2 611 977 T3

Gln Ala Ala Lys Glu Pro Trp Glu Leu Val Ser Tyr Asp Asp Ser Ala
 225 230 235 240

Leu Ala Ala Lys Val Glu Glu Leu Cys Tyr Asp Asn Phe Asp Lys Ala
 245 250 255

Tyr Arg Leu Thr Arg Lys Ala Glu Arg Val Asp Ala Leu Ser Lys Ala
 260 265 270

Lys Ala Val Leu Asp Glu Ala Phe Pro Glu Ala Asp Pro Thr Glu Lys
 275 280 285

Leu Arg Ile Gln Lys Leu Ala Lys Lys Leu Glu Ala Lys Ile Val Arg
 290 295 300

Thr Ala Ile Leu Lys Glu Gly Arg Arg Ile Asp Gly Arg Asp Leu Lys
 305 310 315 320

Thr Val Arg Pro Ile Arg Ser Gln Val Gly Phe Leu Pro Arg Thr His
 325 330 335

Gly Ser Ala Leu Phe Thr Arg Gly Glu Thr Gln Ala Leu Val Ser Thr
 340 345 350

Thr Leu Gly Thr Ala Asp Ala Glu Gln Met Ile Asp Gly Leu Thr Gly
 355 360 365

Leu His Tyr Glu Arg Phe Met Leu His Tyr Asn Phe Pro Pro Tyr Ser
 370 375 380

Val Gly Glu Val Gly Arg Phe Gly Ala Pro Gly Arg Arg Glu Ile Gly
 385 390 395 400

His Gly Lys Leu Ala Trp Arg Ala Leu His Pro Val Leu Pro Ser Lys
 405 410 415

Ala Asp Phe Pro Tyr Thr Ile Arg Val Leu Ser Asp Ile Thr Glu Ser
 420 425 430

Asn Gly Ser Ser Ser Met Ala Thr Val Cys Gly Gly Cys Leu Ala Leu
 435 440 445

Met Asp Ala Gly Val Pro Leu Thr Arg Pro Val Ser Gly Ile Ala Met
 450 455 460

Gly Leu Ile Leu Glu Lys Asp Gly Phe Ala Ile Leu Ser Asp Ile Met

ES 2 611 977 T3

<220>

<223> proteína de fusión manipulada genéticamente

<400> 10

```

Met Phe Asp Ile Lys Arg Gln Glu Ile Asp Trp Gly Gly Lys Lys Leu
1           5           10           15

Thr Leu Glu Thr Gly Gln Val Ala Arg Gln Ala Asp Gly Ala Val Ile
          20           25           30

Ala Thr Leu Gly Glu Thr Val Val Leu Cys Ala Val Thr Ala Ala Lys
          35           40           45

Thr Val Lys Glu Gly Gln Asp Phe Phe Pro Leu Thr Val His Tyr Gln
          50           55           60

Glu Lys Tyr Ser Ala Ala Gly Arg Ile Pro Gly Gly Phe Phe Lys Arg
65           70           75           80

Glu Arg Gly Ala Thr Glu Arg Glu Thr Leu Ile Ser Arg Leu Ile Asp
          85           90           95

Arg Pro Ile Arg Pro Leu Phe Pro Glu Gly Phe Tyr Asn Glu Thr Leu
          100          105          110

Val Ile Ala Gln Val Met Ser Tyr Asp Gly Glu Asn Glu Pro Asp Ile
          115          120          125

Leu Ala Met Ile Ala Ala Ser Ala Ala Leu Ala Leu Ser Gly Val Pro
          130          135          140

Phe Leu Gly Pro Ile Gly Ala Ala Arg Val Gly Tyr Gln Asp Gly Glu
145           150          155          160

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

Leu Val Val Gly Ala Thr Arg Asp Ala Val Met Met Val Glu Ser Glu
          180          185          190

Ala Asn Glu Leu Pro Glu Glu Val Met Leu Asn Ala Val Ser Phe Ala

```

ES 2 611 977 T3

195	200	205													
His	Glu	Ser	Leu	Gln	Pro	Val	Ile	Lys	Ala	Ile	Ile	Asn	Leu	Ala	Glu
	210					215					220				
Gln	Ala	Ala	Lys	Glu	Pro	Trp	Glu	Leu	Val	Ser	Tyr	Asp	Asp	Ser	Ala
225					230					235					240
Leu	Ala	Ala	Lys	Val	Glu	Glu	Leu	Cys	Tyr	Asp	Asn	Phe	Asp	Lys	Ala
				245					250					255	
Tyr	Arg	Leu	Thr	Arg	Lys	Ala	Glu	Arg	Val	Asp	Ala	Leu	Ser	Lys	Ala
			260					265					270		
Lys	Ala	Val	Leu	Asp	Glu	Ala	Phe	Pro	Glu	Ala	Asp	Pro	Thr	Glu	Lys
		275					280					285			
Leu	Arg	Ile	Gln	Lys	Leu	Ala	Lys	Lys	Leu	Glu	Ala	Lys	Ile	Val	Arg
290						295					300				
Thr	Ala	Ile	Leu	Lys	Glu	Gly	Arg	Arg	Ile	Asp	Gly	Arg	Asp	Leu	Lys
305					310					315					320
Thr	Val	Arg	Pro	Ile	Arg	Ser	Gln	Val	Gly	Phe	Leu	Pro	Arg	Thr	His
				325					330					335	
Gly	Ser	Ala	Leu	Phe	Thr	Arg	Gly	Glu	Thr	Gln	Ala	Leu	Val	Ser	Thr
			340					345					350		
Thr	Leu	Gly	Thr	Ala	Asp	Ala	Glu	Gln	Met	Ile	Asp	Gly	Leu	Thr	Gly
		355					360					365			
Leu	His	Tyr	Glu	Arg	Phe	Met	Leu	His	Tyr	Asn	Phe	Pro	Pro	Tyr	Ser
370						375					380				
Val	Gly	Glu	Val	Gly	Arg	Phe	Gly	Ala	Pro	Gly	Arg	Arg	Glu	Ile	Gly
385					390					395					400
His	Gly	Lys	Leu	Ala	Trp	Arg	Ala	Leu	His	Pro	Val	Leu	Pro	Ser	Lys
				405					410					415	
Ala	Asp	Phe	Pro	Tyr	Thr	Ile	Arg	Val	Leu	Ser	Asp	Ile	Thr	Glu	Ser
			420					425					430		
Asn	Gly	Ser	Ser	Ser	Met	Ala	Thr	Val	Cys	Gly	Gly	Cys	Leu	Ala	Leu
		435					440					445			

ES 2 611 977 T3

Met Asp Ala Gly Val Pro Leu Thr Arg Pro Val Ser Gly Ile Ala Met
450 455 460

Gly Leu Ile Leu Glu Lys Asp Gly Phe Ala Ile Leu Ser Asp Ile Met
465 470 475 480

Gly Asp Glu Asp His Leu Gly Asp Met Asp Phe Lys Val Ala Gly Thr
485 490 495

Glu Lys Gly Ile Thr Ser Leu Gln Met Asp Ile Lys Val Ala Gly Ile
500 505 510

Thr Glu Glu Ile Met Gln Lys Ala Leu Glu Gln Ala Lys Gly Gly Arg
515 520 525

Ala His Ile Leu Gly Glu Met Ser Lys Ala Leu Gly Glu Val Arg Ser
530 535 540

Glu Ile Ser Asn Leu Ala Pro Arg Ile Glu Thr Met Ser Val Pro Lys
545 550 555 560

Asp Lys Ile Arg Asp Val Ile Gly Thr Gly Gly Lys Val Ile Arg Glu
565 570 575

Ile Val Ala Thr Thr Gly Ala Lys Val Asp Ile Glu Asp Asp Gly Thr
580 585 590

Val Arg Leu Ser Ser Ser Asp Pro Ala Asn Ile Glu Ala Ala Arg Glu
595 600 605

Trp Ile Asn Gly Ile Val Glu Glu Pro Glu Val Gly Lys Ile Tyr Asn
610 615 620

Gly Lys Val Val Asn Ile Val Asp Phe Gly Ala Phe Val Asn Phe Met
625 630 635 640

Gly Gly Arg Asp Gly Leu Val His Val Ser Glu Ile Lys Asn Glu Arg
645 650 655

Val Asn Lys Val Ser Asp Val Leu Ser Glu Gly Gln Glu Val Lys Val
660 665 670

Lys Val Leu Glu Ile Asp Asn Arg Gly Lys Val Arg Leu Ser Met Arg
675 680 685

Val Val Asp Gln Glu Thr Gly Leu Val
690 695

- <210> 11
- <211> 378
- <212> PRT
- <213> secuencia artificial

ES 2 611 977 T3

<220>

<223> proteína de fusión manipulada genéticamente

<400> 11

```

Met Phe Asp Ile Lys Arg Gln Glu Ile Asp Trp Gly Gly Lys Lys Leu
1           5           10           15

Thr Leu Glu Thr Gly Gln Val Ala Arg Gln Ala Asp Gly Ala Val Ile
          20           25           30

Ala Thr Leu Gly Glu Thr Val Val Leu Cys Ala Val Thr Ala Ala Lys
          35           40           45

Thr Val Lys Glu Gly Gln Asp Phe Phe Pro Leu Thr Val His Tyr Gln
          50           55           60

Glu Lys Tyr Ser Ala Ala Gly Arg Ile Pro Gly Gly Phe Phe Lys Arg
65           70           75           80

Glu Arg Gly Ala Thr Glu Arg Glu Thr Leu Ile Ser Arg Leu Ile Asp
          85           90           95

Arg Pro Ile Arg Pro Leu Phe Pro Glu Gly Phe Tyr Asn Glu Thr Leu
          100          105          110

Val Ile Ala Gln Val Met Ser Tyr Asp Gly Glu Asn Glu Pro Asp Ile
          115          120          125

Leu Ala Met Ile Ala Ala Ser Ala Ala Leu Ala Leu Ser Gly Val Pro
          130          135          140

Phe Leu Gly Pro Ile Gly Ala Ala Arg Val Gly Tyr Gln Asp Gly Glu
145           150          155          160

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

Leu Val Val Gly Ala Thr Arg Asp Ala Val Met Met Val Glu Ser Glu
          180          185          190

Ala Asn Glu Leu Pro Glu Glu Val Met Leu Asn Ala Val Ser Phe Ala
          195          200          205

```

ES 2 611 977 T3

His Glu Ser Leu Gln Pro Val Ile Lys Ala Ile Ile Asn Leu Ala Glu
 210 215 220

Gln Ala Ala Lys Glu Pro Trp Glu Leu Val Ser Tyr Asp Asp Ser Ala
 225 230 235 240

Leu Ala Ala Lys Val Glu Glu Leu Cys Tyr Asp Asn Phe Asp Lys Ala
 245 250 255

Tyr Arg Leu Thr Arg Lys Ala Glu Arg Val Asp Ala Leu Ser Lys Ala
 260 265 270

Lys Ala Val Leu Asp Glu Ala Phe Pro Glu Ala Asp Pro Thr Glu Lys
 275 280 285

Leu Arg Ile Gln Lys Leu Ala Lys Lys Leu Glu Ala Lys Ile Val Arg
 290 295 300

Thr Ala Ile Leu Lys Glu Gly Arg Arg Ile Asp Gly Arg Asp Leu Lys
 305 310 315 320

Thr Val Arg Pro Ile Arg Ser Gln Val Gly Phe Leu Pro Arg Thr His
 325 330 335

Gly Ser Ala Leu Phe Thr Arg Gly Glu Thr Gln Ala Leu Val Ser Thr
 340 345 350

Thr Leu Gly Thr Ala Asp Ala Glu Gln Met Ile Asp Gly Leu Thr Gly
 355 360 365

Thr Ser Val Thr Glu Asp His Phe Ala Glu
 370 375

- <210> 12
- <211> 49
- <212> PRT
- <213> secuencia artificial

5

- <220>
- <223> proteína de fusión manipulada genéticamente
- <400> 12

Met Phe Asp Ile Lys Arg Gln Glu Ile Asp Trp Gly Gly Lys Lys Leu
 1 5 10 15

Thr Leu Glu Thr Gly Gln Val Ala Arg Gln Ala Asp Gly Ala Val Ile
 20 25 30

Asn Ala Glu Pro Val Ser Tyr Thr His Leu Asn Pro Glu Ala Leu Val
 35 40 45

10 Gly

ES 2 611 977 T3

<210> 13
 <211> 474
 <212> ADN
 <213> Zymomonas mobilis

5 <400> 13

```

gtgacctctg ctgtgccatc aaatacgaaa aaaaagctgg tgattgcttc cgatcacgca      60
gcatttgagt tgaatcaac cttgattact tggctgaaag agcttgggtca tgaggtcgaa      120
gaccttggcc ctcatgaaaa ccattcagtc gattatcccg attacgggta taagctggct      180
gtcgctatcg cagaaaaaac cgctgatttc ggtattgctt tatgtggctc gggaatcggg      240
atctcgatcg ctgtcaatcg ccatccggct gcccgttgcg ctttgattac ggataacctt      300
accgcccgtt tggcaagaga acataacaat gccaatgtta tcgctatggg tgcgagattg      360
atcggcattg aaaccgctaa ggattgtatt tcagctttcc ttgcaacgcc gtttggagggt      420
gaacgtcatg ttcgccgtat cgataaactt tcgaatcctc agttcaatat ctag          474
    
```

<210> 14
 <211> 157
 <212> PRT
 <213> Zymomonas mobilis

10

<400> 14

```

Met Thr Ser Ala Val Pro Ser Asn Thr Lys Lys Lys Leu Val Ile Ala
1          5          10          15

Ser Asp His Ala Ala Phe Glu Leu Lys Ser Thr Leu Ile Thr Trp Leu
          20          25          30

Lys Glu Leu Gly His Glu Val Glu Asp Leu Gly Pro His Glu Asn His
          35          40          45

Ser Val Asp Tyr Pro Asp Tyr Gly Tyr Lys Leu Ala Val Ala Ile Ala
          50          55          60

Glu Lys Thr Ala Asp Phe Gly Ile Ala Leu Cys Gly Ser Gly Ile Gly
65          70          75          80

Ile Ser Ile Ala Val Asn Arg His Pro Ala Ala Arg Cys Ala Leu Ile
          85          90          95

Thr Asp Asn Leu Thr Ala Arg Leu Ala Arg Glu His Asn Asn Ala Asn
          100          105          110

Val Ile Ala Met Gly Ala Arg Leu Ile Gly Ile Glu Thr Ala Lys Asp
          115          120          125

Cys Ile Ser Ala Phe Leu Ala Thr Pro Phe Gly Gly Glu Arg His Val
          130          135          140

Arg Arg Ile Asp Lys Leu Ser Asn Pro Gln Phe Asn Ile
145          150          155
    
```

ES 2 611 977 T3

<210> 15
 <211> 660
 <212> ADN
 <213> Escherichia coli

5 <400> 15

```

atgacgcagg atgaattgaa aaaagcagta ggatgggcgg cacttcagta tgttcagccc      60
ggcaccattg ttggtgtagg tacaggttcc accgccgcac actttattga cgcgctcggg      120
acaatgaaag gccagattga aggggccggt tccagttcag atgcttccac tgaaaaactg      180
aaaagcctcg gcattcacgt ttttgatctc aacgaagtcg acagccttgg catctacggt      240
gatggcgcag atgaaatcaa cggccacatg caaatgatca aaggcggcgg cgcggcgctg      300
acccgtgaaa aaatcattgc ttcggttgca gaaaaattta tctgtattgc agacgcttcc      360
aagcaggttg atattctggg taaattcccg ctgccagtag aagttatccc gatggcacgt      420
agtgcagtgg cgcgtcagct ggtgaaactg ggcggtcgtc cggaataccg tcagggcgtg      480
gtgaccgata atggcaacgt gatcctcgac gtccacggca tggaaatcct tgacccgata      540
gcgatggaaa acgccataaa tgcgattcct ggcgtggtga ctgttggtt gtttgctaac      600
cgtggcgcgg acgttgcgct gattggcaca cctgacggtg tcaaaacat tgtgaaatga      660
  
```

<210> 16
 <211> 219
 <212> PRT
 <213> Escherichia coli

10

<400> 16

```

Met Thr Gln Asp Glu Leu Lys Lys Ala Val Gly Trp Ala Ala Leu Gln
1           5           10           15

Tyr Val Gln Pro Gly Thr Ile Val Gly Val Gly Thr Gly Ser Thr Ala
          20           25           30

Ala His Phe Ile Asp Ala Leu Gly Thr Met Lys Gly Gln Ile Glu Gly
          35           40           45

Ala Val Ser Ser Ser Asp Ala Ser Thr Glu Lys Leu Lys Ser Leu Gly
  
```

ES 2 611 977 T3

50		55		60
Ile His Val Phe Asp Leu Asn Glu Val Asp Ser Leu Gly Ile Tyr Val				
65		70		75
				80
Asp Gly Ala Asp Glu Ile Asn Gly His Met Gln Met Ile Lys Gly Gly				
		85		90
				95
Gly Ala Ala Leu Thr Arg Glu Lys Ile Ile Ala Ser Val Ala Glu Lys				
		100		105
				110
Phe Ile Cys Ile Ala Asp Ala Ser Lys Gln Val Asp Ile Leu Gly Lys				
		115		120
				125
Phe Pro Leu Pro Val Glu Val Ile Pro Met Ala Arg Ser Ala Val Ala				
		130		135
				140
Arg Gln Leu Val Lys Leu Gly Gly Arg Pro Glu Tyr Arg Gln Gly Val				
		145		150
				155
				160
Val Thr Asp Asn Gly Asn Val Ile Leu Asp Val His Gly Met Glu Ile				
		165		170
				175
Leu Asp Pro Ile Ala Met Glu Asn Ala Ile Asn Ala Ile Pro Gly Val				
		180		185
				190
Val Thr Val Gly Leu Phe Ala Asn Arg Gly Ala Asp Val Ala Leu Ile				
		195		200
				205
Gly Thr Pro Asp Gly Val Lys Thr Ile Val Lys				
		210		215

<210> 17

<211> 1185

<212> ADN

5 <213> Actinoplanes missouriensis

<400> 17

atgagtgtcc aagctacacg cgaagataaa tttagctttg gcttgtggac agttgggtgg	60
caggcccgtg acgcctttgg tgatgccacg cgtaccgcc ttgatccggt cgaagcagtt	120
cataaacttg ccgaaatcgg agcctatggc attacatttc atgatgatga cttggtcccc	180
ttcggttcgg atgctcagac acgggatggt atcattgctg gtttcaaaaa agcccttgat	240
gaaaccggtt tgatcgttcc tatggttacc acaaatttgt ttacgcatcc ggtttttaa	300
gacggcgggt ttacctctaa cgatcgttct gtgcgccggt atgctatccg gaaagttttg	360
agacagatgg atcttgggtgc ggaactgggg gccaaaacgc ttgtottatg ggcggcaga	420

ES 2 611 977 T3

gaaggtgctg aatattgattc agcaaaagat gttagtgcgc ccttgatcgc ttatcgcgaa 480
gcaactgaatc ttctggcaca atatagcgaa gacagaggct atggacttcg ttttgcaatt 540
gaaccgaaac ctaatgaacc acgtggcgat attctgctgc ctaccgcagg ccatgctatt 600
gcctttgtgc aagaattaga acggcccgaa ctttttgcca tcaatccaga aaccgggcat 660
gaacagatgt caaacttgaa ttttaccagc gggattgctc aggctttgtg gcataaaaaa 720
ctttttcata ttgatttgaa tggacaacat ggtccaagt ttgatcagga ccttgccttt 780
ggatcatggtg accttttaaa tgcctttagc ctggctgatt tgtagaaaaa tggctcagat 840
gggtccccgc cttatgatgg cccgcgccat tttgattata aaccatctcg tactgaagat 900
tatgatggcg tttgggaatc agcgaagcc aatatccgta tgtatctttt attaaaagaa 960
cgtgccaaag cgttttagagc tgatccggaa gttcaggaag cactggcagc aagcaaagtt 1020
gccgaattga aaacccaac gttgaatcct ggtgaaggct atgcagaact gttagcagat 1080
cgcagtgcct ttgaagatta tgatccgat gctgttggtg cgaaaggttt tggtttcgtg 1140
aaattgaacc aattggccat tgaacattta ttaggtgccc gctag 1185

<210> 18

<211> 394

<212> PRT

5 <213> Actinoplanes missouriensis

<400> 18

Met Ser Val Gln Ala Thr Arg Glu Asp Lys Phe Ser Phe Gly Leu Trp
1 5 10 15

Thr Val Gly Trp Gln Ala Arg Asp Ala Phe Gly Asp Ala Thr Arg Thr
20 25 30

Ala Leu Asp Pro Val Glu Ala Val His Lys Leu Ala Glu Ile Gly Ala
35 40 45

Tyr Gly Ile Thr Phe His Asp Asp Asp Leu Val Pro Phe Gly Ser Asp
50 55 60

Ala Gln Thr Arg Asp Gly Ile Ile Ala Gly Phe Lys Lys Ala Leu Asp
65 70 75 80

Glu Thr Gly Leu Ile Val Pro Met Val Thr Thr Asn Leu Phe Thr His
85 90 95

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

Arg Tyr Ala Ile Arg Lys Val Leu Arg Gln Met Asp Leu Gly Ala Glu

ES 2 611 977 T3

115	120	125																		
Leu	Gly	Ala	Lys	Thr	Leu	Val	Leu	Trp	Gly	Gly	Arg	Glu	Gly	Ala	Glu					
130						135					140									
Tyr	Asp	Ser	Ala	Lys	Asp	Val	Ser	Ala	Ala	Leu	Asp	Arg	Tyr	Arg	Glu					
145					150					155					160					
Ala	Leu	Asn	Leu	Leu	Ala	Gln	Tyr	Ser	Glu	Asp	Arg	Gly	Tyr	Gly	Leu					
			165						170					175						
Arg	Phe	Ala	Ile	Glu	Pro	Lys	Pro	Asn	Glu	Pro	Arg	Gly	Asp	Ile	Leu					
			180					185					190							
Leu	Pro	Thr	Ala	Gly	His	Ala	Ile	Ala	Phe	Val	Gln	Glu	Leu	Glu	Arg					
		195					200					205								
Pro	Glu	Leu	Phe	Gly	Ile	Asn	Pro	Glu	Thr	Gly	His	Glu	Gln	Met	Ser					
	210					215					220									
Asn	Leu	Asn	Phe	Thr	Gln	Gly	Ile	Ala	Gln	Ala	Leu	Trp	His	Lys	Lys					
225					230					235					240					
Leu	Phe	His	Ile	Asp	Leu	Asn	Gly	Gln	His	Gly	Pro	Lys	Phe	Asp	Gln					
			245						250					255						
Asp	Leu	Val	Phe	Gly	His	Gly	Asp	Leu	Leu	Asn	Ala	Phe	Ser	Leu	Val					
			260					265					270							
Asp	Leu	Leu	Glu	Asn	Gly	Pro	Asp	Gly	Ala	Pro	Ala	Tyr	Asp	Gly	Pro					
		275					280					285								
Arg	His	Phe	Asp	Tyr	Lys	Pro	Ser	Arg	Thr	Glu	Asp	Tyr	Asp	Gly	Val					
	290					295					300									
Trp	Glu	Ser	Ala	Lys	Ala	Asn	Ile	Arg	Met	Tyr	Leu	Leu	Leu	Lys	Glu					
305					310					315					320					
Arg	Ala	Lys	Ala	Phe	Arg	Ala	Asp	Pro	Glu	Val	Gln	Glu	Ala	Leu	Ala					
			325						330					335						
Ala	Ser	Lys	Val	Ala	Glu	Leu	Lys	Thr	Pro	Thr	Leu	Asn	Pro	Gly	Glu					
			340					345					350							
Gly	Tyr	Ala	Glu	Leu	Leu	Ala	Asp	Arg	Ser	Ala	Phe	Glu	Asp	Tyr	Asp					
		355					360					365								

ES 2 611 977 T3

Ala Asp Ala Val Gly Ala Lys Gly Phe Gly Phe Val Lys Leu Asn Gln
 370 375 380

Leu Ala Ile Glu His Leu Leu Gly Ala Arg
 385 390

<210> 19
 <211> 663
 <212> ADN
 <213> Zymomonas mobilis

5

<400> 19
 atgagtaa at taccctgat tgetccctct atcctttcgg cggattttgc ccatttggga 60
 gatgaggtcg cggcgataga tcaggccggg gccgattgga tccatattga tgtgatggat 120
 ggccatttcg tgccgaatat caccataggg cccatggttg tgaaggcttt gcgtccctat 180
 agccaaaagc cttttgatgt ccatttgatg attgcgcctg tcgatcaata tatcgaggct 240
 ttttctgaag cgggtgctga tattatcagt ttccatcccg aagcggggcg gcaccccat 300
 cgcactattc agcatatcaa atcattgggc aaaaaagcgg gattagtttt taatccggcg 360
 accccttta gctggcttga ttatcta atg gatgatcttg atctgattat ggtgatgagc 420
 gttaaccccg gttttggcgg ccaaaaattt atcaaaacc aattagaaaa gattaaagat 480
 atccgtcaaa gaattaccgc ctctggggcg gatatccgct tggaagtgga tggcggaatt 540
 gatgccacga ctgcaccgct tgccgtcgaa gccggtgccc atgttttggc cgcgggaacg 600
 gccagcttta aaggcggcgc aacatgttac accgataata tcaggatatt gcgtaaatca 660
 tga 663

<210> 20
 <211> 220
 <212> PRT
 <213> Zymomonas mobilis

10

<400> 20
 Met Ser Lys Leu Pro Leu Ile Ala Pro Ser Ile Leu Ser Ala Asp Phe
 1 5 10 15

Ala His Leu Gly Asp Glu Val Ala Ala Ile Asp Gln Ala Gly Ala Asp
 20 25 30

Trp Ile His Ile Asp Val Met Asp Gly His Phe Val Pro Asn Ile Thr
 35 40 45

Ile Gly Pro Met Val Val Lys Ala Leu Arg Pro Tyr Ser Gln Lys Pro
 50 55 60

Phe Asp Val His Leu Met Ile Ala Pro Val Asp Gln Tyr Ile Glu Ala

ES 2 611 977 T3

	<400>	22	
		ccatggcgag ctcgttcgat caacaacccg aatcctatcg taatgatggt ttgcccgatc	60
		agcctcaatc gacaatttta cgcgtttoga tgaagcagg gacgacaatt ggctgggaac	120
		ggatatactgg aataaatgggt cttcgttatg gtattgatgt ttttgggtgca tcggccccgg	180
		cgaatgatct atatgctcat ttcggcttga ccgagtcgg catcacgaac aaggtggttg	240
		ccgcgatcgc cggtaagtcg gcacgttaa aaatagctat ggaatataat agctacttaa	300
		taagttagga gaataaacgt gacctctgct gtgccatcaa atacgaaaa aaagctgggtg	360
		attgcttccg atcacgcagc atttgagttg aaatcaacct tgattacttg gctgaaagag	420
		cttggctcatg aggtcgaaga ccttggccct catgaaaacc attcagtcga ttatcccgat	480
		tacggttata agctggctgt cgctatcgca gaaaaacccg ctgatttcgg tattgcttta	540
		tgtggctcgg gaatcgggat ctcgatcgct gtcaatcgcc atccggctgc ccgttgcgct	600
		ttgattacgg ataaccttac cgccgtttg gcaagagaac ataacaatgc caatggtatc	660
		gctatgggtg cgagattgat cggcattgaa accgctaagg attgtatttc agctttcctt	720
		gcaacgccgt ttggaggtga acgtcatggt cgccgatcgc ataaactttc gaatcctcag	780
		ttcaatatct agctcgaggc ggctgaacg tactgcaagt cctgacgtca ctgtgcagtc	840
		cgttggcccg gttatcggta gcgataccgg gcattttttt aaggaacgat cgatagcggc	900
		cgc	903
	<210>	23	
	<211>	38	
5	<212>	ADN	
	<213>	secuencia artificial	
	<220>		
	<223>	cebador	
	<400>	23	
10		tcactcatgg ccggcctgcg tataatatt gccatgg	38
	<210>	24	
	<211>	37	
	<212>	ADN	
	<213>	secuencia artificial	
15	<220>		
	<223>	cebador	
	<400>	24	
		ctactcatcc tgcaggctga tgaatgctca tccggaa	37
	<210>	25	
20	<211>	10250	
	<212>	ADN	
	<213>	secuencia artificial	
	<220>		
	<223>	plásmido construido	
25	<400>	25	

ES 2 611 977 T3

ctagtgttcg atcaacaacc cgaatcctat cgtaatgatg ttttgcccga tcagcctcaa 60
 tcgacaattt tacgcgtttc gatcgaagca gggacgacaa ttggctggga acggtatact 120
 tgaataaatg gtcttcgtta tggattgat gtttttggtg catcgccccc ggcgaatgat 180
 ctatatgctc atttcggctt gaccgcagtc ggcacacga acaagggtgtt ggccgcgac 240
 gccggtaagt cggcacgtta aaaaatagct atggaatata atagctacta ataagttagg 300
 agaataaaca tgagtgtcca agctacacgc gaagataaat ttagctttgg cttgtggaca 360
 gttgggtggc aggcccgta cgcctttggt gatgccacgc gtaccgccct tgatccggtc 420
 gaagcagttc ataaacttgc cgaaatcgga gcctatggca ttacatttca tgatgatgac 480
 ttggtcccct tcggttcgga tgctcagaca cgggatggta tcattgctgg tttcaaaaaa 540
 gcccttgatg aaaccggttt gatcgttctt atggttacca caaatttgtt tacgcatccg 600
 gtttttaaaag acggcgggtt tacctctaac gatcgttctg tgccgcggtg tgctatccgg 660
 aaagttttga gacagatgga tcttgggtgcg gaactggggg ccaaaaacgct tgtcttatgg 720
 ggcggcagag aaggtgcgga atatgattca gcaaaagatg ttagtgccgc cttggatcgt 780
 tatcgcgaag cactgaatct tctggcacia tatagcgaag acagaggcta tggacttctg 840
 tttgcaattg aaccgaaacc taatgaacca cgtggcgata ttctgctgcc taccgcaggc 900
 catgctattg cctttgtgca agaattagaa cggcccgaac tttttggcat caatccagaa 960
 accgggcatg aacagatgtc aaacttgaat tttaccagc ggattgctca ggctttgtgg 1020
 cataaaaaac tttttcatat tgatttgaat ggacaacatg gtcccagtt tgatcaggac 1080
 cttgtctttg gtcattggtga ccttttaaat gcctttagcc tggctgattt gttagaaaat 1140
 ggtccagatg gtgccccggc ttatgatggc ccgcgccatt ttgattataa accatctcgt 1200
 actgaagatt atgatggcgt ttgggaatca gcgaaagcca atatccgtat gtatctttta 1260
 ttaaaagaac gtgccaagc gtttagagct gatccggaag ttcaggaagc actggcagca 1320
 agcaaagttg ccgaattgaa aaccccaacg ttgaatcctg gtgaaggcta tgcagaactg 1380
 ttagcagatc gcagtgcttt tgaagattat gatgccgatg ctggttgggtc gaaaggtttt 1440
 ggtttcgtga aattgaacca attggccatt gaacatttat taggtgcccg ctagtctaga 1500
 tgactgtata aaaccacagc caatcaaacg aaaccaggct atactcaagc ctggtttttt 1560

ES 2 611 977 T3

gatggatddd cagcgtggcg caggcagggt ttatctaac ccgacactgg cgggacaccc 1620
cgcaagggac agaagtctcc ttctggctgg cgacggacaa cgggccaagc ttccagttac 1680
tcaatacgtc acaataatca gtttatccta actatagaat cgcattgagaa gcgataacgt 1740
ttcaccataa gcaatatatt cattgcaaca gtggaattgc cttatgcgtc aaggaaggat 1800
agatcattga cggactgagt tcaaaaagag actggtctaa aagattdtaa gaaaggtdtc 1860
gatatgtata tcgggataga tcttggcacc tcgggcgtaa aagtdatddd gctcaacgag 1920
cagggtgagg tggtdgctgc gcaaacggaa aagctgaccg tttcgcgccc gcatccactc 1980
tggctggaac aagacccgga acagtgggtg caggcaactg atcgcgcaat gaaagctctg 2040
ggcgatcagc attctctgca ggacgttaa gcattgggtg ttgccggcca gatgcacgga 2100
gcaaccttgc tggatgctca gcaacgggtg ttacgccctg ccatttdtg gaaacgaggg 2160
cgctgtgccc aagagtgcac tttgctgga gcgcgagtdc cgcaatcgcg ggtgattacc 2220
ggcaacctga tgatgcccgg atttactgcg cctaaattgc tatgggttca gcggcatgag 2280
ccggagatat tccgtcaaat cgacaaagta ttattaccga aagattactt gcgtctgct 2340
atgacggggg agtdtgccag cgatatgtct gacgcagctg gcaccatgtg gctggatgtc 2400
gcaaagcgtg actggagtga cgtcatgctg caggcttgcg acttatctcg tgaccagatg 2460
cccgcattat acgaaggcag cgaaattact ggtgcttdgt tacttgaagt tgcgaaagcg 2520
tggggtatgg cgacggtgcc agtdgtcgca ggcggtggcg acaatgcagc tgggtgcagtd 2580
ggtgtgggaa tggtdgatgc taatcaggca atgttatcgc tggggacgtd gggggtctat 2640
tdtgctgtca gcgaagggtd cttaaagcaag ccagaaagcg ccgtacatag ctdtdgcat 2700
gcgctaccgc aacgttggca ttdaatgtct gtgatgctga gtgcagcgtc gtgtctggat 2760
tgggccgcca aattaaccgg cctgagcaat gtcccagctt taatcgtctc agctcaacag 2820
gctgatgaaa gtgccgagcc agtdtggtdt ctgccttacc ttdccggcga gcgtacgcca 2880
cacaataatc cccaggcga gggggttdtc ttdgtdtga ctcatcaaca tggcccaat 2940
gaactggcgc gagcagtdct ggaaggcgtg ggttatgcgc tggcagatgg catggatgtc 3000
gtgcatgcct gcggtattaa accgcaaagt gttacgttga ttdggggcgg ggcgcgtagt 3060
gagtactggc gtcagatgct ggcggatacc agcggctcagc agctcgatta ccgtacgggg 3120
ggggatgtgg ggccagcact gggcgcagca aggttggcgc agatcgcggc gaatccagag 3180
aaatcgtcct ttdaattgtd gccgcaacta ccgttagaac agtcgcatct accagatgcg 3240
cagcgttatg ccgcttatca gccacgacga gaaacgttdc gtcgcctcta tcagcaactt 3300
ctgccattaa tggcgtaaac gttatcccct gcctgaccgg ttdggggata attcacatct 3360
atatactcct gtaattaatt aatatttagt acgaattdat tctgaaaatc atttdgtaat 3420
ggcatttdtc agtdtdgtct ttdgttggtd actcgtaatg tctcgtggt agatatggag 3480

