

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 599 579**

51 Int. Cl.:

C12N 1/21 (2006.01)

C12P 7/06 (2006.01)

C12P 21/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **10.06.2010 PCT/US2010/038121**

87 Fecha y número de publicación internacional: **23.12.2010 WO10147835**

96 Fecha de presentación y número de la solicitud europea: **10.06.2010 E 10789964 (3)**

97 Fecha y número de publicación de la concesión europea: **20.07.2016 EP 2443228**

54 Título: **Zymomonas con utilización de arabinosa mejorada**

30 Prioridad:

18.06.2009 US 218164 P
18.06.2009 US 218166 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
02.02.2017

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:

YANG, JIANJUN

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 599 579 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

Zymomonas con utilización de arabinosa mejorada

Declaración de derechos gubernamentales

- 5 La presente invención se realizó con el apoyo del gobierno de Estados Unidos con el n.º de Contrato DE-FC36-07GO17056 otorgado por el Departamento de Energía. El gobierno estadounidense tiene determinados derechos con respecto a la presente invención.

Campo de la invención

- 10 La invención refiere se refiere a los campos de la microbiología y la fermentación. Más específicamente, se describe el diseño genético de cepas de *Zymomonas* para conferir utilización de arabinosa mejorada, y métodos para elaborar etanol usando las cepas.

Antecedentes de la invención

- 15 La producción de etanol mediante microorganismos proporciona una fuente de energía alternativa a los combustibles fósiles y, por lo tanto, un área importante de investigación actual. Es conveniente que los microorganismos que producen etanol, así como otros productos útiles, puedan usar xilosa y arabinosa como fuentes de carbono ya que son azúcares pentosa predominantes en materiales lignocelulósicos hidrolizados, que pueden proporcionar una fuente de sustrato de carbono disponible en abundancia y de bajo costo para que los biocatalizadores lo usen en la fermentación.

- 20 *Zymomonas mobilis* y otros etanológenos bacterianos que no utilizan xilosa y arabinosa de forma natural pueden diseñarse genéticamente para que utilicen estos azúcares. Para proporcionar la utilización de xilosa, las cepas han sido diseñadas genéticamente para que expresen los genes que codifican las siguientes proteínas: 1) xilosa isomerasa, que cataliza la conversión de xilosa en xilulosa; 2) xilulocinasa, que fosforila xilulosa para formar xilulosa 5-fosfato; 3) transcetolasa; y 4) transaldolasa (US 5514583, US 6566107; Zhang et al. (1995) Science 267:240-243). Para proporcionar la utilización de arabinosa, se han introducido genes adicionales que codifican las siguientes proteínas: 1) L-arabinosa isomerasa para convertir L-arabinosa en L-ribulosa, 2) L-ribulocinasa para convertir L-ribulosa en L-ribulosa-5-fosfato, y 3) L-ribulosa-5-fosfato-4-epimerasa para convertir L-ribulosa-5-fosfato en D-xilulosa (US5843760).

- 30 Aunque algunas cepas de *Z mobilis* se han diseñado genéticamente para la utilización de arabinosa, típicamente solo un bajo porcentaje de arabinosa presente en un medio de fermentación es utilizado por estas cepas diseñadas genéticamente. Todavía existe una necesidad de mejorar la utilización de arabinosa en *Zymomonas* y otros etanológenos bacterianos para mejorar la producción de etanol cuando la fermentación se da en un medio que contiene arabinosa.

Compendio de la invención

- 35 La presente invención se refiere a cepas de *Zymomonas* y *Zymobacter* que están diseñadas genéticamente para que presenten una capacidad mejorada para utilizar arabinosa mediante la introducción de un gen para la expresión de un simportador de protones de arabinosa, y para producir etanol usando estas cepas. Estas cepas presentan una producción mejorada de etanol cuando se cultivan en un medio que contiene arabinosa.

Por consiguiente, la invención proporciona un microorganismo recombinante del género *Zymomonas* o *Zymobacter* que utiliza arabinosa para producir etanol, dicho microorganismo comprende al menos un gen heterólogo que codifica un simportador de protones de arabinosa, donde dicho simportador es expresado por dicho microorganismo.

- 40 Además, la invención proporciona un proceso para generar un microorganismo recombinante del género *Zymomonas* o *Zymobacter* que tiene mayor utilización de arabinosa que comprende:

a) proporcionar una cepa recombinante de *Zymomonas* o *Zymobacter* que utilice arabinosa para producir etanol en condiciones adecuadas; y

- 45 b) introducir al menos un gen que codifique un simportador de protones de arabinosa heterólogo en la cepa de (a) para la expresión de dicho simportador.

En otra realización, la invención proporciona un proceso para producir etanol que comprende:

a) proporcionar una cepa recombinante de *Zymomonas* o *Zymobacter* que utiliza arabinosa para producir etanol, dicha cepa comprende al menos un gen heterólogo que codifica un simportador de protones de arabinosa, donde dicho simportador es expresado por dicho microorganismo;

- 50 b) cultivar la cepa de (a) en un medio que comprende arabinosa a través del cual la arabinosa es convertida por dicha cepa en etanol.

En otra realización, la invención proporciona un método para mejorar la utilización de arabinosa mediante un microorganismo que utiliza arabinosa que comprende:

(a) proporcionar un microorganismo que utiliza arabinosa, donde dicho microorganismo se selecciona del grupo que consiste en una cepa recombinante de *Zymomonas* o *Zymobacter* que utiliza arabinosa para producir etanol;

5 (b) introducir en el genoma de dicho microorganismo al menos un gen heterólogo que codifica un simportador de protones de arabinosa, donde dicho simportador es expresado por dicho microorganismo; y

(c) poner en contacto el microorganismo de (b) con un medio que comprende arabinosa, donde dicho microorganismo metaboliza dicha arabinosa con una mayor velocidad en comparación con dicho microorganismo que no presenta el simportador de protones de arabinosa.

10 Breve descripción de las figuras y descripciones de secuencias

La invención se entenderá mejor a partir de la siguiente descripción detallada, las Figuras y las descripciones de secuencias adjuntas que forman parte de la presente solicitud.

La Figura 1 muestra un diagrama de la vía de fermentación de etanol en *Zymomonas* diseñada genéticamente para la utilización de xilosa y arabinosa, donde glf significa transportador de difusión facilitado por glucosa.

15 La Figura 2 es un dibujo de un mapa de plásmidos de pARA205.

La Figura 3 es un dibujo de un mapa de plásmidos de pARA354.

La Figura 4 muestra gráficas de perfiles de crecimiento y metabolitos de ZW705 (A), ZW705-ara354 (B), y ZW705-ara354A7 (C) en MRM3A5 durante un transcurso de tiempo de 96 horas.

20 La Figura 5 muestra gráficas de perfiles de crecimiento y metabolitos de ZW705 (A), ZW705-ara354 (B), y ZW705-ara354A7 (C) en MRM3A2.5X2.5G5 durante un transcurso de tiempo de 96 horas.

La Figura 6 es un dibujo de un mapa de plásmidos de pARA112.

La Figura 7 es un dibujo de un mapa de plásmidos de pARA113.

La Figura 8 muestra gráficas de perfiles de crecimiento y metabolitos de ZW705-ara354A7 (A), ZW705-ara354A7-ara112-2 (B), y ZW705-ara354A7-ara112-3 (C) en MRM3A5 durante un transcurso de tiempo de 96 horas.

25 La Figura 9 muestra gráficas de perfiles de crecimiento y metabolitos de ZW705-ara354A7 (A), ZW705-ara354A7-ara112-2 (B), y ZW705-ara354A7-ara112-3 (C) en MRM3A2.5X2.5G5 durante un transcurso de tiempo de 96 horas.

La Figura 10 muestra gráficas de perfiles de crecimiento y metabolitos de ZW705-ara354 (A), ZW705-ara354-ara112-1 (B), y ZW705-ara354-ara112-2 (C) en MRM3A5 durante un transcurso de tiempo de 96 horas.

30 La Figura 11 muestra gráficas de perfiles de crecimiento y metabolitos de ZW705-ara354 (A), ZW705-ara354-ara112-1 (B), y ZW705-ara354-ara112-2 (C) en MRM3A2.5X2.5G5 durante un transcurso de tiempo de 96 horas.

La Figura 12 muestra gráficas de perfiles de crecimiento y metabolitos de ZW801-ara354 (A), ZW801-ara354-ara112-5 (B), y ZW801-ara354-ara112-6 (C) en MRM3A5 durante un transcurso de tiempo de 96 horas.

La Figura 13 muestra gráficas de perfiles de crecimiento y metabolitos de ZW801-ara354 (A), ZW801-ara354-ara112-5 (B), y ZW801-ara354-ara112-6 (C) en MRM3A2.5X2.5G5 durante un transcurso de tiempo de 96 horas.

35 Las siguientes secuencias cumplen con los artículos 1.821-1.825 del Título 37 del C.F.R. (Código de Regulaciones Federales) ("Requisitos para las solicitudes de patente que contienen descripciones de secuencias de nucleótidos y/o secuencia de aminoácidos - las regulaciones de secuencia") y son consistentes con la Norma ST.25 (1998) Organización Mundial de la Propiedad Intelectual (OMPI) y los requisitos del listado de secuencias de la EPO y PCT (Normas 5.2 y 49.5(a-bis), y Sección 208 y Anexo C de las Instrucciones Administrativas). Los símbolos y formato
40 utilizado para los datos de secuencias de aminoácidos y nucleótidos cumplen con las regulaciones establecidas en el artículo 1.822 del Título 37 del C.F.R.

Tabla 1. SEQ ID NO de región codificante y de proteína para los simportadores de protones de arabinosa codificados por *araE*

Organismo	Región codificante de SEQ ID NO:	Péptido de SEQ ID NO:
<i>E. coli</i>	1	2
<i>Shigella flexneri</i>	3	4

Organismo	Región codificante de SEQ ID NO:	Péptido de SEQ ID NO:
<i>Shigella boydii</i>	5	6
<i>Shigella dysenteriae</i>	7	8
<i>Salmonella typhimurium</i>	9	10
<i>Salmonella enterica</i>	11	12
<i>Klebsiella pneumoniae</i>	13	14
<i>Klebsiella oxytoca</i>	15	16
<i>Enterobacter cancerogenus</i>	17	18
<i>Bacillus amyloliquefaciens</i>	19	20

Las SEQ ID NO:21 y 22 son la secuencia de aminoácidos y región codificante, respectivamente, para el gen araA de *E. coli*.

Las SEQ ID NO:23 y 24 son la secuencia de aminoácidos y región codificante, respectivamente, para el gen araB de *E. coli*.

- 5 Las SEQ ID NO:25 y 26 son la secuencia de aminoácidos y región codificante, respectivamente, para el gen araD de *E. coli*.

La SEQ ID NO:27 es la secuencia de nucleótidos del producto de PCR del fragmento de ADN de araB-araA.

Las SEQ ID NO:28 y 29 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR del fragmento de ADN de araB-araA.

- 10 La SEQ ID NO:30 es la secuencia de nucleótidos del producto de PCR del fragmento de ADN de araD, que incluye RBS y 3' UTR.

Las SEQ ID NO:31 y 32 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR del fragmento de ADN de araD, que incluye RBS y 3' UTR.

La SEQ ID NO:33 es la secuencia de nucleótidos del promotor Pgap de *Z. mobilis*.

- 15 Las SEQ ID NO:34 y 35 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR del fragmento del promotor Pgap.

La SEQ ID NO:36 es la secuencia de nucleótidos del producto de PCR del fragmento de ADN del promotor Pgap.

Las SEQ ID NO:37 y 38 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR del cassette de resistencia a espectinomicina.

- 20 Las SEQ ID NO:39 y 40 son las secuencias de nucleótidos de los cebadores para la mutagénesis de Pgap para retirar el sitio de NcoI agregado.

La SEQ ID NO:41 es la secuencia de nucleótidos del plásmido pARA205.

Las SEQ ID NO:42 y 43 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR del fragmento de ADN de LDH-L.

- 25 La SEQ ID NO:44 es la secuencia de nucleótidos del producto de PCR del fragmento de ADN de LDH-L.

Las SEQ ID NO:45 y 46 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR del fragmento de ADN de LDH-R.

La SEQ ID NO:47 es la secuencia de nucleótidos del producto de PCR del fragmento de ADN de LDH-R.

- 30 La SEQ ID NO:48 es la secuencia de nucleótidos del producto de PCR del fragmento de ADN de LoxPw-aadA-LoxPw.

La SEQ ID NO:49 es la secuencia de nucleótidos del plásmido pARA354.

Las SEQ ID NO:50 y 51 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR para verificar la integración en 5' de P_{gap}-araBAD-aadA.

Las SEQ ID NO:52 y 53 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR para verificar la integración en 3' de P_{gap}-araBAD-aadA.

Las SEQ ID NO:54 y 55 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR del fragmento de ADN de la región codificante de araE.

- 5 La SEQ ID NO:56 es la secuencia de nucleótidos del producto de PCR del fragmento de ADN de araE.

Las SEQ ID NO:57 y 58 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR del fragmento de ADN de araFGH.

La SEQ ID NO:59 es la secuencia de nucleótidos del producto de PCR del fragmento de ADN de araFGH.

- 10 Las SEQ ID NO:60 y 61 son las secuencias de nucleótidos de los cebadores para la amplificación de PCR del fragmento de ADN de *Actinoplanes missouriensis* P_{gi}.