ES 2 611 977 T3

atcgttatga aaacctcaaa gactgtggca aaactattat ttgttgcg ggcgctggtt 3540
 tatctgggtg ggctatggat ctcatgcca ttgtaagtg gaaaaggcta ttttcttggc 3600
 gtgtaaatga cagcaacttt tggcaactat gcgaattcgc gatcgcataa cttcgtataa 3660
 tgtatgctat acgaagttat gcggcgcag cacaggatga cgcctaacaa ttcattcaag 3720
 ccgacaccgc ttcgcggcgc ggcttaattc aggagttaa catcatgagg gaagcgggtga 3780
 tcgccgaagt atcgactcaa ctatcagagg tagttggcgt catcgagcgc catctcgaac 3840
 cgacgttgct ggccgtacat ttgtacggct ccgcagtgga tggcggcctg aagccacaca 3900
 gtgatattga tttgctgggt acggtgactg taaggcttga tgaacaacg cggcgagctt 3960
 tgatcaacga ccttttgaa acttcggctt ccctggaga gagcgagatt ctccgcgctg 4020
 tagaagtcac cattggttg caccgacaca tcattccgtg gcgttatcca gctaagcgcg 4080
 aactgcaatt tggagaatgg cagcgaatg acattcttgc aggtatcttc gagccagcca 4140
 cgatcgacat tgatctggct atcttgctga caaaagcaag agaacatagc gttgccttgg 4200
 taggtccagc ggcggaggaa ctctttgatc cggttcctga acaggatcta tttgaggcgc 4260
 taaatgaaac cttaacgcta tggaaactgc cgcgccactg ggctggcgat gagcgaatg 4320
 tagtgcttac gttgtcccgc atttggtaca gcgcagtaac cggcaaaatc gcgccgaagg 4380
 atgtcgctgc cgactgggca atggagcgc tcgccgcca gtatcagccc gtcatacttg 4440
 aagctaggca ggcttatctt ggacaagaag atcgcttggc ctcgcgcgca gatcagttgg 4500
 aagaatttgt tcaactacgtg aaaggcgaga tcaccaaggt agtcggcaaa taatgtctaa 4560
 caattcgctt aagccgacgc cgcttcgagg cgcggttaa ctcaagcgtt agagagctgg 4620
 ggaagactat gcgogatctg ttgaagggtg ttctaagcct cgtacttgcg atggcatcgg 4680
 ggcaggcact tgctgacctg ccttaattaa ataacttctg ataatgtatg ctatacgaag 4740
 ttatggccgg cccgttactt gctgggtgaa gaaccgatcg aagtccgtgc ttacacctac 4800
 agcgatccga atgatgaacg tttcgttgaa gtcgaagatc gtattatttg gcagatgcgc 4860
 ttcagaagcg gtgctctgtc tcatggtgca tcttcttatt cgaccacgac gacttcacgt 4920
 ttctcgggtc agggcgacaa agctgttctg ttgatggatc cggctaccgg atattatcag 4980
 aatttgattt ctgtccagac cccaggccat gctaaccagt cgatgatgcc acagttcatc 5040
 atgccagcga acaaccagtt ctctgcacag ttggatcatc tggctgaagc cgtcatcaat 5100
 aacaaaccag ttcgtagccc ggggaagaa ggtatgcagg atgtgcgcct gattcagccc 5160
 atttatgaag cagctcgtac cggctgcccc gtcaacacgg attggggtta tgtccgctag 5220
 ggtggttatt gattctgact taacctattt gggtaaaaca gacttatttt tcctgtttta 5280
 ggaaaatagt taaaaaggcg tcattggttc ttccaatgac gccttttttt ataaacaaaa 5340

ES 2 611 977 T3

aaatcctttt gtcggtttta taaaaaact tcatattttg ataagccgtc ttaaaaatat 5400
 aataaathtt tataatattt atccgatcaa aggaccctt tatgctagaa gtcattatat 5460
 cggcattact accgattata attactttaa tgataggttt tttegetggc tggcgtggtg 5520
 aatttacggc aatcaagcc tcgacctga ataaaatggt cttacgctat gccttaccta 5580
 tgactttatt ctctgggatt ttatcacttc caaaaacaca gattttatcg tcgggttctg 5640
 ccgcaattat tttactttta gccatggctg gcggctatct aattacactt gggataggat 5700
 attttgtctg ccagegccca gtgaatgaat ctgctctttt agctctttct gttagcgcac 5760
 ctgcagttcc ttttgttggc ataacagttc tagggcattt atttggcact gccagcacga 5820
 tattggtttc aatatgtagc ctgatgatga acctcgtcca ggttcccgtt acctttttct 5880
 tttgtcagcg tattctccaa aaaagaatac tgacaaaata gccacggata gttctttttt 5940
 ttctcatatc agacatgctt ttaccgaacc tattgttatt gccctattc tggctcttat 6000
 ctgtgtcagt ctctctattc ctttccctga aaccttaaaa tcttctttaa tgctactagg 6060
 aaaagcgacc ggaggcgttg cgcttttctc ttctggtata attttatttt ctcgaaaagt 6120
 tattttaagt aaaacagtag catctttagt tttatcaaaa aatattatta ttccaacagc 6180
 ggtattggtt cttgcgtcgg tacccaattc gccctatagt gagtcgtatt acgcgcgctc 6240
 actggccgtc gttttacaac gtctgactg ggaaaaccct ggcgttacc aacttaatcg 6300
 ccttgacgca catccccctt tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg 6360
 cccttcccaa cagttgcgca gcctgaatgg cgaatgggac gcgccctgta gcggcgcatt 6420
 aagcgcggcg ggtgtggtgg ttacgcgcag cgtgaccgct acaactgcca gcgccctagc 6480
 gccgcctcct ttogctttct tcccttccct tctgcacaag ttgcgcggtt ttccccgta 6540
 agctctaaat cgggggctcc ctttaggggt ccgatttagt gctttacgac acctcgacct 6600
 caaaaaactt gattagggtg atggttcacg tagtgggcca tcgccctgat agacggtttt 6660
 tcgccctttg acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac 6720
 aacactcaac cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc 6780
 ctattggtta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaatatt 6840
 aacgcttaca atttaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttgttta 6900
 tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt 6960
 caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc 7020
 ttttttgcg ctttttgcct tcctgttttt gctcaccagc aaacgctggt gaaagtaaaa 7080
 gatgctgaag atcagttggg tgcacgagtg gttacatcg aactggatct caacagcgg 7140
 aagatccttg agagttttcg ccccgaagaa cgttttccaa tgatgagcac ttttaagtt 7200
 ctgctatgtg gcgcggtatt atcccgattt gacgcggggc aagagcaact cggtcgccc 7260

ES 2 611 977 T3

atacactatt ctcagaatga cttggttgag tactcaccag tcacagaaaa gcatcttacg 7320
 gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga taacactgcg 7380
 gccaaacttac ttctgacaac gatcggagga ccgaaggagc taaccgcttt ttgacacaac 7440
 atgggggatc atgtaactcg ccttgatcgt tgggaaccgg agctgaatga agccatacca 7500
 aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta 7560
 actggcgaac tacttactct agcttcccgg caacaattaa tagactggat ggaggcggat 7620
 aaagttgcag gaccacttct gcgctcggcc cttccggctg gctggtttat tgctgataaa 7680
 tctggagccg gtgagcgtgg gtctcgggt atcattgcag cactggggcc agatggtaag 7740
 ccctcccgta tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat 7800
 agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt 7860
 tactcatata tactttagat tgatttaaaa cttcattttt aattttaaag gatctaggtg 7920
 aagatccttt ttgataatct catgaccaaa atcccttaac gtgagttttc gttccactga 7980
 gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta 8040
 atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttgttt gccggatcaa 8100
 gagctaccaa ctctttttcc gaaggttaact ggcttcagca gagcgcagat accaaatact 8160
 gtccttctag tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca 8220
 tacctcgctc tgctaatacct gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt 8280
 accgggttg actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg 8340
 ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag 8400
 cgtgagctat gagaaagcgc cacgcttccc gaaggagaaa aggcggacag gtatccggta 8460
 agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctggtat 8520
 ctttatagtc ctgtcgggtt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg 8580
 tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc 8640
 ttttgctggc cttttgctca catgttcttt cctgcgttat ccctgattc tgtggataac 8700
 cgtattaccg cttttgagtg agctgatacc gctcgcgcga gccgaacgac cgagcgcagc 8760
 gagtcaatga gcgaggaagc ggaagagcgc ccaatacga aaccgcctct ccccgcgct 8820
 tggccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag 8880
 cgcaacgcaa ttaatgtgag ttagctcact cattaggcac cccaggcttt acactttatg 8940
 cttccggctc gtatggttg tggaattgtg agcggataac aatttcacac aggaaacagc 9000
 tatgaccatg attaogccaa gcgcgcaatt aaccctcact aaagggaaaca aaagctggag 9060
 ctcgccaga aaagacagca ttccttctca ataaagaaat attatttttt gtttttgaaa 9120

ES 2 611 977 T3

aatttttcca aaatctagaa tgctacatta aatatacaaa aatattatta taaaaataag 9180
gcttttaa at acccatat tttagaat ctttacaag aaacatgta aatatagatt 9240
tagagattaa tatkagccat ttttatcaaa aattcttttt ttgttttata atattatgct 9300
gcaaaactaa taaaaacgcc ctttcgaaat taacgatcac ccacaagaaa taattatctg 9360
acagcgctta ccaatcaatt attgccgaac gcagagtcgc gtattaggac ggtcaacaat 9420
ctaaaccgtt tttcagaaaa tattgcttta taagcctcaa aacttaaaag ctgcggtatt 9480
ttaatatacc aaaat tttct ggaaaagccg gcgaatcaga taacagttcc gcacaggtga 9540
gaaccacgac ggatcttctc tgaattgttg gttagttaaag aaagaaacaa ggattatgac 9600
gaacaaaatc tegtcttcag ataacttttc caatgctgtt tcagcaacgg atgacaacgc 9660
ttcccgtacg ccaaatctga cccgtcgcgc tctcgttggg ggtggtgttg gactggccgc 9720
agctggcgcc ttagccagtg gtcttcaggc agcgacgctt cctgctggtg ccagccaggt 9780
tccgaccacg cctgcaggtc gcccgatgcc ttacgcgatc cgcgccgatc cggaagatcg 9840
tcgtttcggg tatgctatcg tcggtctggg taaatatgcc cttaaccaga ttttaccggg 9900
ttttgccgga tgccagcatt cccgcacgca agctttggtc agcggtaacg ctgaaaaagc 9960
taaaatcgtt gccgctgaat atggcgtcga tccccgtaaa atttatgatt acagcaactt 10020
cgacaagatc gctaaagatc caaaaatcga cgctgtttac atcattttgc caaactcttt 10080
gcatgctgaa tttgctatcc gtgctttcaa agccggcaag catgttatgt gtgaaaagcc 10140
gatggcaacc tctggtgctg attgtcagcg gatgatcgat gcagccaagc ctgctaataa 10200
aaagctgatg atcgggtacc gttgccacta tgatccaatg aaccgtgcaa 10250

<210> 26
<211> 1186
<212> ADN
5 <213> secuencia artificial

<220>
<223> fragmento construido

<400> 26
gtccagaaaa gacagcattc cttctcaata aagaaatatt attttttgtt tttgaaaaat 60
ttttccaaaa tctagaatgc tacattaaat atacaaaaat attattatac aaataaggct 120
tttaaatacc catat tttttt agaatttctt tacaagaaa catgttaaat atagatttag 180
agattaatat cagccatttt tatcaaaaat tctttttttg ttttataata ttatgctgca 240
aaactaataa aaacgccctt tcgaaattaa cgatcaccca caagaaataa ttatctgaca 300
gcgcttacca atcaattatt gccgaacgca gagtcccgta ttaggacggt caacaatcta 360
aaccgttttt cagaaaaatat tgctttataa gcctcaaac ttaaaagctg cggat tttta 420
atataccaaa attttctgga aaagccggcg aatcagataa cagttccgca caggtgagaa 480

ES 2 611 977 T3

ccacgacgga tcttctctga attgttgggt agttaagaaa gaaacaagga ttatgacgaa 540
 caaaatctcg tcttcagata atctttccaa tgctgtttca gcaacggatg acaacgcttc 600
 ccgtacgcca aatctgaccc gtcgcgctct cgttgggtgg ggtggtggac tggccgcagc 660
 tggcgcctta gccagtggtc ttcaggcagc gacgcttctt gctggtgcca gccaggttcc 720
 gaccacgcct gcaggtcgcc cgatgcctta cgcgatccgc ccgatgccgg aagatcgtcg 780
 tttcggttat gctatcgtcg gtctgggtaa atatgccctt aaccagattt taccgggttt 840
 tgccggatgc cagcattccc gcatogaagc tttggtcagc ggtaacgctg aaaaagctaa 900
 aatcgttgcc gctgaatatg gcgtcgatcc ccgtaaaatt tatgattaca gcaacttcga 960
 caagatcgtc aaagatccaa aaatcgacgc tgtttacatc attttgccaa actctttgca 1020
 tgctgaattt gctatccgtg ctttcaaagc cggcaagcat gttatgtgtg aaaagccgat 1080
 ggcaacctct gttgctgatt gtcagcggat gatcgatgca gccaaaggctg ctaataaaaa 1140
 gctgatgatc ggttaccggt gccactatga tccaatgaac cgtgca 1186

<210> 27
 <211> 1302
 <212> ADN
 <213> Zymomonas mobilis

5

<400> 27
 atgacgaaca aaatctcgtc ttcagataat ctttccaatg ctgtttcagc aacggatgac 60
 aacgcttccc gtacgcaaaa tctgaccogt cgcgctctcg ttggtgggtg tgttgactg 120
 gccgcagctg gcgccttagc cagtggctct caggcagcga cgcttcctgc tggtgccagc 180
 caggttccga ccacgcctgc aggtcgcccg atgccttacg cgatccgccc gatgccggaa 240
 gatcgtcgtt toggttatgc tatcgtcggg ctgggtaaat atgcccttaa ccagatthta 300
 ccgggttttg ccggatgcca gcattccogc atcgaagctt tggtcagcgg taacgctgaa 360
 aaagctaaaa tcggtgcccg tgaatatggc gtcgatcccc gtaaaattta tgattacagc 420
 aacttcgaca agatcgctaa agatccaaaa atcgacgctg tttacatcat tttgccaaac 480
 tctttgcatg ctgaatttgc tatccgtgct ttcaaagccg gcaagcatgt tatgtgtgaa 540
 aagccgatgg caacctctgt tgctgattgt cagcggatga tcgatgcagc caaggctgct 600
 aataaaaagc tgatgatcgg ttaccgttgc cactatgatc caatgaaccg tgcagcggta 660
 aaattgatcc gtgaaaacca gttgggtaaa ctgggcatgg ttaccaccga caactcagac 720
 gttatggatc agaacgatcc tgcacagcag tggcgtctgc gtcgtgaact cgccgggtggc 780
 ggttctttga tggatatcgg tatttatggc ttgaacggta cccgttactt gctgggtgaa 840
 gaaccgatcg aagtccgtgc ttacacctac agcgatccga atgatgaacg tttcgttgaa 900
 gtcgaagatc gtattatttg gcagatgcgc ttcagaagcg gtgctctgtc tcatgggtgca 960

ES 2 611 977 T3

tcttcttatt cgaccacgac gacttcacgt ttctcgggtc agggcgacaa agctggtctg 1020
 ttgatggatc cggctaccgg atattatcag aatttgattt ctgtccagac cccaggccat 1080
 gctaaccagt cgatgatgcc acagttcatc atgccagcga acaaccagtt ctctgcacag 1140
 ttggatcatc tggctgaagc cgtcatcaat aacaaaccag ttcgtagccc gggggaagaa 1200
 ggtatgcagg atgtgcccct gattcaggcc atttatgaag cagctcgtac cggtcgcccc 1260
 gtcaacacgg attgggggta tgtccgtcag ggtggttatt ga 1302

<210> 28
 <211> 1446
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> fragmento construido

<400> 28

cgttacttgc tgggtgaaga accgatogaa gtccgtgctt acacctacag cgatccgaat 60
 gatgaacggt tcggtgaagt cgaagatogt attatctggc agatgcgctt cagaagcggg 120
 gctctgtctc atgggtgcac ttcttattcg accacgacga cttcacggtt ctcggtgcag 180
 ggcgacaaaag ctggtctggt gatggatccg gctaccggat attatcagaa tttgatttct 240
 gtccagaccc caggccatgc taaccagtgc atgatgccac agttcatcat gccagcgaac 300
 aaccagttct ctgcacagtt ggatcatctg gctgaagccg tcatcaataa caaaccagtt 360
 cgtagcccgg gtgaagaagg tatgcaggat gtgcgcctga ttcaggccat ttatgaagca 420
 gctcgtaccg gtcgccccgt caacacggat tgggggtatg tccgtcaggg tgggttattga 480
 ttctgactta acctatctgg gttaaacaga cttatcttct ctgcttttagg aaaatagtta 540
 aaaaggcgtc attggttctt ccaatgagc ctttttttat aaacaaaaaa atccttttgt 600
 cggttttata aaaatacttc atatcttgat aagccgtctt aaaaatataa taaatcttta 660
 taatatttat cggatcaaag gacccttta tgctagaagt cattatatcg gcattactac 720
 cgattataat tactttaatg ataggttttt tcgctggctg gcgtggtgaa tttacggcaa 780
 atcaagcctc gaccttgaat aaaatggtct tacgctatgc cttacctatg actttattct 840
 ctgggatttt atcaactccc aaaacacaga tttatcgtc ggggtctgcc gcaattattt 900
 tacttttagc catggctggc ggctatctaa ttacacttgg gataggatat tttgtctgcc 960
 agcggccagt gaatgaatct gctcttttag ctctttctgt tagcgcacct gcagttcctt 1020
 ttgttgcat aacagttcta gggcatttat ttggcactgc cagcacgata ttggtttcaa 1080
 tatgtagcct gatgatgaac ctcgtccagg ttcccgttac ctttttcttt tgtcagcgta 1140
 ttctccaaaa aagaatactg acaaaatagc cacggatagt tctttttttt ctcatatcag 1200

ES 2 611 977 T3

acatgctttt accgaaccta ttgttattgc ccctattctg gctcttatct gtgtcagtct 1260
 ctctattcct ttccctgaaa ccttaaaatc ttctttaatg ctactaggaa aagcgaccgg 1320
 aggcggttgcg cttttctctt ctggtataat tttatcttct cgaaaagtta ttttaagtaa 1380
 aacagtagca tcttttagttt tatcaaaaaa tattattatt ccaacagcgg tattggttct 1440
 tgcgtc 1446

- 5 <210> 29
 <211> 1661
 <212> ADN
 <213> secuencia artificial

- <220>
 <223> fragmento construido
 <400> 29

gttcgatcaa caaccggaat cctatcgtaa tgatgttttg cccgatcagc ctcaatcgac 60
 aattttacgc gtttcgatcg aagcagggac gacaattggc tgggaacggt atacttgaat 120
 aatgggtctt cgttatggta ttgatgtttt tgggtgatcg gccccggcga atgatctata 180
 tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtgttggccg cgatcgccgg 240
 taagtcggca cgttaaaaaa tagctatgga atataatagc tactaataag ttaggagaat 300
 aaacatgagt gtccaagcta cacgcgaaga taaatttagc tttggcttgt ggacagttgg 360
 gtggcaggcc cgtgacgcct ttggtgatgc cacgcgtacc gcccttgatc cggtcgaagc 420
 agttcataaa cttgccgaaa tcggagccta tggcattaca tttcatgatg atgacttgg 480
 ccccttcggt tcggatgctc agacacggga tggatcatt gctggtttca aaaaagccct 540
 tgatgaaacc ggtttgatcg ttcctatgg 1 taccacaaat ttgtttacgc atccggtttt 600
 taaagacggc gggtttacct ctaacgatcg ttctgtgcgc cggatgcta tccggaaagt 660
 tttgagacag atggatcttg gtgcggaact gggggccaaa acgcttgtct tatggggcgg 720
 cagagaaggt gcggaatatg attcagcaaa agatgttagt gccgccttgg atcgttatcg 780
 cgaagcactg aatcttctgg cacaatatag cgaagacaga ggctatggac ttcgttttgc 840
 aattgaaccg aaacctaatg aaccacgtgg cgatattctg ctgcctaccg caggccatgc 900
 tattgccttt gtgcaagaat tagaacggcc cgaacttttt ggcatcaatc cagaaaccgg 960
 gcatgaacag atgtcaaact tgaattttac ccaggggatt gctcaggctt tgtggcataa 1020
 aaaacttttt catattgatt tgaatggaca acatggtccc aagtttgatc aggaccttgt 1080
 ctttggtcat ggtgaccttt taaatgcctt tagcctggtc gatttgtagt aaaatggtcc 1140
 agatggtgcc cggccttatg atggcccgcg ccattttgat tataaacat ctcgtactga 1200
 agattatgat ggcgtttggg aatcagcga agccaatc cgtatgtatc ttttattaaa 1260
 agaacgtgcc aaagcgttta gagctgatcc ggaagttcag gaagcactgg cagcaagcaa 1320

ES 2 611 977 T3

agttgccgaa ttgaaaaccc caacgttgaa tcctggtgaa ggctatgcag aactgttagc 1380
 agatcgcagt gcttttgaag attatgatgc cgatgctggt ggtgcgaaag gttttggttt 1440
 cgtgaaattg aaccaattgg ccattgaaca tttattaggt gcccgctagt ctagatgact 1500
 gtataaaacc acagccaatc aaacgaaacc aggctatact caagcctggt tttttgatgg 1560
 attttcagcg tggcgcaggc aggttttattc ttaacccgac actggcgggga caccocgcaa 1620
 gggacagaag tctccttctg gctggcgacg gacaacgggc c 1661

<210> 30
 <211> 1960
 <212> ADN
 5 <213> secuencia artificial

<220>
 <223> fragmento construido que contiene un P_{eno} de 191 pb, una secuencia de codificación xylB de pb de *E. coli* de 1.455 pb (ECxylB) y un xylB 3'UTR de *E. coli* de 314 pb.

<400> 30
 ccagttactc aatacgtaac aataatcagt ttatcctaac tatagaatcg catgagaagc 60
 gataacgttt caccataagc aatatattca ttgcaacagt ggaattgcct tatgcgctcaa 120
 ggaaggatag atcattgacg gactgagttc aaaaagagac tgggtctaaaa gattttaaga 180
 aaggtttcga tatgtatatac gggatagatc ttggcacctc gggcgtaaaa gttattttgc 240
 tcaacgagca gggtagagtg gttgctgocg aaacggaaaa gctgaccggt tcgocgcccg 300
 atccactctg gtcggaacaa gaccocggaac agtggtgcca ggcaactgat cgcgcaatga 360
 aagctctggg cgatcagcat tctctgcagg acgttaaagc attgggtatt gccggccaga 420
 tgcacggagc aaccttgctg gatgctcagc aacgggtggt acgccctgcc attttgtgga 480
 acgacgggcg ctgtgcgcaa gagtgcactt tgctggaagc gcgagttccg caatcgcggg 540
 tgattaccgg caacctgatg atgcccgat ttactgcgcc taaattgcta tgggttcagc 600
 ggcattgagc ggagatattc cgtcaaactc acaaagtatt attaccgaaa gattacttgc 660
 gtctgcgtat gacgggggag tttgccagcg atatgtctga cgcagctggc accatgtggc 720
 tggatgtcgc aaagcgtgac tggagtgcag tcatgctgca ggcttgcgac ttatctcgtg 780
 accagatgcc cgcattatac gaaggcagcg aaattactgg tgctttgtta cctgaagttg 840
 cgaaagcgtg gggtagggcg acggtgccag ttgtcgcagg cggtagcgac aatgcagctg 900
 gtgcagttgg tgtgggaatg gttgatgcta atcaggcaat gttatcgctg gggacgtcgg 960
 gggctctattt tgctgtcagc gaagggttct taagcaagcc agaaagcgcc gtacatagct 1020
 tttgccatgc gctaccgcaa cgttggcatt taatgtctgt gatgctgagt gcagcgtcgt 1080
 gtctggattg ggccgcgaaa ttaacccggc tgagcaatgt cccagcttta atcgtcgcag 1140
 10 ctcaacaggc tgatgaaagt gccgagccag tttggtttct gccttatctt tccggcgagc 1200

ES 2 611 977 T3

gtacgccaca caataatccc caggcgaagg gggttttcct tggtttgact catcaacatg 1260
 gccccaatga actggcgcga gcagtgctgg aaggcgtggg ttatgctgctg gcagatggca 1320
 tggatgtcgt gcatgcctgc ggtattaaac cgcaaagtgt tacgttgatt gggggcgggg 1380
 cgcgtagtga gtactggcgt cagatgctgg cggatatcag cggtcagcag ctogattacc 1440
 gtacgggggg ggatgtgggg ccagcactgg gcgcagcaag gctggcgcag atcggcggca 1500
 atccagagaa atcgcctcatt gaattggtgc cgcaactacc gttagaacag tcgcatctac 1560
 cagatgctga gcgttatgcc gcttatcagc cagcagcaga aacgttccgt cgcctctatc 1620
 agcaacttct gccattaatg gcgtaaacgt tatcccctgc ctgaccgggt gggggataat 1680
 tcacatctat atatctcagt aattaattaa tatttagtac gaatttattc tgaaaatcat 1740
 ttgttaatgg catttttcag ttttgtcttt cgttggttac tcgtaatgta tcgctggtag 1800
 atatggagat cgttatgaaa acctcaaaga ctgtggcaaa actattattt gttgtcgggg 1860
 cgctggttta tctgggtggg ctatggatct catgccatt gttaagtgga aaaggctatt 1920
 ttcttggcgt gttaatgaca gcaacttttg gcaactatgc 1960

<210> 31
 <211> 1014
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> fragmento construido con gen de resistencia a la spec ligado por sitios de recombinación lox

<400> 31

agcacaggat gacgcctaac aattcattca agccgacacc gcttcgcggc gcggcttaat 60
 tcaggagtta aacatcatga ggaagcggg gatcgccgaa gtatcgactc aactatcaga 120
 ggtagttggc gtcacgcgac gccatctcga accgacgttg ctggccgtac atttgtacgg 180
 ctccgcagtg gatggcggcc tgaagccaca cagtgatatt gatttgctgg ttacgggtgac 240
 tgtaaggctt gatgaaacaa cgcggcgcgc tttgatcaac gaccttttg aaacttcggc 300
 ttcccctgga gagagcgaga ttctccgcgc tgtagaagtc accattggtg tgcacgcgca 360
 catcattccg tggcgttatc cagctaagcg cgaactgcaa tttggagaat ggcagcgcga 420
 tgacattctt gcaggatatc tcgagccagc cagcatcgac attgatctgg ctatottgct 480
 gacaaaagca agagaacata gcgttgccctt ggtaggtcca gcggcggagg aactctttga 540
 tccggttcct gaacaggatc tatttgaggc gctaaatgaa accttaacgc tatggaactc 600
 gccgcccagc tgggctggcg atgagcgaaa tgtagtgctt acgttgctcc gcatttggtg 660
 cagcgcagta accggcaaaa tcgcgcgcaa ggatgtcgtt gccgactggg caatggagcg 720
 cctgccggcc cagtatcagc ccgtcatact tgaagctagg caggcttatt ttggacaaga 780

ES 2 611 977 T3

agatcgcttg gcctcgcgcg cagatcagtt ggaagaatth gttcactacg tgaaaggcga 840
gatcaccaag gtagtcggca aataatgtct aacaattcgt tcaagccgac gccgcttcgc 900
ggcgcggcctt aactcaagcg ttagagagct ggggaagact atgcgcgatc tgttgaaggt 960
ggttctaagc ctcgacttg cgatggcatc ggggcaggca cttgctgacc tgcc 1014

<210> 32
<211> 12198
<212> ADN
5 <213> secuencia artificial

<220>
<223> plásmido construido
<400> 32

ctagtgttcg atcaacaacc cgaatcctat cgtaatgatg ttttgcccga tcagcctcaa 60
tcgacaattt tacgcgtttc gatcgaaaca gggacgacaa ttggctggga acggtatact 120
ggaataaatg gtcttcgtta tggattgat gtttttggtg catcggcccc ggcgaatgat 180
ctatatgctc atttoggcct gaccgcagtc ggcacacga acaaggtgtt ggccgcgatc 240
gccggtaagt cggcacgtta aaaaatagct atggaatata atagctacta ataagttagg 300
agaataaaca tgacggacaa attgacctc cttcgtcagt acaccaccgt agtggccgac 360
actggggaca tcgoggcaat gaagctgtat caaccgcagg atgccacaac caacccttct 420
ctcattctta acgcagcgca gattccggaa taccgtaagt tgattgatga tgctgtcgcc 480
tgggcgaaac agcagagcaa cgatcgcgcg cagcagatcg tggacgcgac cgacaaactg 540
gcagtaaata ttggtctgga aatcctgaaa ctggttccgg gccgtatctc aactgaagtt 600
gatgcgcgcc tttcctatga caccgaagcg tcaattgcga aagcaaacg cctgatcaaa 660
ctctacaacg atgctggtat tagcaacgat cgtattctga tcaaactggc ttctacctgg 720
cagggtatcc gtgctgcaga acagctggaa aaagaaggca tcaactgtaa cctgaccctg 780
ctgttctcct tcgctcaggc tcgtgcttgt gcggaagcgg gcgtgttcct gatctcgccg 840
tttgttggcc gtattcttga ctggtacaaa gcgaataccg ataagaaaga gtacgctccg 900
gcagaagatc cgggcgtggg ttctgtatct gaaatctacc agtactacaa agagcacggt 960
tatgaaaccg tggttatggg cgcaagcttc cgtaacatcg gcgaaattct ggaactggca 1020
ggctgcgacc gtctgaccat cgcaccggca ctgctgaaag agctggcgga gagcgaaggg 1080
gctatcgaac gtaaactgtc ttacaccggc gaagtgaaag cgcgtccggc gcgtatcact 1140
gagtccgagt tcctgtggca gcacaaccag ggtccaatgg cagtagataa actggcggaa 1200
ggtatccgta agtttgctat tgaccaggaa aactggaaa aatgatcgg cgatctgctg 1260
taatctagac gatctggagt caaatgtcc tcacgtaaag agcttgccaa tgctattcgt 1320

ES 2 611 977 T3

gcgctgagca tggacgcagt acagaaagcc aaatccggtc acccggtgc ccctatgggt 1380
atggctgaca ttgccgaagt cctgtggcgt gatttctga aacacaacc gcagaatccg 1440
tcctgggctg accgtgaccg cttcgtgctg tccaacggcc acggctccat gctgatctac 1500
agcctgctgc acctcaccgg ttacgatctg ccgatggaag aactgaaaaa cttccgtcag 1560
ctgcaactcta aaactccggg tcacccggaa gtgggttaca ccgctggtgt ggaaaccacc 1620
accggtccgc tgggtcaggg tattgccaac gcagtcggtg tggcgattgc agaaaaaacg 1680
ctggcggcgc agtttaaccg tccgggccac gacattgtcg accactacac ctacgccttc 1740
atgggcgacg gctgcatgat ggaaggcatc tcccacgaag tttgctctct ggccgggtacg 1800
ctgaagctgg gtaaactgat tgcattctac gatgacaacg gtatttctat cgatggtcac 1860
gttgaaggct ggttcaccga cgacaccgca atgcgttctg aagcttacgg ctggcacggt 1920
attcgcgaca tcgacgggtca tgacgggca tctatcaaac gcgcagtaga agaagcgcgc 1980
gcagtgactg acaaaccttc cctgctgatg tgcaaaacca tcatcggttt cggttccccg 2040
aacaagccg gtaccacga ctcccacggt gcgccgctgg gcgacgctga aattgccctg 2100
accgcgaac aactgggctg gaaatatgcg ccgttcgaaa tcccgtctga aatctatgct 2160
cagtgggatg cgaaagaagc aggccaggcg aaagaatccg catggaacga gaaattcgct 2220
gcttacgca aagcttatcc gcaggaagcc gctgaattta cccgccgtat gaaaggcga 2280
atgccgtctg acttcgacgc taaagcgaag gagttcatcg ctaaactgca ggctaaccg 2340
gcgaaaatcg ccagccgtaa agcgtctcag aatgctatog aagcgttcgg tccgctgtt 2400
ccggaattcc tcggcgggtc tgctgacctg gcgccgtcta acctgacct gtggtctggt 2460
tctaaagcaa tcaacgaaga tgctgagggt aactacatcc actacggtgt tcgagagttc 2520
ggatgaccg cgattgctaa cggtatctcc ctgcacggtg gcttcctgcc gtacacctcc 2580
accttctga tgttcgtgga atacgcacgt aacgccgtac gtatggctgc gctgatgaaa 2640
cagcgtcagg tgatggttta caccacgac tccatcggtc tgggcgaaga cggcccact 2700
caccagccgg ttgagcaggc cgcttctctg cgcgtaacct cgaacatgtc tacatggcgt 2760
ccgtgtgacc aggttgaatc cgcggtcgcg tggaaatacg gtggtgagcg tcaggacggc 2820
ccgaccgcac tgatcctctc ccgtcagaac ctggcgcagc aggaacgaac tgaagagcaa 2880
ctggcaaca tcgcgcgcgg tggttatgtg ctgaaagact gcgccggtca gccggaactg 2940
atcttcatcg ctaccgggtc agaagttgaa ctggctgttg ctgcctacga aaaactgact 3000
gccgaaggcg tgaaagcgcg cgtggtgtcc atgccgtcta ccgacgcatt tgacaagcag 3060
gatgctgctt accgtgaatc cgtactgccg aaagcggta ctgcacgcgt tgctgtagaa 3120
gcgggtattg ctgactactg gtacaagtat gttggcctga accggtgctat cgtcggtatg 3180
accaccttcg gtgaatctgc tccggcagag ctgctgtttg aagagttcgg cttcactggt 3240