La SEQ ID NO:62 es la secuencia de nucleótidos del promotor GI de *Actinoplanes missouriensis* en el plásmido usado como la plantilla de PCR.

La SEQ ID NO:63 es la secuencia de nucleótidos del producto de PCR del fragmento de ADN de *Actinoplanes missouriensis* P_{gi}.

- 15 La SEQ ID NO:64 es la secuencia de nucleótidos del marcador de resistencia a cloranfenicol.

La SEQ ID NO:65 es la secuencia de nucleótidos del plásmido pARA112.

La SEQ ID NO:66 es la secuencia de nucleótidos del plásmido pARA113.

Descripción detallada

- 20 La presente invención describe cepas de *Zymomonas* o *Zymobacter* recombinantes con utilización de arabinosa mejorada que se diseñan genéticamente de manera adicional para que expresen un simportador de protones de arabinosa, y un proceso para diseñar genéticamente las cepas mediante la introducción de un gen que codifica un simportador de protones de arabinosa. En otros aspectos, la presente invención describe procesos para mejorar la utilización de arabinosa y para producir etanol en medio que comprende arabinosa, usando dichas cepas. Las cepas que utilizan arabinosa que expresan un simportador de protones de arabinosa tienen utilización de arabinosa mejorada y son útiles para producir etanol en medio que comprende arabinosa.
- 25

El etanol producido por las cepas de la presente con utilización de arabinosa mejorada se puede utilizar como una fuente de energía alternativa a los combustibles fósiles.

Las siguientes abreviaturas y definiciones se utilizarán para la interpretación de la memoria descriptiva y las reivindicaciones.

- 30 Como se emplea en esta memoria, se pretende que los términos “comprende”, “que comprende”, “incluye”, “que incluye”, “tiene”, “que tiene”, “contiene” o “que contiene” o cualquier otra variación de estos, cubran una inclusión no exclusiva. Por ejemplo, una composición, una mezcla, un proceso, un método, un artículo o un aparato que comprende una lista de elementos, no se encuentra necesariamente limitado a esos elementos únicamente sino que puede incluir otros elementos que no se enumeran expresamente o que son inherentes a dicha composición, mezcla, proceso, método, artículo o aparato. Adicionalmente, a menos que se indique lo contrario, “o” hace referencia a un o inclusivo y no a un o exclusivo. Por ejemplo, una condición A o B se cumple 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 ambos A y B son verdaderos (o presentes).
- 35

- 40 También, los artículos indefinidos “un” y “una” que preceden a un elemento o componente de la invención no pretenden ser restrictivos en cuanto a la cantidad de instancias (es decir, apariciones) del elemento o componente. Por lo tanto, debe interpretarse que “un” o “una” incluyen al menos uno y la forma en singular del elemento o componente también incluye el plural a menos que la cantidad indique obviamente el singular.

- 45 “Gen” hace referencia a un fragmento de ácido nucleico que expresa una proteína específica, que puede incluir secuencias reguladoras que preceden (secuencias no codificantes 5') y que siguen (secuencias no codificantes 3') a la secuencia codificante. “Gen natural” o “gen de tipo salvaje” hace referencia a un gen tal como se encuentra naturalmente con sus propias secuencias reguladoras. “Gen quimérico” hace referencia a cualquier gen que no es un gen natural, que comprende secuencias reguladoras y codificantes que no se encuentran juntas de forma natural. Por consiguiente, un gen quimérico puede comprender secuencias reguladoras y secuencias codificantes derivadas de diferentes fuentes o secuencias reguladoras y secuencias codificantes derivadas de la misma fuente, pero organizadas de manera diferente a la que se encuentran de forma natural. “Gen endógeno” se refiere a un gen natural en su ubicación natural dentro del genoma de un organismo. Un gen “foráneo” hace referencia a un gen no encontrado normalmente en el organismo hospedador, pero que se introduce en el organismo hospedador mediante
- 50

transferencia génica. Los genes foráneos comprenden genes naturales insertados en un organismo no natural o genes quiméricos.

El término “*araE*” se refiere a un gen o construcción genética que codifica una proteína simportadora de protones de arabinosa bacteriana que es una transportadora de arabinosa de baja afinidad y alta capacidad con una K_m de $1,25 \times 10^{-4}$ M. Los genes que codifican la proteína simportadora de protones de arabinosa pueden aislarse de múltiples bacterias y aquellos de bacterias entéricas, tal como *Escherichia*, *Klebsiella*, *Salmonella* y *Shigella* son particularmente útiles en la presente invención.

El término “utilización de arabinosa” cuando se utiliza en el contexto de un microorganismo se refiere a la capacidad que presente ese microorganismo de utilizar arabinosa para la producción de productos, particularmente etanol.

El término “cepa adaptada” se refiere a un microorganismo que se ha seleccionado para cultivarse en una fuente de carbono particular con el fin de mejorar su capacidad de usar esa fuente de carbono para la producción de productos. Una “cepa adaptada de arabinosa” por ejemplo, es una cepa de un microorganismo que se ha seleccionado para cultivarse en altas concentraciones de arabinosa.

El término “construcción genética” se refiere a un fragmento de ácido nucleico que codifica para la expresión de una o más proteínas específicas. En la construcción genética el gen puede ser natural, quimérico o foráneo. Típicamente, una construcción genética comprenderá una “secuencia codificante”. Una “secuencia codificante” hace referencia a una secuencia de ADN que codifica una secuencia de aminoácidos específica.

“Promotor” o “regiones de control de iniciación” se refiere a una secuencia de ADN capaz de controlar la expresión de una secuencia codificante o ARN funcional. En general, una secuencia codificante se ubica 3' respecto de una secuencia promotora. Los promotores pueden derivar en su totalidad de un gen natural, o estar compuestos por diferentes elementos derivados de diferentes promotores que se encuentran en la naturaleza, o incluso comprender segmentos de ADN sintéticos. Los expertos en la técnica entienden que diferentes promotores pueden dirigir la expresión de un gen en diferentes tipos de células o tejidos, o en diferentes etapas de desarrollo, o en respuesta a diferentes condiciones ambientales. Los promotores que provocan que un gen se exprese en la mayoría de los tipos de células en la mayoría de los casos, se denominan comúnmente “promotores constitutivos”.

El término “expresión” como se emplea en esta memoria, se refiere a la transcripción y acumulación estable de ARN sentido (ARNm) o antisentido derivado de un gen. Expresión también puede hacer referencia a la traducción de ARNm en un polipéptido. “Inhibición antisentido” se refiere a la producción de transcripciones de ARN antisentido capaces de suprimir la expresión de la proteína objetivo. “Sobreexpresión” hace referencia a la producción de un producto génico en organismos transgénicos que excede los niveles de producción en organismos normales o no transformados. “Supresión conjunta” se refiere a la producción de fragmentos o transcripciones de ARN sentido capaces de suprimir la expresión de genes endógenos o foráneos idénticos o sustancialmente similares (U.S. 5,231,020).

El término “transformación”, como se emplea en esta memoria, se refiere a la transferencia de un fragmento de ácido nucleico en un organismo hospedador que resulta en herencia genéticamente estable. El ácido nucleico transferido puede encontrarse en forma de un plásmido mantenido en la célula hospedadora, o algún ácido nucleico transferido puede integrarse al genoma de la célula hospedadora. Los organismos hospedadores que contienen los fragmentos de ácido nucleico transformados se denominan organismos “transgénicos” o “recombinantes” o “transformados”.

Los términos “plásmido” y “vector”, como se emplea en esta memoria, se refieren a un elemento extracromosómico que a menudo porta genes que no son parte del metabolismo central de la célula, y normalmente tiene forma de moléculas de ADN de cadena doble circular. Dichos elementos pueden ser secuencias que se replican autónomamente, secuencias que integran genomas, secuencias de nucleótidos o fagos, lineales o circulares, de un ADN o ARN de cadena doble o simple, derivadas de cualquier fuente, donde se unió o recombinó una cantidad de secuencias de nucleótidos en una única construcción que es capaz de introducir un fragmento de promotor y secuencia de ADN para un producto génico seleccionado junto con la secuencia no traducida 3' adecuada en una célula.

El término “unido de forma operativa” se refiere a la asociación de secuencias de ácido nucleico en un único fragmento de ácido nucleico de forma que la función de uno se vea afectado por el otro. Por ejemplo, un promotor está unido de forma operativa a una secuencia codificante cuando es capaz de afectar la expresión de esa secuencia codificante (es decir, que la secuencia codificante está bajo el control transcripcional del promotor). Las secuencias codificantes pueden estar unidas de forma operativa a las secuencias reguladoras en orientación sentido o antisentido.

El término “marcador seleccionable” significa un factor identificador, a menudo un gen de resistencia química o antibiótico, que puede seleccionarse en función del efecto del gen marcador, es decir, la resistencia a un antibiótico, donde el efecto se utiliza para monitorear la herencia de un ácido nucleico de interés y/o para identificar una célula u organismo que ha heredado el ácido nucleico de interés.

Como se emplea en esta memoria, el término “degeneración de codones” hace referencia a la naturaleza en el código genético que permite la variación de la secuencia de nucleótidos sin afectar la secuencia de aminoácidos de un polipéptido codificado. El experto en la técnica conoce la “preferencia codónica” exhibida por una célula hospedadora específica en el uso de codones de nucleótidos para especificar un aminoácido dado. Por lo tanto,

5 cuando se sintetiza un gen para mejorar la expresión en una célula hospedadora, es conveniente diseñar el gen de modo que su frecuencia del uso de codones se aproxime a la frecuencia del uso de codones preferida de la célula hospedadora.

El término “optimizado por codones” cuando se refiere a genes o regiones codificantes de moléculas de ácido nucleico para la transformación de varios hospedadores, hace referencia a la alteración de codones en el gen o

10 regiones codificantes de las moléculas de ácido nucleico para que reflejen el uso de codones típico del organismo hospedador sin alterar el polipéptido codificado por el ADN.

El término “fuente de carbono” se refiere a azúcares tales como oligosacáridos y monosacáridos que se pueden utilizar por un microorganismo en un proceso de fermentación (“azúcar fermentable”) para producir un producto tal como etanol. Un microorganismo puede tener la capacidad de usar una sola fuente de carbono para la producción

15 de un producto y, como tal, la fuente de carbono se denomina en la presente memoria fuente de carbono “única”.

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

El término “celulósico” hace referencia a una composición que comprende celulosa y componentes adicionales, que incluyen hemicelulosa.

20 El término “sacarificación” se refiere a la producción de azúcares fermentables o fuentes de carbono de polisacáridos.

El término “biomasa pretratada” significa biomasa que se ha sometido a pretratamiento antes de la sacarificación.

“Biomasa” se refiere a cualquier material celulósico o lignocelulósico e incluye materiales que comprenden celulosa, y opcionalmente comprenden además hemicelulosa, lignina, almidón, oligosacáridos y/o monosacáridos. La biomasa

25 también puede comprender componentes adicionales, tal como proteína y/o lípido. La biomasa puede derivar 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 forraje de maíz, o una mezcla de pasto y hojas. La biomasa incluye, pero no se limita a, cultivos de bioenergía, residuos agrícolas, desechos sólidos municipales, desechos sólidos industriales, aguas residuales de la fabricación de papel, desechos de jardín, desechos de madera y

30 forestación. Los ejemplos de biomasa incluyen, pero no se limitan a, mazorcas de maíz, residuos de cultivos tal como cáscaras de maíz, forraje de maíz, hierbas, trigo, paja de trigo, paja de cebada, heno, paja de arroz, pasto varilla, desechos de papel, bagazo de caña de azúcar, bagazo o forraje de sorgo, forraje de soja, componentes obtenidos a partir de la molienda de granos, árboles, ramas, raíces, hojas, virutas de madera, aserrín, arbustos y matorrales, vegetales, frutas, flores y estiércol de animales.

35 “Hidrolizado de biomasa” se refiere al producto resultante de la sacarificación de la biomasa. La biomasa también se puede pretratar o preprocesar antes de la sacarificación.

El término “heterólogo” significa no encontrado de forma natural en la ubicación de interés. Por ejemplo, un gen heterólogo hace referencia a un gen que no se encuentra naturalmente en el organismo hospedador, pero que se introduce en el organismo hospedador mediante transferencia génica. Por ejemplo, una molécula de ácido nucleico

40 heteróloga que está presente en un gen quimérico es una molécula de ácido nucleico que no se encuentra asociada de forma natural a los otros segmentos del gen quimérico, tal como las moléculas de ácido nucleico que tienen la región codificante y segmentos de promotor no asociados de forma natural entre sí.