ES 2 611 977 T3

gataacgttg ttgcgaaagc aaaagaactg ctgtaattag catttcgggt aaaaaaggct 3300
gcttcggcga ccttttttat taccttgata atgtccgttt gcgcggcgcg ccccagttac 3360
tcaatacgtg acaataatca gtttatocta actatagaat cgcgatgagaa gcgataacgt 3420
ttcaccataa gcaatatatt cattgcaaca gtggaattgc cttatgcgtc aaggaaggat 3480
agatcattga cggactgagt tcaaaaagag actggtctaa aagattttaa gaaaggtttc 3540
gatatgacct ctgctgtgcc atcaaatacg aaaaaaagc tggtgattgc ttccgatcac 3600
gcagcatttg agttgaaatc aaccttgatt acttggtgga aagagcttgg tcatgaggtc 3660
gaagacctg gccctcatga aaaccattca gtcgattatc ccgattacgg ttataagctg 3720
gctgtcgtc tgcagaaaa aaccgctgat ttccgtattg ctttatgtgg ctogggaatc 3780
ggatctcga tcgctgtcaa tcgccatccg gctgcccgtt gcgctttgat tacggataac 3840
cttaccgccc gtttggcaag agaacataac aatgccaatg ttatcgctat ggggtgcgaga 3900
ttgatcggca ttgaaaccgc taaggattgt atttcagctt tccttgcaac gccgtttgga 3960
ggatgaacgtc atgttcgccc tatcgataaa ctttcgaaac ctcagttcaa tatctagata 4020
agttaggaga ataaacatga gtaaattacc cctgattgct ccctctatcc tttcggcgga 4080
ttttgccat ttgggagatg aggtcgcggc gatagatcag gccggtgccg attggatcca 4140
tattgatgtg atggatggcc atttcgtgcc gaatatcacc ataggcccca tggttgtgaa 4200
ggctttgct ccctatagcc aaaagccttt tgatgtccat ttgatgattg cgctgtcga 4260
tcaatatac gaggcctttt ctgaagcggg tgctgatatt atcagtttc atcccgaagc 4320
gggcgcgcat ccccatcgca ctattcagca tatcaaatca ttgggcaaaa aagcgggatt 4380
agtttttaat cggcgaccc ctttaagctg gcttgattat ctaatggatg atcttgatct 4440
gattatggtg atgagcgtta accccggttt tggcggccaa aaatttatca aaaccaatt 4500
agaaaagatt aaagatatcc gtcaaagaat taccgcctct gggcgggata tccgcttga 4560
agtggatggc ggaattgatg ccacgactgc accgcttgc gtcgaagccg gtgccgatgt 4620
tttggtcgcg ggaacggcca gctttaaggg cggcgcaaca tgttacaccg ataatatcag 4680
gatattgctg aatcatgat taattaactc gagcggcct gaacgtactg caagtcctga 4740
cgtcactgtg cagtccgttg gcccggttat cggtagcgat accgggcatt tttttaagga 4800
acgatcgata gaattcgcgg gccggccaag cttgaattca tggttttggt gccaatgtta 4860
tcgcctataa accgcatcca gacccogaat tggcgaaaaa ggtcggtttc cgcttcacct 4920
ctctcgatga agtgatcgag accagcgaca tcatttcgct tcaactgtccg ctcacgccag 4980
aaaatcatca catgattaat gaagaaacac tggcaagggc aaaaaaggc ttttacctcg 5040
tcaataccag tcggggcggc ttggttgata ccaaggcggg gattaatcgt ctgaaagcca 5100

ES 2 611 977 T3

aacatctcgg cggttatgcg gcggatgttt acgaagagga ggggccttta ttcttcgaaa 5160
atcacgctga cgatattatc gaagatgata ttctcgaaag gttgatcgct ttcccgaatg 5220
tggttttcac gggacatcag gcctttttga cgaaagaggc cttatcaaac attgctcaca 5280
gtattctaca agatatacgc gatgccgaag ctggaaaaga aatgccggat gcgcttgttt 5340
agtagacaag cgacaattaa ctttttgaag atcataatga tcaaattttt gggtaattc 5400
ggtagttatg gcataggcta ttacgcgcta attgatataca aaaaaagca tagccggaca 5460
tcataccggc tatgtttttt attaggaaaa aatttccttt caccttgctt agccatcgcc 5520
gcattattta atcaatatgc cgagtttttc ttgaaatccc tatcttacac caaggccaac 5580
aagggaatca tccatactcg gtgtcctatc ctatgacttt ttaaattttc tccaaattta 5640
ctaaaatcac gccatctcag cggctgctat tttcaaaaag cgcctctcaa aacogctttt 5700
tcctgctcaa atatcggatc ccaaaattcc ctcaaaaaag gcagggtatt ttttacaaaa 5760
tcgcccctaa tatctctcaa tccgctgcct tgttcatatg tttttgcaa tgatttttat 5820
taaacttttt taggcgtatt tttatcaaga aaatttaaat aatcacattt ttattatttt 5880
agatttaagt attgatataca gtgatataca taaatgtttt tataactttc tggatogtaa 5940
tcggctggca atcgtttttc ctatattcgc aagatgtatg tcagccgcgc gccgcctaa 6000
ttccggatga gcattcatca ggcgggcaag aatgtgaata aaggccggat aaaacttggtg 6060
cttatttttc tttacgggtc ttaaaaaggc cgtaatatcc agctgaacgg tctggttata 6120
ggtagattga gcaactgact gaaatgcctc aaaatgttct ttacgatgcc attgggatat 6180
atcaacggtg gtatatccag tgattttttt ctccatttta gcttccttag ctctgaaaa 6240
tctcgataac tcaaaaaata cgcccggtag tgatcttatt tcattatggt gaaagttgga 6300
acctcttacg tgccgatcaa cgtctcattt tcgccaaaag ttggcccagg gcttcccggg 6360
atcaacaggg acaccaggat ttattttattc tgcgaagtga tcttccgtca caggatttta 6420
ttcggcgcaa agtgcgctcg gtgatgctgc caacttactg atttagtgta tgatggtggt 6480
tttgaggtgc tccagtggtc tctgtttcta tcagctgtcc ctctgttca gctactgacg 6540
gggtggtgcg taacggcaaa agcaccgccc gacatcagcg ctagcggagt gtatactggc 6600
ttactatggt ggcactgatg aggggtgcag tgaagtgcct catgtggcag gagaaaaaag 6660
gctgcaccgg tgcgtcagca gaatatgtga tacaggatat attccgcttc ctgctcact 6720
gactcgctac gctcggctcg tgcactgcgg cgagcggaaa tggcttacga acggggcgga 6780
gatttcctgg aagatgccag gaagatactt aacagggaag tgagagggcc gcggcaaagc 6840
cgtttttcca taggctccgc cccctgaca agcatcacga aatctgacgc tcaaatcagt 6900
ggtagcgaaa cccgacagga ctataaagat accaggcgtt tcccctggc ggctccctcg 6960
tgcgctctcc tgttctgccc tttcggttta ccggtgtcat tccgctgtta tggccgcggt 7020

ES 2 611 977 T3

tgtctcattc cacgcctgac actcagttcc gggtaggcag ttcgctccaa gctggactgt	7080
atgcacgaac cccccgttca gtccgaccgc tgcgccttat cgggtaacta tcgtcttgag	7140
tccaaccocgg aaagacatgc aaaagcacca ctggcagcag ccactggtaa ttgatttaga	7200
ggagttagtc ttgaagtcac gcgcccggta aggctaaact gaaaggacaa gttttggtga	7260
ctgcgctcct ccaagccagt tacctcgggt caaagagttg gttagctcaga gaaccttcga	7320
aaaaccgccc tgcaaggcgg ttttttcggt ttcagagcaa gagattacgc gcagaccaa	7380
acgatctcaa gaagatcacc ttattaatca gataaaatat ttctagattt cagtgcgaatt	7440
tatctcttca aatgtagcac ctgaagtcag cccatacga tataagttgt aattctcatg	7500
tttgacagct tatcatcgat gtgacggaag atcacttcgc agaataaata aatcctgggtg	7560
tcctgtttga taccgggaag ccctgggcca acttttggcg aaaatgagac gttgatcggc	7620
acgtaagagg ttccaacttt caccataatg aaataagatc actaccgggc gtattttttg	7680
agttatcgag attttcagga gctaaggaag ctaaaatgga gaaaaaatc actggatata	7740
ccaccgttga tatatcccaa tggcatcgta aagaacattt tgaggcattt cagtccagttg	7800
ctcaatgtac ctataaccag accgttcagc tggatattac ggccttttta aagaccgtaa	7860
agaaaaataa gcacaagttt tatccggcct ttattcacat tcttgcccgc ctgatgaatg	7920
ctcatccgga attccgtatg gcaatgaaag acggtgagct ggtgatatgg gatagtgttc	7980
acccttgtta caccgttttc catgagcaaa ctgaaacggt ttcacgctc tggagtgaat	8040
accacgacga tttccggcag tttctacaca tatattcgca agatgtggcg tgttacggtg	8100
aaaacctggc ctatttccct aaagggttta ttgagaatat gtttttcgtc tcagccaatc	8160
cctgggtgag tttcaccagt tttgatttaa acgtggccaa tatggacaac ttcttcgccc	8220
ccgtttttcac catgggcaaa tattatacgc aaggcgacaa ggtgctgatg ccgctggcga	8280
ttcaggttca tcatgccgtt tgtgatggct tccatgtcgg cagaatgctt aatgaattac	8340
aacagtactg cgatgagtgg caggcggggg cgtaattttt ttaaggcagt tattggtgcc	8400
cttaaacgcc tggttgctac gcctgaataa gtcgacctt gtagtcttgg cctgttgtgt	8460
gcatgagcaa atcaatggca ccacccctc ctttttgagc tgaatggtca taaaatttat	8520
aattatctat cgtaattcgg aatctatggt cagggtctcg ccattgcttt ttgtctgctg	8580
ggtcaagttc catgcctaag gtttttaaga catcagaaag aggtattgca cgcagctat	8640
cagcttttct tctagctaat gacaggcctt cctctgctct atctgctcgt ttttttctt	8700
ccacatatct cgcgcctttg tcagccagcg gctgtattac ggaaagtgcc gatttttggg	8760
cttttagggc ttotttttct gccattctt ccttatttgt aaaaattgag ggtgggatgg	8820
gtgcctgaat cttgggatct agctgtaaag ttttgttgat atttccgtaa tgtctttgga	8880

ES 2 611 977 T3

ctctttgatg cgttgctttt gaacctttta cgctctggc cagccctaga ggctccatag 8940
aagccgcata atccgtctgg agggcagaaa gggcttttcg accatcaaac catctcgatg 9000
cgtttaaacg gcctgtatcg gggctctctag gcaccataaa gccgggtaag tgggggtgtg 9060
tttcatcagc atgtagctga agagatacaa ggttgttttc tccaaagggt tgttccgccc 9120
attgctgggt gattgttttc cagtgttcga gtttttcagg agtggcctgt tttgaccatt 9180
ctggagacat accaaagaac agttctatgg cctgcacacc gttttttcta agaggctttc 9240
ccgtttcttt ctgaatttta ttcagcatag atttaacatc tgctgatggg tcagtagagc 9300
ctttgagtat ttogtttagt tcttttctat ctgggtcagc gttttgtgtt tcgcggcctc 9360
gcgtcatatg caggctcgcg gctttaatcg tgccaactgt tttatgtttt tcaaacctaa 9420
agattgcata gttcggcatg ttttaactgc ttaatttga gaaaagacca gaggaaataa 9480
tccagcctat atttctttcc ctagtagcga actggaattg tttttccgaa ggaaaaaagc 9540
aattccgtag tgagtactga atttattctg attcgtcttg cttttggagc gtctttttgc 9600
gttctataac tgttgtgaaa gctacgcggg cgcattgaa aacgaaatta ggattaataa 9660
aataccatcc ttggcgaaca tgctttgcaa tgattttagc tttttctaata tcggctagac 9720
ctcttgcaaa ggtagcttga gatagtcca gtttttttc ttgtgcgta agaaagtcct 9780
ctaaaacgaa tttgtctaaa gggacgaggt ctttctgat gcctttgtct tgaagtatcc 9840
aaaccagaac gctgaaagct tttattccag cggtccttag ttcaaaagt agcgcgatat 9900
tggtgctaaa taattttaca aattcttcac tatcaacacg tctgtaagtc gtcacatgag 9960
tgccttgcac ctcaccagtg gcttgattga ccagaatgtt atcatctcgt cctaactcag 10020
ataactgaac cctctgactt ttaactggca caaccatacc ttcgatgaaa ggattctcgt 10080
catatctgat tggctgcttt ctcaattttg tcgccatatt tgataaacct ttaatcaaaa 10140
aaaccacatt ttttgattat acctattcat cgaatgaggc aaggctctac aattttacc 10200
ctttttttga tagacggttt aatcaatatt gatagacccc ttcacagatt ctgaaaatcg 10260
acttcctat tttagggata ttttcacgat tccctttctt agttcttcct agtggggaaa 10320
ttcgttgaat cctgcctcgg aaaaacctg agaaagctgt tggttatata cacgggcaaa 10380
gccaccctat ttttagctac tggggaaaaga gataaggcag ggtatttcta aaattaaaac 10440
cggatttttc gctttacggg ttgttttaggc gcaactgtct ttttaagacc gcgtttaacc 10500
atcaaaagat cgttccaatc ttttcogtgt atcatctggt ctttaggtgg gagccagttt 10560
tcaacttttt ttggttgaaa cgcggcttta atcgtccga ctaatagcga tgctgctctt 10620
tgtcctacag catcccaatc ataggcaata tggacagaag atgccttttc aacgattttt 10680
cggagagttt tagtaagaga cgttcttacg ccgctgggtc ttaataattt tacgccagct 10740
ttaatttttt ctgggcttaa aaagcogact actgaaatcg cgtctatcgc actttcagcg 10800

ES 2 611 977 T3

atataaagat catacttttc gtcatttttt acattgatgc tgccagtaaa atgggcttcg 10860
cgactgcttc ccaaggctaa ccctttaaaa ccactgcttg ttccgcgtaa ttctgcgccc 10920
tgaagtgtat ctttatcgtc atacatcaag aaggctacat taccgcgatc atctgttcgg 10980
atagagtcag gaatattggt aaatgatatt cctcgggcag cggtgggtcc tggccacggg 11040
tgcgcatgat cgtgctcctg tcgttgagga cccggctagc ctggcggggg tgccttactg 11100
gttagcagaa tgaatcaccg atacgcgagc gaacgtgaag cgactgctgc tgcaaaacgt 11160
ctgcgacctg agcaacaaca tgaatggtct tcggtttccg tgtttcgtaa agtctggaaa 11220
cgcggaagtc ccctacgtgc tgctgaagtt gcccgcaaca gagagtggaa ccaaccgggtg 11280
ataccacgat actatgactg agagtcaacg ccatgggagc tcgtttttct atccccatca 11340
cctcggtttt gttgacaaaa aaaggtggcc actaaattgg ctttcgcac cgatgggatg 11400
atTTTTattc tttgctattc ttctctcttt gcccaattca ttaaaagcgg aaatcatcac 11460
caaagataga agacgcagcc ttcaccattt cagattgcc ttctcgggca ttttctgctg 11520
ctagaatcct cttaaaaaata ttaaattcca ctctattggt aatatgtttc cctctttagg 11580
gaacaaataa agcccttctt tgttctataa aagttagctt accgatttta caaaaaataa 11640
taccgcttca ttcaatcggg aatacatatc tttttcttc aaaaaacttt tcaagagggt 11700
gtctatgctc gtgcgaatat tcagttccaa aaactatgac catcattcta ttgaaaaga 11760
aatgaacat tatggccatg accttgtttt tctgaatgag cggcttacca aagagacagc 11820
agaaaaagcc aaagacgcag aagctgtttg tatctttgtg aatgacgaag ccaatgccga 11880
agtgctggaa attttggcag gcttaggcat caagttggtt gctcttcggt gcgccggtta 11940
taacaatgct gatctcgatg cggccaaaaa gctgaatatc aaggttgtgc gcgtgcctgc 12000
ctattcggcc tattcgggtg ccgaatatgc agtagggatg ttgctcacc tgaatcggca 12060
aatttcacgc ggtttgaagc ggttcggga aaataacttc tccttgaag gtttgattgg 12120
ccttgatgtg catgacaaaa cagtcggcat tatcgggtgt ggtcatatcg ggagcgtctt 12180
tgcccatatt atgacca 12198

<210> 33

<211> 875

<212> ADN

5 <213> secuencia artificial

<220>

<223> fragmento construido para dirigir la integración en el locus ldh de Z. mobilis

<400> 33

gtttttctat ccccatcacc tcggttttgt tgacaaaaaa aggtggccac taaattggct 60
ttccgcaccg atgggatgat ttttattctt tgctattctt cgtctttgc ccaattcatt 120

ES 2 611 977 T3

aaaagcggaa atcatcacca aagatagaag acgcagcctt caccatttca gattgccctt 180
 ctctgggcatt ttctgctgct agaatcctct taaaaatatt aaattccact ctattggtaa 240
 tatgtttccc tctttaggga acaaataaag cccttctttg ttctataaaa gttagcttac 300
 cgattttaca aaaaataata ccgcttcatt caatcggtaa tacatatctt ttttcttcaa 360
 aaaacttttc aagaggggtgt ctatgcgcgt cgcaatattc agttccaaaa actatgacca 420
 tcattctatt gaaaaagaaa atgaacatta tggccatgac cttgtttttc tgaatgagcg 480
 gcttaccaaa gagacagcag aaaaagccaa agacgcagaa gctgtttgta tctttgtgaa 540
 tgacgaagcc aatgccgaag tgctggaaat tttggcaggc ttaggcatca agttggttgc 600
 tcttcgttgc gccggttata acaatgtcga tctcgatgcg gccaaaaagc tgaatatcaa 660
 gggtgtgctg gtgcctgcct attcgccta ttcggttgc gaatatgcag tagggatgtt 720
 gctcaccctg aatcggcaaa tttcacgctg tttgaagcgg gttcgggaaa ataacttctc 780
 cttggaaggt ttgattggcc ttgatgtgca tgacaaaaca gtcggcatta tcggtgttgg 840
 tcatatcggg agcgtctttg cccatattat gacct 875

<210> 34
 <211> 1149
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> fragmento construido para integración en el locus Idh de Z. mobilis
 <400> 34

atggtttttg tgccaatggt atcgcctata aaccgcatcc agaccccga tggcgaaaa 60
 aggtcggttt ccgcttcacc tctctogatg aagtgatoga gaccagcgac atcatttcgc 120
 ttcactgtcc gctcacgcca gaaaatcatc acatgattaa tgaagaaaca ctggcaaggg 180
 caaaaaaagg cttttacctc gtcaatacca gtcgcgctgg cttggttgat accaaggcgg 240
 tgattaaatc gctgaaagcc aaacatctcg gcggttatgc ggcggatggt tacgaagagg 300
 aggggccttt attcttcgaa aatcacgctg acgatattat cgaagatgat attctcgaaa 360
 ggttgatcgc tttcccgaat gtggttttca cgggacatca ggcctttttg acgaaagagg 420
 ccttatcaaa cattgctcac agtattctac aagatattcag cgatgccgaa gctggaaaag 480
 aaatgccgga tgcgcttggt tagtagacaa gcgacaatta accttttgaa gatcataatg 540
 atcaaatttt tgggttaatt cggtagtatt ggcataggct attacgcgct aattgatatc 600
 aaaaaaagc atagccggac atcataccgg ctatgttttt tattaggaaa aaatttcctt 660
 tcaccttgct tagccatcgc cgcattattt aatcaatatg ccgagttttt cttgaaatcc 720
 ctatcttaca ccaaggccaa caaggaatc atccatactc ggtgtcctat cctatgactt 780

ES 2 611 977 T3

tttaaatttt ctccaaattt actaaaatca cgccatctca gcggtgcta ttttcaaaaa 840
 ggcctctca aaaccgcttt ttctgctca aatcggat cccaaaattc cctcaaaaa 900
 ggcagggtat tttttacaaa atcgcccta atatctctca atccgctgcc ttgttcatat 960
 gtttttgcaa atgattttta ttaaactttt ttaggcgtat ttttatcaag aaaatttaaa 1020
 taatcacatt tttattattt tagatttaag tattgataca agtgatatct ataatgttt 1080
 ttataacttt ctggatcgta atcggtggc aatcgttttc cctatattcg caagatgtat 1140
 gtcagccgc 1149

<210> 35
 <211> 996
 <212> ADN
 <213> Zymomonas mobilis

5

<400> 35
 atgctgctcg caatattcag ttccaaaaac tatgaccatc attctattga aaaagaaaat 60
 gaacattatg gccatgacct tgtttttctg aatgagcggc ttaccaaaga gacagcagaa 120
 aaagccaaag acgcagaagc tgtttgtatc tttgtgaatg acgaagccaa tgccgaagtg 180
 ctggaaattt tggcaggctt aggcatacaag ttggttgctc ttcggtgctc cggttataac 240
 aatgtcgcgc tcgatgcggc caaaaagctg aatatcaagg ttgtgcgcgt gcctgcctat 300
 tcgccctatt cggttgccga atatgcagta gggatgttgc tcaccctgaa tcggcaaatt 360
 tcacgcggtt tgaagcgggt tcgggaaaat aacttctcct tggaaggttt gattggcctt 420
 gatgtgcatg acaaaacagt cggcattatc ggtgttggtc atatcgggag cgtctttgcc 480
 catattatga cccatggttt tggtgccaat gttatgcct ataaaccgca tccagacccc 540
 gaattggcga aaaaggctcg tttccgcttc acctctctcg atgaagtgat cgagaccagc 600
 gacatcattt cgcttactg tccgctcagc ccagaaaatc atcacatgat taatgaagaa 660
 aactggcaa gggcaaaaaa aggcttttac ctctcaata ccagtcgcgg cggcttggtt 720
 gataccaagg cgggtgattaa atcgctgaaa gccaaacatc tcggcgggta tgcggcggat 780
 gtttacgaag aggaggggccc tttattcttc gaaaatcacg ctgacgatat tatcgaagat 840
 gatattctcg aaaggttgat cgctttcccg aatgtggttt tcacgggaca tcaggccttt 900
 ttgacgaaag aggccttctc aaacattgct cacagtatc tacaagatat cagcgtgccc 960
 gaagctggaa aagaaatgcc ggatgcgctt gtttag 996

<210> 36
 <211> 3339
 <212> ADN
 <213> secuencia artificial

10

<220>
 <223> fragmento construido que contiene un mutante T de 304 pb del promotor GAP de *Z. mobilis* (P_{GAP^T}), una región de codificación Tal de *E. coli* de 954 pb (ECTal), una región de codificación Tkt de *E. coli* de 1.992 pb y una Tkt 3'UTR de *E. coli* de 68 pb (ECTkt 3'UTR).

15

<400> 36

ES 2 611 977 T3

gttcgatcaa caaccgaat cctatcgtaa tgatgtttg cccgatcagc ctcaatcgac 60
aattttacgc gtttcgatcg aaacagggac gacaattggc tgggaacggt atactggaat 120
aatgggtctt cgttatggta ttgatgtttt tggatgatcg gccccggcga atgatctata 180
tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtggtggccg cgatcgccgg 240
taagtccgca cgtaaaaaa tagctatgga atataatagc tactaataag ttaggagaat 300
aaacatgacg gacaaattga cctcccttcg tcagtacacc accgtagtgg ccgacactgg 360
ggacatcgcg gcaatgaagc tgtatcaacc gcaggatgcc acaaccaacc cttctctcat 420
tcttaacgca gcgcagattc cggaataccg taagttgatt gatgatgctg tcgcctgggc 480
gaaacagcag agcaacgatc gcgcgcagca gatcgtggac gcgaccgaca aactggcagt 540
aatatttggc ctggaaatcc tgaaactggc tccgggcccgt atctcaactg aagttgatgc 600
gcgcctttcc tatgacaccg aagcgtcaat tgcgaaagca aaacgcctga tcaaactcta 660
caacgatgct ggtattagca acgatcgtat tctgatcaaa ctggcttcta cctggcaggg 720
tatccgtgct gcagaacagc tggaaaaaga aggcatcaac tgtaacctga ccctgctgtt 780
ctccttcgct caggctcgtg cttgtgcgga agcgggcccgt ttctgatct cgccgtttgt 840
tggccgtatt cttgactggc acaaagcga taccgataag aaagagtacg ctccggcaga 900
agatccgggc gtggtttctg tatctgaaat ctaccagtac tacaagagc acggttatga 960
aacctgtggt atgggocgca gcttccgtaa catcggcgaa attctggaac tggcaggctg 1020
cgaccgtctg accatcgcac cggcactgct gaaagagctg gcggagagcg aaggggctat 1080
cgaacgtaaa ctgtcttaca ccggcgaagt gaaagcgcgt ccggcgcgta tcaactgagtc 1140
cgagttcctg tggcagcaca accagggctc aatggcagta gataaactgg cggaaggtat 1200
ccgtaagttt gctattgacc aggaaaaact ggaaaaaatg atcggcgatc tgctgtaatc 1260
tagacgatct ggagtcaaaa tgcctcaccg taaagagctt gccaatgcta ttcgtgcgct 1320
gagcatggac gcagtacaga aagccaaatc cggtaaccgg ggtgccccta tgggtatggc 1380
tgacattgcc gaagtcctgt ggcgtgattt cctgaaacac aaccgcaga atccgtcctg 1440
ggctgaccgt gaccgcttcg tgctgtccaa cggccacggc tccatgctga tctacagcct 1500
gctgcacctc accggttacg atctgccgat ggaagaactg aaaaacttcc gtcagctgca 1560
ctctaaaact ccgggtcacc cggaaagtgg ttacaccgct ggtgtggaaa ccaccaccgg 1620
tccgctgggt cagggatttg ccaacgcagt cggatggcg attgcagaaa aaacgctggc 1680
ggcgcagttt aaccgtccgg gccacgacat tgtcgaccac tacacctacg ccttcatggg 1740

ES 2 611 977 T3

cgacggctgc atgatggaag gcatctocca cgaagtttgc tctctggcgg gtacgctgaa 1800
gctgggtaaa ctgattgcat tctacgatga caacggtatt tctatcgatg gtcacgttga 1860
aggctggttc accgacgaca ccgcaatgcg ttctgaagct tacggctggc acgttattcg 1920
cgacatcgac ggtcatgacg cggcatctat caaacgcgca gtagaagaag cgcgcgcagt 1980
gactgacaaa ccttcctcgc tgatgtgcaa aaccatcadc ggtttcgggt ccccgaacaa 2040
agccgggtacc cacgactccc acggtgcgcc gctggggcgc gctgaaattg ccctgacccg 2100
cgaacaactg ggctggaaat atgcgccggt cgaaatcccg tctgaaatct atgctcagtg 2160
ggatgcgaaa gaagcaggcc aggcgaaaga atccgcatgg aacgagaaat tcgctgctta 2220
cgcgaaagct tatccgcagg aagccgctga atttaccgcg cgtatgaaag gcgaaatgcc 2280
gtctgacttc gacgctaaag cgaaagagtt catcgctaaa ctgcaggcta atccggcgaa 2340
aatcgccagc cgtaaagcgt ctcagaatgc tatcgaagcg ttcgggtccgc tgttgccgga 2400
attcctcggc ggttctgctg acctggcgcc gtctaacctg accctgtggt ctggttctaa 2460
agcaatcaac gaagatgctg cgggtaacta catccactac ggtgttcgcg agttcggtat 2520
gaccgcgatt gctaaccgta tctccctgca cgggtggcttc ctgccgtaca cctccacett 2580
cctgatgttc gtggaatacg cacgtaacgc cgtacgtatg gctgcgctga tgaaacagcg 2640
tcaggtgatg gtttacaccc acgactccat cggctctgggc gaagacggcc cgactcacca 2700
gccggttgag caggtcgctt ctctgcgcgt aaccccgaa atgtctacat ggcgtccgtg 2760
tgaccaggtt gaatccgcgg tcgcgtggaa atacggtggt gagcgtcagg acggccccgac 2820
cgcactgatc ctctcccgtc agaacctggc gcagcaggaa cgaactgaag agcaactggc 2880
aaacatcgcg cgcgggtggtt atgtgctgaa agactgcgcc ggtcagccgg aactgatttt 2940
catcgctacc ggttcagaag ttgaactggc tgttgcctgc tacgaaaaac tgactgccga 3000
aggcgtgaaa gcgcgcgtgg tgtccatgcc gtctaccgac gcatttgaca agcaggatgc 3060
tgcttaccgt gaatccgtac tgccgaaagc ggttactgca cgcggttgcg tagaagcggg 3120
tattgctgac tactggtaaca agtatggtgg cctgaacggt gctatcgtcg gtatgaccac 3180
cttcggtgaa tctgctccgg cagagctgct gtttgaagag ttcggcttca ctggtgataa 3240
cgttggtgcg aaagcaaaaag aactgctgta attagcattt cgggtaaaaa aggtcgcctc 3300
ggcgaccttt tttattacct tgataatgtc cgtttgcgc 3339

<210> 37
<211> 304
<212> ADN

5 <213> secuencia artificial

<220>

<223> promotor mutante Pgap con un cambio de "G" a "A" en la posición 83 en SEQ ID NO : 21

<400> 37

ES 2 611 977 T3

gttcgatcaa caaccogaat cctatcgtaa tgatgttttg cccgatcagc ctcaatcgac 60
aattttacgc gtttcgatcg aaacagggac gacaattggc tgggaacggt atactggaat 120
aatgggtcct cgttatggta ttgatgtttt tggatgcatcg gccccggcga atgatctata 180
tgctcatttc ggcttgaccg cagtcggcat cacgaacaag gtggtggccg cgatcgccgg 240
taagtccgca cgtaaaaaa tagctatgga atataatagc tactaataag ttaggagaat 300
aaac 304

<210> 38
<211> 1443
<212> ADN

5 <213> secuencia artificial

<220>

<223> fragmento construido que contiene un P_{eno} de 191 pb, una secuencia de codificación Rpi de *Z. mobilis* (ZMRpi) de 471 pb Rpi, una secuencia de codificación de Rpe de *Z. mobilis* (ZMRpe) de 663 pb, y un xylA 3'UTR de *E. coli* de 35 pb (ECxylA 3'UTR)

10 <400> 38

ccagttactc aatacgtaac aataatcagt ttatcctaac tatagaatcg catgagaagc 60
gataacgttt caccataagc aatatattca ttgcaacagt ggaattgcct tatgcgtaa 120
ggaaggatag atcattgacg gactgagttc aaaaagagac tggctctaaa gattttaaga 180
aaggtttcga tatgacctct gctgtgccat caaatcagaa aaaaagctg gtgattgctt 240
ccgatcacgc agcatttgag ttgaaatcaa ccttgattac ttggctgaaa gagcttggtc 300
atgaggtcga agaccttggc cctcatgaaa accattcagt cgattatccc gattacgggt 360
ataagctggc tgtcgctatc gcagaaaaa ccgctgattt cggtattgct ttatgtggct 420
cgggaatcgg tatctcgatc gctgtcaatc gccatccggc tgcccgttgc gctttgatta 480
cggataacct taccgccctg ttggcaagag aacataaaa tgccaatggt atcgctatgg 540
gtgagagatt gatcggcatt gaaaccgcta aggattgtat ttcagctttc cttgcaacgc 600
cgtttgagg tgaacgtcat gttcgcgta tcgataaact ttogaatcct cagttcaata 660
tctagataag ttaggagaat aaacatgagt aaattacccc tgattgctcc ctctatcctt 720
tcggcggatt ttgccattt gggagatgag gtcgcgccga tagatcaggc cggtgccgat 780
tggatccata ttgatgtgat ggatggccat ttcgtgccga atatcaccat aggccccatg 840
gttgtgaagg ctttgcgctc ctatagccaa aagccttttg atgtccattt gatgattgog 900
cctgtcgatc aatatatcga ggctttttct gaagcgggtg ctgatattat cagtttccat 960
cccgaagcgg gcgcgcaccc ccacgcact attcagcata tcaaatcatt gggcaaaaaa 1020
gcgggattag tttttaatcc ggcgaccctt ttaagctggc ttgattatct aatggatgat 1080

ES 2 611 977 T3

cttgatctga ttatggtgat gacggttaac cccggttttg gcggccaaaa atttatcaaa 1140
 acccaattag aaaagattaa agatatccgt caaagaatta ccgcctctgg gcgggatatc 1200
 cgcttggaag tggatggcgg aattgatgcc acgactgcac cgcttgccgt cgaagccggt 1260
 gccgatgttt tggtcgcggg aacggccagc tttaaaggcg gcgcaacatg ttacaccgat 1320
 aatatcagga tattgcgtaa atcatgatta attaactcga ggcggcctga acgtactgca 1380
 agtcctgacg tcaactgtgca gtccgttggc cccggttatcg gtagcgatac cgggcatttt 1440
 ttt 1443