Como se emplea en esta memoria, una “molécula de ácido nucleico aislada” es un polímero de ARN o ADN de cadena doble o simple, que opcionalmente contiene bases de nucleótidos alteradas, no naturales o sintéticas. Una

45 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 puede “hibridarse” con otro fragmento de ácido nucleico, tal como un ADNc, ADN genómico o molécula de ARN, cuando una forma monocatenaria de un fragmento de ácido nucleico puede hibridarse con otro fragmento de ácido nucleico en las condiciones adecuadas de temperatura y fuerza iónica de la

50 solución. Las condiciones de hibridación y lavado son conocidas y se ejemplifican en Sambrook, J., Fritsch, E. F. y Maniatis, T. Molecular Cloning: A Laboratory Manual, 2.^a ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989), particularmente el Capítulo 11 y la Tabla 11.1 en este. Las condiciones de temperatura y fuerza iónica determinan la “rigurosidad” de la hibridación. Las condiciones de rigurosidad pueden ajustarse para seleccionar fragmentos moderadamente similares (tal como secuencias homólogas de organismos con parentesco lejano), a

55 fragmentos altamente similares (tal como genes que duplican enzimas funcionales de organismos con parentesco cercano). Los lavados posteriores a la hibridación determinan las condiciones de rigurosidad. Un conjunto de condiciones preferidas utiliza una serie de lavados comenzando con 6X SSC, SDS al 0,5% a temperatura ambiente

durante 15 min, luego repetir con 2X SSC, SDS al 0,5% a 45°C durante 30 min, y luego repetir dos veces con 0,2X SSC, SDS al 0,5% a 50°C durante 30 min. Un conjunto de condiciones de rigurosidad más preferidas utiliza temperaturas mayores en las que los lavados son idénticos a los anteriores excepto por la temperatura de los dos lavados finales de 30 min en 0,2X SSC, SDS al 0,5% que se aumentó a 60°C. Otro conjunto preferido de condiciones de alta rigurosidad utiliza dos lavados finales en 0,1 X SSC, SDS al 0,1% a 65°C. Un conjunto adicional de condiciones de rigurosidad incluye la hibridación a 0,1 X SSC, SDS al 0,1%, 65°C y lavados con 2X SSC, SDS al 0,1% seguido de 0,1 X SSC, SDS al 0,1%, por ejemplo.

La hibridación requiere que los dos ácidos nucleicos contengan secuencias complementarias aunque, dependiendo de la rigurosidad de la hibridación, son posibles los malapareamientos entre bases. La rigurosidad adecuada para la hibridación de ácidos nucleicos depende de la longitud de los ácidos nucleicos y el grado de complementación, variables que son conocidas en la técnica. Cuanto mayor sea el grado de similitud u homología entre dos secuencias de nucleótidos, mayor será el valor de T_m para los híbridos de ácidos nucleicos que tienen esas secuencias. La estabilidad relativa (que corresponde a mayor T_m) de hibridaciones de ácido nucleico disminuye en el siguiente orden: ARN:ARN, ADN:RNA, ADN:ADN. Para híbridos con una longitud mayor que 100 nucleótidos, se han derivado las ecuaciones para calcular T_m (véase Sambrook et ál., supra, 9.50-9.51). Para la hibridación con ácidos nucleicos más cortos, es decir, oligonucleótidos, la posición de malapareamientos se vuelve más importante, y la longitud del oligonucleótido determina su especificidad (véase Sambrook et ál., supra, 11.7-11.8). En una 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, el experto en la técnica reconocerá que la temperatura y la concentración de sal en solución de lavado pueden ajustarse según sea necesario según factores tal como longitud de la sonda.

Una "parte sustancial" de una secuencia de aminoácidos o nucleótidos es esa parte que comprende lo suficiente de la secuencia de aminoácidos de un polipéptido o la secuencia de nucleótidos de un gen para identificar de manera putativa ese polipéptido o gen, ya sea mediante evaluación manual de la secuencia por parte de un experto en la técnica o mediante comparación e identificación de secuencias automatizada por computadora usando algoritmos tal como BLAST (Altschul, S. F., et ál., J. Mol. Biol., 215:403-410 (1993)). En general, una secuencia de diez o más aminoácidos contiguos o treinta o más nucleótidos es necesaria para identificar de manera putativa una secuencia de ácido nucleico o polipéptido como homóloga a una proteína o gen conocido. Además, con respecto a secuencias de nucleótidos, se pueden utilizar sondas de oligonucleótidos específicas del gen que comprenden 20-30 nucleótidos contiguos en métodos que dependen de la secuencia de identificación génica (p. ej., hibridación Southern) y aislamiento (p. ej., hibridación in situ de colonias bacterianas o placas bacteriófagas). Además, los oligonucleótidos cortos de 12-15 bases se pueden utilizar como cebadores de amplificación en PCR con el fin de obtener un fragmento de ácido nucleico particular que comprenda los cebadores. Por consiguiente, una "parte sustancial" de una secuencia de nucleótidos comprende lo suficiente de la secuencia para identificar y/o aislar específicamente un fragmento de ácido nucleico que comprenda la secuencia. La presente memoria descriptiva describe la secuencia de nucleótidos y aminoácidos completa que codifica proteínas fúngicas particulares. El experto en la técnica que cuente con el beneficio de las secuencias indicadas en la presente memoria, ahora puede usar la totalidad o una parte sustancial de las secuencias descritas con propósitos conocidos por los expertos en esta técnica.

El término "complementariedad" se utiliza para describir la relación entre bases de nucleótidos que pueden hibridarse entre sí. Por ejemplo, con respecto al ADN, adenosina es complementaria a timina y citosina es complementaria a guanina.

Los términos "homología" y "homólogo" se utilizan de manera intercambiable en la presente memoria. Se refieren a fragmentos de ácido nucleico en donde los cambios en una o más bases de nucleótidos no afectan la capacidad del ácido nucleico de mediar la expresión génica o de producir un determinado fenotipo. Estos términos también hacen referencia a modificaciones de los fragmentos de ácido nucleico descritos en la presente memoria tal como eliminación o inserción de uno o más nucleótidos que no alteran sustancialmente las propiedades funcionales del fragmento de ácido nucleico resultante con respecto al fragmento inicial no modificado. Se entiende, por lo tanto, tal como lo entenderán los expertos en la técnica, que se puede utilizar otras además de las secuencias de ejemplo específicas en las cepas y métodos de la invención.

Además, el experto en la técnica reconoce que las secuencias de ácido nucleico homólogas descritas en la presente memoria también se definen por su capacidad de hibridarse, en condiciones de rigurosidad moderada (p. ej., 0,5 X SSC, SDS al 0,1%, 60°C) con las secuencias ejemplificadas en la presente memoria, o con cualquier parte de las secuencias de nucleótidos descritas en la presente memoria y que son funcionalmente equivalentes a cualquiera de las secuencias de nucleótidos descritas en la presente memoria.

El término "porcentaje de identidad", tal como se conoce en la técnica, es una relación entre dos o más secuencias de polipéptidos o dos o más secuencia de polinucleótidos, tal como se determina mediante comparación de secuencias. En la técnica, "identidad" también implica el grado de relación de secuencia entre las secuencias de polipéptidos o polinucleótidos, según sea el caso, determinada por la coincidencia entre cadenas de dichas secuencias. La "identidad" y "similitud" pueden calcularse fácilmente mediante métodos conocidos, que incluyen

pero no se limitan a 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 están diseñados para proporcionar la mayor coincidencia entre las secuencias analizadas. Los métodos para determinar la identidad y similitud se codifican en programas informáticos de disponibilidad pública. Se pueden realizar alineaciones de secuencias y cálculos de porcentajes de identidad mediante el uso del programa MegAlign™ del paquete informático LASERGENE de bioinformática (DNASTAR, Inc., Madison, WI). El alineamiento múltiple de secuencias se realiza usando el "Método Clustal de alineamiento" que abarca diversas variedades del algoritmo que incluye el "Método Clustal V de alineamiento" que corresponde al método de alineamiento denominado Clustal V (descrito por Higgins y Sharp, CABIOS. 5:151-153 (1989); Higgins, D.G. et ál., Comput. Appl. Biosci., 8:189-191 (1992)) y encontrado en el programa MegAlign™ del paquete informático LASERGENE de bioinformática (DNASTAR Inc.). Para alineamientos múltiples, los valores predeterminados corresponden a PENALIZACIÓN DE HUECO=10 y PENALIZACIÓN DE LONGITUD DE HUECO=10. Los parámetros predeterminados para los alineamientos en pares y el cálculo del porcentaje de identidad de secuencias de proteínas usando el método Clustal son KTUPLE=1, PENALIZACIÓN DE HUECO=3, VENTANA=5 y DIAGONALES GUARDADAS=5. Para los ácidos nucleicos estos parámetros son KTUPLE=2, PENALIZACIÓN DE HUECO=5, VENTANA=4 y DIAGONALES GUARDADAS=4. Después del alineamiento de las secuencias usando el programa Clustal V, es posible obtener un "porcentaje de identidad" al observar la tabla de "distancias de secuencia" en el mismo programa. De manera adicional, el "Método Clustal W de alineamiento" está disponible y corresponde al método de alineamiento denominado Clustal W (descrito por Higgins y Sharp, CABIOS. 5:151-153 (1989); Higgins, D.G. et ál., Comput. Appl. Biosci. 8:189-191 (1992)) y encontrado en el programa MegAlign™ v6.1 del paquete informático LASERGENE de bioinformática (DNASTAR Inc.). Los parámetros predeterminados para el alineamiento múltiple (PENALIZACIÓN DE HUECO=10, PENALIZACIÓN DE LONGITUD DE HUECO=0,2, Retardo de las secuencias divergentes(%)=30, Peso de transición de ADN=0,5, Matriz de peso de proteína=Serie Gonnet, Matriz de peso de ADN=IUB). Después del alineamiento de las secuencias usando el programa Clustal W, es posible obtener un "porcentaje de identidad" al observar la tabla de "distancias de secuencia" en el mismo programa.

Los expertos en la técnica entienden que muchos niveles de identidad de secuencia son útiles en la identificación de polipéptidos, de otras especies, en donde dichos polipéptidos tienen la misma función o actividad o una similar. Los ejemplos útiles de porcentaje de identidad incluyen, pero no se limitan a: 24%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% o 95%, o cualquier porcentaje entero de 24% 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 solo tienen las homologías anteriores sino que típicamente codifican un polipéptido que tiene al menos 50 aminoácidos, preferiblemente al menos 100 aminoácidos, más preferiblemente al menos 150 aminoácidos, aun más preferiblemente al menos 200 aminoácidos, y lo más preferiblemente al menos 250 aminoácidos.

El término "software de análisis de secuencias" se refiere a cualquier algoritmo informático o programa de software que es útil para el análisis de secuencias de aminoácidos o nucleótidos. El "software de análisis de secuencias" puede encontrarse comercialmente disponible o puede desarrollarse de forma independiente. Un software de análisis de secuencias típico incluirá, pero no se limita a: 1.) el paquete de programas GCG (Wisconsin Package Versión 9.0, Genetics Computer Group (GCG), Madison, WI); 2.) BLASTP, BLASTN, BLASTX (Altschul et ál., 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). En el contexto de la presente solicitud, se comprenderá que en los casos en que se usa software de análisis de secuencias para el análisis, los resultados del análisis se basarán en los "valores predeterminados" del programa al que se hace referencia, a menos que se especifique lo contrario. Como se emplea en esta memoria, "valores predeterminados" significa cualquier conjunto de valores o parámetros que se carga originalmente con el software cuando se inicia por primera vez.

Las técnicas de clonación molecular y ADN recombinante estándar utilizadas en la presente memoria son conocidas en la técnica y fueron descritas por Sambrook, J., Fritsch, E. F. y Maniatis, T. Molecular Cloning: A Laboratory Manual, 2.^a ed.; Cold Spring Harbor Laboratory: Cold Spring Harbor, New York, 1989 (de aquí en adelante "Maniatis"); y por Silhavy, T. J., Brennan, M. L. y Enquist, L. W. Experiments with Gene Fusions; Cold Spring Harbor Laboratory: Cold Spring Harbor, Nueva York, 1984; y por Ausubel, F. M. et ál., In Current Protocols in Molecular Biology, publicado por Greene Publishing y Wiley-Interscience, 1987.

La presente invención se refiere a cepas diseñadas genéticamente de *Zymomonas* o *Zymobacter* que utilizan arabinosa que tienen utilización de arabinosa mejorada cuando se fermentan en un medio que contiene arabinosa, y

a procesos para la producción de etanol usando estas cepas. Un desafío para mejorar la producción de etanol mediante la fermentación de un biocatalizador en un medio que incluye un hidrolizado de biomasa, producido típicamente mediante pretratamiento y sacarificación de la biomasa, es obtener una utilización eficaz de arabinosa. La arabinosa es una de las azúcares pentosa predominantes en materiales lignocelulósicos hidrolizados, la otra es xilosa. Los solicitantes han descubierto que la expresión de un simportador de protones de arabinosa conduce a una mayor eficacia en la utilización de arabinosa mediante cepas que utilizan arabinosa y, por lo tanto, a mayores rendimientos de etanol cuando la fermentación se produce en un medio que contiene arabinosa.

Cepa hospedadora que utiliza arabinosa

Cualquier cepa de *Zymomonas* o *Zymobacter* que pueda utilizar arabinosa como una fuente de carbono puede utilizarse como un hospedador para preparar las cepas de la presente invención. Las cepas de *Zymomonas*, tal como *Z. mobilis* que han sido diseñadas genéticamente para fermentación de arabinosa en etanol son particularmente útiles. *Zymomonas* ha sido diseñada genéticamente para utilización de arabinosa mediante la introducción de genes que codifican 1) L-arabinosa isomerasa para convertir L-arabinosa en L-ribulosa, 2) L-ribulocinasa para convertir L-ribulosa en L-ribulosa-5-fosfato, y 3) L-ribulosa-5-fosfato-4-epimerasa para convertir L-ribulosa-5-fosfato en D-xilulosa (US5843760 y descrito en los Ejemplos 1 y 2 en la presente memoria; véase diagrama en la Figura 1). Las secuencias de ADN que codifican estas enzimas pueden obtenerse a partir de microorganismos que tienen la capacidad de metabolizar la arabinosa. Las fuentes para las regiones codificantes incluyen *Klebsiella*, *Escherichia*, *Rhizobium*, *Agrobacterium* y *Salmonella*. Las regiones codificantes de *E. coli* son particularmente útiles, que son para L-arabinosa isomerasa: región codificante de araA (región codificante SEQ ID NO:21; proteína SEQ ID NO:22), para L-ribulocinasa: región codificante de araB (región codificante SEQ ID NO:23; proteína SEQ ID NO:24), y para L-ribulosa-5-fosfato-4-epimerasa: región codificante de araD (región codificante SEQ ID NO:25; proteína SEQ ID NO:26). Estas proteínas y sus regiones codificantes pueden identificarse fácilmente en otros microorganismos que utilizan arabinosa, tal como los indicados anteriormente, por un experto en la técnica mediante el uso de bioinformática o métodos experimentales tal como se describe más adelante para araE.