<210> 39
 <211> 12704
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> plásmido construido
 <400> 39

ctagtgttcg atcaacaacc cgaatcctat cgtaatgatg ttttgcccga tcagcctcaa 60
 tcgacaattt tacgcgtttc gatcgaagca gggacgacaa ttggctggga acggtatact 120
 ggaataaatg gtcttcgtta tggattgat gtttttgggt catcggcccc ggcgaatgat 180
 ctatatgctc atttcggctt gaccgcagtc ggcacacga acaaggtgtt ggccgcgatc 240
 gccggtaagt cggcacgtta aaaaatagct atggaatata atagctacta ataagttagg 300
 agaataaaca tgacggacaa attgacctc ctctcgtcagt acaccaccgt agtggccgac 360
 actggggaca tcgcggaat gaagctgtat caaccgcagg atgccacaac caacccttct 420
 ctctattcta acgcagcgca gattccggaa taccgtaagt tgattgatga tgctgtogcc 480
 tgggcgaaac agcagagcaa cgatcgcgag cagcagatcg tggacgcgac cgacaaactg 540
 gcagtaaata ttggtctgga aatcctgaaa ctggttccgg gccgtatctc aactgaagtt 600
 gatgcgcgctc tttcctatga caccgaagcg tcaattgcga aagcaaacg cctgatcaaa 660
 ctctacaacg atgctggtat tagcaacgat cgtattctga tcaaactggc ttctacctgg 720
 cagggtatcc gtgctgcaga acagctggaa aaagaaggca tcaactgtaa cctgaccctg 780
 ctgttctcct tcgctcaggc tcgtgcttgt gcggaagcgg gcgtgttcct gatctcgccg 840
 tttgttggcc gtattcttga ctggtacaaa gcgaataccg ataagaaaga gtacgctccg 900
 gcagaagatc cgggcgtggt ttctgtatct gaaatctacc agtactaaa agagcacggt 960
 tatgaaaccg tggttatggg cgcaagctc cgtaacatcg gcgaaattct ggaactggca 1020
 ggctgcgacc gtctgaccat cgcaccggca ctgctgaaag agctggcggg gagcgaaggg 1080
 gctatcgaac gtaaactgtc ttacaccggc gaagtgaaag cgcgtccggc gcgtatcact 1140

ES 2 611 977 T3

gagtccgagt tcctgtggca gcacaaccag gatccaatgg cagtagataa actggcggaa 1200
ggtatccgta agtttgctat tgaccaggaa aaactggaaa aatgatcgg cgatctgctg 1260
taatctagac gatctggagt caaaatgtcc tcacgtaaag agcttgccaa tgctattcgt 1320
gcgctgagca tggacgcagt acagaaagcc aaatccggtc acccgggggc ccctatgggt 1380
atggctgaca ttgccgaagt cctgtggcgt gatttcctga aacacaacc gcagaatccg 1440
tcctgggctg accgtgaccg ctctgtgctg tccaacggcc acggctccat gctgatctac 1500
agcctgctgc acctcaccgg ttacgatctg ccgatggaag aactgaaaaa ctcccgctcag 1560
ctgcaactta aaactccggg tcacccggaa gtgggttaca ccgctgggtg ggaaaccacc 1620
accggtccgc tgggtcaggg tattgccaac gcagtcggta tggcgattgc agaaaaaacg 1680
ctggcggcgc agtttaaccg tccgggccac gacattgtcg accactacac ctacgccttc 1740
atgggcgacg gctgcatgat ggaaggcatc tcccacgaag tttgctctct ggccgggtacg 1800
ctgaagctgg gtaaactgat tgcattctac gatgacaacg gtatttctat cgatggtcac 1860
gttgaaggct ggttcaccga cgacaccgca atgcgtttcg aagcttacgg ctggcacggt 1920
attcgcgaca tcgacgggtca tgacgggca tctatcaaac gcgcagtaga agaagcgcgc 1980
gcagtgactg acaaaccttc cctgctgatg tgcaaaacca tcatcggttt cggttccccg 2040
aacaagccg gtaccacga ctcccacggt gcgcgctgg gcgacgctga aattgccctg 2100
accgcgaac aactgggctg gaaatatgcg ccgttcgaaa tcccgtctga aatctatgct 2160
cagtgggatg cgaaagaagc aggccaggcg aaagaatccg catggaacga gaaattcgct 2220
gcttacgca aagcttatcc gcaggaagcc gctgaattta cccgccgat gaaaggcgaa 2280
atgccgtctg acttcgacgc taaagcgaag gagttcatcg ctaaactgca ggctaaccg 2340
gcgaaaatcg ccagccgtaa agcgtctcag aatgctatcg aagcgttcgg tccgctgttg 2400
ccggaattcc tcggcgggtc tgctgacctg gcgccgtcta acctgacct gtggtctggt 2460
tctaaagcaa tcaacgaaga tgctgagggt aactacatcc actacggtgt tcgagagttc 2520
ggtatgaccg cgattgctaa cggtatctcc ctgcaagggt gcttcctgcc gtacacctcc 2580
accttcctga tgttcgtgga atacgcacgt aacgccgtac gtatggctgc gctgatgaaa 2640
cagcgtcagg tgatggttta caccacgac tccatcggtc tgggcgaaga cgggccgact 2700
caccagccgg ttgagcaggt cgcttctctg cgcgtaacc cgaacatgtc tacatggcgt 2760
ccgtgtgacc aggttgaatc cgcggtcgcg tggaaatacg gtggtgagcg tcaggacggc 2820
ccgaccgcac tgatcctctc ccgtcagaac ctggcgcagc aggaacgaac tgaagagcaa 2880
ctggcaaca tcgcgcggc tggttatgtg ctgaaagact gcgccggtca gccggaactg 2940
atcttcatcg ctaccgggtc agaagttgaa ctggtctgtg ctgcctacga aaaactgact 3000
gccgaaggcg tgaaagcgcg cgtggtgtcc atgtcgtcta ccgacgcatt tgacaagcag 3060

ES 2 611 977 T3

gatgctgctt accgtgaatc cgtactgccg aaagcgggta ctgcacgcgt tgctgtagaa 3120
gctgggtattg ctgactactg gtacaagtat gttggcctga acgggtgctat cgtcgggatg 3180
accaccttcg gtgaatctgc tccggcagag ctgctgtttg aagagttcgg cttcactggt 3240
gataacgctg ttgcgaaagc aaaagaactg ctgtaattag catttcgggt aaaaaaggtc 3300
gcttcggcga ccttttttat taccttgata atgtccggtt gcgcggcgcg ccccagttac 3360
tcaatacgtg acaataatca gtttatccta actatagaat cgcatagaga gcgataacgt 3420
ttcaccataa gcaatatatt cattgcaaca gtggaattgc cttatgcgtc aaggaaggat 3480
agatcattga cggactgagt tcaaaaagag actcgtctaa aagattttta gaaaggtttc 3540
gatatgacct ctgctgtgcc atcaaatagc aaaaaaagc tggtgattgc ttccgatcac 3600
gcagcatttg agttgaaatc aaccttgatt acttggctga aagagcttgg tcatgaggtc 3660
gaagaccttg gccctcatga aaaccattca gtcgattatc ccgattacgg ttataagctg 3720
gctgtcgcga tcgcagaaaa aaccgctgat ttcggattg ctttatgtgg ctogggaatc 3780
ggtatctcga tcgctgtcaa tcgccatccg gctgcccgtt gcgctttgat tacggataac 3840
cttaccgccc gtttggcaag agaacataac aatgccaatg ttatcgtat gggtgcgaga 3900
ttgatcggca ttgaaaccgc taaggattgt atttcagctt tccttgcaac gccgtttgga 3960
ggtgaacgtc atgttcgccg tatcgataaa ctttcgaatc ctcagttcaa tatctagata 4020
agttaggaga ataaacatga gtaaattacc cctgattgct ccctctatcc tttcggcgga 4080
ttttgcccat ttgggagatg aggtcgcggc gatagatcag gccggtgccg attggatcca 4140
tattgatgtg atggatggcc atttcgtgcc gaatatcacc ataggcccca tggttgtgaa 4200
ggctttgctt ccctatagcc aaaagccttt tgatgtccat ttgatgattg cgccctgtcga 4260
tcaatatatc gaggcctttt ctgaagcggg tgctgatatt atcagtttcc atcccgaagc 4320
gggcgcgcat ccccatcgca ctattcagca tatcaaatca ttgggcaaaa aagcgggatt 4380
agtttttaat ccggcgaccc ctttaagctg gcttgattat ctaatggatg atcttgatct 4440
gattatggtg atgagcgtta accccggttt tggcggccaa aaatttatca aaaccaatt 4500
agaaaagatt aaagatatcc gtcaaagaat taccgcctct gggcgggata tccgcttggg 4560
agtggatggc ggaattgatg ccacgactgc accgcttgc gtcgaagccg gtgccgatgt 4620
tttggctcgc ggaacggcca gctttaagg cggcgcaaca tgttacaccg ataatatcag 4680
gatattgctt aatcatgat taattaactc gaggcggcct gaacgtactg caagtcctga 4740
cgtcactgtg cagtccgttg gcccggttat cggtagcgat accgggcatt tttttaagga 4800
acgatcgata gaattcgcgg ccggcccggc aagacgtgat atggaaccgg aatttgctcc 4860
ggcattcctg cgcaaagata gctaatatct ttcataatct gtatcgaaaa aggagggtct 4920

ES 2 611 977 T3

ttaaagatcc tccttttttt tgcataaaaa gaaggccata gaacaaacag tgataaagac 4980
 agtctcaaac tgtcttttta tagaaaatac cagaatattg tatctggggg aggatgcatg 5040
 gtcttaatcc ggaatacccc ggtcatgcac aggatgttag agcttttgcc tttatggcaa 5100
 aataaacat ggctcgggaa tatctgcgct ttgatttttg taggatgtgc cttccttgtc 5160
 cgtagtatta ttgggcattt ttaccggca ggttatcctt tcgtgacctt tatgccgaca 5220
 atgcttgtgg ttactttcct ctttgggaca agaccgggta ttatcgcggc tattccttagc 5280
 ttgatggttg cgccttattt tatcgaagaa ggaagccgat ttaacggtgt attggtctgg 5340
 tttctttgcc tgctagaaac agtcactgat atgggattgg tgattgcgct acagcaaggt 5400
 aattaccgcc tccagaaaaa gcgtgcctat aatcagatgc tggctgaacg caatgagttg 5460
 ctgtttcatg aattacagca tcgcatttca aataacttac aggttattgc gtcattattg 5520
 cggatgcaaa gccgcagcat caccgatgaa aaagccaagg aagctattga tgacctgtt 5580
 cgtcggattc atatgatcgg tgaattacag cgggcgcttt atattaataa cgggaatcag 5640
 cttggggcaa aattgatcct tgatcgcttg atcaaagagg tcattgctgc cagtaatctc 5700
 ccgaacatcc gctataaaat agaagctgaa gacctgatct taccgtcaga tatggcaatc 5760
 cctttagcgc ttgtatctgc tgaatcogtt tcaaacgcgt tagagcatgg ctttaaaggc 5820
 gatcataaag acgcgtttat tgaattaag cttcaaaaaa ttagcgggca aatcgaactt 5880
 accatttcca ataattggcaa acctcttccc caaggctttt cccttgaaaa ggtcgatagc 5940
 ttaggcctga aaattgcggc tatgtttgcc cgacaattca aaggaaaatt caccttaagt 6000
 aatcagccta accgttatgt ggtttctagc cttattttgc cttgcgggta ggcggccgcc 6060
 taattccgga tgagcattca tcagcggggc aagaatgtga ataaaggccg gataaaactt 6120
 gtgcttattt ttctttacgg tctttaaaaa ggccgtaata tccagctgaa cggctctggtt 6180
 ataggtacat tgagcaactg actgaaatgc ctcaaatgt tctttacgat gccattggga 6240
 tatatcaacg gtggtatatac cagtgatttt tttctccatt ttagcttccct tagctcctga 6300
 aatctcgat aactcaaaaa atacgcccgg tagtgatcct atttcattat ggtgaaagtt 6360
 ggaacctctt acgtgccgat caacgtctca ttttcgcaa aagttggccc agggcttccc 6420
 ggtatcaaca gggacaccag gatttattta ttctgcgaag tgatcttccg tcacaggtat 6480
 ttattcggcg caaagtgcgt cgggtgatgc tgccaactta ctgatttagt gtatgatggt 6540
 gtttttgagg tgctccagtg gcttctggtt ctatcagctg tccctcctgt tcagctactg 6600
 acggggtggt gcgtaacggc aaaagcaccg ccggacatca gcgctagcgg agtgtatact 6660
 ggcttactat gttggcactg atgaggggtg cagtgaatg cttcatgtgg caggagaaaa 6720
 aaggctgcac cgggtcgtca gcagaatatg tgatacagga tatattcgc ttcctcgctc 6780
 actgactcgc tacgctcggg cgttcgactg cggcgagcgg aaatggctta cgaacggggc 6840

ES 2 611 977 T3

ggagatttcc tggaaatgc caggaagata cttaacaggg aagtgagagg gccgcggcaa 6900
 agccgttttt ccataggctc cgccccctg acaagcatca cgaaatctga cgctcaaatc 6960
 agtgggtggcg aaacccgaca ggactataaa gataaccaggc gtttccccct ggccggctccc 7020
 tcgtgcgctc tcctgttccct gcctttcggg ttaccgggtg cattccgctg ttatggccgc 7080
 gtttgtctca ttccacgcct gacactcagt tccgggtagg cagttcgctc caagctggac 7140
 tgtatgcacg aacccccctg tcagtcogac cgctgcgcct tatccggtaa ctatcgtctt 7200
 gagtccaacc cggaaagaca tgcaaaagca ccaactggcag cagccactgg taattgattt 7260
 agaggagtta gtcttgaagt catgcgcggg ttaaggctaa actgaaagga caagttttgg 7320
 tgactgcgct cctccaagcc agttacctcg gttcaaagag ttggtagctc agagaacctt 7380
 cgaaaaccg ccctgcaagg cggttttttc gttttcagag caagagatta cgcgcagacc 7440
 aaaacgatct caagaagatc atcttattaa tcagataaaa tatttctaga tttcagtgca 7500
 atttatctct tcaaatgtag cacctgaagt cagccccata cgatataagt tgtaattctc 7560
 atgtttgaca gcttatcatc gatgtgacgg aagatcactt cgcagaataa ataaatcctg 7620
 gtgtccctgt tgataccggg aagccctggg ccaacttttg gcgaaaatga gacgttgatc 7680
 ggcacgtaag aggttccaac tttcaccata atgaaataag atcactaccg ggcgtatfff 7740
 ttgagttatc gagatfffca ggagctaagg aagctaaaat ggagaaaaaa atcactggat 7800
 ataccaccgt tgatatatcc caatggcatc gtaaagaaca ttttgaggca tttcagtcag 7860
 ttgctcaatg tacctataac cagaccgttc agctggatat tacggccttt ttaaagaccg 7920
 taaagaaaaa taagcacaag ttttatccgg cctttattca cattcttgc cgcctgatga 7980
 atgctcatcc ggaattccgt atggcaatga aagacgggta gctgggtgata tgggatagtg 8040
 ttcacccttg ttacaccgtt ttccatgagc aaactgaaac gttttcatcg ctctggagtg 8100
 aataccacga cgatttccgg cagtttctac acatatattc gcaagatgtg gcgtgttacg 8160
 gtgaaaacct ggctatfff cctaaagggt ttattgagaa tatgtfffcc gtctcagcca 8220
 atccctgggt gagtttcacc agttttgatt taaacgtggc caatatggac aacttcttcg 8280
 cccccgtfff caccatgggc aaatattata cgcaaggcga caaggtgctg atgccgctgg 8340
 cgattcaggt tcatcatgcc gtttgtgatg gcttccatgt cggcagaatg cttaatgaat 8400
 tacaacagta ctgcgatgag tggcagggcg gggcgtaatt tttttaaggc agttattggg 8460
 gcccttaaac gctggttgc tacgcctgaa taagtcgacc tttgtagtct tggcctgttg 8520
 tgtgcatgag caaatcaatg gcaccacccc ctctttttg agctgaatgg tcataaaatt 8580
 tataattatc tatcgttaatt cggaatctat gttcagggtc tcgccattgc tttttgtctg 8640
 ctgggtcaag ttccatgcct aaggttffta agacatcaga aagaggtatt gcacgcatgc 8700

ES 2 611 977 T3

tatcagcttt	tcttctagct	aatgacaggg	cttctctgct	tctatctgct	cgtttttttt	8760
cttccacata	tctcgccgct	ttgtcagcca	gcggctgtat	tacggaaagt	gccgattttt	8820
gggcttttag	gcgttctttt	tctgcccatt	cttcttatt	tgtaaaaatt	gaggggtggga	8880
tgggtgcctg	aatcttggga	tctagctgta	aagttttggt	gatatttccg	taatgtcttt	8940
ggactctttg	atgcgctgct	tttgaacctt	ttacgcctct	ggccagccct	agaggctcca	9000
tagaagccgc	ataatccgtc	tggagggcag	aaagggcttt	tcgaccatca	aaccatctcg	9060
atgcgtttaa	acggcctgta	tcggggctct	taggcaccat	aaagccggtt	aagtgggggtg	9120
ttgtttcatc	agcatgtagc	tgaagagata	caaggttgtt	ttctccaaag	gtttgttccg	9180
cccattgctg	ggtgattggt	ttccagtgtt	cgagtttttc	aggagtggcc	tgttttgacc	9240
attctggaga	cataccaaag	aacagttcta	tggcctgcac	accgtttttt	ctaagaggct	9300
ttcccgtttc	tttctgaatt	ttattcagca	tagatttaac	atctgctgat	gggtcagtag	9360
agcctttgag	tatttcgttt	agttcttttc	tatctgggtc	agcgttttgt	gtttcgcggc	9420
ctcgcgtcat	atgcaggctc	gcggctttaa	tcgtgccaac	tgttttatgt	ttttcaaacc	9480
taaagattgc	atagttcggc	atgttttaac	tgctttaatt	tgagaaaaga	ccagaggaaa	9540
taatccagcc	tatatttctt	tccttagtag	cgaactggaa	ttgtttttcc	gaaggaaaaa	9600
agcaattccg	tagtgagtac	tgaatttatt	ctgattcgtc	ttgcttttgg	agcgtctttt	9660
tgcgttctat	aactgttgtg	aaagctacgc	ggtcgccatt	gaaaacgaaa	ttaggattaa	9720
taaaatacca	tccttggcga	acatgctttg	caatgatttt	agctttttct	aattcggcta	9780
gacctcttgc	aaaggtagct	tgagatagtg	ccagtttttt	ttcttgtgcg	ttaagaaagt	9840
cctctaaaac	gaatttgtct	aaaggaoga	ggtctttgct	gatgcctttg	tcttgaagta	9900
tccaaaccag	aacgctgaaa	gcttttattc	cagcggctcc	tagttcaaaa	gttagcgcga	9960
tattggtgct	aaataatttt	acaaattctt	cactatcaac	acgtctgtaa	gtcgtcacat	10020
gagtgccttg	catctcacca	gtggcttgat	tgaccagaat	gttatcatct	cgctctaate	10080
gagataactg	aacctctga	cttttaactg	gcacaaccat	accttcgatg	aaaggattct	10140
cgtcatatct	gattggctgc	tttctcaatt	ttgtcgccat	atltgataaa	cctttaatca	10200
aaaaaaccac	atlttttgat	tatacctatt	catcgaatga	ggcaaggtct	atcaatttta	10260
cccctttttt	tgatagacgg	ttaaatcaat	attgatagac	cccttcacag	attctgaaaa	10320
tcgacttccc	tatttttaggg	atattttcac	gattcccttt	cttagttctt	cctagtgggg	10380
aaattcgttg	aatcctgcct	cgaaaaaac	atgagaaagc	tgttggttat	atacacgggc	10440
aaagccacc	tatttttagc	tactggggaa	agagataag	cagggatttt	gtaaaattaa	10500
aaccggattt	ttcgctttac	ggtttgttta	ggcgcaactg	tctttttaag	accgcgttta	10560
accatcaaaa	gatogttcca	atcttttccg	tgtatcatct	gttcttttag	tgggagccag	10620

ES 2 611 977 T3

ttttcaactt tttttggttg aaacgcggct ttaatcgctc cgactaatag cgatgctgct 10680
 ctttgtccta cagcatccca atcataggca atatggacag aagatgcctt ttcaacgatt 10740
 tttcggagag ttttagtaag agacgttctt acgccgctgg tgcttaataa ttttacgcca 10800
 gctttaattt tttctgggct taaaaagccg actactgaaa tcgctgctat cgcactttca 10860
 gcgatataaa gatcatactt ttctgcattt tttacattga tgctgccagt aaaatgggct 10920
 tcgcgactgc ttccaaggc taacccttta aaaccactgc ttgttccgcg taattctgcg 10980
 ccctgaagtg tatctttatc gtcatacatc aagaaggcta cattaccgcg atcatctggt 11040
 cggatagagt caggaatatt gttaaagat attcctcggg cagcgttggg tcctggccac 11100
 ggggtgcgcat gatcgtgctc ctgtcgttga ggaccocggct aggctggcgg ggttgcctta 11160
 ctggttagca gaatgaatca ccgatacgcg agcgaacgtg aagcgactgc tgctgcaaaa 11220
 cgtctgcgac ctgagcaaca acatgaatgg tcttcggttt ccgtgtttcg taaagtctgg 11280
 aaacgcggaa gtcccctacg tgctgctgaa gttgcccgca acagagagtg gaaccaaccg 11340
 gtgataccac gatactatga ctgagagtca acgccatggg agctccctat cgtctgactc 11400
 gcaaggctga acgtgttgac gccttgagca aggccaaagc ggttcttgac gaagccttcc 11460
 cagaagctga tccgacagaa aagctgcgca tccagaagct tgccaagaag ctggaagcaa 11520
 aaatcgtccg caccgccatt ctgaaagaag gccggagaat tgacggacgc gatctgaaaa 11580
 cagttcggcc gatccgctct caggttgat tcttgccccg cacgcatggt tctgcccctgt 11640
 ttacgcgtgg tgaaacacag gctttggttt caaccaccct tggaacggcg gatgctgaac 11700
 agatgatcga cggtttaacc ggccttcatt atgaacgctt catgctgcat tacaacttcc 11760
 ccccatattc ggtcggtgaa gttggtcgtt ttggtgctcc gggtcgtcgt gaaatcggcc 11820
 atggtaaaact ggcattggcgt gcgcttcac cggttttgcc gagcaaggct gatttcccg 11880
 ataccatccg tgttttgtcg gatatacccg aatctaagtg ttctctctcc atggcaaccg 11940
 tttgcggtgg ctgccttgca ttgatggatg ccggtgttcc cttaacgcgt ccggtttccg 12000
 gtatcgccat gggctcttatt ctggaaaaag acggcttcgc tattttgtcc gatatcatgg 12060
 gtgatgaaga tcacttgggt gatatggact ttaaggtcgc cggtagcgaa aaaggtatca 12120
 ccagcctcca gatggacatc aaggttgctg gcattaccga agaaatcatg cagaaagctt 12180
 tggaacaggc taaaggtggc cgtgctcata tcttgggtga aatgtccaaa gcgctgggtg 12240
 aagtccgctc cgaaatttct aatttggcac cgcgcattga aacaatgagc gtacccaaaag 12300
 acaaaatccg tgatgttatc ggaacgggog gaaaagttat ccgtgaaatc gtggcgacca 12360
 caggtgccaa ggtcgatatac gaagatgacg gcacggttcg totgtcttct tctgatccgg 12420
 ccaatattga agcagcccgt gaatggatca atggtattgt tgaagaaccg gaagtaggca 12480

ES 2 611 977 T3

aaatctataa cggtaaagtc gtcaatatcg ttgatttcgg tgccttcgta aacttcatgg 12540
 gtggccgtga cggcttggtgta catgtttcgg aatcaagaa cgaacgtgtc aacaaggtca 12600
 gcgatgtcct gtccgaaggt caggaagtca aagtcaaggt tcttgaaatt gacaaccgtg 12660
 gcaaggttcg cctgtctatg cgtgttgcg atcaggaaac cgga 12704

<210> 40
 <211> 1318
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> fragmento construido para integración en el gen pnp de Z. m pnp
 <400> 40

cctatcgtct gactcgcaag gctgaacgtg ttgacgcctt gagcaaggcc aaagcggttc 60
 ttgacgaagc cttcccagaa gctgatccga cagaaaagct gcgcatccag aagcttgoga 120
 agaagctgga agcaaaaatc gtccgcaccg ccattctgaa agaaggccgg agaattgacg 180
 gacgcgatct gaaaacagtt cgcccgatcc gctctcaggt tggattcttg ccccgcacgc 240
 atggttctgc cctgtttacg cgtggtgaaa cacaggcttt ggtttcaacc acccttgaa 300
 cggcgggatgc tgaacagatg atcgacgggt taaccggcct tcattatgaa cgcttcatgc 360
 tgcattacaa cttcccccca tattcggctg gtgaagttgg tcgttttggg gctccgggtc 420
 gtcgtgaaat cggccatggt aaactggcat ggcgtgcgct tcatccggtt ttgccgagca 480
 aggctgattt cccgtatacc atccgtggtt tgtcggatat caccgaatct aatggttcct 540
 cttccatggc aaccgtttgc ggtggctgcc ttgcattgat ggatgccggg gttcccttaa 600
 cgcgtccggg ttccgggtatc gccatgggtc ttattctgga aaaagacggc ttcgctattt 660
 tgtccgatat catgggtgat gaagatcact tgggtgatat ggactttaag gtcgccggtg 720
 ccgaaaaagg tatcaccagc ctccagatgg acatcaaggt tgctggcatt accgaagaaa 780
 tcatgcagaa agctttggaa caggctaaag gtggccgtgc tcatatcttg ggtgaaatgt 840
 ccaaagcgtt ggggtgaagtc cgctccgaaa tttctaattt ggcaccgcgc attgaaacaa 900
 tgagcgtacc aaaagacaaa atccgtgatg ttatcggaac gggcggaaaa gttatccgtg 960
 aaatcgtggc gaccacaggt gccaaaggtc atatcgaaga tgacggcacg gttcgtctgt 1020
 cttcttctga tccggccaat attgaagcag cccgtgaatg gatcaatggt attgttgaag 1080
 aaccggaagt aggcaaaaatc tataacggta aagtcgtcaa tatcgttgat ttcggtgcct 1140
 tcgtaaactt catgggtggc cgtgacggct tggtagatgt ttoggaaatc aagaacgaa 1200
 gtgtcaacaa ggtcagcgat gtcctgtccg aaggtcagga agtcaaagtc aaggttcttg 1260
 aaattgacaa ccgtggcaag gttcgcctgt ctatgcgtgt tgcgatcag gaaaccgg 1318

10 <210> 41
 <211> 1225
 <212> ADN
 <213> secuencia artificial

<220>

ES 2 611 977 T3

<223> fragmento construido para la integración en el gen pnp de Z. m

<400> 41

```

cggcaagacg tgatatggaa ccggaatttg ctccggcatt cctgcgcaaa gatagctaat      60
atctttcata ttttgtatcg aaaaaggagg gtctttaaag atcctccttt tttttgcata      120
aaaagaaggc catagaacaa acagtgataa agacagtctc aaactgtctt tttatagaaa      180
ataccagaat attgtatctg ggggaggatg catggtctta atccggaata ccccggtcat      240
gcacaggatg ttagagcttt tgcctttatg gcaaaataaa ccatggctcg ggaatatctg      300
cgctttgatt tttgtaggat gtgccttcct tgtccgtagt attattgggc attttttacc      360
ggcaggttat cctttcgtga cctttatgcc gacaatgctt gtggttactt tcctctttgg      420
gacaagaccg ggtattatcg cgctattctt tagcttgatg gttgcgcctt attttatcga      480
agaaggaagc cgatttaacg gtgtattggt ctggtttctt tgcctgctag aaacagtcac      540
tgatatggga ttggtgattg cgctacagca aggtaattac cgcctccaga aaaagcgtgc      600
ctataatcag atgctggctg aacgcaatga gttgctgttt catgaattac agcatcgcac      660
ttcaaataac ttacaggtta ttgcgctcatt attgcggatg caaagccgca gcatcaccga      720
tgaaaaagcc aaggaagcta ttgatgcctc tgttcgctcg attcatatga tcggtgaatt      780
acagcgggcg ctttatatta aaaacgggaa tcagcttggg gcaaaattga tccttgatcg      840
cttgatcaaa gaggtcattg cgtccagtaa tctcccgaac atccgctata aaatagaagc      900
tgaagacctg atcttaccgt cagatatggc aatcccttta gcgcttgtat ctgctgaatc      960
cgtttcaaac gcggttagagc atggctttaa aggcgatcat aaagacgcgt ttattgaaat     1020
taagcttcaa aaaattagcg ggcaaatcga acttaccatt tccaataatg gcaaacctct     1080
tccccaggc ttttcccttg aaaaggctga tagcttaggc ctgaaaattg cggctatggt     1140
tgcccgacaa ttcaaaggaa aattcacctt aagtaatcag cctaaccggt atgtggtttc     1200
tagccttatt ttgccttgcg gttag                                             1225

```

<210> 42

5 <211> 25

<212> ADN

<213> secuencia artificial

<220>

<223> cebador

10 <400> 42

```

cggcttcaat cggattgta gcagg                                             25

```

<210> 43

<211> 38

<212> ADN

15 <213> secuencia artificial

<220>

<223> cebador

<400> 43

```

cgtgtagctt ggacactcat gttattctc ctaacta                               38

```

ES 2 611 977 T3

	<210> 44	
	<211> 20	
	<212> ADN	
	<213> secuencia artificial	
5	<220>	
	<223> cebador sintético	
	<400> 44	
	ccagtatcag cccgtcatac	20
10	<210> 45	
	<211> 20	
	<212> ADN	
	<213> secuencia artificial	
	<220>	
	<223> cebador sintético	
15	<400> 45	
	ccagcatggt tggatggct	20
	<210> 46	
	<211> 20	
	<212> ADN	
20	<213> secuencia artificial	
	<220>	
	<223> cebador sintético	
	<400> 46	
	gcctgggct ttaaagcct	20
25	<210> 47	
	<211> 22	
	<212> ADN	
	<213> secuencia artificial	
	<220>	
30	<223> cebador sintético	
	<400> 47	
	gagaagggt ggtgtggca tc	22
	<210> 48	
	<211> 22	
35	<212> ADN	
	<213> secuencia artificial	
	<220>	
	<223> cebador sintético	
	<400> 48	
40	gtggatggcg gaattgatgc ca	22
	<210> 49	
	<211> 26	
	<212> ADN	
	<213> secuencia artificial	
45	<220>	
	<223> cebador sintético	
	<400> 49	
	tctcgagag atagaggtca gtcgac	26

ES 2 611 977 T3

	<210> 50	
	<211> 23	
	<212> ADN	
	<213> secuencia artificial	
5	<220>	
	<223> cebador sintético	
	<400> 50	
	cagctatgat gacagcgcat tgg	23
10	<210> 51	
	<211> 23	
	<212> ADN	
	<213> secuencia artificial	
	<220>	
	<223> cebador sintético	
15	<400> 51	
	gggcgggttcg atccatagaa agg	23
	<210> 52	
	<211> 34	
	<212> ADN	
20	<213> secuencia artificial	
	<220>	
	<223> cebador sintético	
	<400> 52	
	gggagctcac tagtgtgacg gaagatcact tcgc	34
25	<210> 53	
	<211> 20	
	<212> ADN	
	<213> secuencia artificial	
	<220>	
30	<223> cebador sintético	
	<400> 53	
	tcgtcatgga caacaaacgg	20
	<210> 54	
	<211> 33	
35	<212> ADN	
	<213> secuencia artificial	
	<220>	
	<223> cebador sintético	
	<400> 54	
40	gcaattggcc ggccatttc aggcgtagca acc	33
	<210> 55	
	<211> 5548	
	<212> ADN	
	<213> secuencia artificial	
45	<220>	
	<223> plásmido construido	
	<400> 55	

ES 2 611 977 T3

cgaattcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcaacaattc	60
cacacaacat acgagccgga agcataaagt gtaaagcctg gggcgcctaa tgagtgagct	120
aactcacatt aattgcggtg cgctcaactgc ccgctttcca gtcgggaaac ctgtcgtgcc	180
agctgcatta atgaatcggc caacgcgcgg ggagagcggg tttgcgtatt gggcgcctct	240
ccgcttcctc gctcaactgac tcgctgcgct cggcgttcg gctgcggcga gcggtatcag	300
ctcaactcaa ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca	360
tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcggtg ctggcgtttt	420
tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc	480
gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc ctcgtgcgct	540
ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg	600
tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgtcca	660
agctgggctg tgtgcacgaa cccccggtc agcccgaccg ctgcgcctta tccggtaact	720
atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta	780
acaggattag cagagcgagg tatgtaggag gtgctacaga gttcttgaag tggcggccta	840
actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag ccagttacct	900
tcggaaaaag agttggtagc tcttgatccg gcaaaaaaac caccgctggt agcggtggtt	960
tttttgttg caagcagcag attacgcgca gaaaaaagg atctcaagaa gatcctttga	1020
tcttttctac ggggtctgac gctcagtgga acgaaaactc acgtaaggg attttggtca	1080