Además, las actividades de transcetolasa y transaldolasa se utilizan en la vía biosintética de arabinosa a etanol (véase la Figura 1). La transcetolasa y transaldolasa son dos enzimas de la vía de la pentosa fosfato que convierten xilulosa 5-fosfato en intermedios que acoplan el metabolismo pentosa a la vía glucolítica de Entner-Doudoroff lo que permite el metabolismo de arabinosa o xilosa en etanol. Estas pueden ser actividades endógenas, o las actividades endógenas pueden complementar las actividades introducidas para estas enzimas.

Típicamente, la *Zymomonas* que utiliza arabinosa también está diseñada genéticamente para utilización de xilosa. Típicamente, se han introducido cuatro genes en *Z. mobilis* para la expresión de cuatro enzimas que participan en el metabolismo de xilosa (Figura 1) tal como se describe en US 5514583. Estas incluyen genes que codifican transcetolasa y transaldolasa tal como se describió anteriormente, así como también xilosa isomerasa, que cataliza la conversión de xilosa en xilulosa y xilulocinasa, que fosforila xilulosa para formar xilulosa 5-fosfato (véase la Figura 1). Las secuencias de ADN que codifican estas enzimas pueden obtenerse de numerosos microorganismos que son capaces de metabolizar xilosa, tal como bacterias entéricas, y algunas levaduras y hongos. Las fuentes para las regiones codificantes incluyen *Xanthomonas*, *Klebsiella*, *Escherichia*, *Rhodobacter*, *Flavobacterium*, *Acetobacter*, *Gluconobacter*, *Rhizobium*, *Agrobacterium*, *Salmonella*, *Pseudomonads* y *Zymomonas*. Las regiones codificantes de *E. coli* son particularmente útiles.

Para la expresión, las secuencias de ADN codificantes para proteínas que utilizan arabinosa y proteínas que utilizan xilosa están unidas de forma operativa a promotores que se expresan en células de *Z. mobilis*, y terminadores de transcripción. Los ejemplos de promotores que se pueden utilizar incluyen los promotores del gen codificador de gliceraldehído-3-fosfato deshidrogenasa de *Z. mobilis* (promotor GAP; Pgap), del gen codificador de enolasa de *Z. mobilis* (promotor ENO; Peno), y del gen codificador de xilosa isomerasa de *Actinoplanes missouriensis* (promotor GI, Pgi). Las regiones codificantes pueden expresarse individualmente a partir de un promotor, típicamente como un gen quimérico, o dos o más regiones codificantes pueden unirse en un operón con expresión a partir del mismo promotor. Los operones y/o genes quiméricos resultantes típicamente se construyen en o se transfieren a un vector para manipulaciones adicionales.

Los vectores son ampliamente conocidos en la técnica. Para la expresión en *Zymomonas* son particularmente útiles los vectores que pueden replicarse tanto en *E. coli* como en *Zymomonas*, tal como pZB188 que se describe en la patente de EE.UU. n.º 5.514.583. Los vectores pueden incluir plásmidos para la replicación autónoma en una célula, y plásmidos para llevar construcciones que se integrarán al genoma celular. Los plásmido para la integración de ADN pueden incluir transposones, regiones de secuencia de ácido nucleico homólogas al genoma celular objetivo, secuencias de integración dirigidas al sitio u otras secuencias que apoyan la integración. En la recombinación homóloga, las secuencias de ADN que flanquean un sitio de integración objetivo se colocan unidas a un gen de resistencia a espectinomicina, u otro marcador seleccionable, y el gen quimérico conduce la inserción del marcador seleccionable y gen quimérico en el sitio genómico objetivo tal como se describe en el Ejemplo 2 en la presente memoria. Además, el marcador seleccionable puede estar unido mediante sitios de recombinación específicos del sitio, de modo que luego de la expresión de la recombinasa específica del sitio correspondiente, el gen de recombinación pueda extraerse del genoma.

Las cepas que utilizan xilosa que son de uso particular incluyen CP4(pZB5) (US 5514583), ATCC31821/pZBS (US 6566107), 8b (US 20030162271; Mohagheghi et ál., (2004) Biotechnol. Lett. 25; 321-325), y ZW658 con derivados ZW800 y ZW801-4 (publicación de patente de EE.UU., de titularidad conjunta y en tramitación con la presente, #US20080286870; depositado, ATCC # PTA-7858). También se puede utilizar ZW705, que se describe en la publicación de solicitud de patente de EE.UU., de titularidad conjunta y en tramitación con la presente, #US 20110014670. Las cepas que utilizan arabinosa que se pueden utilizar se describen en US5843760, y también se describen en la presente memoria en los Ejemplos 1 y 2.

Adaptación para utilización de arabinosa

Los solicitantes descubrieron que una cepa de *Z. mobilis* diseñada genéticamente para utilización de xilosa y arabinosa tal como se describió anteriormente utiliza aproximadamente 33% de arabinosa en un medio donde la arabinosa es la única fuente de carbono (a 50 g/L), y aproximadamente 68% de arabinosa en un medio que incluye azúcares mezclados de 25 g/L de arabinosa, 25 g/L de xilosa y 50 g/L de glucosa en condiciones de cultivo de prueba. En un intento de derivar una cepa con utilización de arabinosa mejorada, los solicitantes adaptaron células de la cepa que utiliza xilosa y arabinosa mediante cultivo en serie en un medio con 50 g/L de arabinosa como la única fuente de carbono tal como se describe en la presente memoria en el Ejemplo 2. Usando este proceso, se obtuvieron cepas aisladas que tenían una mejora sustancial en la utilización de arabinosa en un medio donde la arabinosa es la única fuente de carbono, que son cepas adaptadas a arabinosa. Por ejemplo, una cepa utilizó aproximadamente 83% de arabinosa donde 50 g/L de arabinosa es la única fuente de carbono. En el medio de azúcares mezclados que contenía 25 g/L de arabinosa, 25 g/L de xilosa y 50 g/L de glucosa, hubo una mejora menor: se utilizó aproximadamente 74% de arabinosa. También en el medio de azúcares mezclados, la utilización de arabinosa se retrasó en comparación con la utilización de glucosa y xilosa.

Para obtener cepas con utilización de arabinosa mejorada, las cepas diseñadas genéticamente para la expresión de genes con utilización de arabinosa tal como se describió anteriormente pueden adaptarse mediante cultivo en serie en un medio que contiene arabinosa como la única fuente de carbono en concentraciones entre aproximadamente 20 g/L y 100 g/L, o mayores. La adaptación puede realizarse en concentraciones de arabinosa, pero con cultivo inicial en aproximadamente 20 g/L o más. El cultivo en serie es típicamente para al menos aproximadamente 25 duplicaciones. La adaptación puede realizarse antes o después de introducir un simportador de protones de arabinosa heterólogo, que se describe más adelante, en una cepa que utiliza arabinosa. Además, las células pueden adaptarse tanto antes como después de la introducción de un simportador de protones de arabinosa heterólogo.

Descubrimiento para diseñar genéticamente la utilización de arabinosa mejorada

Los solicitantes diseñaron genéticamente cepas de *Zymomonas* que utilizan xilosa y arabinosa para la expresión de los dos sistemas de transporte de arabinosa diferentes presentes en *E. coli*. Los dos sistemas son 1) un transportador ABC que consiste en tres proteínas codificadas por araFGH: proteína de unión a arabinosa periplasmática de 33 kD codificada por araF, ATPasa unida a la membrana de 55 kD codificada por araG, y proteína unida a la membrana de 34 kD codificada por araH; y 2) un simportador de protones de arabinosa que consiste en una proteína: simportador de protones de arabinosa de 52 kD codificado por araE. El transportador ABC es un transportador de arabinosa de alta afinidad y baja capacidad con una Km de 3×10^{-4} M, mientras el simportador de protones de arabinosa es un transportador de arabinosa de baja afinidad y alta capacidad con una Km of $1,25 \times 10^{-4}$ M. Los solicitantes descubrieron que la expresión del transportador ABC en realidad dio como resultado una utilización de arabinosa reducida en medio solamente con arabinosa. La expresión del simportador de protones de arabinosa aumentó la utilización de arabinosa tanto en el medio solamente con arabinosa como en el medio con azúcares mezclados. Por lo tanto, los solicitantes han descubierto que el transportador ABC de *E. coli* no mejora la utilización de arabinosa, mientras el simportador de protones de arabinosa mejora la utilización de arabinosa en *Zymomonas*. Con la expresión del simportador de protones de arabinosa, la utilización de arabinosa aumentó ampliamente tanto en el medio solamente con arabinosa como en el medio con azúcares mezclados.

La expresión de un simportador de protones de arabinosa aumentó la utilización de arabinosa en todas las cepas analizadas. Estas incluyen una cepa de *Z. mobilis* que utiliza arabinosa y xilosa sin adaptación, una cepa de *Z. mobilis* que utiliza arabinosa y xilosa que ha sido adaptada para utilización de xilosa en condiciones de estrés (descrita en la publicación de solicitud de patente de EE.UU., de titularidad conjunta y en tramitación con la presente, #US 20110014670), y una cepa de *Z. mobilis* que utiliza arabinosa y xilosa que ha sido adaptada para utilización de xilosa en condiciones de estrés y también para utilización de arabinosa tal como se describió anteriormente en la presente memoria y en el Ejemplo 2. En cepas sin adaptación de arabinosa, la utilización de arabinosa aumentó en al menos aproximadamente 28% en el medio solamente con arabinosa así como también en el medio con azúcares mezclados. También en una cepa adaptada a arabinosa, la utilización de arabinosa aumentó en al menos aproximadamente 28% en el medio con azúcares mezclados. En el medio solamente con arabinosa el nivel de utilización de arabinosa en la cepa original adaptada a arabinosa sin expresión del simportador de protones de arabinosa ya se encuentra a aproximadamente 80% y, por lo tanto, el aumento en la utilización de arabinosa no puede exceder 20%, y es de aproximadamente 18%.

Por lo tanto, cualquier cepa de *Zymomonas* o *Zymobacter* que sea capaz de utilizar arabinosa, también denominada cepa que utiliza arabinosa, se puede utilizar para crear las cepas de la presente. Son particularmente útiles las

cepas que utilizan adicionalmente xilosa y glucosa. En estas cepas la utilización de arabinosa se mejora en al menos aproximadamente 10% mediante la expresión de un simportador de protones de arabinosa. La utilización de arabinosa puede mejorarse en al menos aproximadamente 10%, 12%, 16%, 18%, 20%, 24%, 28% o más. El % de mejora puede variar dependiendo de las condiciones de cultivo utilizadas que incluyen el tipo de medio y el microorganismo original utilizado para diseñar genéticamente la expresión del simportador de protones de arabinosa, así como también la cepa específica diseñada genéticamente resultante. Los factores que provocan la variación incluyen el nivel de expresión del simportador de protones de arabinosa introducido y el nivel de actividad del transportador resultante, que puede variar entre transformantes.

Expresión de un simportador de protones de arabinosa

En las células diseñadas genéticamente de *Zymomonas* o *Zymobacter* de la presente puede expresarse cualquier simportador de protones de arabinosa para proporcionar utilización de arabinosa mejorada. Las proteínas bacterianas simportadoras de protones de arabinosa y sus secuencias codificantes para la expresión en *Zymomonas* o *Zymobacter* son heterólogas, ya que no se encuentran de forma natural en *Zymomonas* o *Zymobacter*. Los ejemplos de proteínas simportadoras de protones de arabinosa y secuencias codificantes que pueden expresarse incluyen las codificadas por los genes araE de *E. coli* (región codificante SEQ ID NO:1; proteína SEQ ID NO:2), *Shigella flexneri* (región codificante SEQ ID NO:3; proteína SEQ ID NO:4), *Shigella boydii* (región codificante SEQ ID NO:5; proteína SEQ ID NO:6), *Shigella dysenteriae* (región codificante SEQ ID NO:7; proteína SEQ ID NO:8), *Salmonella typhimurium* (región codificante SEQ ID NO:9; proteína SEQ ID NO:10), *Salmonella enterica* (región codificante SEQ ID NO:11; proteína SEQ ID NO:12), *Klebsiella pneumoniae* (coding region SEQ ID NO:13; proteína SEQ ID NO:14), *Klebsiella oxytoca* (región codificante SEQ ID NO:15; proteína SEQ ID NO:16), *Enterobacter cancerogenus* (región codificante SEQ ID NO:17; proteína SEQ ID NO:18) y *Bacillus amyloliquefaciens* (región codificante SEQ ID NO:19; proteína SEQ ID NO:20).

Dado que las secuencias de las regiones codificantes del simportador de protones de arabinosa y las proteínas codificadas son ampliamente conocidas, tal como se ejemplifica en las SEQ ID NO indicadas anteriormente en la Tabla 1, un experto en la técnica puede identificar fácilmente simportadores de protones de arabinosa adecuados adicionales en función de la similitud de secuencia usando enfoques de bioinformática. Típicamente, la búsqueda BLAST (descrita anteriormente) de bases de datos de disponibilidad pública con secuencias de aminoácidos del simportador de protones de arabinosa, tal como las proporcionadas en la presente memoria, se utiliza para identificar simportadores de protones de arabinosa adicionales, y sus secuencias codificantes que se pueden utilizar en las cepas de la presente. Estas proteínas pueden tener al menos aproximadamente 80-85%, 85%-90%, 90%-95% o 95%-99% de identidad de secuencia con cualquiera de los simportadores de protones de arabinosa de las SEQ ID NO:2, 4, 6, 8, 10, 12, 14, 16, 18 o 20 y al mismo tiempo presentan actividad del simportador de protones de arabinosa. Las identidades se basan en el método de alineamiento Clustal W usando los parámetros predeterminados de PENALIZACIÓN DE HUECO=10, PENALIZACIÓN DE LONGITUD HUECO=0,1, y serie Gonnet 250 de matriz de peso de proteína.

Además de usar la secuencia de región codificante o de proteína y los métodos de bioinformática para identificar simportadores de protones de arabinosa adicionales, se pueden utilizar las secuencias descritas en la presente memoria o las indicadas en la técnica para identificar experimentalmente otros homólogos en la naturaleza. Por ejemplo, cada uno de los fragmentos de ácido nucleico que codifican simportadores de protones de arabinosa descritos en la presente memoria se puede utilizar para aislar los genes que codifican proteínas homólogas. El aislamiento de genes homólogos usando protocolos que dependen de la secuencia es ampliamente conocido en la técnica. Los ejemplos de protocolos que dependen de la secuencia incluyen, pero no se limitan a: 1.) métodos de hibridación de ácido nucleico; 2.) métodos de amplificación de ADN y ARN, tal como se ejemplifica mediante diversos usos de las tecnologías de amplificación de ácido nucleico [por ejemplo, reacción en cadena de la polimerasa (PCR, por sus siglas en inglés), Mullis et ál., patente de EE.UU. 4.683.202; reacción en cadena de la ligasa (LCR, por sus siglas en inglés), Tabor, S. et ál., Proc. Acad. Sci. USA 82:1074 (1985); o amplificación por desplazamiento de hebra (SDA, por sus siglas en inglés), Walker, et ál., Proc. Natl. Acad. Sci. U.S.A., 89:392 (1992)]; y 3.) métodos de construcción de biblioteca y selección mediante complementación.

Por ejemplo, las regiones codificantes para proteínas o polipéptidos similares a las de las secuencias que codifican el simportador de protones de arabinosa descritas en la presente memoria pueden aislarse directamente usando la totalidad o una parte de los fragmentos de ácido nucleico de la presente como sondas de hibridación de ADN para seleccionar bibliotecas a partir de cualquier organismo deseado usando metodología conocida por los expertos en la técnica. Las sondas de oligonucleótidos específicas basadas en las secuencias de ácido nucleico descritas pueden diseñarse y sintetizarse mediante métodos conocidos en la técnica (Maniatis, *supra*). Además, las secuencias enteras se pueden utilizar directamente para sintetizar sondas de ADN mediante métodos conocidos por el experto en la técnica (p. ej., etiquetado de ADN de cebadores aleatorios, traducción de mellas y técnicas de etiquetado de extremos), o cebadores de ARN que utilizan sistemas de transcripción *in vitro* disponibles. Además, pueden diseñarse cebadores específicos y utilizarse para amplificar una parte (o la longitud completa) de las secuencias de la presente. Los productos de amplificación resultantes pueden etiquetarse directamente durante las reacciones de amplificación o etiquetarse después de las reacciones de amplificación, y utilizarse como sondas para aislar fragmentos de ADN de longitud completa en condiciones de rigurosidad adecuada.

Típicamente, en las técnicas de amplificación tipo PCR, los cebadores tienen diferentes secuencias y no son complementarias entre sí. Dependiendo de las condiciones de prueba deseadas, las secuencias de los cebadores deberían diseñarse para proporcionar una replicación eficaz y fiel del ácido nucleico objetivo. Los métodos de diseño de sonda de PCR son comunes y conocidos en la técnica (Thein y Wallace, "The use of oligonucleotides as specific hybridization probes in the Diagnosis of Genetic Disorders", in Human Genetic Diseases: A Practical Approach, K. E. Davis Ed., (1986) pp 33-50, IRL: Hemdon, VA; y Rychlik, W., In Methods in Molecular Biology, White, B. A. Ed., (1993) Vol. 15, pp 31-39, PCR Protocols: Current Methods and Applications. Humana: Totowa, NJ).

Generalmente, dos segmentos cortos de las secuencias descritas se pueden utilizar en los protocolos de reacción en cadena de la polimerasa para amplificar fragmentos de ácido nucleico más largos que codifican genes homólogos de ADN o ARN. La reacción en cadena de la polimerasa también se puede realizar en una biblioteca de fragmentos de ácido nucleico clonados en donde la secuencia de un cebador deriva de los fragmentos de ácido nucleico descritos, y la secuencia del otro cebador se aprovecha de la presencia de los tramos de ácido poliadenílico en el extremo 3' del precursor de ARNm que codifica genes microbianos.

De manera alternativa, la segunda secuencia de cebador puede basarse en las secuencias derivadas del vector de clonación. Por ejemplo, el experto en la técnica puede seguir el protocolo RACE (Frohman et ál., PNAS USA 85:8998 (1988)) para generar ADNc mediante el uso de PCR para amplificar copias de la región entre un solo punto en la transcripción y el extremo 3' o 5'. Los cebadores orientados en las direcciones 3' y 5' pueden diseñarse a partir de las secuencias de la presente. Usando los sistemas RACE 3' o RACE 5' comercialmente disponibles (p. ej., BRL, Gaithersburg, MD), se pueden aislar fragmentos de ADNc 3' o 5' específicos (Ohara et ál., PNAS USA 86:5673 (1989); Loh et ál., Science 243:217 (1989)).

De manera alternativa, las secuencias que codifican el simportador de protones de arabinosa descritas pueden emplearse como agentes de hibridación para la identificación de homólogos. Los componentes básicos de una prueba de hibridación de ácido nucleico incluyen una sonda, una muestra que se sospecha que contiene el gen o fragmento génico de interés y un método de hibridación específico. Típicamente, las sondas son secuencias de ácido nucleico de cadena simple que son complementarias a las secuencias de ácido nucleico que se detectarán. Las sondas son "hibridables" con las secuencias de ácido nucleico que se detectarán. La longitud de la sonda puede variar de 5 bases a decenas de miles de bases, y dependerá de la prueba específica que se realizará. Típicamente, la longitud de una sonda de aproximadamente 15 bases a aproximadamente 30 bases es adecuada. Solo una parte de la molécula de la sonda debe ser complementaria a la secuencia de ácido nucleico que se detectará. Además, la complementariedad entre la sonda y la secuencia objetivo no es necesariamente perfecta. La hibridación no se produce entre moléculas de complementariedad imperfecta con el resultado de que una fracción determinada de las bases en la región hibridada no se empareja con la base complementaria adecuada.

Los métodos de hibridación están bien definidos. Típicamente, la sonda y la muestra deben mezclarse en condiciones que permitirán la hibridación de ácido nucleico. Esto implica que la sonda y la muestra entren en contacto en presencia de una sal orgánica o inorgánica en las condiciones de concentración y temperatura adecuadas. Los ácidos nucleicos de la sonda y la muestra deben estar en contacto por un tiempo lo suficientemente prolongado para que pueda producirse cualquier hibridación posible entre el ácido nucleico de la sonda y la muestra. La concentración de la sonda u objetivo en la mezcla determinará el tiempo necesario para que se produzca la hibridación. Cuanto mayor sea la concentración de la sonda u objetivo, menor será el tiempo de incubación de hibridación necesario. De manera opcional, se puede incluir un agente caotrópico. El agente caotrópico estabiliza los ácidos nucleicos al inhibir la actividad de la nucleasa. Además, el agente caotrópico permite una hibridación sensible y rigurosa de las sondas de oligonucleótidos cortas a temperatura ambiente (Van Ness y Chen, Nucl. Acids Res. 19:5143-5151 (1991)). Los agentes caotrópicos adecuados incluyen cloruro de guanidinio, tiocianato de guanidinio, tiocianato de sodio, tetracloroacetato de litio, perclorato de sodio, tetracloroacetato de rubidio, yoduro de potasio y trifluoroacetato de cesio, entre otros. Típicamente, el agente caotrópico estará presente a una concentración final de aproximadamente 3 M. Si se desea, se puede agregar formamida a la mezcla de hibridación, típicamente 30-50% (v/v).

Se pueden emplear varias soluciones de hibridación. Típicamente, estas comprenden de aproximadamente 20 a 60% en volumen, preferiblemente 30%, de un solvente orgánico polar. Una solución de hibridación común emplea aproximadamente 30-50% v/v de formamida, aproximadamente 0,15 a 1 M de cloruro de sodio, aproximadamente 0,05 a 0,1 M de amortiguadores (p. ej., citrato de sodio, Tris-HCl, PIPES o HEPES (intervalo de pH de aproximadamente 6-9)), aproximadamente 0,05 a 0,2% de detergente (p. ej., dodecilsulfato de sodio), o entre 0,5-20 mM de EDTA, FICOLL (Pharmacia Inc.) (aproximadamente 300-500 kdal), polivinilpirrolidona (aproximadamente 250-500 kdal) y albúmina en suero. En la solución de hibridación típica también se incluirán los ácidos nucleicos portadores no etiquetados de aproximadamente 0,1 a 5 mg/mL, ADN nucleico fragmentado (p. ej., ADN de timo de becerro o esperma de salmón, o ARN de levadura), y opcionalmente de aproximadamente 0,5 a 2% p/vol de glicina. También se pueden incluir otros aditivos, tal como agentes de exclusión de volumen que incluyen una variedad de agentes polares solubles en agua o dilatables (p. ej., polietilenglicol), polímeros aniónicos (p. ej., poliácido o polimetilacrilato) y polímeros sacarídicos aniónicos (p. ej., sulfato de dextrano).

La hibridación de ácido nucleico se puede adaptar a una variedad de formatos de ensayos. Uno de los más adecuados es el formato de ensayo tipo sándwich. El ensayo tipo sándwich se puede adaptar particularmente a la

hibridación en condiciones no desnaturalizantes. Un componente principal de un ensayo tipo sándwich es un soporte sólido. El soporte sólido se ha adsorbido o unido de forma covalente a su sonda de ácido nucleico inmovilizado que no está etiquetada y es complementaria a una parte de la secuencia.

- 5 La expresión del simportador de protones de arabinosa se logra mediante la transformación con una secuencia que codifica un simportador de protones de arabinosa. Tal como se conoce en la técnica, puede haber variaciones en secuencias de ADN que codifican una secuencia de aminoácidos debido a la degeneración del código genético. La secuencia codificante puede estar optimizadas por codones para lograr la máxima expresión en la célula hospedadora de *Zymomonas* o *Zymobacter* objetivo, tal como lo conocen los expertos en la técnica. Típicamente, un gen quimérico que incluye un promotor activo en células de *Zymomonas* que está unido de forma operativa a la
- 10 región codificante deseada, así como también un terminador de transcripción, se utiliza para la expresión. Se puede utilizar cualquier promotor que sea activo en células de *Zymomonas*, tal como los ejemplos indicados anteriormente para la expresión de proteínas para la utilización de arabinosa. Un gen quimérico construido con un promotor y región codificante de simportador de arabinosa es un gen heterólogo para la expresión en *Zymomonas* o *Zymobacter* ya que la región codificante es de un organismo diferente al descrito anteriormente. Los vectores para la expresión y/o integración son tal como se describieron anteriormente para la expresión de proteínas para la
- 15 utilización de arabinosa.

Producción mejorada de etanol

- Las cepas de la presente tienen utilización de arabinosa mejorada en medio con arabinosa como la única fuente de carbohidratos y en medio con azúcares mezclados que incluyen arabinosa. Las cepas de la presente también
- 20 presentan mejor producción de etanol. En comparación con la cepa original antes de la introducción de un gen de expresión del simportador de protones de arabinosa, la producción de la cepa que expresa un simportador de protones de arabinosa se ve aumentada. El aumento en la producción de etanol puede variar dependiendo del medio y las condiciones de cultivo utilizadas en la fermentación así como también el simportador de protones de arabinosa que expresa la cepa utilizada como el biocatalizador. Típicamente, la producción de etanol puede
- 25 aumentar en al menos aproximadamente 10%, puede aumentar al menos aproximadamente 10%, 12%, 16%, 18%, 20%, 24%, 28% o más.

Fermentación de la cepa que utiliza arabinosa mejorada

- Se puede utilizar una cepa diseñada genéticamente que utiliza arabinosa que expresa un simportador de protones de arabinosa y genes u operones para la expresión de L-arabinosa isomerasa, L-ribulocinasa, L-ribulosa-5-fosfato-4-
- 30 epimerasa, transaldolasa y transcetolasa en la fermentación para producir un producto que sea un producto natural de la cepa, o un producto cuya cepa fue diseñada genéticamente para que produzca. Por ejemplo, *Zymomonas mobilis* y *Zymobacter palmarum* son etanológenos naturales. Se prefieren las cepas que también utilizan xilosa y están diseñadas genéticamente además para la expresión de xilosa isomerasa y xilulocinasa. Como ejemplo, se describe la producción de etanol mediante una cepa de *Z. mobilis* de la invención, que utiliza xilosa y arabinosa. *Z. mobilis*
- 35 también utiliza glucosa de forma natural.