ES 2 611 977 T3

tgagattatc	aaaaaggatc	ttcacctaga	tccttttaaa	ttaaaaatga	agtttttaaat	1140
caatctaaag	tatatatgag	taaacttggt	ctgacagtta	ccaatgctta	atcagtgagg	1200
cacctatctc	agcgatctgt	ctatttcggt	catccatagt	tgctgactc	cccgtcgtgt	1260
agataactac	gatacgggag	ggcttaccat	ctggccccag	tgctgcaatg	ataccgcgag	1320
accacgctc	accggctcca	gatttatcag	caataaacca	gccagccgga	agggccgagc	1380
gcagaagtgg	tctgcaact	ttatccgct	ccatccagtc	tattaattgt	tgccgggaag	1440
ctagagtaag	tagttcgcca	gttaatagtt	tgcgcaacgt	tgttgccatt	gctacaggca	1500
tcgtgggtgc	acgctcgtcg	tttggatgg	cttcattcag	ctccggttcc	caacgatcaa	1560
ggcgagttac	atgatcccc	atgttggtgca	aaaagcggg	tagctccttc	ggtcctccga	1620
tcgttgctcag	aagtaagttg	gccgcagtgt	tatcactcat	ggttatggca	gactgcata	1680
attctcttac	tgcatgcca	tccgtaagat	gcttttctgt	gactggtgag	tactcaacca	1740
agtcattctg	agaatagtgt	atgcggcgac	cgagttgctc	ttgcccggcg	tcaatacggg	1800
ataataccgc	gccacatagc	agaacttta	aagtgctcat	cattggaaaa	cgttcttcgg	1860
ggcgaaaact	ctcaaggatc	ttaccgctgt	tgagatccag	ttcgatgtaa	cccactcgtg	1920
caccaactg	atcttcagca	tcttttactt	tcaccagcgt	ttctgggtga	gcaaaaacag	1980
gaaggcaaaa	tgccgcaaaa	aaggaataa	ggcgacacg	gaaatggtga	atactcatac	2040
tcttcctttt	tcaatattat	tgaagcattt	atcagggtta	ttgtctcatg	agcggataca	2100
tatttgaatg	tatttagaaa	aataaacaaa	taggggttcc	gcgcacattt	cccogaaaag	2160
tgccacctga	cgtotaagaa	accattatta	tcatgacatt	aacctataaa	aataggcgta	2220
tcacgaggcc	ctttcgtctc	gcgcgtttcg	gtgatgacgg	tgaaaacctc	tgacacatgc	2280
agctcccgga	gacggtcaca	gcttgtctgt	aagcggatgc	cgggagcaga	caagcccgtc	2340
agggcgctc	agcgggtggt	ggcgggtgct	ggggctggct	taactatgcg	gcatcagagc	2400
agattgtact	gagagtgcac	catatgcggt	gtgaaatacc	gcacagatgc	gtaaggagaa	2460
aataccgcat	cagggcccat	tcgccattca	ggctgcgcaa	ctggtgggaa	gggcgatcgg	2520
tgccggcctc	ttcgctatta	cggcagctgg	cgaaagggg	atgtgctgca	agcggattaa	2580
gttgggtaac	gccagggttt	tcccagtcac	gacgttgtaa	aacgacggcc	agtgccaaagc	2640
ttgcatgcct	gcaggtcgac	gacctctttg	atcaagcgat	caaggatcaa	ttttgcccc	2700
agctgattcc	cgtttttaat	ataaagcgcc	cgctgtaatt	caccgatcat	atgaatccga	2760
cgaacagagg	catcaatagc	ttccttggt	ttttcatcgg	tgatgctgcg	gctttgcatc	2820
cgcaataatg	acgcaataac	ctgtaagtta	tttgaaatgc	gatgctgtaa	ttcatgaaac	2880
agcaactcat	tgcgttcagc	cagcatctga	ttataggcac	gctttttctg	gaggcggtaa	2940

ES 2 611 977 T3

ttaccttgct gtagcgcaat caccaatccc atatcagtgga ctgtttctag caggcaaaga 3000
 aaccagacca atacaccggt aaatcggcct ccttcttcga taaaataagg cgcaaccatc 3060
 aagctaagaa tagccgcat aatacccggt cttgtcccaa agaggaaagt aaccacaagc 3120
 attgtcggca taaaggcac gaaaggataa cctgcccgtg aaaaatgccc aataatacta 3180
 cggacaagga aggcacatcc tacaanaatc aaagcgcaga tattcccagc ccatgggtta 3240
 ttttgccata aaggcaaaag ctctaacatc ctgtgcatga ccgggggtatt ccggattaag 3300
 accatgcatc ctccccaga tacaatatcc tggatatttc tataaaaaaga cagtttgaga 3360
 ctgtctttat cactgtttgt tctatggcct tctttttatg caaaaaaag gaggatcttt 3420
 aaagaccctc ctttttcgat acaaaatatg aaagatatta gctatctttg cgcaggaatg 3480
 ccggagcaaa ttccggttcc atatcacgtc ttgccggaag atcgcagat tcccagcag 3540
 ggccccgatc acgacgagga cgctcaccac cgcgacgttc tgcgtctgtc tcttatacac 3600
 atctcaacca tcatcgatga attttctcgg gtgttctcgc atattggctc gaattctacc 3660
 tgcaggctga tgaatgctca tccggaatta gcggccgctt aattaactta ttcaggcgt 3720
 gcaaccaggc gtttaagggc accaataact gccttaaaaa aattacgccc cgccttgcca 3780
 ctcatcgcag tactgttgta attcattaag cattctgcog acatggaagc catcacaac 3840
 ggcatgatga acctgaatcg ccagcggcat cagcaccttg tgccttgcg tataatattt 3900
 gcccatggtg aaaacggggg cgaagaagtt gtccatattg gccacgttta aatcaaaact 3960
 ggtgaaactc acccagggat tggctgagac gaaaaacata ttctcaataa acccttagg 4020
 gaaataggcc aggttttcac cgtaacacgc cacatcttgc gaatatatgt gtagaaactg 4080
 ccggaaatcg tcgtggtatt cactccagag cgatgaaaac gtttcagttt gctcatggaa 4140
 aacgggtgaa caagggtgaa cactatccca tatcaccagc tcaccgtctt tcattgcat 4200
 acggaattcc ggatgagcat tcatcaggcg ggcaagaatg tgaataaagg ccggataaaa 4260
 cttgtgctta tttttcttta cggctctttaa aaagccgta atatccagct gaacggctc 4320
 gttataggta cattgagcaa ctgactgaaa tgcctcaaaa tgttctttac gatgccattg 4380
 ggatataatca acggtggtat atccagtgat tttttctcc attttagctt ccttagctcc 4440
 tgaaaatctc gataactcaa aaaatacgcc cggtagtgat cttatttcat tatggtgaaa 4500
 gttggaacct cttacgtgcc gatcaacgtc tcattttcgc caaaagtgg ccagggtctt 4560
 cccggtatca acagggacac caggatttat ttattctgog aagtgatctt ccgtcacgog 4620
 gccgcataac ttcgtatagc atacattata cgaagttatg cgatcgcaag cttgccaacg 4680
 actacgcact agccaacaag agcttcaggg ttgagatgtg tataagagac aggttctgog 4740
 ttctcacgtg gcggacgggt atcatccagc tetgcgcccg tttctgatc gacaacacgc 4800
 atagacaggc gaaccttgcc acggttgca atttcaagaa ccttgacttt gacttctg 4860

ES 2 611 977 T3

ccttcggaca ggacatcgct gaccttgttg acacgttcgt tcttgatttc cgaaacatgt 4920
 accaagccgt cacggccacc catgaagttt acgaaggcac cgaaatcaac gatattgacg 4980
 actttaccgt tatagatttt gcctaacttc ggttcttcaa caataccatt gatccattca 5040
 cgggctgctt caatattggc cggatcagaa gaagacagac gaaccgtgcc gtcattctcg 5100
 atatcgacct tggcacctgt ggtcgccacg atttcacgga taacttttcc gcccgttccg 5160
 ataacatcac ggatthttgtc ttttggtacg ctcatgtttt caatgctcgg tgccaaatta 5220
 gaaatttcgg agcggacttc acccagcgcct ttggacattt cacccaagat atgagcacgg 5280
 ccacctttag cctgttccaa agctttctgc atgatttctt cggtaatgcc agcaaccttg 5340
 atgtccatct ggaggctggg gatacctttt tcggtaccgg cgaccttaa gtccatatca 5400
 cccaagtgat cttcatcacc catgatatcg gacaaaatag cgaagccgtc tttttccaga 5460
 ataagacca tggcgatacc ggaaaccgga cgcgttaagg gaacaccggc atccatgtcg 5520
 actctagagg atccccgggt accgagct 5548

<210> 56
 <211> 891
 <212> ADN
 5 <213> secuencia artificial

<220>
 <223> fragmento construido para integración en el locus pnp de *Z. mobilis*

<400> 56

ataacttcgt atagcataca ttatacgaag ttatgctgac gcaagcttgc caacgactac 60
 gcactagcca acaagagctt cagggttgag atgtgtataa gagacaggtt ctgctgttctc 120
 acgtggcggg cggttatcat ccagctctgc gccggtttcc tgatcgaaa cacgcataga 180
 caggcgaacc ttgccacggt tgtcaatttc aagaaccttg actttgactt cctgaccttc 240
 ggacaggaca tcgctgacct tgttgacacg ttctgtcttg atttccgaaa catgtacca 300
 gccgtcacgg ccacctatga agtttacgaa ggcaccgaaa tcaacgatat tgacgacttt 360
 accgttatag attttgctta cttccggttc ttcaacaata ccattgatcc attcacgggc 420
 tgcttcaata ttggccggat cagaagaaga cagacgaacc gtgccgtcat ctccgatatc 480
 gaccttggca cctgtggctc ccacgatttc acggataact tttccgcccg ttccgataac 540
 atcacggatt ttgtcttttg gtacgctcat tgtttcaatg cgcggtgcca aattagaaat 600
 ttcggagcgg acttcacca gcgctttgga catttcaccc aagatatgag cacggccacc 660
 tttagcctgt tccaaagctt tctgcatgat ttcttcggtg atgccagcaa ccttgatgtc 720
 catctggagg ctggtgatac ctttttcggg accggcgacc ttaaagtcca tatcacccaa 780
 gtgatcttca tcacctatga tatcggacaa aatagcgaag ccgtcttttt ccagaataag 840
 10 acccatggcg ataccgaaa ccggacgcgt taagggaaca ccggcatcca t 891

<210> 57
 <211> 1030
 <212> ADN
 <213> secuencia artificial

ES 2 611 977 T3

<220>

<223> fragmento construido para integración en el locus pnp de Z. mobilis

<400> 57

```

gacctctttg atcaagcgat caaggatcaa ttttgcccca agctgattcc cgtttttaat      60
ataaagcgcc cgctgtaatt caccgatcat atgaatccga cgaacagagg catcaatagc      120
ttccttggct ttttcatcgg tgatgctgcg gctttgcatc cgcaataatg acgcaataac      180
ctgtaagtta tttgaaatgc gatgctgtaa ttcataaagc agcaactcat tgcggtcagc      240
cagcatctga ttataggcac gctttttctg gaggcggtaa ttaccttgct gtagcgcaat      300
caccaatccc atatcagtga ctgtttctag caggcaaaga aaccagacca atacaccggt      360
aaatcggcct ccttcttcga taaaataagg cgcaaccatc aagctaagaa tagcccgcat      420
aataccgggt cttgtcccaa agaggaaagt aaccacaagc attgtcggca taaaggctac      480
gaaaggataa cctgccggta aaaaatgccc aataatacta cggacaagga aggcacatcc      540
tacaaaaatc aaagcgcaga tattcccgag ccatggttta ttttgccata aaggcaaaag      600
ctctaacatc ctgtgcatga ccgggggtatt ccggattaag accatgcatc ctccccaga      660
tacaatattc tggatatttc tataaaaaga cagtttgaga ctgtctttat cactgtttgt      720
tctatggcct tctttttatg caaaaaaag gaggatcttt aaagaccctc ctttttcgat      780
acaaaatatg aaagatatta gctatctttg cgcaggaatg ccggagcaaa ttccggttcc      840
atatcacgtc ttgccggacg atcgccagat tcccgacgag ggccccgatc acgacgagga      900
cgctcaccac cgcgacggtc tgcgtctgtc tcttatacac atctcaacca tcatcgaatg      960
atcttctcgg gtgttctcgc atattggctc gaattctacc tgcaggctga tgaatgctca     1020
tccggaatta                                     1030

```

- 5 <210> 58
- <211> 6471
- <212> ADN
- <213> secuencia artificial

- 10 <220>
- <223> plásmido construido

<400> 58

```

cgaattcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcaaatc      60
cacacaacat acgagccgga agcataaagt gtaaagcctg ggggtgcctaa tgagtgaact      120
aactcacatt aattgcggtg cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc      180

```

ES 2 611 977 T3

agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggcgcctctt 240
 ccgcttcctc gctcactgac tcgctgcgct cggtcggtcg gctgcggcga gcggtatcag 300
 ctactcaaa ggcggaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 360
 tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcggtg ctggcgtttt 420
 tccataggct ccgccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc 480
 gaaacccgac aggactataa agataccagg cgtttcccc tggaagctcc ctcgctgcgt 540
 ctctgttcc gaccctgccg cttaccgat acctgtccgc ctttctccct tcgggaagcg 600
 tggcgctttc tcatagctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca 660
 agctgggctg tgtgcacgaa cccccgttc agcccgaccg ctgcgcctta tccggtaact 720
 atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta 780
 acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtagccta 840
 actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag ccagttacct 900
 tcggaaaaag agttggtagc tcttgatccg gcaaaaaac caccgctggt agcggtggtt 960
 tttttgtttg caagcagcag attacgcgca gaaaaaag atctcaagaa gatcctttga 1020
 tcttttctac ggggtctgac gctcagtga acgaaaactc acgttaaggg attttggtca 1080
 tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat 1140
 caatctaaag tatatatgag taaacttggg ctgacagtta ccaatgctta atcagtgagg 1200
 cacctatctc agcgatctgt ctatttcggt catccatagt tgccctgactc cccgtcgtgt 1260
 agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 1320
 acccacgctc accggtcca gatttatcag caataaacca gccagccgga agggccgagc 1380
 gcagaagtgg tcctgcaact ttatccgct ccatccagtc tattaattgt tgccgggaag 1440
 ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 1500
 tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 1560
 ggcgagttac atgatcccc atgttggtgca aaaaagcggg tagctccttc ggtcctcga 1620
 tcgttgctcag aagtaagttg gccgcagtgt tatcactcat ggttatggca gcactgcata 1680
 attctcttac tgtcatgcc tccgtaagat gcttttctgt gactggtgag tactcaacca 1740
 agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 1800
 ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 1860
 ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttogatgtaa cccactcgtg 1920
 cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 1980
 gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac 2040

ES 2 611 977 T3

tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca 2100
tatttgaatg tatttagaaa aataa caaaa taggggttcc gcgcacattt ccccgaaaag 2160
tgccacctga cgtctaagaa accattatta tcatgacatt aacctataaa aataggcgta 2220
tcacgaggcc ctttcgtctc gcgcgtttcg gtgatgacgg tgaaaacctc tgacacatgc 2280
agctcccgga gacggtcaca gcttgtctgt aagcggatgc cgggagcaga caagcccgtc 2340
agggcgcgtc agcgggtggt ggcgggtgtc ggggctggct taactatgcg gcatcagagc 2400
agattgtact gagagtgcac catatgcggt gtgaaatacc gcacagatgc gtaaggagaa 2460
aataccgcat caggcgccat tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg 2520
tgcgggcctc ttcgctatta cgccagctgg cgaaaggggg atgtgctgca aggcgattaa 2580
gttgggtaac gccagggttt tcccagtcac gacgttgtaa aacgacggcc agtgccaagc 2640
ttgcatgcct gcaggctgac catcaatgca aggcagccac cgcaaaccgt tgccatggaa 2700
gaggaacct tagattcgggt gatatccgac aaaacacgga tggatatacgg gaaatcagcc 2760
ttgctcggca aaaccggatg aagcgcacgc catgccagtt taccatggcc gatttcacga 2820
cgaccgggag caccaaaacg accaacttca ccgaccgaat atggggggaa gttgtaatgc 2880
agcatgaagc gttcataatg aaggccgggt aaaccgtcga tcatctgttc agcatccgcc 2940
gttccaaggg tggttgaaac caaagcctgt gttcaccac gcgtaaacag ggcagaacca 3000
tgcgtgcggg gcaagaatcc aacctgagag cggatcgggc gaactgtttt cagatcgcgt 3060
ccgtcaattc tccggccttc tttcagaatg gcggtgcgga cgatthttgc ttccagcttc 3120
ttcgcaagct tctggatgcg cagctthttct gtcggatcag cttctgggaa ggcttcgtca 3180
agaaccgctt tggccttgct caaggcgtca acacgttcag ccttgcgagt cagacgatag 3240
gccttatcga aattgtcgtg gcagagttct tcgactttgg cagccaatgc gctgtcatca 3300
tagctgacca gttcccaagg ctcttttagcg gcctgttctg ccagattgat gatagctttg 3360
ataaccggct gtaaagattc atgcgcaaaa gaaacggcat tgagcatgac ttcttcggga 3420
agctcattcg cttccgattc aaccatcatc acggcatcac gggtagcccc gacaaccaga 3480
tcaagatcac ttttttcaag ctgttccaag gtcgggttaa gaatgaactc gccatcttga 3540
taaccacac gggcagcacc gatggggccc aagaaaggca caccggaaag agcaagagcc 3600
gcagaagcag cgatcatcgc caagatatcc ggttcattct cgccgtcata ggacatgacc 3660
tgcgcaatga ccaaggtttc gttatagaaa ccttccggaa acagaggacg gattggacgg 3720
tcgattaacc gtgaaatcag cgtttcccgt tcggttgcgc cacgttcacg cttgaaaaag 3780
ccaccgggaa tacggcctgc tgctgaatat tttcctgat aatggacggt caaaggaaag 3840
aaatcctgac cttcttttac cgtttttgct gccgttacgg cgcataatac gaccgtttca 3900
cctaaggtcg cgttctgcgt ctgtctotta tacacatctc aaccatcatc gatgaattht 3960

ES 2 611 977 T3

ctcggtgtt ctgcatatt ggctcgaatt ctacctgcag gctgatgaat gctcatccgg 4020
 aattagcggc cgcttaatta acttattcag gcgtagcaac caggcgтта agggcaccaa 4080
 taactgcctt aaaaaaatta cgccccgccc tgccactcat cgcagtactg ttgtaattca 4140
 ttaagcattc tgccgacatg gaagccatca caaacggcat gatgaacctg aatcgccagc 4200
 ggcacagca ccttgtcgcc ttgctataa tatttgcca tggtgaaaac gggggcgaag 4260
 aagttgtcca tattggccac gtttaaatca aaactgggta aactcaccca gggattggct 4320
 gagacgaaaa acatattctc aataaacctt ttagggaaat aggccaggtt ttcaccgtaa 4380
 cagccacat cttgcaata tatgtgtaga aactgccgga aatcgctctg gtattcactc 4440
 cagagcgatg aaaacgtttc agtttgotca tggaaaacgg tgtaacaagg gtgaacta 4500
 tcccatatca ccagctcacc gtctttcatt gccatacggg attccggatg agcattcatc 4560
 aggggggcaa gaatgtgaat aaaggccgga taaaacttgt gcttattttt ctttacggtc 4620
 tttaaaagg ccgtaatatc cagctgaacg gtctggttat aggtacattg agcaactgac 4680
 tgaaatgcct caaaatgttc ttacgatgc cattgggata tatcaacggt ggtatatcca 4740
 gtgatttttt tctccatttt agcttcotta gctcctgaaa atctcgataa ctcaaaaaat 4800
 acgcccggta gtgatcttat ttcattatgg tgaaagtgg aacctctac gtgccgatca 4860
 acgtctcatt ttogccaaaa gttggcccag ggcttcccgg tatcaacagg gacaccagga 4920
 tttatttatt ctgogaagtg atcttccgtc acgcgccgcg ataacttcgt atagcataca 4980
 ttatacgaag ttatgcgatc gcaagcttgc caacgactac gcactagcca acaagagctt 5040
 cagggttgag atgtgtataa gagacaggtt ctgcttaat gacggcgcca tctgcctgac 5100
 gggcaacctg tccggtttcc agtgtcagtt tttttccgcc ccaatcgatt tccgtggctt 5160
 taatatcgaa cattcaattt ccttctggct ccgacctat gtcaggagct gcctttttcc 5220
 ggattcggcg gattgcgaga taaccggatg cgaaacggcc tcccgataag gaggccgcat 5280
 cgtcttttac ttacgaagac ccagtcttgc gataaggctg ctgtagcggg ctttgtcgtt 5340
 attcttgacg taatcaagca ggctacgacg cttgttaacc aaacggagca aaccacgacg 5400
 ggaatggta tccttagcgt gggttttgaa atgttcggtc aggttaacaa tccgttcggt 5460
 cagaatagca acctggactt ccggagaacc cgtatgcct tcgacgcgtg catattcttt 5520
 gattaacgct tctttgcgct ctgctgtgat cgacatcagt cgtcctttct attttaaaaa 5580
 ttgaaaccgc gaaaaaccg agcttcttgc ccagtgactt caataagagc aaccggaatt 5640
 tgtgattcca ttgcatatt caagccggat tctgcgttaa ttccgatcaa tttctgcccc 5700
 tgacggagtg cttttgcctg atcggaggaa actgctaaag ccgggatgct gtccagccct 5760
 gctgttaaag gaaaaagcat ttgatacagc tttccagctt gcgccatttc cattaatttg 5820

ES 2 611 977 T3

tccagtghaa ttgcctgttc actcgtaaat ggccccgatt tcaatcgacg caaatagctg 5880
 acatgacat aacaattcag tttatgggca atgtcgcggg caaggcttct aatataggtg 5940
 cctttagaaa ccttagcgat aaaagtagcc tgatcggggc taaaatcgtc taaggtcaat 6000
 tcatggatag taatatcgcg tcctttgacg gtcaccgttt ctccggcgcg gacgcggtcg 6060
 caggcgcgtt tcccattgat tttcaaggca gaataaacgg gggggatttg gtgtatttca 6120
 cctcggagc aagatagaac agattccact tcttcccttg tgggacgttt atctgattga 6180
 gcggtgatgc tcccctctct atccaagcta tccgtttccg ctccaaaagc gagggtaaag 6240
 gcatagactt tgtcggaatt aagcgcatag cccgcccaatt tcgtcgcctc gcctaaggca 6300
 acaggtaaaa cgccagaggc aaggggatct aacgtcccgc catgaccgac ttttactttc 6360
 ggtaatccgg cgatacggat agcgcgtttg acagcagaga cggcttgggt tgagcctaag 6420
 ccttcttgc tgtcgagaat gatccagcgg cgcgcccccg ggtaccgagc t 6471

<210> 59
 <211> 1369
 <212> ADN

5 <213> secuencia artificial

<220>
 <223> fragmento construido para integración en el locus pnp de *Z. mobilis*

<400> 59

aatgacggcg ccatctgcct gacgggcaac ctgtccggtt tccagtgtca gttttttcc 60
 gcccgaatcg atttctggc gttaaatatc gaacattcaa tttccttctg gctccgacc 120
 tatgtcagga gctgcctttt tccggattcg ccggattgcg agataaccgg atgcgaaacg 180
 gcctccgat aaggaggccg catcgtcttt tacttacgaa gaccagctct tgcgataagg 240
 ctgctgtagc ggtctttgtc gttattcttg acgtaataca gcaggctacg acgcttgta 300
 accaaacgga gcaaaccacg acgggaatgg ttatccttag cgtgggtttt gaaatgttcg 360
 gtcaggtaaa caatccgttc ggtcagaata gcaacctgga cttccggaga acccgatcgc 420
 ccttcgacgc gtgcatattc tttgattaac gcttctttgc gttctgctgt gatcgacatc 480
 agtcgtcctt tctattttaa aaattgaaac cgcgcaaac ccgagcttct tgcccagtg 540
 cttcaataag agcaaccgga atttgtgatt ccattgccat attcaagccg gattctgcgt 600
 taattccgat caatttctgc ccctgacgga gtgcttttgc ctgatcggag gaaactgcta 660
 aagccgggat gtcgtccagc cctgctgtta aaggcaaaag catttgatac agctttccag 720
 cttgcgccat ttccattaat ttgtccagtg aaattgcctg ttcactcgtg aatggccccg 780
 atttcaatcg acgcaaatag ctgacatgac cataacaatt cagtttatgg gcaatgtcgc 840
 gggcaaggct tctaataatag gtgcctttag aaaccttagc gataaaagta gcctgatcgg 900
 ggctaaaatc gtotaaggtc aattcatgga tagtaaatc gcgtcctttg acggtcaccg 960

ES 2 611 977 T3

tttctccggc gcggaacgcg tcgcaggcgc gtttccatt gattttcaag gcagaataaa 1020
 cgggggggat ttggtgtatt tcacctcgga agcaagatag aacagattcc acttcttccc 1080
 ttgtgggacg tttatctgat tgagcgggta tgctcccctc tctatccaag ctatccgttt 1140
 ccgctccaaa agcgagggta aaggcataga ctttgtcggg attaagcgca tagcccgcca 1200
 atttcgtcgc ctgcctaag gcaacaggta aaacgccaga ggcaagggga tctaactgcc 1260
 cgccatgacc gacttttact ttcggtaatc cggcgatagc gatagcgcgt ttgacagcag 1320
 agacggcctg ggttgagcct aagccttctt gcttgtcagc aatgatcca 1369

<210> 60
 <211> 1251
 <212> ADN
 <213> secuencia artificial

5

<220>
 <223> fragmento construido para integración en el locus pnp de Z. mobilis

<400> 60

catcaatgca aggcagccac cgaaaacggt tgccatggaa gaggaacat tagattcggg 60
 gatatccgac aaaacacgga tggatatacgg gaaatcagcc ttgctcggca aaacccggatg 120
 aagcgcacgc catgccagtt taccatggcc gatttcacga cgaccggag caccaaacg 180
 accaacttca ccgaccgaat atggggggaa gttgtaatgc agcatgaagc gttcataatg 240
 aaggccgggt aaaccgtcga tcatctgttc agcatccgcc gttccaaggg tggttgaaac 300
 caaagcctgt gtttcaccac gcgtaaacag ggcagaacca tgcgtgcggg gcaagaatcc 360
 aacctgagag cggatcgggc gaactgtttt cagatcgcgt ccgtcaattc tccggccttc 420
 tttcagaatg gcggtgcgga cgatttttgc ttccagcttc ttcgcaagct tctggatgcg 480
 cagcttttct gtcggatcag cttctgggaa ggcttcgtca agaaccgctt tggccttgct 540
 caaggcgtca acacgttcag ccttgccgagt cagacgatag gccttatcga aattgtcgtg 600
 gcagagttct tcgactttgg cagccaatgc gctgtcatca tagctgacca gttccaagg 660
 ctcttttagc gcctgttctg ccagattgat gatagctttg ataaccggct gtaaagattc 720
 atgcgcaaaa gaaacggcat tgagcatgac ttcttcggga agctcattcg cttccgattc 780
 aaccatcatc acggcatcac gggtagcccc gacaaccaga tcaagatcac ttttttcaag 840
 ctgttccaag gtcgggttaa gaatgaactc gccatcttga taaccacac gggcagcacc 900
 gatggggccc aagaaaggca caccggaaag agcaagagcc gcagaagcag cgatcatcgc 960
 caagatatcc ggttcattct cgcctcata ggacatgacc tgcgcaatga ccaaggtttc 1020
 gttatagaaa ccttccggaa acagaggacg gattggacgg tcgattaacc gtgaaatcag 1080
 cgtttcccgt tcggttgccg cacgttcacg cttgaaaaag ccaccgggaa tacggcctgc 1140
 tgctgaatat ttttctgat aatggacggg caaaggaaag aaatcctgac cttcttttac 1200
 cgtttttgct gccgttaccg cgcataatac gaccgtttca octaaggctcg c 1251

10

<210> 61
 <211> 6342

ES 2 611 977 T3

<212> ADN
 <213> secuencia artificial
 <220>
 <223> plásmido construido

5 <400> 61

```

ctagtgtagac ggaagatcac ttcgcagaat aaataaatcc tgggtgtccct gttgataccg      60
ggaagccctg ggccaacttt tggcgaaaat gagacggtga tcggcacgta agaggttcca      120
actttcacca taatgaaata agatcactac cgggcgtatt ttttgagtta tcgagatfff      180
caggagctaa ggaagctaaa atggagaaaa aaatcactgg atataccacc gttgatatat      240
cccaatggca tcgtaaagaa cattttgagg catttcagtc agttgctcaa tgtacctata      300
accagaccgt tcagctggat attacggcct ttttaagac cgtaaagaaa aataagcaca      360
agttttatcc ggcctttatt cacattcttg cccgcctgat gaatgctcat ccggaattcc      420
gtatggcaat gaaagacggg gagctggtga tatgggatag tgttcaccct tgttacaccg      480
ttttccatga gcaaactgaa acgttttcat cgctctggag tgaataccac gacgatffcc      540
ggcagtttct acacatatat tcgcaagatg tggcgtgta cggtgaaaac ctggcctatt      600
tccctaaagg gtttattgag aatatgtfff tcgtctcagc caatccctgg gtgagtttca      660
ccagttttga tttaaacgtg gccaatatgg acaacttctt cggccccgtt ttcacatgg      720
gcaaatatta tacgcaaggc gacaaggtgc tgatgccgct ggcgattcag gttcatcatg      780
ccgtttgtga tggcttccat gtcggcagaa tgcttaatga attacaacag tactgcatg      840
agtggcaggg cggggcgtaa tttttttaag gcagttattg gtgcccttaa acgcctgggt      900
gctacgcctg aataagggcc ggcccggcaa gacgtgatat ggaaccggaa tttgctccgg      960
cattcctgcy caaagatagc taatatcttt catatfffgt atcgaaaaag gagggctfff      1020
aaagatcctc cttttttttg cataaaaaga aggccataga acaaacagtg ataaagacag      1080
tctcaaactg tctttttata gaaaatacca gaatattgta tctgggggag gatgcatggg      1140
cttaatccgg aataccccgg tcatgcacag gatgtagag cttttgcctt tatggcaaaa      1200
taaaccatgg ctcgggaata tctgcgcttt gattfffgtg ggatgtgcct tccttgtccg      1260
tagtattatt gggcattfff taccggcagg ttatccttcc gtgaccttta tgccgacaat      1320
gcttgtgggt actttcctct ttgggacaag accgggtatt atcgcggcta ttcttagcct      1380
gatggttgcy ccttattffa tcgaagaagg aagccgattt aacgggtgat tggctcgggt      1440
tctttgcctg ctagaaacag tcactgatat gggattgggt attgcgctac agcaaggtaa      1500
    
```

ES 2 611 977 T3

ttaccgcctc cagaaaaagc gtgcctataa tcagatgctg gctgaacgca atgagttgct 1560
 gtttcatgaa ttacagcatc gcatttcaaa taacttacag gttattgctg cattattgctg 1620
 gatgcaaagc cgcagcatca ccgatgaaaa agccaaggaa gctattgatg cctctgttcg 1680
 tcggattcat atgatcggtg aattacagcg ggcgctttat attaaaaacg ggaatcagct 1740
 tggggcaaaa ttgatccttg atcgcttgat caaagaggtc attgctcca gtaatctccc 1800
 gaacatccgc tataaaatag aagctgaaga cctgatctta ccgtcagata tggcaatccc 1860
 tttagcgctt gtatctgctg aatccgtttc aaacgcgcta gagcatggct ttaaaggcga 1920
 tcataaagac gcgtttattg aaattaagct tcaaaaaatt agcgggcaaa tcgaacttac 1980
 catttccaat aatggcaaac ctcttcccca aggcttttcc cttgaaaagg tcgatagctt 2040
 aggcctgaaa attgctgcta tgtttgcccg acaattcaaa ggaaaattca ccttaagtaa 2100
 tcagcctaac cgttatgtgg tttctagcct tattttgcct tgcggttagg cggccgctgg 2160
 tacccaattc gccctatagt gagtcgtatt acgcgcgctc actggccgctc gttttacaac 2220
 gtcgtgactg ggaaaaccct ggcgttacc aacttaatcg ccttgacgca catccccctt 2280
 tcgccagctg gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca 2340
 gcctgaatgg cgaatgggac gcgccctgta gcggcgcaat aagcgcggcg ggtgtggtgg 2400
 ttacgcgcag cgtgaccgct aacttgcca gcgccctagc gcccgctcct ttcgctttct 2460
 tccttctctt tctcgccacg ttcgcccggct tccccgta agctctaaat cgggggctcc 2520
 ctttagggtt ccgatttagt gctttacggc acctcgacc caaaaactt gattagggtg 2580
 atggttcacg tagtgggcca tcgccctgat agacggtttt tcgcccttg acgttggagt 2640
 ccacgttctt taatagtgga ctctgttcc aaactggaac aacactcaac cctatctcgg 2700
 tctattcttt tgatttataa gggattttgc cgatttcggc ctattggtta aaaaatgagc 2760
 tgatttaaca aaaatttaac gcgaatttta acaaaatatt aacgcttaca atttaggtgg 2820
 cacttttcgg ggaaatgtgc gcggaacccc tattgttta ttttctaaa tacattcaaa 2880
 tatgtatccg ctcatgagac aataaccctg ataaatgctt caataatatt gaaaaaggaa 2940
 gagtatgagt attcaacatt tccgtgtcgc ccttattccc ttttttgcgg cattttgcct 3000
 tcctgttttt gctcaccagc aaacgctggt gaaagtaaaa gatgctgaag atcagttggg 3060
 tgcacgagtg ggttacatcg aactggatct caacagcggg aagatccttg agagttttcg 3120
 ccccgaagaa cgttttccaa tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt 3180
 atcccgatt gacgccgggc aagagcaact cggtcgccgc atacactatt ctcagaatga 3240
 cttggttgag tactcaccag tcacagaaaa gcactctacg gatggcatga cagtaagaga 3300
 attatgcagt gctgccataa ccatgagtga taacactgcg gccaaacttac ttctgacaac 3360