Para la producción de etanol, *Z. mobilis* recombinante que utiliza xilosa y arabinosa que expresa un simportador de protones de arabinosa se pone en contacto con un medio que contiene arabinosa. Típicamente, el medio contiene azúcares mezclados que incluyen azúcar, xilosa y glucosa. El medio puede contener hidrolizado de biomasa que incluye estos azúcares que derivan de biomasa lignocelulósica o celulósica tratada.

- 40 Cuando la concentración de azúcares mezclados es tan alta que inhibe el crecimiento, el medio incluye sorbitol, manitol, o una mezcla de estos tal como se describe en la publicación de patente de EE.UU., de titularidad conjunta y en tramitación con la presente, #US20080081358 A1. Galactitol o ribitol pueden reemplazarse o combinarse con sorbitol o manitol. La *Z. mobilis* crece en el medio donde ocurre la fermentación y se produce etanol. La fermentación ocurre sin complemento de aire, oxígeno u otros gases (que puede incluir condiciones tal como
- 45 fermentación anaeróbica, microaeróbica o microaerofílica), durante al menos 24 horas, y puede producirse durante 30 horas o más. El tiempo para alcanzar la máxima producción de etanol es variable, dependiendo de las condiciones de fermentación. Típicamente, si hay inhibidores presentes en el medio, se necesita un período de fermentación mayor. La fermentación puede producirse a temperaturas que se encuentran entre aproximadamente 30° C y aproximadamente 37° C, a un pH de aproximadamente 4,5 a aproximadamente 7,5.
- 50 La *Z. mobilis* de la presente puede crecer en medio que contiene azúcares mezclados que incluyen arabinosa en fermentadores de laboratorio, y en fermentación a gran escala, 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 de las cepas de *Z. mobilis* de la presente pueden producirse mediante metodologías de cultivo continuo o en lotes. Un método clásico de cultivo en lotes es un sistema cerrado donde la
- 55 composición del medio se establece al comienzo del cultivo y no se somete a alteraciones artificiales durante el proceso de cultivo. Por lo tanto, al comienzo del proceso de cultivo el medio se inocula con el organismo deseado y se deja que se produzca la actividad metabólica o crecimiento sin agregarle nada al sistema. Típicamente, sin embargo, un cultivo "en lotes" es un lote con respecto a la adición de una fuente de carbono y a menudo se intenta controlar factores tales como el pH y la concentración de oxígeno. En los sistemas en lotes, las composiciones de

biomasa y metabolito del sistema cambian de manera constante hasta el momento en que se termina el cultivo. Dentro de los cultivos en lotes, las células moderan a través de una fase latente estática a una fase logarítmica de alto crecimiento y, finalmente, a una fase estacionaria donde la velocidad de crecimiento disminuye o se detiene. Si no se tratan, las células en la fase estacionaria morirán eventualmente. Las células en la fase logarítmica a menudo son responsables de la producción a granel del producto final o intermedio en algunos sistemas. La producción de fase posexponencial o estacionaria puede obtenerse en otros sistemas.

Una variación del sistema en lotes estándar es el sistema de lote alimentado. Los procesos de cultivo de lote alimentado también son adecuados para el crecimiento de las cepas de *Z. mobilis* de la presente y comprenden un sistema en lotes típico con la excepción de que el sustrato se agrega en incrementos a medida que el cultivo evoluciona. Los sistemas de lote alimentado son útiles cuando la represión de catabolitos está apta para inhibir el metabolismo de las células y cuando se desea tener cantidades limitadas de sustrato en el medio. La medición de la concentración real del sustrato en los sistemas de lote alimentado es difícil y, por lo tanto, se calcula en base a los cambios de los factores medibles tales como el pH y la presión parcial de los gases residuales tal como CO₂. Los métodos de cultivo en lote y en lote alimentado son comunes y ampliamente 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 también puede lograrse con un cultivo continuo. Los cultivos continuos son sistemas abiertos donde un medio de cultivo definido se agrega continuamente a un biorreactor y simultáneamente se retira una cantidad igual de medio acondicionado para su procesamiento. Los cultivos continuos generalmente mantienen las células a una densidad de fase líquida alta constante, donde las células se encuentran principalmente en crecimiento de fase logarítmica. De manera alternativa, el cultivo continuo puede realizarse con células inmovilizadas donde se agrega carbono y nutrientes de manera continua, y se retiran productos valiosos, subproductos o productos residuales de manera continua de la masa celular. La inmovilización celular puede realizarse usando una amplia variedad de soportes sólidos compuestos de materiales naturales y/o sintéticos conocidos por el experto en la técnica.

El cultivo continuo o semicontinuo permite la modulación de un factor o cualquier cantidad de factores que afectan el crecimiento celular o la concentración del producto final. Por ejemplo, un método mantendrá a un nutriente limitante tal como la fuente de carbono o nivel de nitrógeno a una velocidad fija y permitirá la moderación de todos los otros parámetros. En otros sistemas, se puede alterar de manera continua una cantidad de factores que afectan el crecimiento mientras la concentración celular, medida por la turbidez del medio, se mantiene constante. Los sistemas continuos se esfuerzan para mantener condiciones de crecimiento de estado estable y, por lo tanto, la pérdida de células debido a la extracción del medio debe equilibrarse contra la velocidad de crecimiento celular en el cultivo. Los métodos para modular los nutrientes y los factores de crecimiento para los procesos de cultivo continuo así como otras técnicas para maximizar la velocidad de formación del producto son ampliamente conocidas en la técnica de microbiología industrial y Brock, *supra*, detalla una variedad de métodos.

El régimen de fermentación a continuación es particularmente adecuado para la producción de etanol. La cepa de *Z. mobilis* deseada de la presente invención se cultiva en matraces de agitación en medio semicomplejo a aproximadamente 30°C a aproximadamente 37°C con agitación a aproximadamente 150 rpm en matraces orbitales y luego se transfiere a un fermentador de siembra de 10 L que contiene un medio similar. El cultivo de siembra crece en el fermentador de siembra anaeróbicamente hasta que la OD₆₀₀ se encuentre entre 3 y 6, momento en el que se transfiere al fermentador de producción donde los parámetros de fermentación se optimizan para la producción de etanol. Los volúmenes de inoculación típicos transferidos del tanque de siembra al tanque de producción oscilan de aproximadamente 2% a aproximadamente 20% v/v. Un medio de fermentación típico contiene componentes de medio mínimos tal 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 compleja de nitrógeno tal como extracto de levadura o productos a base de soja (0-10 g/L). Una concentración final de sorbitol o manitol de aproximadamente 5 mM está presente en el medio. Los azúcares mezclados que incluyen arabinosa y al menos un azúcar adicional tal como glucosa (o sacarosa), que proporcionan una fuente de carbono, se agregan de forma continua al recipiente de fermentación tras el agotamiento de la fuente de carbono inicial en lotes (50-200 g/l) para maximizar la titulación y velocidad del etanol. Las velocidades de alimentación de la fuente de carbono se ajustan de forma dinámica para garantizar que el cultivo no esté acumulando glucosa en exceso, lo que podría provocar la acumulación de subproductos tóxicos tal como ácido acético. Con el fin de maximizar el rendimiento del etanol producido a partir de un sustrato utilizado, el crecimiento de la biomasa es restringido por la cantidad de fosfato que se colocó en lotes inicialmente o que es alimentado durante el transcurso de la fermentación. La fermentación se controla a pH 5,0-6,0 usando solución cáustica (tal como hidróxido de amonio, hidróxido de potasio o hidróxido de sodio) y ácido fosfórico o sulfúrico. La temperatura del fermentador se controla a 30°C-35°C. Para minimizar el espumado, se agregan agentes antiespumantes (de cualquier clase, con base de silicio, con base orgánica, etc) al recipiente según sea necesario. Un antibiótico, para el cual hay un marcador resistente a antibióticos en la cepa, tal como kanamicina, se puede utilizar opcionalmente para minimizar la contaminación.

Además, la fermentación puede darse al mismo tiempo que la sacarificación usando un proceso de SSF (sacarificación y fermentación simultánea). En este proceso, se producen azúcares a partir de la biomasa a medida que son metabolizadas por el biocatalizador de producción.

- 5 Cualquier conjunto de condiciones descrito anteriormente y las variaciones adicionales de estas condiciones que son ampliamente conocidas en la técnica, son condiciones adecuadas para la producción de etanol mediante una cepa de *Zymomonas* o *Zymobacter* recombinante que utiliza arabinosa, que está diseñada genéticamente para que exprese un simportador de protones de arabinosa mediante la introducción de una región codificante heteróloga de un simportador de protones de arabinosa.

Ejemplos

- 10 La presente invención se define adicionalmente en los siguientes Ejemplos. Debería entenderse que estos Ejemplos, aunque indican realizaciones preferidas de la invención, se proporcionan únicamente a modo ilustrativo. A partir de lo indicado anteriormente y de estos Ejemplos, un experto en la técnica puede determinar las características esenciales de esta invención y puede realizar varios cambios y modificaciones de la invención para adaptarla a los diversos usos y condiciones que se encuentran dentro del alcance de las reivindicaciones adjuntas.

15 Métodos Generales

- Las técnicas de clonación molecular y ADN recombinante estándar utilizadas en la presente memoria son conocidas en la técnica y fueron descritas por Sambrook, J., Fritsch, E. F. y Maniatis, T. *Molecular Cloning: A Laboratory Manual*, 2.^a ed., Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989) (de aquí en adelante "Maniatis"); y por Silhavy, T. J., Brennan, 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 ál., *Current Protocols in Molecular Biology*, publicado por Greene Publishing Assoc. y Wiley-Interscience, Hoboken, NJ (1987).

- 25 Los significados de las abreviaciones son los siguientes: "kb" significa kilobase/s, "pb" significa pares de base, "nt" significa nucleótido/s, "hr" significa hora/s, "min" significa minuto/s, "s" significa segundo/s, "d" significa día/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, "Cm^r" significa resistente a cloranfenicol, "Cm^s" significa sensible a cloranfenicol, "Sp^r" significa resistencia a espectinomicina, "Sp^s" significa sensible a espectinomicina, "UTR" significa región no traducida, "RBS" significa sitio de unión a ribosoma.

Los cebadores fueron sintetizados por Sigma (St. Luis, MO) a menos que se especifique lo contrario.

30 Ejemplo 1

Construcción y expresión de operón para proteínas con utilización de arabinosa en *Zymomonas*

- Para diseñar genéticamente *Zymomonas mobilis* para utilización de arabinosa, las regiones codificantes araA, araB y araC de *E. coli* se construyeron en un operón con un promotor de *Z. mobilis* y se expresó en un plásmido en células de *Z. mobilis*. AraB, araA y araD codifican las proteínas L-ribulosa cinasa, L-arabinosa isomerasa y L-ribulosa-5-fosfato-4-epimerasa, respectivamente, que proporcionan una vía de asimilación de arabinosa, junto con actividades de transcetolasa y transaldolasa (véase la Figura 1).

1. Clonación de secuencias codificantes araBAD de *E. coli* y promotor P_{gap} de *Z. mobilis*

- Las regiones codificantes araB, araA y araD de *E. coli* (SEQ ID NO:23, 21 y 25, respectivamente) están presentes en el operón araBAD. Se preparó un fragmento de ADN de araB-araA (araBA; SEQ ID NO:27) usando los cebadores de oligonucleótido ara1 (SEQ ID NO:28) y ara2 (SEQ ID NO:29) que son cebadores directo e inverso, respectivamente. El cebador ara1 agrega los nucleótidos CC antes del codón de inicio ATG de la región codificante araB para crear un sitio NcoI. El cebador ara2 agrega un sitio XbaI después del codón de terminación de la región codificante araA. Se preparó un fragmento de ADN de araD (SEQ ID NO:30) usando los cebadores de oligonucleótido ara3 (SEQ ID NO:31) y cebador ara4 (SEQ ID NO:32) que son cebadores directo e inverso, respectivamente. El cebador ara3 agrega un sitio XbaI en el extremo 5' de la secuencia 5' del sitio de unión al ribosoma (RBS, por sus siglas en inglés) con respecto a la región codificante araD. El cebador ara4 agrega un sitio HindIII después de la región no traducida 3' (UTR) que se encuentra en 3' con respecto a la región codificante araD. Cada par de cebadores se utilizó en una reacción PCR estándar, que incluye 50 μL de mezcla AccuPrime Pfx SuperMix (Invitrogen, Carlsbad, CA), 1 μL de cebadores directo e inverso 10 μM y 2 μL (aprox. 50 a 100 ng) de ADN genómico de *E. coli* preparado a partir de MG1655 (ATCC# 700926; una cepa de K12) usando un kit de purificación de ADN genómico Wizard (Promega, Madison, WI). Se llevó a cabo una reacción usando los cebadores ara1 y ara2 durante 5 min a 95°C, seguido de 35 ciclos de 30 s a 95°C/30 s a 56°C/3,5 min a 68°C, y finalizó durante 7 min a 68°C. Dio como resultado un fragmento de araB-araA de 3226 pb con un sitio NcoI 5' y un sitio XbaI 3' (SEQ ID NO:27). Se llevó a cabo otra reacción usando los cebadores ara3 y ara4 usando un programa similar, excepto que la extensión de tiempo a 68°C se redujo a 1,5 min. Produjo un fragmento de araD de 889 pb (que incluye la 3' UTR de araD) con un sitio XbaI 5' y un sitio HindIII 3' (SEQ ID NO:30).