ES 2 611 977 T3

gatcggagga ccgaaggagc taaccgcttt tttgcacaac atgggggatc atgtaactcg 3420
ccttgatcgt tgggaaccgg agctgaatga agccatacca aacgacgagc gtgacaccac 3480
gatgcctgta gcaatggcaa caacgttgcg caaactatta actggcgaac tacttactct 3540
agcttcccgg caacaattaa tagactggat ggaggcggat aaagttgcag gaccacttct 3600
gcgctcggcc cttccggctg gctggtttat tgctgataaa tctggagccg gtgagcgtgg 3660
gtctcgcggt atcattgcag cactgggggc agatggtaag cccctcccga tcgtagttat 3720
ctacacgacg gggagtcagg caactatgga tgaacgaaat agacagatcg ctgagatagg 3780
tgcctcactg attaagcatt ggtaactgtc agaccaagtt tactcatata tactttagat 3840
tgatttaaaa cttcattttt aatttaaaag gatctaggtg aagatccttt ttgataatct 3900
catgaccaa atcccttaac gtgagttttc gttccactga gcgtcagacc ccgtagaaaa 3960
gatcaaagga tcttcttgag atcctttttt tctgcgcgta atctgctgct tgcaaacaaa 4020
aaaaccaccg ctaccagcgg tggtttgtt gccggatcaa gagctaccaa ctctttttcc 4080
gaaggttaact ggcttcagca gagcgcagat accaaatact gtccttctag tgtagccgta 4140
gttaggccac cacttcaaga actctgtagc accgcctaca tacctcgctc tgctaatact 4200
gttaccagtg gctgctgcca gtggcgataa gtcgtgtctt accgggttg actcaagacg 4260
atagttaccg gataaggcgc agcggtcggg ctgaacgggg gggtcgtgca cacagcccag 4320
cttgagcga acgaactaca ccgaactgag atacctacag cgtgagctat gagaaagcgc 4380
cacgcttccc gaagggagaa aggcggacag gtatccgta agcggcaggg tcggaacagg 4440
agagcgcacg agggagcttc cagggggaaa cgcctggtat ctttatagtc ctgtcgggtt 4500
tcgccacctc tgacttgagc gtcgattttt gtgatgctcg tcaggggggc ggagcctatg 4560
gaaaaacgcc agcaacgcgg cttttttacg gttcctggcc ttttgctggc cttttgctca 4620
catgttcttt cctgcgttat cccctgattc tgtggataac cgtattaccg cctttgagtg 4680
agctgatacc gctcgcgca gccgaacgac cgagcgcagc gagtcagtga gcgaggaagc 4740
ggaagagcgc ccaatacgca aaccgcctct ccccgcgctg tggccgattc attaatgcag 4800
ctggcacgac aggtttcccg actggaaagc gggcagtgag cgcaacgcaa ttaatgtgag 4860
ttagctcact cattaggcac cccaggcttt acactttatg cttccggctc gtatgttggtg 4920
tggaattgtg agcggataac aatttcacac aggaacagc tatgaccatg attacgcaa 4980
gcgcgcaatt aaccctcact aaaggaaca aaagctggag ctccctatcg tctgactcgc 5040
aaggctgaac gtgttgacgc cttgagcaag gccaaagcgg ttcttgacga agccttccca 5100
gaagctgatc cgacagaaaa gctgcgcac cagaagcttg cgaagaagct ggaagcaaaa 5160
atcgtccgca ccgccattct gaaagaaggc cggagaattg acggacgcga tctgaaaaca 5220
gttcgcccga tccgctctca ggttgattc ttgccccgca cgcattggtc tgccctgttt 5280

ES 2 611 977 T3

acgcgtggtg aaacacaggc tttggtttca accacccttg gaacggcgga tgctgaacag 5340
atgatcgacg gtttaaccgg ccttcattat gaacgcttca tgctgcatta caacttcccc 5400
ccatattcgg tcggtgaagt tggtcgtttt ggtgctccgg gtcgctcgtga aatcggccat 5460
ggtaaactgg catggcgtgc gcttcatccg gttttgccga gcaaggctga tttcccgtat 5520
accatccgtg ttttgtcggga tatcaccgaa tctaattggtt cctcttccat ggcaaccggt 5580
tgcggtggct gccttgcatl gatggatgcc ggtgttccct taacgcgtcc ggtttccggt 5640
atcgccatgg gtcttattct ggaaaagac ggcttcgcta ttttgtccga tatcatgggt 5700
gatgaagatc acttgggtga tatggacttt aaggctcggc gtaccgaaaa aggtatcacc 5760
agcctccaga tggacatcaa ggttgcctggc attaccgaag aatcatgca gaaagctttg 5820
gaacaggcta aagggtggccg tgctcatatc ttgggtgaaa tgtccaaagc gctgggtgaa 5880
gtccgctccg aaatttctaa tttggcaccg cgcattgaaa caatgagcgt accaaaagac 5940
aaaatccgtg atgttatcgg aacgggcgga aaagttatcc gtgaaatcgt ggcgaccaca 6000
ggtgccaaag togatataca agatgacggc acggttcgtc tgtcttcttc tgatccggcc 6060
aatattgaag cagcccgtga atggatcaat ggtattgttg aagaaccgga agtaggcaaa 6120
atctataacg gtaaagtcgt caatatacgtt gatttcggtg ccttcgtaaa cttcatgggt 6180
ggccgtgacg gcttggtaaa tgtttcggaa atcaagaacg aacgtgtcaa caaggtcagc 6240
gatgtcctgt ccgaaggtca ggaagtcaaa gtcaaggttc ttgaaattga caaccgtggc 6300
aaggttcggc tgtctatgcg tgttgtogat caggaaaccg ga 6342

<210> 62
<211> 6322
<212> ADN
5 <213> secuencia artificial

<220>
<223> plásmido construido
<400> 62

ctagtgtgac ggaagatcac ttcgcagaat aaataaatcc tgggtgtccct gttgataccg 60
ggaagccctg ggccaacttt tggcgaaaat gagacgttga tcggcacgta agaggttcca 120
actttcacca taatgaaata agatcactac cgggogtatt ttttgagtta togagatttt 180
caggagctaa ggaagctaaa atggagaaaa aatcactgg atataccacc gttgatatat 240
cccaatggca tcgtaaagaa cattttgagg catttcagtc agttgctcaa tgtacctata 300
accagaccgt tcagctggat attacggcct ttttaaagac cgtaaagaaa aataagcaca 360
agttttatcc ggcctttatt cacattcttg cccgctgat gaatgctcat ccggaattcc 420
gtatggcaat gaaagacggg gagctggtga tatgggatag tgttcaccct tgttacaccg 480

ES 2 611 977 T3

ttttccatga gcaaactgaa acgttttcat cgctctggag tgaataccac gacgatttcc 540
 ggcagtttct acacatatat tgcgaagatg tggcgtgta cggtgaaaac ctggcctatt 600
 tccctaaagg gtttattgag aatatgtttt tctctcagc caatccctgg gtgagtttca 660
 ccagttttga tttaaacgtg gccaatatgg acaacttctt cgccccggtt ttcacatg 720
 gcaaataatta tacgcaaggc gacaagggtc tgatgccgct ggcgattcag gttcatcatg 780
 ccgtttgtga tggcttccat gtcggcagaa tgcttaatga attacaacag tactgcatg 840
 agtggcaggg cggggcgtaa tttttttaag gcagttattg gtgcccttaa acgcctggtt 900
 gctacgcctg aataagggcc ggccttcatt atgaaagcct catgctgcat tacaacttcc 960
 ccccatattc ggtcggtgaa gttggtcgtt ttggtgctcc gggtcgtcgt gaaatcggcc 1020
 atggtaaaact ggcattggcgt gcgcttcac cggttttgcc gagcaaggct gatttcccgt 1080
 ataccatccg tgttttgcg gatatcaccg aatctaagtg ttctcttcc atggcaaccg 1140
 tttgcgggtg ctgccttgca ttgatggatg ccggtgttcc cttaacgcgt ccggtttccg 1200
 gtatcgccat gggctcttatt ctggaaaaag acggcttgcg tattttgtcc gatatcatg 1260
 gtgatgaaga tcacttgggt gatatggact ttaaggctgc cggtagcgaa aaaggtatca 1320
 ccagcctcca gatggacatc aaggttgctg gcattaccga agaaatcatg cagaaagctt 1380
 tggaacaggc taaaggtggc cgtgctcata tcttgggtga aatgtccaaa gcgctgggtg 1440
 aagtccgctc cgaaatttct aatttggcac cgcgcattga aacaatgagc gtaccaaag 1500
 acaaaatccg tgatgttatt ggaacgggcg gaaaagttat ccgtgaaatc gtggcgacca 1560
 caggtgccaa ggtcgatatt gaagatgacg gcacggttcg tctgtcttct tctgatccgg 1620
 ccaatattga agcagcccgt gaatggatca atggtattgt tgaagaaccg gaagtaggca 1680
 aaatctataa cggtaaagtc gtcaatatcg ttgatttcgg tgcttctgta aacttcatg 1740
 gtggccgtga cggcttggta catgtttcgg aatcaagaa cgaacgtgac aacaaggtca 1800
 gcgatgtcct gtccgaaggc caggaagtca aagtcaaggc tcttgaatt gacaaccgtg 1860
 gcaaggttcg cctgtctatg cgtgttgcg atcaggaaac cggcgcagag ctggatgata 1920
 accgtccgcc acgtgagaac gcagaacgac gcggtgggtg gcgtcctcgt cgtgatcggg 1980
 gccctcgtcg ggaatctggc gatcgtccgg caagacgtga tatggaaccg gaatttgcct 2040
 cggcattcct gcgcaaagat agctaataatc tttcatattt tgtatcgaat aaggagggtc 2100
 tttaaagatc ctctttttt ttgcataaaa agaaggccat agaacaaaca gtgataaaga 2160
 cagtctcaaa ctgtctttt atagaaaata ccagaatatt gtatctgggg gaggatgcat 2220
 ggtcttaatc cggaaatccc cggtcattgc gccgctggta cccaattcgc cctatagtga 2280
 gtcgtattac gcgcgctcac tggccgtcgt ttacaacgt cgtgactggg aaaaccctgg 2340
 cgttacccaa cttaatcgcc ttgcagcaca tcccccttcc gccagctggc gtaatagcga 2400

ES 2 611 977 T3

agaggcccc accgatcgcc ctcccccaaca gttgcgcagc ctgaatggcg aatgggacgc 2460
 gccctgtagc ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac 2520
 acttgccagc gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacggt 2580
 cgccggcctt ccccgctcaag ctctaaatcg ggggctccct ttagggttcc gatttagtgc 2640
 tttacggcac ctcgacccca aaaaacttga ttagggatgat gggtcacgta gtgggccatc 2700
 gccctgatag acggtttttc gccctttgac gttggagtcc acgttcttta atagtggact 2760
 cttgttccaa actggaacaa cactcaaccc tatctcggtc tattcttttg atttataagg 2820
 gattttgccc atttcggcct attggttaaa aatgagctg atttaacaaa aatttaacgc 2880
 gaattttaac aaaatattaa cgcttacaat ttaggtggca cttttcgggg aatgtgcgc 2940
 ggaacccta tttgtttatt tttctaata cattcaaata tgtatccgct catgagacaa 3000
 taaccctgat aatgcttca ataatttga aaaaggaaga gtatgagtat tcaacatttc 3060
 cgtgtcggcc ttattccctt ttttgcggca ttttgccttc ctgtttttgc tcaccagaa 3120
 acgctggtga aagtaaaaga tgctgaagat cagttgggtg cacgagtggg ttacatcgaa 3180
 ctggatctca acagcggtaa gatccttgag agttttcgcc ccgaagaacg ttttccaatg 3240
 atgagcactt ttaaagttct gctatgtggc gcggtattat cccgtattga cgccgggcaa 3300
 gagcaactcg gtcgccgcat acactattct cagaatgact tggttgagta ctcaccagtc 3360
 acagaaaagc atcttacgga tggcatgaca gtaagagaat tatgcagtgc tgccataacc 3420
 atgagtgata aactgcggc caacttactt ctgacaacga tcggaggacc gaaggagcta 3480
 accgcttttt tgcaacaacat ggggatcat gtaactcgcc ttgatcgttg ggaaccggag 3540
 ctgaatgaag ccataccaaa cgacgagcgt gacaccacga tgcctgtagc aatggcaaca 3600
 acgttgcgca aactattaac tggcgaacta ctactctag cttcccggca acaattaata 3660
 gactggatgg aggcggataa agttgcagga ccacttctgc gctcggccct tccggctggc 3720
 tggtttattg ctgataaatc tggagccggt gagcgtgggt ctgcggtat cattgcagca 3780
 ctggggccag atggttaagc ctcccgtatc gtagttatct acacgacggg gagtacggca 3840
 actatggatg aacgaaatag acagatcgct gagatagggt cctcactgat taagcattgg 3900
 taactgtcag accaagttta ctcatatata ctttagattg atttaaaact tcatttttaa 3960
 tttaaaagga tctaggtgaa gatccttttt gataatctca tgacccaaaat cccttaacgt 4020
 gagttttcgt tccactgagc gtcagacccc gtagaaaaga tcaaaggatc ttcttgagat 4080
 cctttttttc tgccgcgtaat ctgctgcttg caaacaaaaa aaccaccgct accagcggtg 4140
 gtttgtttgc cggatcaaga gctaccaact ctttttccga aggtaactgg cttcagcaga 4200
 gcgcagatac caaatactgt ccttctagtg tagccgtagt taggccacca cttcaagaac 4260

ES 2 611 977 T3

tctgtagcac cgctacata cctcgctctg ctaatcctgt taccagtggc tgctgccagt 4320
ggcgataagt cgtgtcttac cgggttggac tcaagacgat agttaccgga taaggcgag 4380
cggtcgggct gaacggggggg ttcgtgcaca cagcccagct tggagcgaac gacctacacc 4440
gaactgagat acctacagcg tgagctatga gaaagcgcca cgcttcccga agggagaaag 4500
gcgacaggt atccggtaag cggcagggtc ggaacaggag agcgcacgag ggagcttcca 4560
gggggaaacg cctggtatct ttatagtcct gtcgggtttc gccacctctg acttgagcgt 4620
cgatTTTTgt gatgctcgtc aggggggagg agcctatgga aaaacgccag caacgcggcc 4680
TTTTTtaggt tcctggcctt ttgctggcct tttgctcaca tgttcttcc tgcgttatcc 4740
cctgattctg tggataaccg tattaccgcc tttgagttag ctgataccgc tcgccgcagc 4800
cgaacgaccg agcgcgagca gtcagttagc gaggaagcgg aagagcggcc aatacgcaaa 4860
ccgcctctcc ccgcgcgttg gccgattcat taatgcagct ggcacgacag gtttcccagc 4920
tggaaagcgg gcagttagcg caacgcaatt aatgtgagtt agctcactca ttaggcaccc 4980
caggctttac actttatgct tccggctcgt atgttgtgtg gaattgtgag cggataacaa 5040
ttcacacag gaaacagcta tgaccatgat tacgccaaag gcgcaattaa ccctcactaa 5100
agggaaacaa agctggagct cttatcggga gcccgtttcg catccggtta tctcgcaatc 5160
cggcgaatcc ggaaaaaggc agctcctgac atagggtcgg agccagaagg aaattgaatg 5220
ttcgatatta aacgccagga aatcgattgg gccggaaaaa aactgacact ggaaaccgga 5280
caggttgcc gtcaggcaga tggcgccgtc attgcgacct taggtgaaac ggtcgtatta 5340
tgcgcggtaa cggcagcaaa aacggtaaaa gaaggtcagg atttcttcc tttgaccgtc 5400
cattatcagg aaaaatattc agcagcaggc cgtattcccg gtggcttttt caagcgtgaa 5460
cgtggcgcaa ccgaacggga aacgctgatt tcacggtaa tcgaccgtcc aatccgtcct 5520
ctgtttccgg aaggtttcta taacgaaacc ttggtcattg cgcaggtcat gtcctatgac 5580
ggcgagaatg aaccggatat cttggcgatg atcgtcgtt ctgcggctct tgctcttcc 5640
ggtgtgcctt tcttgggccc catcgggtgct gcccggtgg gttatcaaga tggcgagttc 5700
attcttaacc cgaccttga acagcttgaa aaaagtgatc ttgatctggt tgtcggggct 5760
accgctgatg ccgtgatgat ggttgaatcg gaagcgaatg agcttcccga agaagtcatg 5820
ctcaatgccg tttcttttgc gcatgaatct ttacagccgg ttatcaaagc tatcatcaat 5880
ctggcagaac aggccgctaa agagccttgg gaactggtca gctatgatga cagcgcattg 5940
gctgccaaaag tcgaagaact ctgctaogac aatttcgata aggcctatcg tctgactcgc 6000
aaggctgaac gtgttgacgc cttgagcaag gccaaaagcgg ttcttgacga agccttccca 6060
gaagctgatc cgacagaaaa gctgcgcac cagaagcttg cgaagaagct ggaagcaaaa 6120
atcgtccgca ccgccattct gaaagaagggc cggagaattg acggacgcga totgaaaaca 6180

ES 2 611 977 T3

	gttcgcccga tccgctctca ggttgatc ttgccccgca cgcattggtc tgcctgttt	6240
	acgcgtggtg aaacacaggc tttggttca accacccttg gaacggcgga tgctgaacag	6300
	atgatcgacg gtttaaccgg ca	6322
	<210> 63	
	<211> 1200	
	<212> ADN	
5	<213> secuencia artificial	
	<220>	
	<223> fragmento construido para integración en el locus pnp de Z. mobilis	
	<400> 63	
	ttatcgggag gccgtttcgc atccggttat ctcgcaatcc ggcgaatccg gaaaaaggca	60
	gctcctgaca tagggtcggg gccagaagga aattgaatgt tcgatattaa acgccaggaa	120
	atcgattggg gcggaaaaaa actgacactg gaaaccggac aggttgcccc tcaggcagat	180
	ggcgccgtca ttgcgacctt aggtgaaacg gtcgtattat gcgcggtaac ggcagcaaaa	240
	acggtaaaag aaggtcagga tttctttcct ttgaccgtcc attatcagga aaaatattca	300
	gcagcaggcc gtattcccgg tggctttttc aagcgtgaac gtggcgcaac cgaacgggaa	360
	acgctgattt cacggttaat cgaccgtcca atccgtcctc tgtttccgga aggtttctat	420
	aacgaaacct tggtcattgc gcaggtcatg tcctatgacg gcgagaatga accggatatac	480
	ttggcgatga tcgctgcttc tgcggtctt gctctttccg gtgtgccttt cttgggcccc	540
	atcgggtgctg cccgtgtggg ttatcaagat ggcgagttca ttottaacct gaccttgaa	600
	cagcttgaaa aaagtgatct tgatctggtt gtcggggcta cccgtgatgc cgtgatgatg	660
	gttgaatcgg aagcgaatga gcttcccga gaagtcattc tcaatgccgt tttttttgcg	720
	catgaatctt tacagccggt tatcaaagct atcatcaatc tggcagaaca ggcgctaaa	780
	gagccttggg aactggtcag ctatgatgac agcgcattgg ctgccaaagt cgaagaactc	840
	tgctacgaca atttcgataa ggcctatcgt ctgactcgca aggctgaacg tgttgacgcc	900
	ttgagcaagg ccaaagcggg tcttgacgaa gccttcccag aagctgatcc gacagaaaag	960
	ctgcgcatcc agaagcttgc gaagaagctg gaagcaaaaa tcgtccgcac cgccattctg	1020
	aaagaaggcc ggagaattga cggacgcgat ctgaaaacag ttcgcccgat ccgctctcag	1080
	gttgattct tgcgccgcac gcatggttct gccctgttta cgcgtggtga aacacaggct	1140
	ttggtttcaa ccacccttgg aacggcggat gctgaacaga tgatcgacgg ttttaaccggc	1200
10	<210> 64	
	<211> 1323	
	<212> ADN	
	<213> secuencia artificial	
	<220>	
15	<223> fragmento construido para integración en el locus pnp de Z. mobilis	
	<400> 64	

ES 2 611 977 T3

ttcattatga acgcttcatg ctgcattaca acttcccccc atattcggtc ggtgaagttg 60
gtcgttttgg tgctccgggt cgtcgtgaaa tcggccatgg taaactggca tggcgtgcgc 120
ttcatccgggt tttgcccagc aaggctgatt tcccgtatac catccgtgtt ttgtcggata 180
tcaccgaatc taatggttcc tcttccatgg caaccgtttg cggtggtgc cttgcattga 240
tggatgccgg tgttccctta acgcgtccgg tttccggtat cgccatgggt cttattctgg 300
aaaaagacgg cttogctatt ttgtccgata tcatgggtga tgaagatcac ttgggtgata 360
tggactttaa ggtcgcgggt accgaaaaag gtatcaccag cctccagatg gacatcaagg 420
ttgctggcat taccgaagaa atcatgcaga aagctttgga acaggctaaa ggtggccgtg 480
ctcatatctt gggtgaaatg tccaaagcgc tgggtgaagt ccgctccgaa atttctaatt 540
tggcaccgcy cattgaaaca atgagcgtac caaaagacaa aatccgtgat gttatcggaa 600
cgggcggaaa agttatccgt gaaatcgtgg cgaccacagg tgccaaggtc gatatcgaag 660
atgacggcac ggttcgtctg tcttcttctg atccggccaa tattgaagca gcccgatgaat 720
ggatcaatgg tattgttgaa gaaccggaag taggcaaaat ctataacggt aaagtcgtca 780
atcgtgttga tttcgggtgcc ttcgtaaact tcatgggtgg ccgtgacggc ttggtacatg 840
tttcggaaat caagaacgaa cgtgtcaaca aggtcagcga tgtcctgtcc gaaggtcagg 900
aagtcaaagt caaggttctt gaaattgaca accgtggcaa ggttcgcctg tctatgctg 960
ttgtcgatca ggaaaccggc gcagagctgg atgataaccg tccgccacgt gagaacgcag 1020
aacgtcgcgg tggtgagcgt cctcgtcgtg atcggggccc tcgtcgggaa tctggcgatc 1080
gtccggcaag acgtgatatg gaaccggaat ttgctccggc attcctgcgc aaagatagct 1140
aatatctttc atattttgta tcgaaaaagg agggctttta aagatcctcc tttttttg 1200
ataaaaagaa ggccatagaa caaacagtga taaagacagt ctcaaactgt ctttttatag 1260
aaaataccag aatattgtat ctgggggagg atgcatggtc ttaatccgga ataccccgg 1320
cat 1323

<210> 65
<211> 474
<212> ADN

5 <213> secuencia artificial

<220>
<223> codón de inicio cambiado a ATG

<400> 65

atgacctctg ctgtgccatc aaatacgaag aaaaagctgg tgattgcttc cgatcacgca 60
gcatttgagt tgaatcaac cttgattact tggctgaaag agcttggtca tgaggtcgaa 120

ES 2 611 977 T3

gaccttgGCC ctcatgaaaa ccattcagtc gattatcccG attacggTta taagctggct	180
gtcgctatcg cagaaaaaac cgctgatttc ggtattgctt tatgtggctc gggaatcggT	240
atctcgatcg ctgtcaatcg ccatccggct gcccgTtgcg ctttgattac ggataacott	300
accgcccgtt tggcaagaga acataacaat gccaatgTta tcgctatggg tgcgagattg	360
atcggcattg aaaccgctaa ggattgtatt tcagctttcc ttgcaacgcc gtttgagggt	420
gaacgtcatg ttcgccgtat cgataaaactt tcgaatcctc agttcaatat ctag	474

REIVINDICACIONES

1. Una célula huésped bacteriana recombinante que comprende:
- a) una ruta metabólica de xilosa que comprende al menos un polipéptido que tiene actividad de xilosa isomerasa;
- 5 b) al menos una modificación genética que incrementa la actividad de ribosa-5-fosfato isomerasa en la célula huésped en comparación con la actividad de ribosa-5-fosfato isomerasa en la célula huésped que carece de dicha modificación genética, en donde dicha al menos una modificación genética que incrementa la actividad de ribosa-5-fosfato isomerasa se elige entre el grupo que consiste en:
- i) sobreexpresión de un gen endógeno que codifica un polipéptido que tiene actividad de ribosa-5-fosfato isomerasa; y/o
- 10 ii) expresión de al menos un gen no endógeno que codifica un polipéptido que tiene actividad de ribosa-5-fosfato isomerasa; y
- c) al menos una modificación genética en la secuencia de un gen endógeno que codifica polinucleótido fosforilasa que acorta la región de codificación lo que da como resultado la expresión de una proteína C-terminal truncada;
- 15 en donde la célula huésped bacteriana utiliza xilosa para producir etanol y la actividad de xilosa isomerasa no es limitante en la célula huésped bacteriana, en donde la actividad de xilosa isomerasa no es limitante cuando un aumento en la actividad de xilosa isomerasa no mejora la utilización de xilosa y la producción de etanol en dicha célula.
2. La célula huésped recombinante según la reivindicación 1, en la que la proteína truncada C-terminal de (c) comprende al menos aproximadamente 350 aminoácidos de la secuencia de aminoácidos N-terminal codificada por el gen endógeno que codifica polinucleótido fosforilasa.
- 20 3. La célula huésped recombinante según la reivindicación 1 ó 2, en la que la proteína truncada C-terminal de (c) es parte de una proteína de fusión que comprende al menos un aminoácido no nativo en el término C.
4. La célula huésped recombinante según una cualquiera de las reivindicaciones 1 a 3, en la que el polipéptido que tiene actividad de ribosa-5-fosfato isomerasa tiene la clasificación EC 5.3.1.6.
- 25 5. La célula huésped recombinante según la reivindicación 4, en la que el polipéptido que tiene actividad de ribosa-5-fosfato isomerasa se selecciona entre el grupo que consiste en ribosa-5-fosfato isomerasa A y ribosa-5-fosfato isomerasa B.
6. La célula huésped recombinante según una cualquiera de las reivindicaciones 1 a 5, en la que el polipéptido que tiene actividad de xilosa isomerasa proporciona actividad de xilosa isomerasa a más de aproximadamente 0,25 micromoles de producto/mg de proteína/minuto.
- 30 7. La célula huésped recombinante según una cualquiera de las reivindicaciones 1 a 6, en la que el polipéptido que tiene actividad de xilosa isomerasa tiene la clasificación EC 5.3.1.5.
8. La célula huésped recombinante según una cualquiera de las reivindicaciones 1 a 7, en la que la actividad de xilosa isomerasa se expresa usando un método seleccionado entre el grupo que consiste entre múltiples genes, entre un promotor altamente activo mutante, entre una xilosa isomerasa del Grupo I y entre una combinación de los mismos.
- 35 9. La célula huésped recombinante según una cualquiera de las reivindicaciones 1 a 8, en la que se hace al menos una modificación genética que aumenta la expresión de ribulosa-fosfato 3-epimerasa en comparación con la actividad de ribulosa-fosfato 3-epimerasa en la célula huésped que carece de dicha modificación genética, en donde dicha al menos una modificación genética que incrementa la actividad de ribulosa-fosfato 3-epimerasa se elige entre el grupo que consiste en:
- i) sobreexpresión de un gen endógeno que codifica un polipéptido que tiene actividad de ribulosa-fosfato 3-epimerasa; y/o
- 40 ii) expresión de al menos un gen no endógeno que codifica un polipéptido que tiene actividad de ribulosa-fosfato 3-epimerasa.
10. La célula huésped recombinante según la reivindicación 9, en la que el polipéptido que tiene actividad de ribulosa-fosfato 3-epimerasa tiene la clasificación EC 5.1.3.1.
11. La célula huésped bacteriana según una cualquiera de las reivindicaciones 1 a 10, en la que la célula se elige entre el grupo que consiste en *Zymomonas* y *Zymobacter*.
- 50 12. Un procedimiento para producir etanol que comprende:

- a) proporcionar la célula huésped bacteriana recombinante según una cualquiera de las reivindicaciones 1 a 11; y
- b) cultivar la célula huésped de (a) en un medio que comprende xilosa, con lo que la xilosa se convierte en etanol.

13. El proceso según la reivindicación 12, en el que el polipéptido que tiene actividad de ribosa-5-fosfato isomerasa según la reivindicación 1 (b) tiene la clasificación EC 5.3.1.6.

- 5 14. El proceso según la reivindicación 12, en el que el medio comprende o bien una mezcla de azúcares que incluye xilosa, o bien xilosa como único azúcar.

Ruta del Metabolismo de Xilosa

Ruta de Entner-Doudoroff

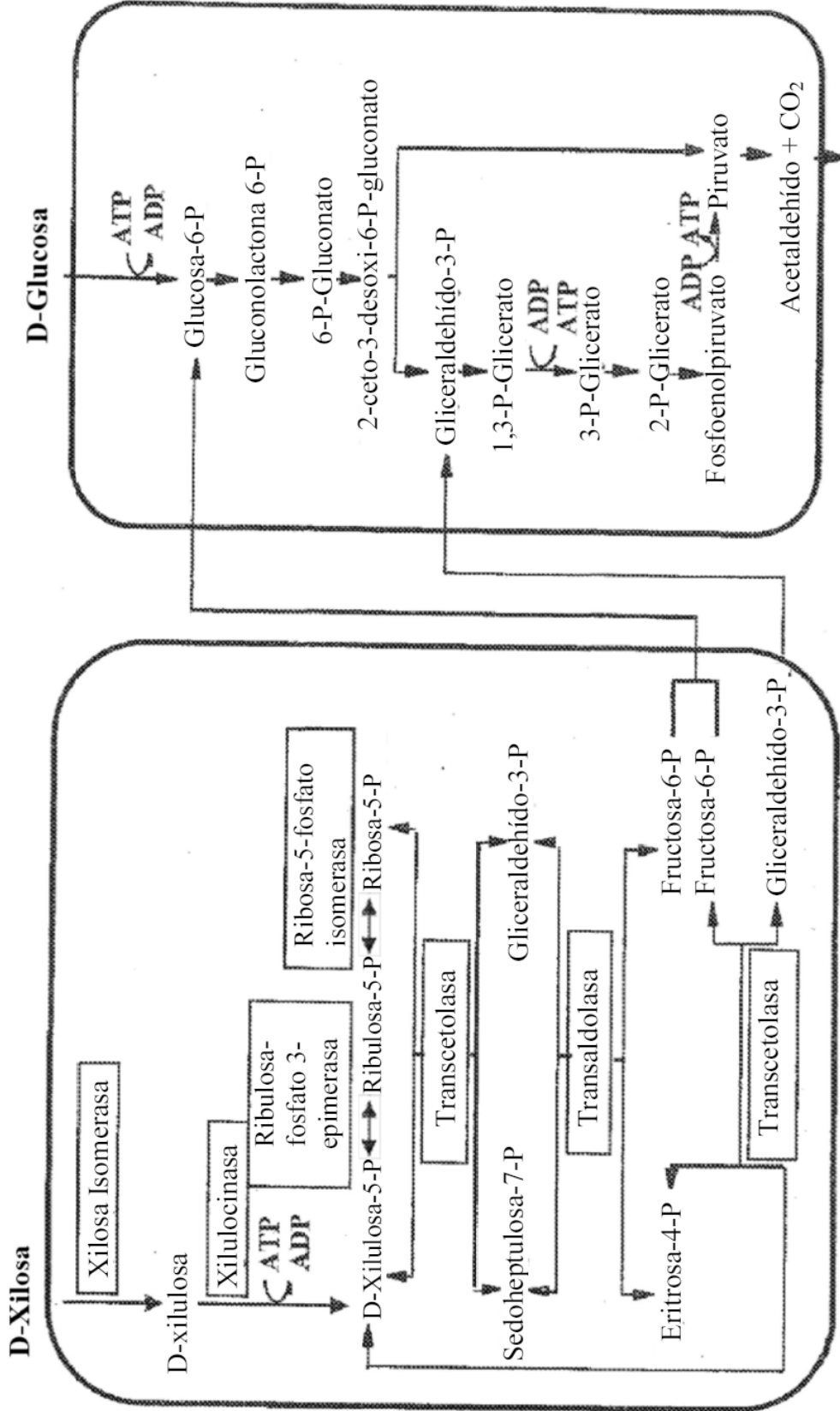


FIG. 1

Etanol

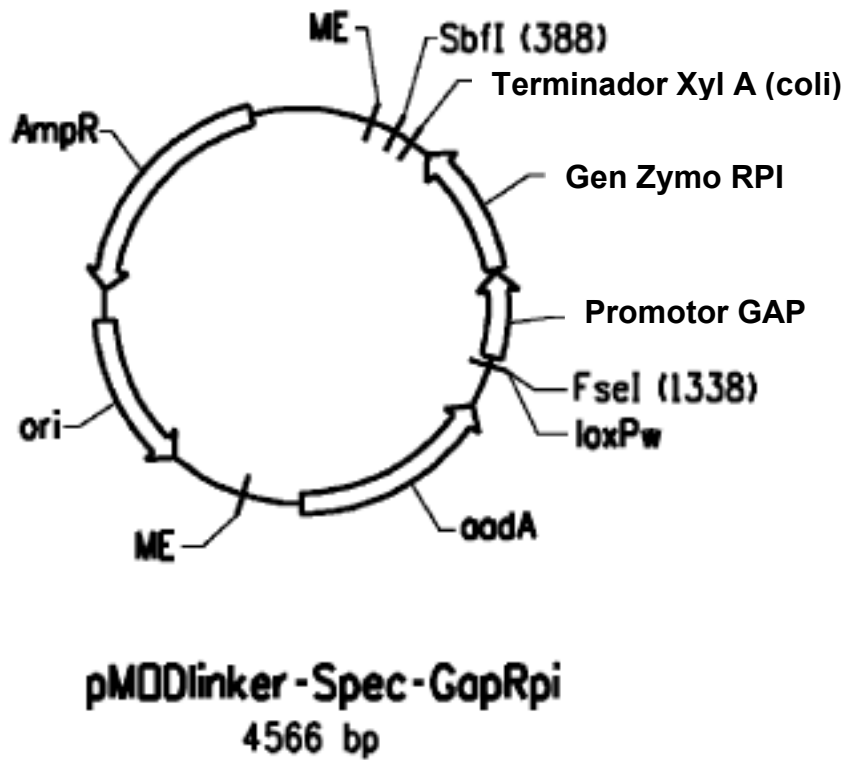


FIG. 2

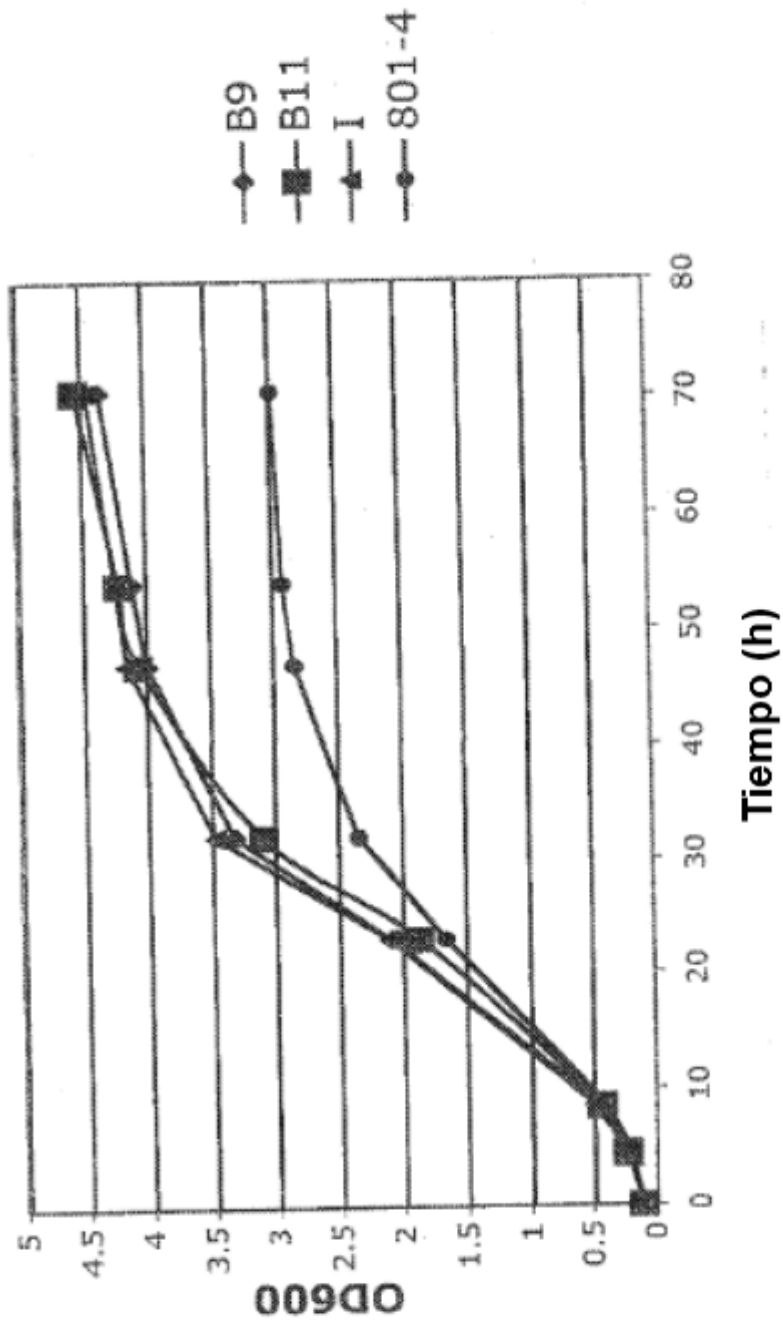


FIG. 3A

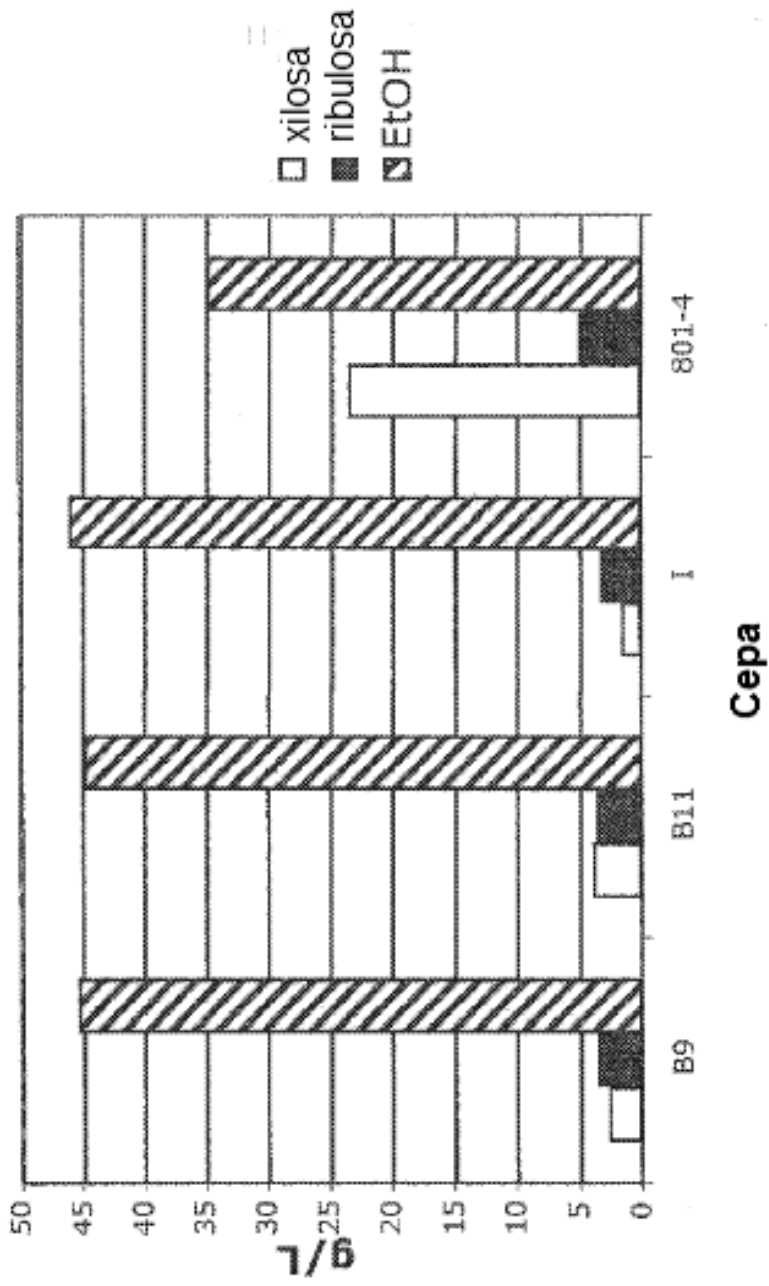


FIG. 3B

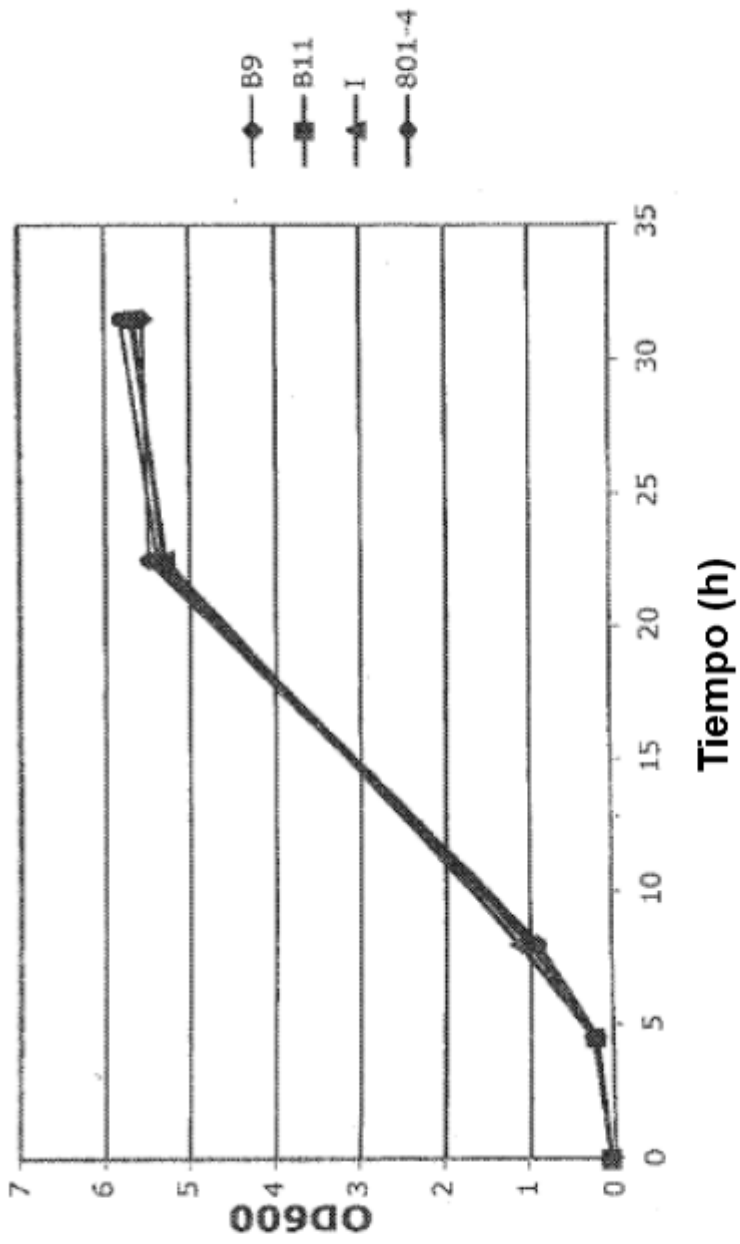


FIG. 3C

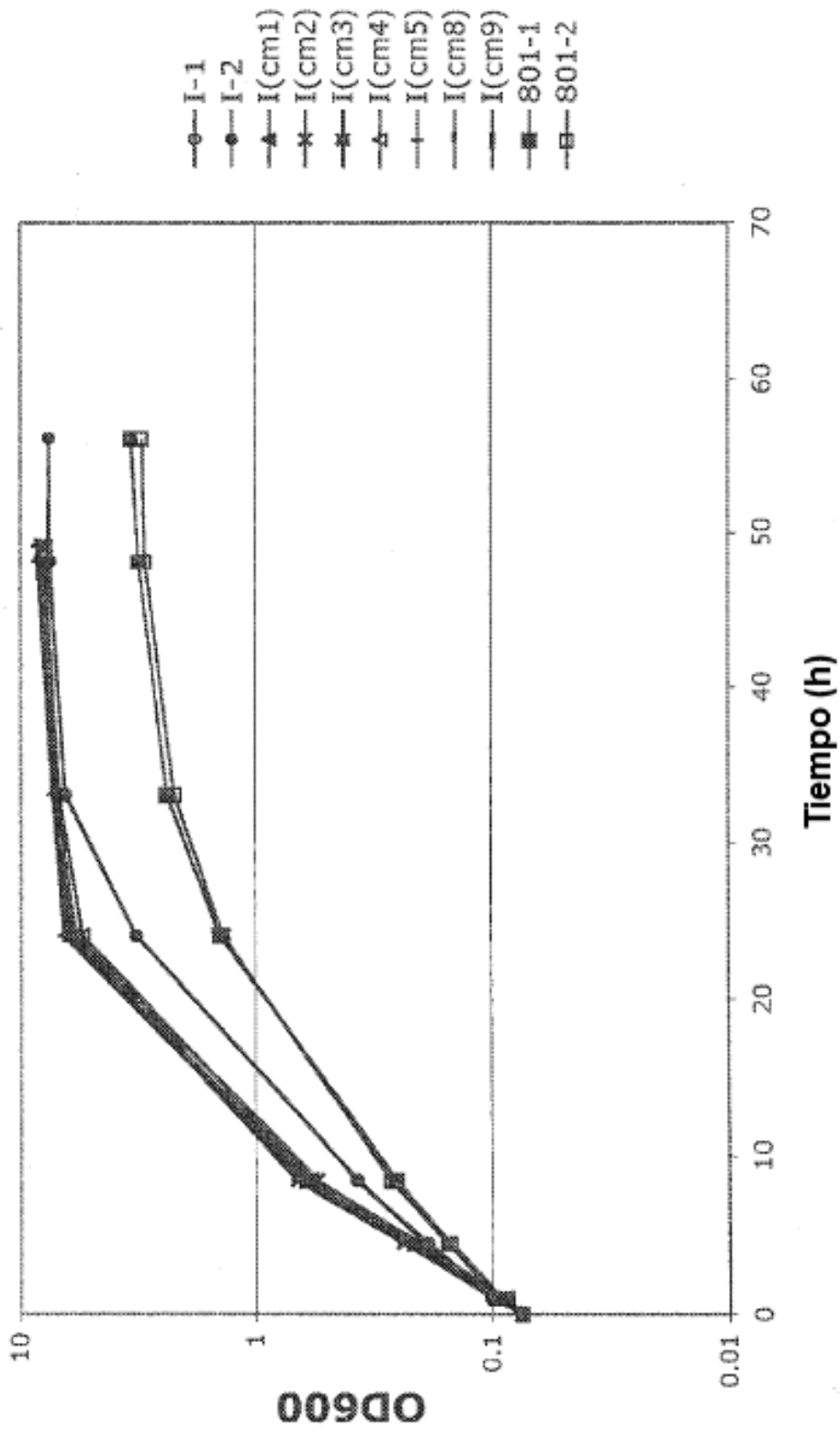


FIG. 4A

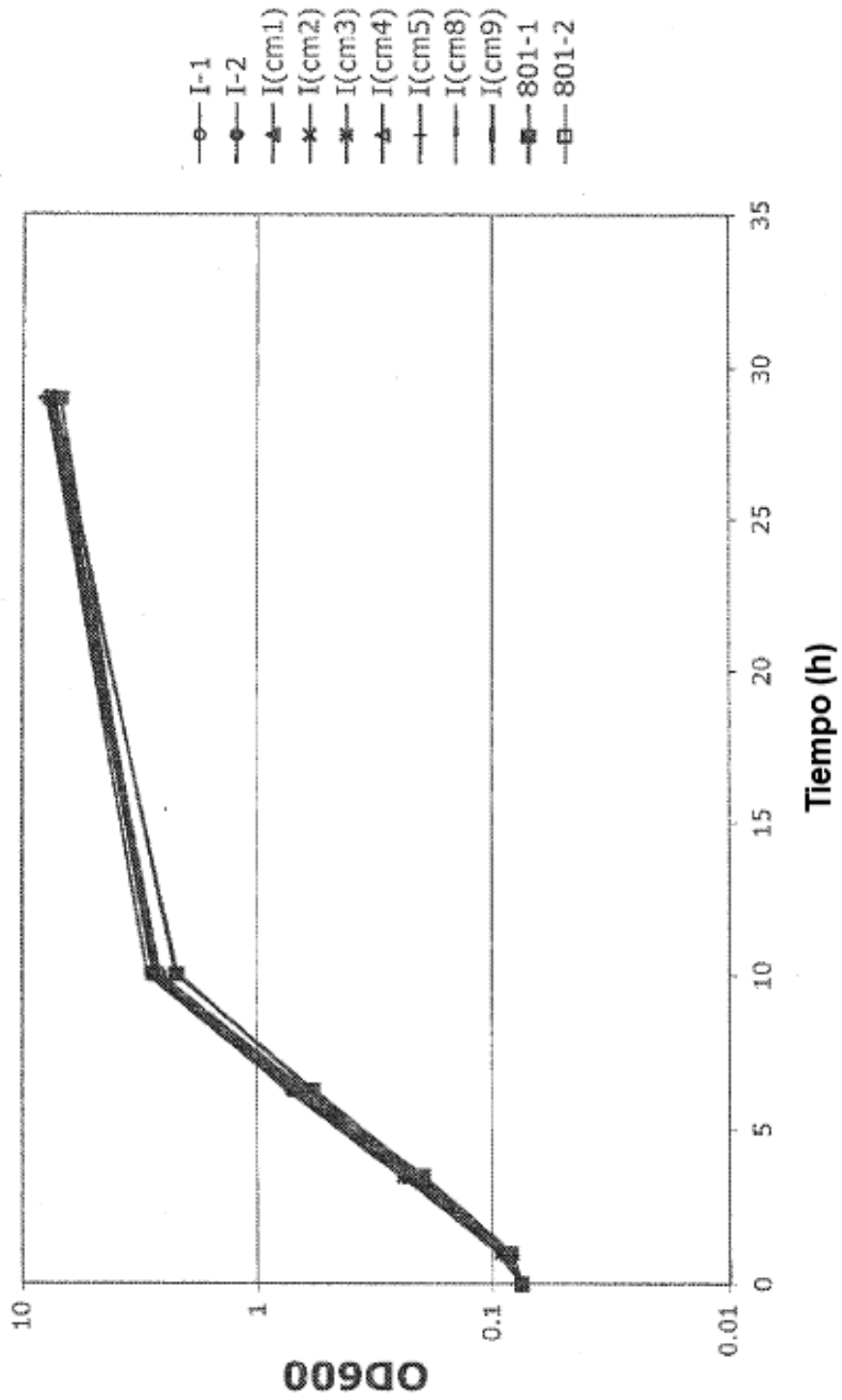


FIG. 4B

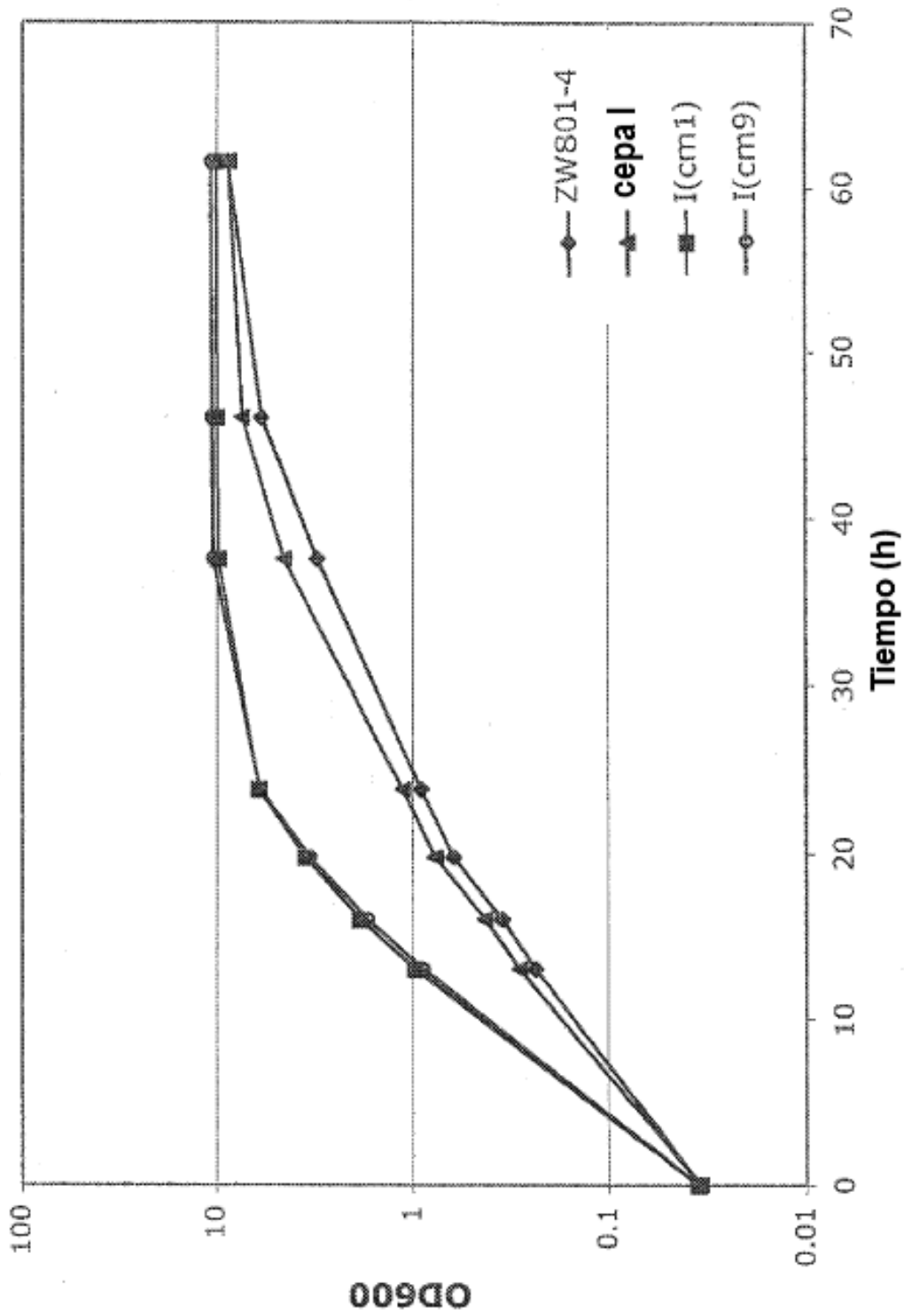


FIG. 5

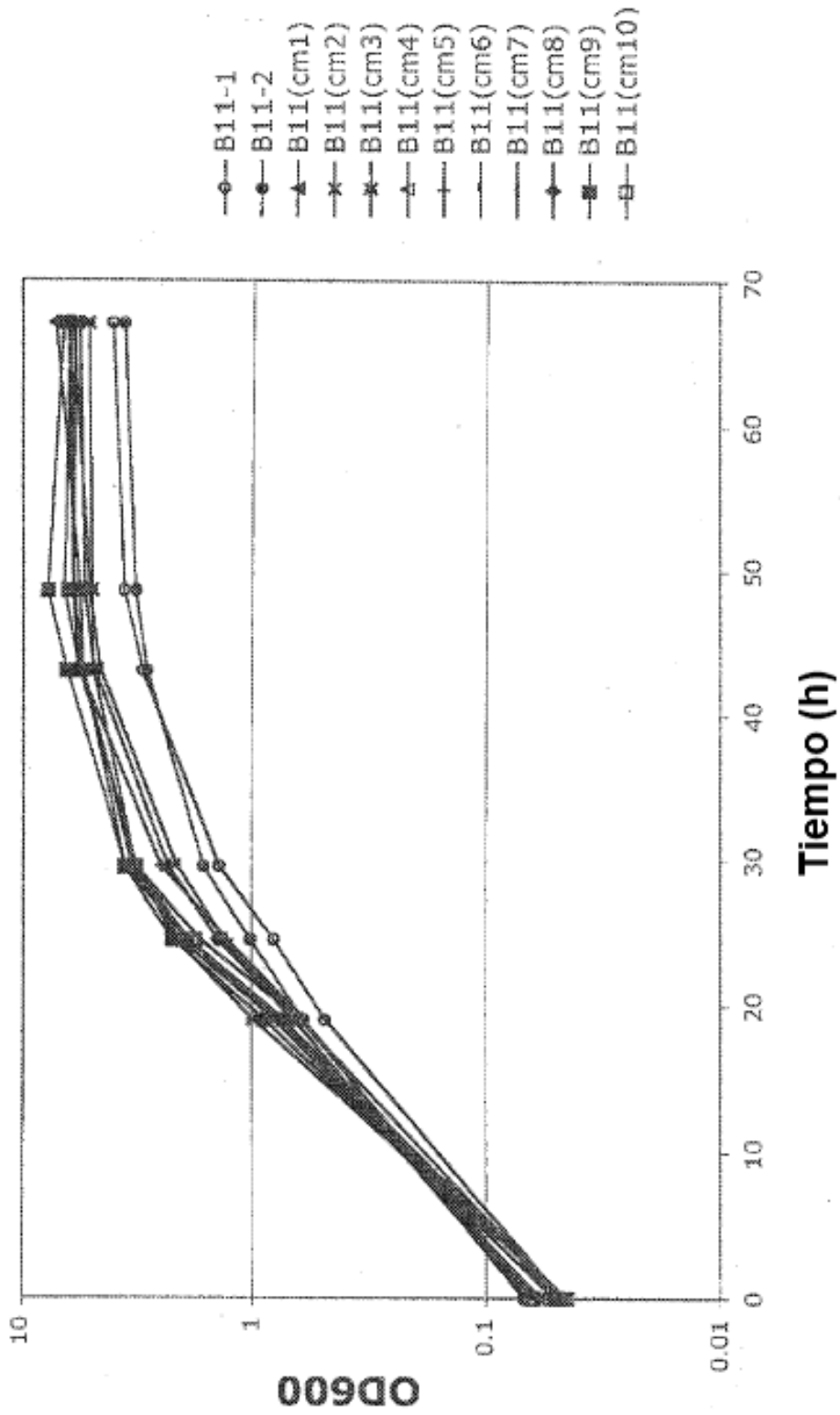


FIG. 6A

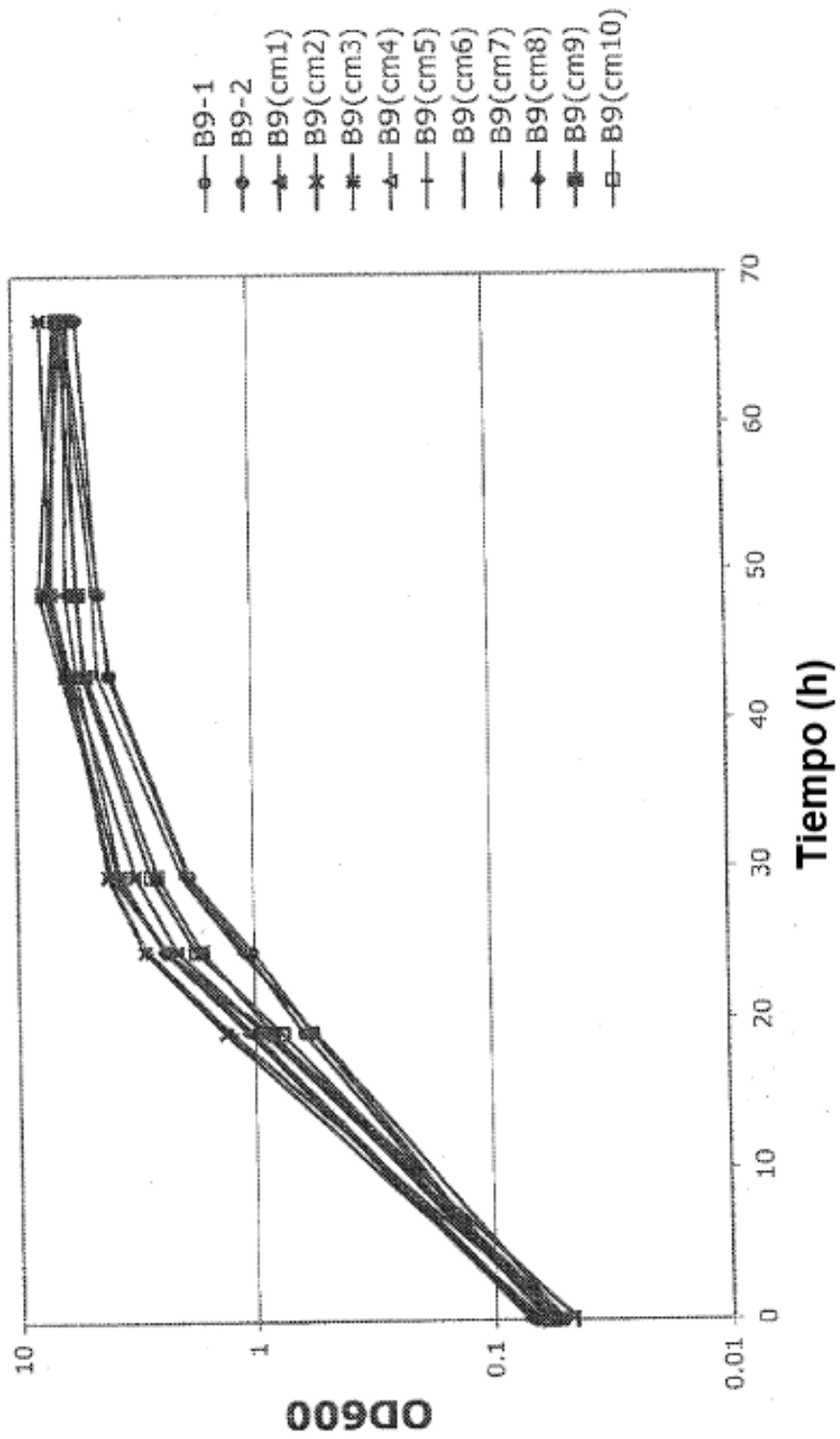


FIG. 6B

I Strain	(1)	MEDIKRQEI	DWGGKKL	ILETGO	VARQAD	GAVIATL	GETVVL	CAVTA	AAKTV	KEGQD	FFPL	TVHYQ	EKYSA	140
Wild Type	(1)	MEDIKRQEI	DWGGKKL	ILETGO	VARQAD	GAVIATL	GETVVL	CAVTA	AAKTV	KEGQD	FFPL	TVHYQ	EKYSA	140
I Strain	(71)	GRIPGGF	KRRER	GATER	ETLIS	RLLIDR	PIRPL	EPG	FYNET	FLVIA	QVMSY	DGENE	PDILAMIA	AASAALAL
Wild Type	(71)	GRIPGGF	KRRER	GATER	ETLIS	RLLIDR	PIRPL	EPG	FYNET	FLVIA	QVMSY	DGENE	PDILAMIA	AASAALAL
I Strain	(141)	SGVFFL	GPIGA	ARVGY	ODGE	FILN	PTLE	QLEK	SDLL	VGATR	DAVM	VESE	ANEL	PEEV
Wild Type	(141)	SGVFFL	GPIGA	ARVGY	ODGE	FILN	PTLE	QLEK	SDLL	VGATR	DAVM	VESE	ANEL	PEEV
I Strain	(211)	SLQPVI	KAIIN	LAEQ	AKEP	WELV	SYD	SSA	LA	AKVE	ELCY	DN	FDK	AYR
Wild Type	(211)	SLQPVI	KAIIN	LAEQ	AKEP	WELV	SYD	SSA	LA	AKVE	ELCY	DN	FDK	AYR
I Strain	(281)	PEADP	TEKLR	IQK	LAK	LEA	KIV	RTAIL	KE	RRID	GR	DLK	TVR	P
Wild Type	(281)	PEADP	TEKLR	IQK	LAK	LEA	KIV	RTAIL	KE	RRID	GR	DLK	TVR	P
I Strain	(351)	STILG	TADAE	Q	MID	GLT	GL	HYE	R	F	M	L	H	N
Wild Type	(351)	STILG	TADAE	Q	MID	GLT	GL	HYE	R	F	M	L	H	N
I Strain	(421)	YTRV	LS	D	I	T	E	S	N	G	S	S	M	A
Wild Type	(421)	YTRV	LS	D	I	T	E	S	N	G	S	S	M	A
I Strain	(491)	FKVAC	TEK	GI	T	S	L	Q	M	I	K	V	A	G
Wild Type	(491)	FKVAC	TEK	GI	T	S	L	Q	M	I	K	V	A	G
I Strain	(561)	DKIRD	VIG	T	G	G	K	V	I	R	E	I	V	A
Wild Type	(561)	DKIRD	VIG	T	G	G	K	V	I	R	E	I	V	A
I Strain	(631)	VDFG	AF	V	N	F	M	G	G	R	D	G	L	V
Wild Type	(631)	VDFG	AF	V	N	F	M	G	G	R	D	G	L	V
I Strain	(701)	NRPP	RE	NAE	P	V	S	T	H	L	N	P	E	A
Wild Type	(701)	NRPP	RE	NAE	P	V	S	T	H	L	N	P	E	A