El promotor de *E. coli* natural para el operón araBAD es un promotor inducible que no es adecuado para la expresión deseada en *Z. mobilis*. Se utilizó el promotor GAP (Gliceraldehídos-3-fosfato deshidrogenasa) (P_{gap} ; SEQ ID NO:33) de *Z. mobilis* ya que es un fuerte promotor constitutivo para la expresión en *Z. mobilis*. Se preparó un fragmento de ADN que contiene el P_{gap} de *Z. mobilis* usando los cebadores de oligonucleótido ara10 y ara11. El cebador ara10 (SEQ ID NO:34) es un cebador directo que agrega un sitio SacI y ApeI en el extremo 5' del fragmento de ADN del promotor. El cebador ara11 (SEQ ID NO:35) es un cebador inverso que cambia los últimos dos nucleótidos del promotor de AC a CC, por lo tanto, agrega un sitio NcoI en el extremo 3' del fragmento de ADN del promotor. Estos dos cebadores se utilizaron en una reacción PCR estándar, tal como se describió anteriormente, usando un plásmido que contiene el P_{gap} como la plantilla de ADN para producir un fragmento de ADN del promotor P_{gap} de 323 pb con sitios SacI y SpeI 5' y un sitio NcoI 3' (SEQ ID NO:36).

Cada uno de estos productos de PCR se clonó en el vector TOPO Blunt Zero (Invitrogen, Calsbad, CA) siguiendo las instrucciones del fabricante. Los plásmidos resultantes pTP-araB-araA, pTP-araD y pTP- P_{gap} se propagaron en las células DH5a de *E. coli* (Invitrogen) y cada una se preparó usando un Kit Qiagen DNA Miniprep. Sus secuencias se confirmaron mediante secuenciamiento de ADN. 2. Ensamblaje de un operón P_{gap} -araBAD en un vector transportador.

Se ensambló un operón P_{gap} -araBAD en un vector transportador de *Zymomonas-E. coli* denominado pZB188aada, que se basa en el vector pZB188 (Zhang et ál. (1995) Science 267:240-243; US 5514583) que incluye un fragmento de ADN genómico de 2582 pb de *Z. mobilis* que contiene una región de replicación que permite que el vector se replique en células de *Zymomonas*. En pZB188aada el cassette de resistencia a tetraciclina (Tc^r -cassette) de pZB188 se reemplazó con un cassette de resistencia a espectinomicina ($Spec^r$ -cassette). El $Spec^r$ -cassette se generó mediante PCR usando el plásmido pHP15578 (Cahoon et ál, (2003) Nature Biotechnology 21: 1082-1087) como la plantilla y los cebadores 1 (SEQ ID NO:32 de CL4236) y 2 (SEQ ID NO:33 de CL4236). El plásmido pHP15578 contiene la secuencia de nucleótidos completa para el $Spec^r$ -cassette y su promotor, que se basa en la secuencia publicada del gen de aadA Transposon Tn7 (número de acceso GenBank X03043) que codifica la 3' (9)-O-nucleotidiltransferasa.

Cebador 1 (SEQ ID NO:37):

CTACTCATTTatcgatGGAGCACAGGATGACGCCT

Cebador 2 (SEQ ID NO:38):

CATCTTACTacgcgtTGGCAGGTCAGCAAGTGCC

Las bases subrayadas del cebador 1 (cebador directo) se hibridan corriente arriba del promotor para el $Spec^r$ -Cassette (a los nts 4-22 del número de acceso GenBank X03043), mientras las letras en minúscula corresponden a un sitio ClaI que se agregó en el extremo 5' del cebador. Las bases subrayadas del cebador 2 (cebador inverso) se hibridan aproximadamente 130 bases corriente abajo del codón de terminación para el $Spec^r$ -Cassette (a los nts 1002-1020 del número de acceso GenBank X03043), mientras las letras en minúscula corresponden a un sitio AflIII que se agregó en el extremo 5' del cebador. El $Spec^r$ -cassette generado por PCR de 1048 pb se digirió doblemente con ClaI y AflIII, y el fragmento de ADN resultante se purificó usando el kit de purificación de PCR QIAquick (Qiagen, Cat. n.º 28104) y el protocolo recomendado del vendedor. El plásmido pZB188 (aislado de *E. coli* SSC110 (dcm⁻, dam⁻) para obtener ADN de plásmido no metilado para cortar con ClaI (que es sensible a la metilación dam) se digirió doblemente con ClaI y BssHII para retirar el Tc^r -cassette, y el gran vector resultante se purificó mediante electroforesis en gel de agarosa. Este fragmento de ADN y el producto de PCR limpio luego se ligaron entre sí, y la mezcla de reacción de transformación se introdujo en JM110 de *E. coli* usando células químicamente competentes que se obtuvieron de Stratagene (Cat. n.º 200239). Debe observarse que BssHII y AflIII generan "extremos pegajosos" compatibles, pero ambos sitios se destruyen cuando están ligados entre sí. Los transformantes se colocaron en placas en medio LB que contenía espectinomicina (100 µg/ml) y se cultivaron a 37°C. Un transformante resistente a espectinomicina que contenía un plásmido con una inserción del tamaño correcto se identificó mediante el análisis de digestión de restricción con NotI y denominado pZB188/aada.

El fragmento pTP- P_{gap} SpeI - NcoI P_{gap} , el fragmento pTP-araB-araA NcoI - XbaI araB-araA y el fragmento pTP-araD XbaI - NotI araD se clonaron en un vector de NotI-SpeI pZB188/aada, formando un vector transportador basado en pZB188aada que contenía un operón P_{gap} -araBAD. El plásmido resultante, denominado pARA201, se propagó en DH5a de *E. coli* y se preparó usando un kit Qiagen DNA Miniprep. Se preparó pARA205 (Figura 2; SEQ ID NO:41) a partir de pARA201 mediante restauración de los nucleótidos en el extremo 3' de P_{gap} desde CC nuevamente hasta los nucleótidos AC originales. Esto se realizó usando un kit de mutagénesis de sitio dirigido QuickChange XL (Stratagene, La Jolla, CA). Para esta mutagénesis, el cebador directo ara31 (SEQ ID NO:30) y el cebador inverso ara32 (SEQ ID NO:40) se utilizaron para realizar los cambios siguiendo las instrucciones del fabricante. pARA205 se propagó en DH5a de *E. coli* y se preparó usando un kit Qiagen DNA Miniprep.

3. Expresión de araBAD en *Z. mobilis*

Para confirmar que P_{gap} -araBAD es un operón funcional en *Z. mobilis*, se introdujo pARA205 en la cepa ZW801-4 de *Z. mobilis* para su expresión. ZW801-4 es una cepa de *Z. mobilis* que utiliza xilosa. La construcción y caracterización

de las cepas ZW658, ZW800 y ZW801-4 se describió en la publicación de solicitud de patente de EE.UU., de titularidad conjunta y en tramitación con la presente US20080286870 A1. Se construyó ZW658 (ATCC # PTA-7858) mediante la integración de dos operones, *P_{gap}xylAB* y *P_{gap}taitkt*, que contienen cuatro genes que utilizan xilosa que codifican la xilosa isomerasa, xilulocinasa, transaldolasa y transcetolasa, dentro del genoma de ZW1 (ATCC #31821) a través de eventos de transposición secuencial, y seguido de la adaptación en medio selectivo que contiene xilosa. ZW800 es un derivado de ZW658 que tiene una inserción de entrecruzamiento doble de un cassette de resistencia a espectinomicina en la secuencia que codifica la enzima glucosa-fructosa oxidoreductasa (GFOR) para inactivar esta actividad. ZW801-4 es un derivado de ZW800 en el cual se ha eliminado el cassette de resistencia a espectinomicina mediante recombinación específica del sitio dejando un codón de terminación en el marco que trunca la proteína de forma prematura.

Las células competentes de ZW801-4 se prepararon mediante cultivo de las células de siembra durante la noche en MRM3G5 (extracto de levadura al 1%, KH₂PO₄ 15 mM, MgSO₄ 4 mM y 50 g/L de glucosa) a 30°C con agitación a 150 rpm, hasta un valor de OD₆₀₀ cercano a 5. Las células se cosecharon y se volvieron a suspender en medio fresco hasta un valor de OD₆₀₀ de 0,05. Se cultivaron adicionalmente en las mismas condiciones hasta la fase logarítmica temprana o media (OD₆₀₀ cercana a 0,5). Las células se cosecharon y se lavaron dos veces con agua helada y luego una vez con glicerol helado al 10%. Las células competentes resultantes se recolectaron y se volvieron a suspender en glicerol helado al 10% hasta un valor de OD₆₀₀ cercano a 100. Dado que la transformación de *Z. mobilis* requiere ADN no metilado, el plásmido pARA205 se transformó en células competentes SCS110 de *E. coli* (Stratagene). Una colonia de células transformadas se cultivó 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, usando un kit Qiagen DNA Miniprep.

Aproximadamente 500 ng de ADN de plásmido pARA205 no metilado se mezcló con 50 µL de células competentes ZW801-4 en una cubeta de electroporación 1 MM (VWR, West Chester, PA). El ADN de plásmido se sometió a electroporación en las células a 2,0 KV usando un BT720 Transporater Plus (BTX-Genetronics, San Diego, CA). Las células transformadas se recuperaron en 1 mL de medio MMG5 (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 de MgSO₄) durante 4 horas a 30°C y se cultivaron en placas MMG5-Spec250 (MMG5 con 250 mg/L de espectinomicina y 15 g/L de agar) durante 2 días a 30°C, dentro de un frasco anaeróbico con AnaeroPack (Mitsubishi Gas Chemical, Nueva York, NY). Las colonias individuales se estrilaron en una placa MMA5-Spec250 (igual que MMG5-Spec250 pero la glucosa se reemplazó con 50 g/L de arabinosa) y una nueva placa MMG5-Spec250 en duplicado. En las mismas condiciones que se describieron anteriormente, las estrías crecieron bien aunque el crecimiento en la placa MMA5-Spec250 llevó más tiempo. Esto indicó la expresión del operón *P_{gap}-araBAD*.

Dos estrías de las células transformadas que crecían en la placa MMG5-Spec250 (ZW801-ara205-4 y ZW801-ara205-5) se seleccionaron para un ensayo de crecimiento de 72 horas. En el ensayo, las células de cada estría se cultivaron durante la noche en 2 mL de MRM3G5-Spec250 (MRM3G5 con 250 mg/L de espectinomicina) a 30°C con agitación de 150 rpm. Las células se cosecharon, se lavaron con MRM3A5 (igual que MRM3G5 pero la glucosa se reemplazó con arabinosa), y se volvieron a suspender en MRM3A5-Spec250 (MRM3A5 que contiene 250 mg/L de espectinomicina) para tener una OD₆₀₀ de inicio a 0,1. Se colocaron cuatro mL de la suspensión en un tubo Falcon con tapa de 14 mL y se cultivaron durante 72 horas a 30°C con agitación de 150 rpm. Al finalizar el crecimiento, se midió la OD₆₀₀. Luego, 1 mL del cultivo se centrifugó a 10.000x g para retirar las células. El sobrenadante se filtró a través de un filtro de tubo de centrifuga Costar Spin-X de 0,22 µm (Corning Inc, Corning, NY) y se analizó mediante una columna de exclusión iónica BioRad Aminex HPX-A7H (BioRad, Hercules, CA) con H₂SO₄ 0,01 N a una velocidad de 0,6 mL/min a 55°C en un sistema de HPLC Agilent 1100 (Agilent Technologies, Santa Clara, CA) para determinar las concentraciones de etanol y azúcar. En paralelo, se cultivó ZW801-4 (sin antibióticos) y se analizó como control. Los resultados proporcionados en la Tabla 2 demuestran que la expresión de *araBAD* permitió que ZW801-4 de *Z. mobilis* creciera y produjera etanol usando arabinosa como la única fuente de carbono.

Tabla 2 Ensayo de crecimiento de 72 horas para cepas ZW801-ara205 en MRM3A5

Cepa	Crecimiento (OD ₆₀₀)	Etanol (g/L)	Arabinosa (g/L)
ZW801-4	0,106	0	51,20
ZW801-ara205-4	1,75	7,22	33,15
ZW801-ara205-5	1,96	10,68	27,16

Ejemplo 2

Integración de operón de utilización de arabinosa en el genoma de *Z. mobilis* y caracterización de las cepas resultantes

Este ejemplo describe la integración estable del operón *P_{gap}-araBAD* en dos cepas de *Z. mobilis* que utilizan xilosa. 1. Construcción del operón *P_{gap}-araBAD* en un vector suicida.