FIG. 7

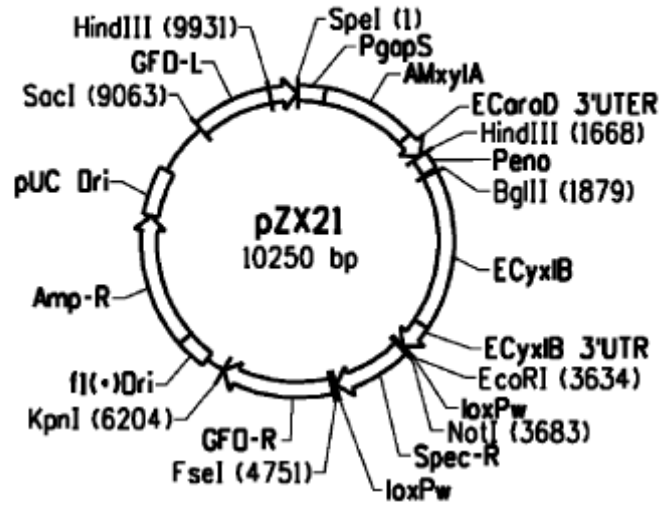


FIG. 8A

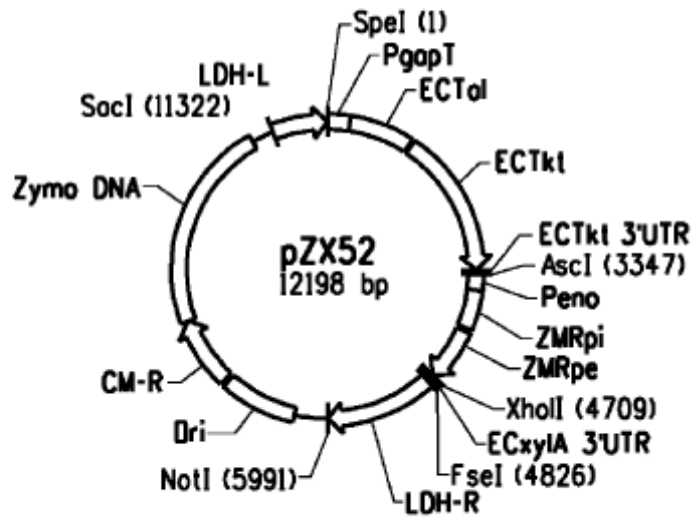


FIG. 8B

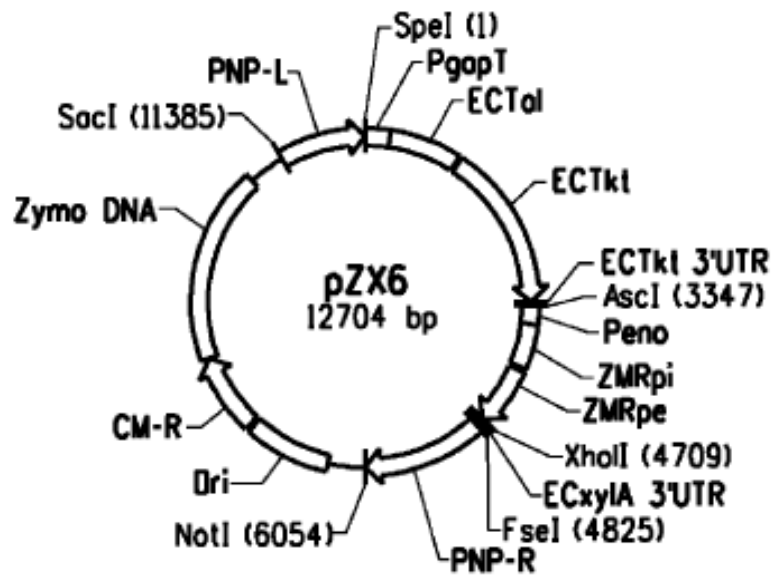


FIG. 8C

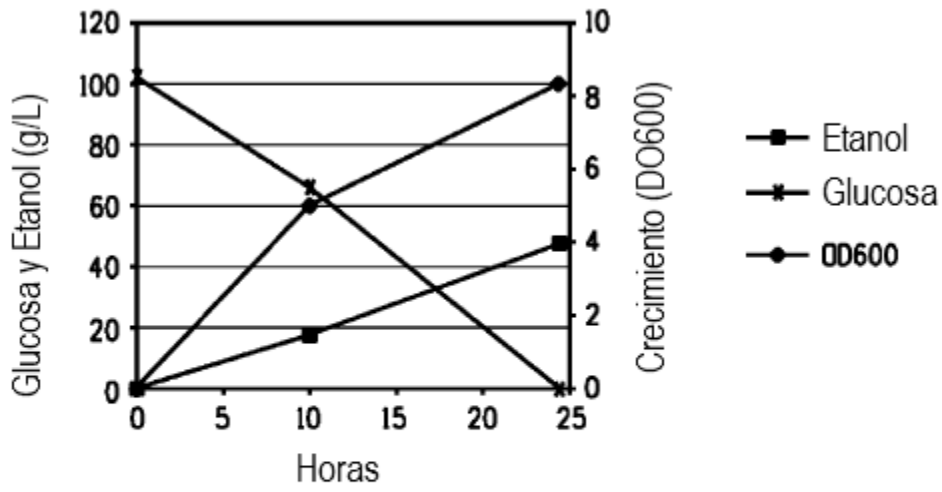


FIG. 9A

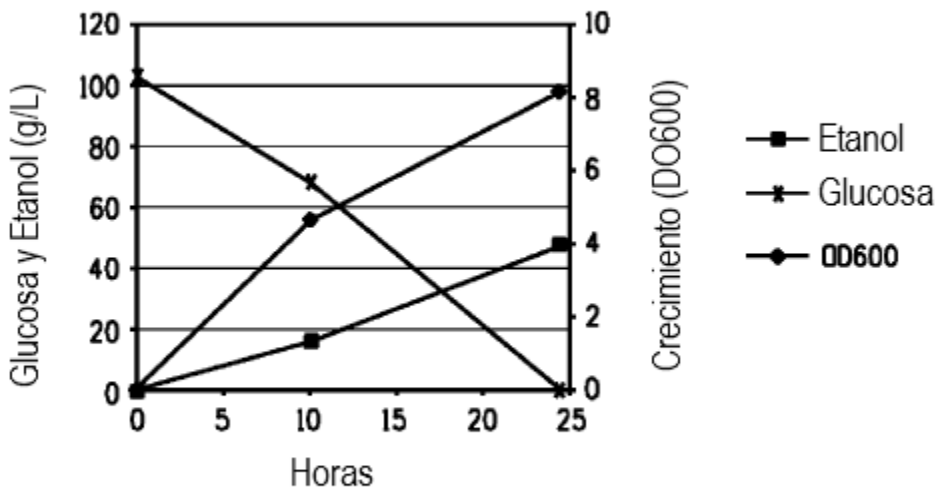


FIG. 9B

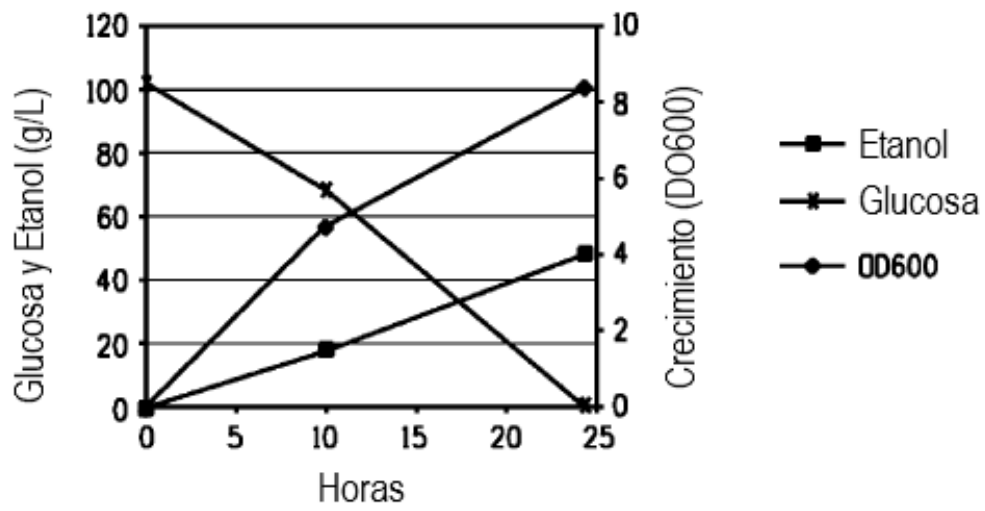


FIG. 9A

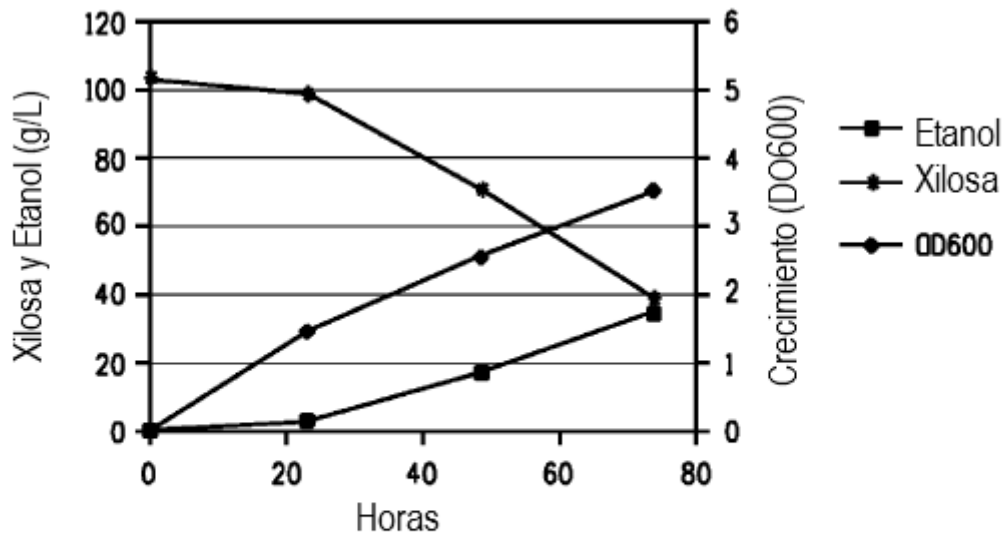


FIG. 10A

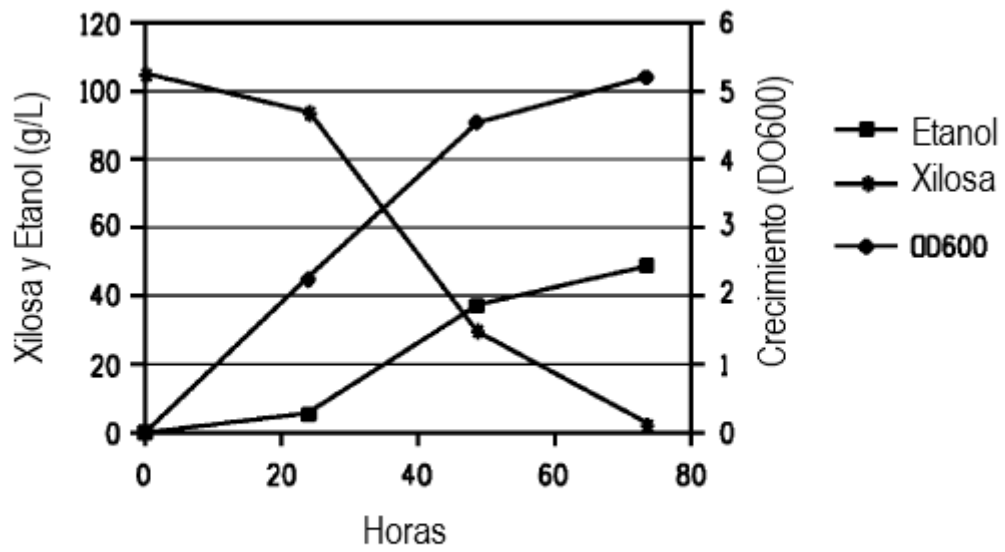


FIG. 10B

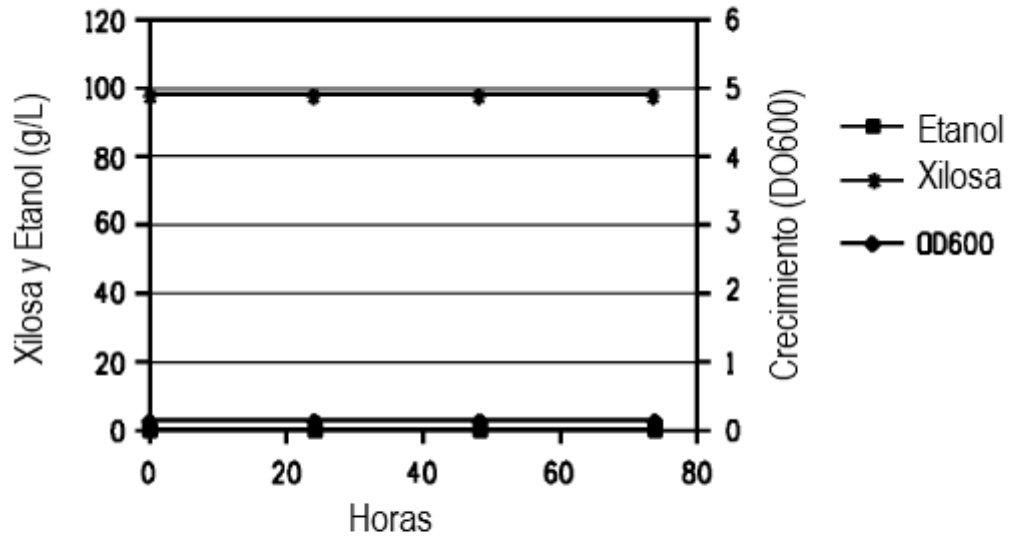


FIG. 10C

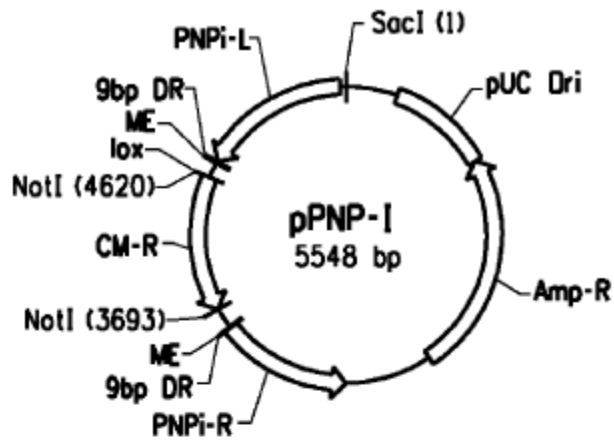


FIG. 11A

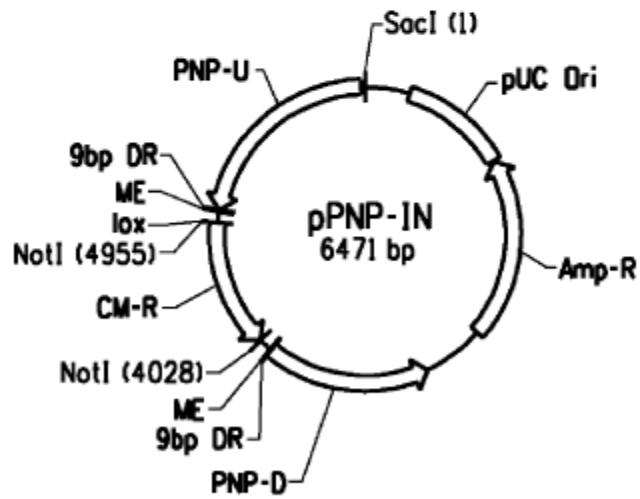


FIG. 11B

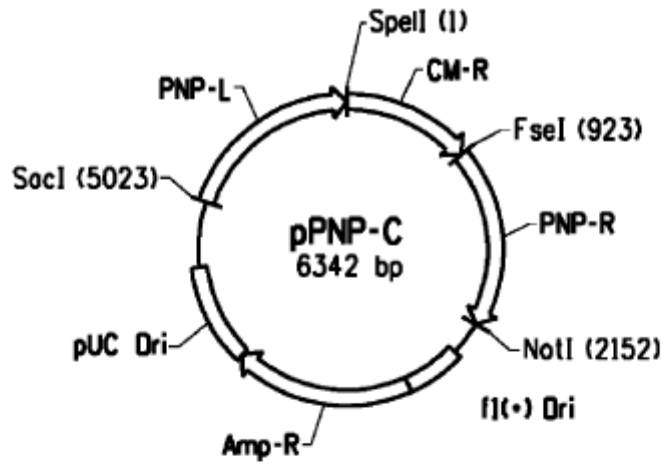


FIG. 11C

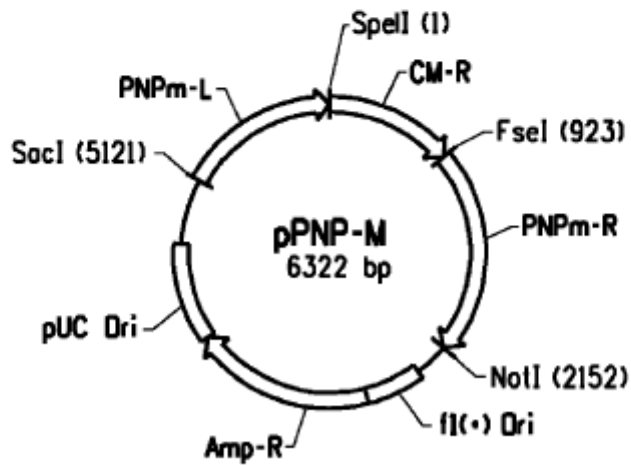


FIG. 11D

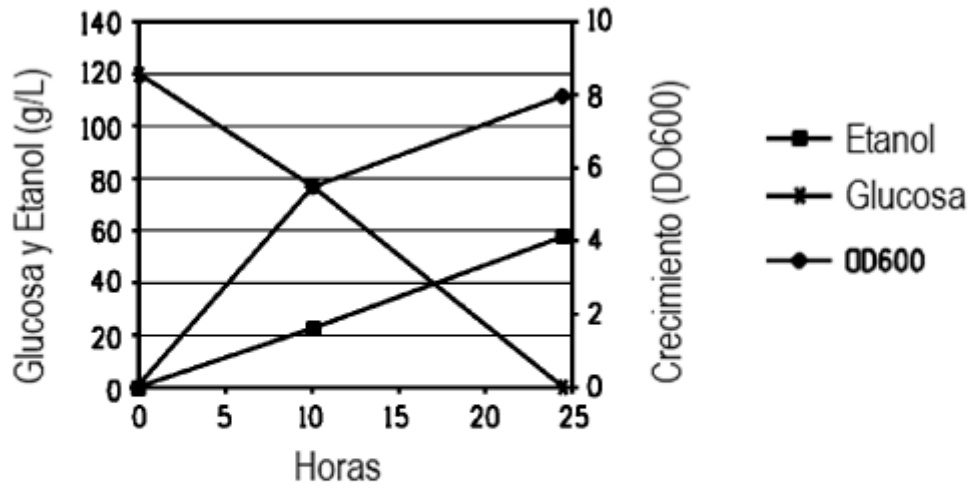


FIG. 12A

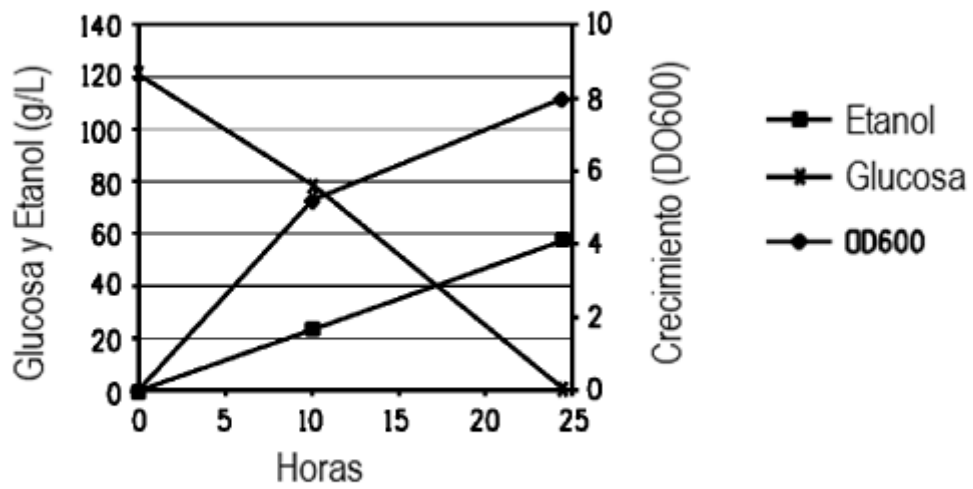


FIG. 12B

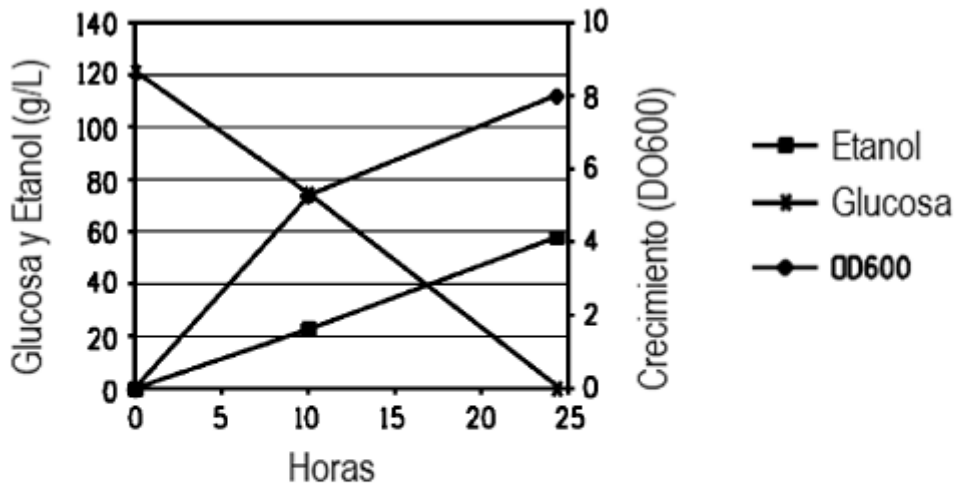


FIG. 12C

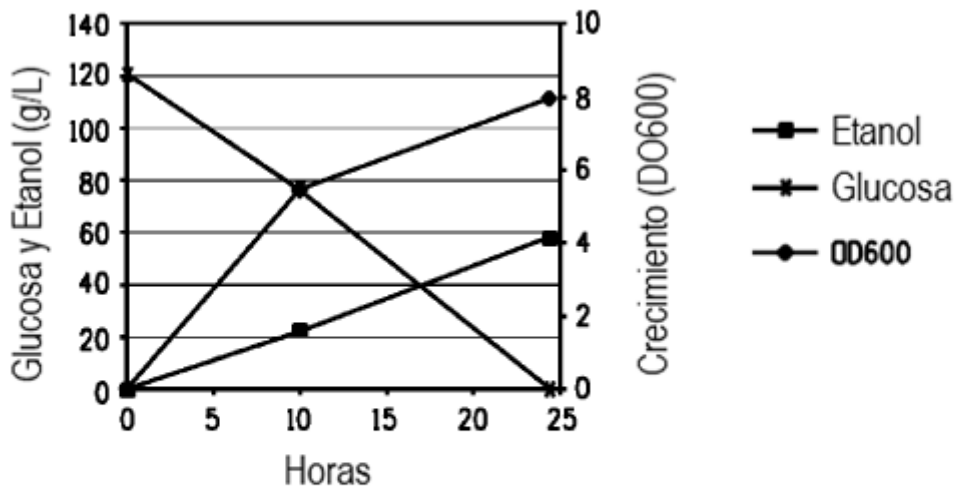


FIG. 12D

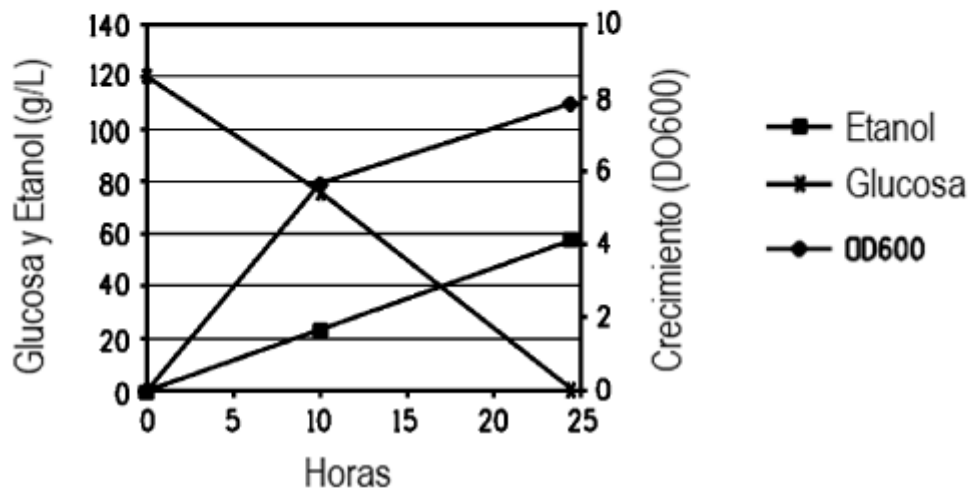


FIG. 12E

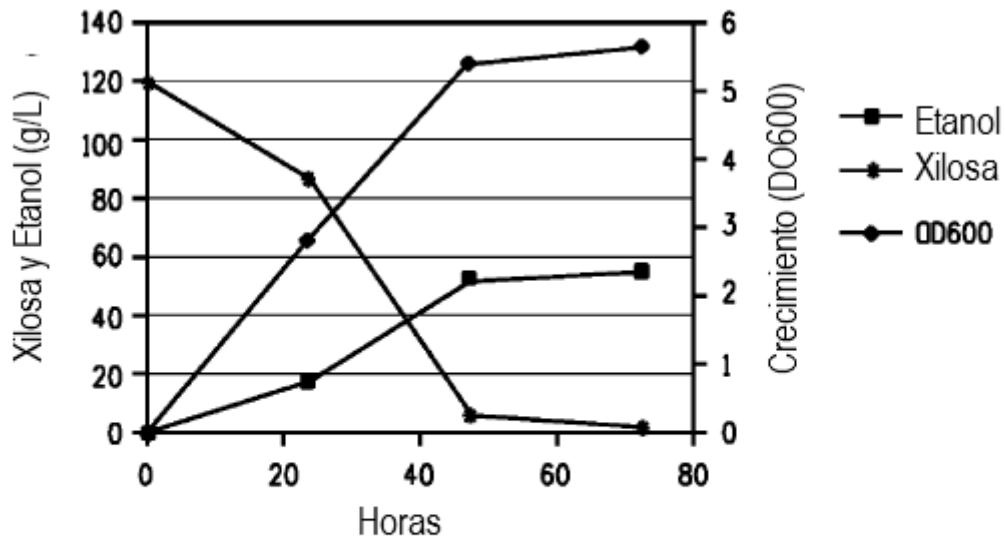


FIG. 13A

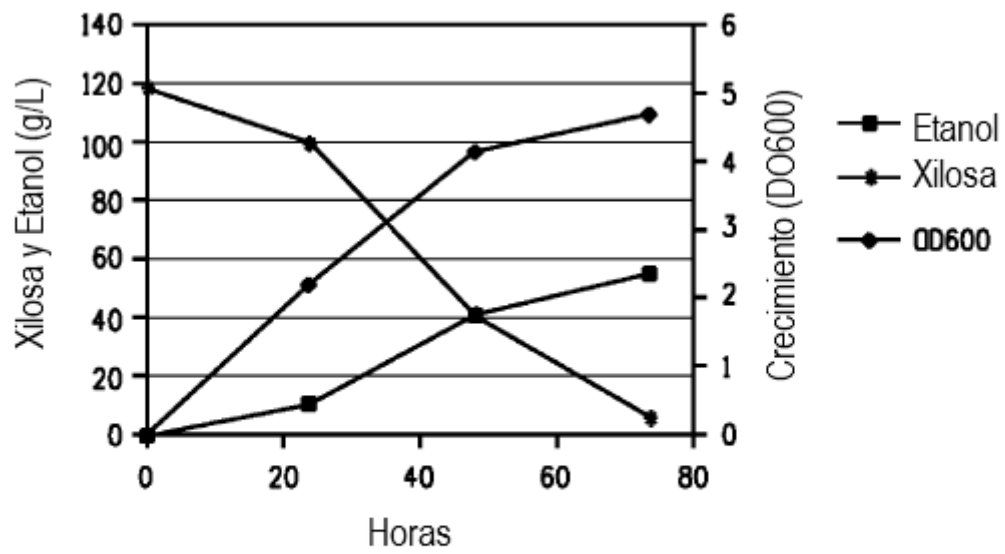


FIG. 13B

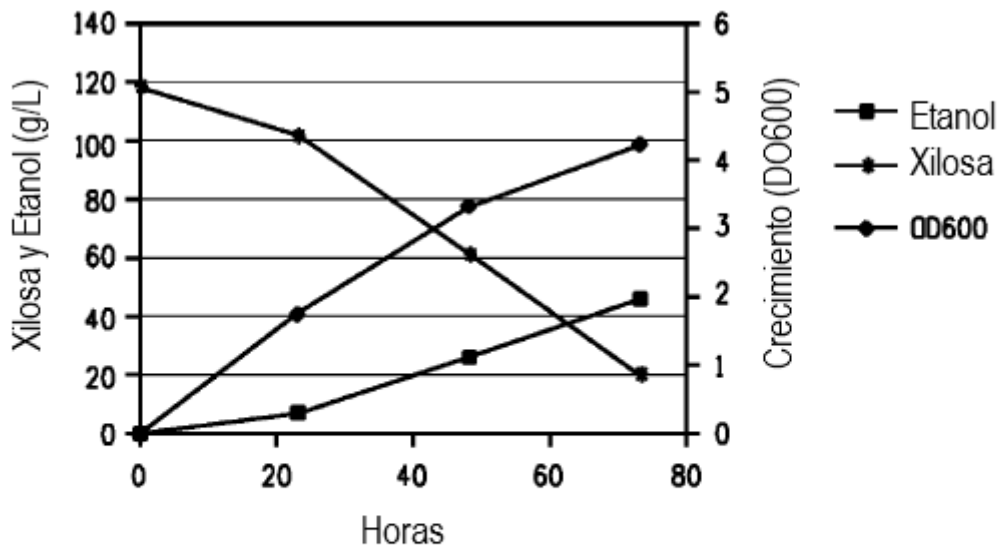


FIG. 13C

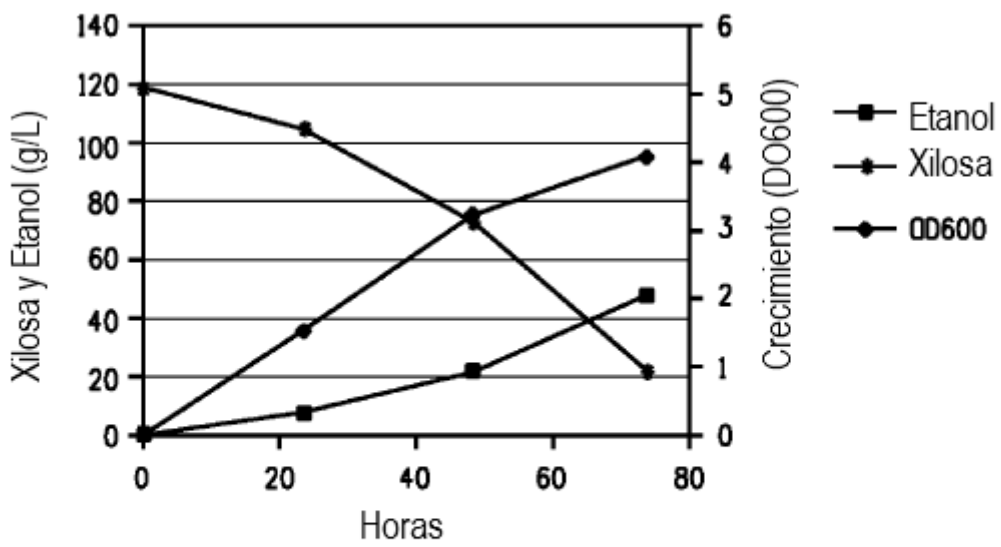


FIG. 13D

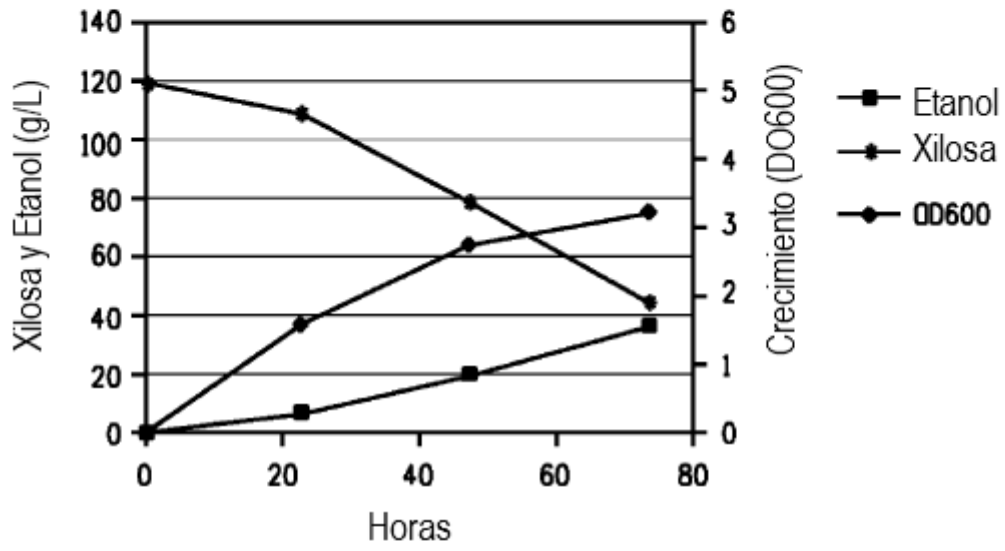


FIG. 13E

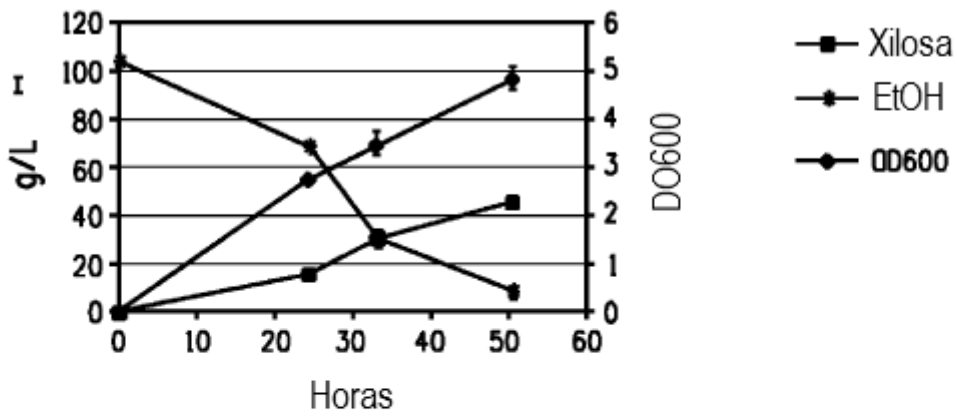


FIG. 14A

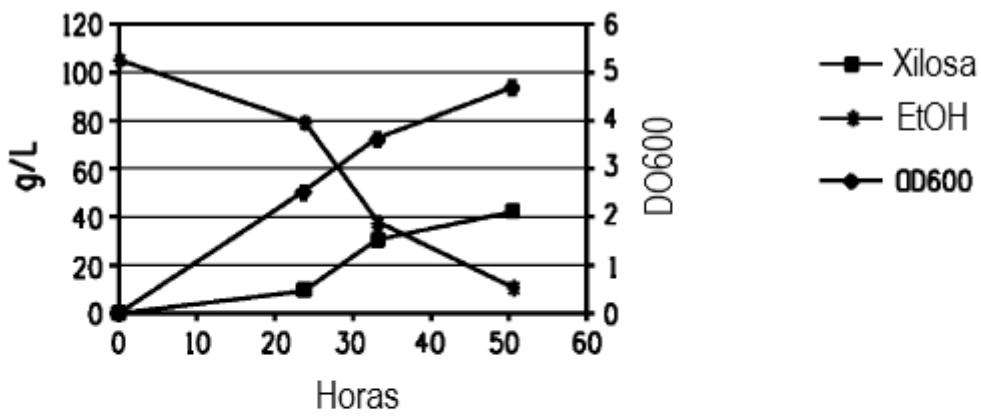


FIG. 14B

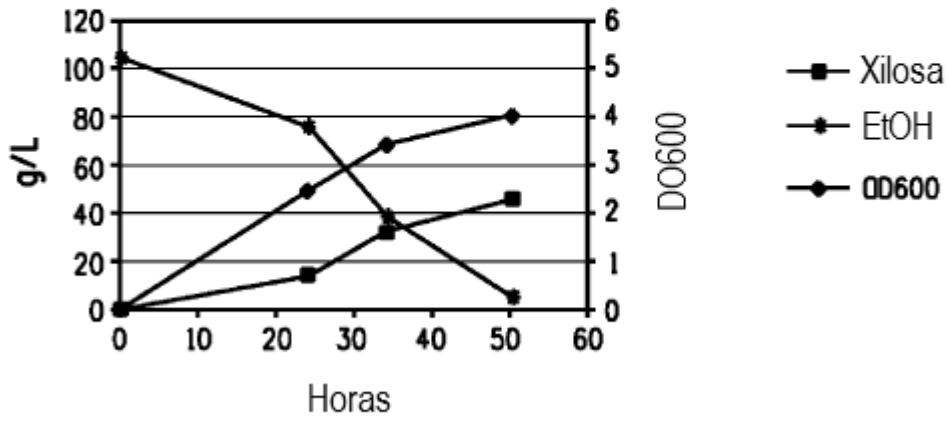


FIG. 14C