Para integrar el operón P_{gap} -araBAD dentro del genoma de *Z. mobilis*, se preparó un vector suicida para recombinación homóloga de DCO (entrecruzamiento doble). Además de P_{gap} -araBAD, este vector incluía fragmentos de recombinación homóloga de DCO para dirigir la integración de P_{gap} -araBAD y un gen *aadA* para proporcionar un marcador selectivo para la resistencia a espectinomicina. Se eligió el locus de *ldhA* como el sitio de inserción. Dos fragmentos de ADN de *ldhA* para DCO, LDH-L y LDH-R, se sintetizaron mediante PCR usando el ADN de ZW801-4 de *Z. mobilis* como plantilla. La reacción utilizó la mezcla AccuPrime y siguió el procedimiento de PCR estándar descrito en el Ejemplo 1. El fragmento de ADN LDH-L DNA se sintetizó utilizando el cebador directo ara20 (SEQ ID NO:42) y el cebador inverso ara21 (SEQ ID NO:43). El producto resultante fue un fragmento de ADN de 895 pb que incluye la secuencia 5' con respecto a la región codificante de *ldhA* y los nucleótidos 1-493 de la región codificante de *ldhA*, con un sitio *SacI* 5' y un sitio *SpeI* 3' (SEQ ID NO:44). El fragmento de ADN LDH-R DNA se sintetizó utilizando el cebador directo ara22 (SEQ ID NO:45) y el cebador inverso ara23 (SEQ ID NO:46). El producto resultante fue un fragmento de 1169 pb que incluye los nucleótidos 494-996 de la región codificante de *ldhA* y la secuencia 3' con respecto a la región codificante de *ldhA*, con un sitio *EcoRI* 5' y un sitio *NotI* 3' (SEQ ID NO:47).

Se utilizó pBS SK(+) (un plásmido Bluescript plasmid; Stratagene) como un vector suicida ya que los vectores pBS no pueden replicarse en *Zymomonas*. Se construyó pARA354 (SEQ ID NO:49) al clonar el operón P_{gap} -araBAD de pARA205, el fragmento LDH-L y el fragmento LDH-R en pBS SK(+). Además, un fragmento de ADN que contiene el marcador *aadA* (para resistencia a espectinomicina) unido mediante sitios *LoxP* de tipo salvaje (fragmento *LoxPw*-*aadA*-*LoxPw*; SEQ ID NO:48) se incluyó en pARA354. pARA354 tiene el operón P_{gap} -araBAD y el fragmento de marcador *LoxPw*-*aadA*-*LoxPw* ubicados entre las secuencias de LDH-L y LDH-R.

La Figura 3 muestra un mapa de pARA354 de 10.441 pb. Tiene un origen de *f1*(+) y un gen de resistencia a ampicilina para la propagación del plásmido en *E. coli*. Dado que LDH-L y LDH-R contenían los primeros 493 pares de bases y los 503 pares de bases restantes de la secuencia codificante de *ldhA*, respectivamente, se diseñó pARA354 para dirigir la inserción de P_{gap} -araBAD y *aadA* en la secuencia codificante de *ldhA* de *Z. mobilis* entre los nucleótidos #493 y #494 mediante recombinación por entrecruzamiento.

2. Desarrollo de las cepas de integración de P_{gap} -araBAD

La cepa ZW705 de *Z. mobilis* es una cepa de *Z. mobilis* diseñada genéticamente, con utilización de xilosa mejorada en condiciones de estrés que derivó de ZW801-4 mediante adaptación en cultivo continuo tal como se describe en la solicitud de patente de EE.UU., de titularidad conjunta y en tramitación con la presente, US 20110014670. Las células de ZW801-4 de *Zymomonas* que utilizan xilosa se cultivaron de forma continua en medio que comprende al menos aproximadamente 50 g/L de xilosa para producir un cultivo que comprende etanol, luego se agregó amoníaco y ácido acético, creando un cultivo de estrés. Las células se cultivaron adicionalmente de forma continua en el cultivo de estrés y se aislaron las células con utilización de xilosa mejorada, incluyendo la cepa ZW705.

Para transformar pARA354 en ambas, las cepas ZW705 y ZW801-4, se sometieron a electroporación 800 ng de ADN de plásmido no metilado en 50 μ L de células competentes preparados de cada cepa. La desmetilación de ADN, la preparación de células competentes y la electroporación se realizaron tal como se describe en el Ejemplo 1. Las colonias de células transformadas de cada cepa se cultivaron en una placa MMG5-Spec250 durante 2 días a 30°C dentro de un frasco anaeróbico con un AnaeroPack. Dado que pARA354 no pudo replicarse en *Z. mobilis*, la resistencia a espectinomicina indicó que estas colonias eran cepas de integración. Las colonias se estrilaron en una nueva placa MMG5-Spec250 y una placa MMA5-Spec250, en duplicado, y se cultivaron durante 2 días y 4 días respectivamente. Su crecimiento en la placa MMA5-Spec250 también indicó la integración. Para demostrar adicionalmente la integración, las uniones entre el fragmento P_{gap} -araBAD-*aadA* y el ADN genómico de *Z. mobilis* se inspeccionaron mediante la reacción PCR de 35 ciclos estándar, que contenía la mezcla PCR Super Mix (Invitrogen), un par de cebadores y las células transformadas analizadas. Un ciclo de PCR incluía 45 segundos de desnaturalización a 95°C, 45 segundos de hibridación a 58°C y 2 minutos de extensión a 72°C. El cebador ara45 (SEQ ID NO:50) y cebador ara42 (SEQ ID NO:51) eran un cebador directo ubicado corriente arriba de la secuencia de LDH-L en el ADN genómico de *Z. mobilis* y un cebador inverso ubicado en el gen *araB* de pARA354, respectivamente. Este par de cebadores amplificó un fragmento de 1694 pb de todas las colonias inspeccionadas mediante PCR. También se utilizó el cebador ara46 (SEQ ID NO:52) y el cebador ara43 (SEQ ID NO:53) que son un cebador directo ubicado en el gen *aadA* de pARA354 y un cebador inverso ubicado corriente abajo de la secuencia de LDH-R en el ADN genómico de *Z. mobilis*, respectivamente. Este par de cebadores amplificó un fragmento de 1521 pb de todas las colonias inspeccionadas mediante PCR. Por lo tanto, el fragmento P_{gap} -araBAD-*aadA* se ha integrado en los genomas de ZW801-4 y ZW705 de manera satisfactoria mediante el enfoque de DCO. Dado que la recombinación homóloga de DCO fue una integración específica objetivo, cada colonia que resulta de la integración en ZW801-4 o ZW705 tendría el genotipo idéntico. Una colonia de cada una de las integraciones se cultivó en 5 mL de MRMG5-Spec250 durante la noche a 30°C con agitación de 150 rpm. Las células se recolectaron mediante centrifugación, se volvieron a suspender en 0,5 mL de glicerol al 50%, y luego se almacenaron a -80°C. Las cepas se denominaron ZW705-ara354 y ZW801-ara354.

Para mejorar aun más la función del operón P_{gap} -araBAD integrado, la cepa ZW705-ara354 se sometió a adaptación. Con este propósito, se recolectó un cultivo de ZW705-ara354 durante la noche mediante centrifugación, se lavó con MRM3A5 y se volvió a suspender en MRM3A5-Spec250 con OD₆₀₀ a 0,1. Cuatro mL de esta suspensión se colocaron en un tubo Falcon con tapa de 14 mL y se cultivaron durante 72 horas en un agitador a 30°C y 150 rpm,

hasta que la OD₆₀₀ quedara por encima de 1. Luego, el cultivo se inoculó en un nuevo tubo falcon que contenía 4 mL de MRM3A5-Spec250 fresco para lograr una OD₆₀₀ de partida cercana a 0,1 para una segunda ronda de crecimiento. En total, se completaron 9 rondas sucesivas. Cada ronda llevó la OD₆₀₀ de aproximadamente 0,1 a por encima 1 y tomó 3 a 4 días, excepto la 4.^a ronda que tomó 6 días ya que las células crecían más lentamente. Para caracterizar las cepas adaptadas, la 9.^a ronda se diluyó 100 veces, y se dispersaron 10 µL de la dilución y se cultivaron en una placa MMA5-Spec250 durante 3 días a 30°C en un frasco anaeróbico con un AnaeroPack. Se seleccionaron colonias individuales (es decir, cepas de adaptación) y se cultivaron durante la noche en 3 mL de MRM3G5-Spec250 en un agitador a 30°C y 150 rpm. Se sometieron al ensayo de crecimiento de 72 horas en MRM3A5-Spec250, tal como se describe en el Ejemplo 1. Se utilizó la cepa ZW705-ara354 como control en el ensayo. Los datos del análisis para 5 cepas de adaptación (ZW705-ara354A4 a A8) se presentan en la Tabla 3, que demuestran que todas las cepas de adaptación tuvieron un mejor rendimiento que ZW705-ara354. ZW705-ara354A7 fue la mejor cepa en términos de crecimiento, producción de etanol y utilización de arabinosa.

Tabla 3. Ensayo de crecimiento de 72 horas para cepas de adaptación de ZW705-ara354 en MRM3A5

Cepa	Crecimiento (OD ₆₀₀)	Etanol (g/L)	Arabinosa (g/L)
ZW705-ara354	1,03	9,10	32,71
ZW705-ara354A4	3,29	19,03	10,31
ZW705-ara354A5	3,71	18,56	10,07
ZW705-ara354A6	3,61	18,47	9,23
ZW705-ara354A7	4,04	19,73	7,36
ZW705-ara354A8	2,96	17,37	12,18

3. Caracterización de perfiles de crecimiento y metabolitos de las cepas de integración de P_{gap}-araBAD, con y sin adaptación.

Las cepas de integración de P_{gap}-araBAD se caracterizaron adicionalmente por su capacidad de utilizar arabinosa para permitir el crecimiento celular y la producción de etanol en medio que contiene arabinosa como la única fuente de carbono y en medio que contiene azúcares mezclados. Para caracterizar estas cepas en medio que contiene arabinosa como la única fuente de carbono, primero se cultivaron células de ZW705-ara354 y ZW705-ara354A7 durante la noche en 2 mL de MRM3G5-Spec250 en un agitador a 30°C y 150 rpm. Las células se cosecharon, se lavaron con MRM3A5 y se volvieron a suspender en MRM3A5-Spec250 a una OD₆₀₀ de partida de 0,1. Veinte mL de la suspensión se colocaron en un tubo de centrifuga VWR con tapa a rosca de 50 mL y se cultivaron a 30°C con agitación de 150 rpm durante un período de tiempo de 96 horas. Durante el período de tiempo, se midió la OD₆₀₀ a las 0, 24, 48, 72 y 96 horas, respectivamente. En cada punto de tiempo, se extrajo 1 mL del cultivo y se centrifugó a 10.000x g para retirar 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 para determinar las concentraciones de azúcar y etanol mediante una columna de exclusión iónica BioRad Aminex HPX-A7H con H₂SO₄ 0,01 N utilizando una velocidad de 0,6 mL/min a 55°C en un sistema de HPLC Agilent 1100. En paralelo, se cultivó ZW705 en medio sin antibióticos y se analizó como control. Los resultados se muestran en la Figura 4. Estos resultados indican que, sin P_{gap}-araBAD, ZW705 no podría metabolizar la arabinosa y no podría crecer cuando la arabinosa es la única fuente de carbono (Figura 4A). Después de la integración of P_{gap}-araBAD, ZW705-ara354 pudo utilizar arabinosa para permitir el crecimiento y producir etanol (Figura 4B). La velocidad máxima de consumo de arabinosa fue 0,2 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 32,8%, a 34 g/L. La adaptación mejoró ampliamente la utilización de arabinosa, el crecimiento celular y la producción de etanol en ZW705-ara354A7. La velocidad máxima de consumo de arabinosa fue 0,73 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 83,4%, a 8,4 g/L.

Para caracterizar las cepas en un medio que contiene azúcares mezclados, se cultivaron ZW705, ZW705-ara354 y ZW705-ara354A7 y se analizaron tal como se describió anteriormente, pero el medio MRM3A5 utilizado en el experimento previo se reemplazó con medio MRM3A2.5X2.5G5 (MRM3 con 25 g/L de arabinosa, 25 g/L de xilosa y 50 g/L de glucosa). Debido al rápido crecimiento en MRM3A2.5X2.5G5, se agregó un punto de tiempo a las 10 horas. El análisis se realizó tal como se describió anteriormente para el experimento usando medio de arabinosa. Los resultados se muestran en la Figura 5. Estos resultados muestran que ZW705 utilizó glucosa y xilosa de forma eficaz para permitir el fuerte crecimiento celular y producción de etanol, pero no pudo metabolizar arabinosa (Figura 5A). Después de la integración of P_{gap}-araBAD, ZW705-ara354 pudo utilizar arabinosa para mejorar el crecimiento celular y la producción de etanol (Figura 5B). La velocidad máxima de consumo de arabinosa fue 0,3 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 67,9%, a 8,8 g/L. En la cepa adaptada ZW705-ara354A7 hubo cierta mejora con respecto a la cepa ZW705-ara354 en utilización de arabinosa, que permitió mejor crecimiento y producción de etanol. La velocidad máxima de consumo de arabinosa fue 0,36 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 74,1%, a 7,1 g/L.

Ejemplo 3

Construcciones para la expresión de dos sistemas de transporte de arabinosa de *E. coli* en *Zymomonas*

Cada uno de los dos sistemas de transporte de arabinosa que están presentes en *E. coli*, codificado por araE o por araFGH, se expresó en *Zymomonas* y se analizó la utilización de arabinosa. araE codifica un simportador de protones de arabinosa mientras araFGH codifica tres proteínas que forman un transportador ABC.

1. Construcción del gen quimérico araE y el operón araFGH para la expresión en *Zymomonas*

Se prepararon fragmentos de ADN de secuencia codificante de araFGH y araE de *E. coli* mediante PCR de 30 ciclos estándar, tal como se describe en el Ejemplo 1, usando el ADN de MG1655 de *E. coli* (una cepa de K12: ATCC #700926) como plantilla. Cada ciclo incluía 45 s de desnaturalización a 94°C, 45 s de hibridación a 60°C y 4 min de extensión a 72°C. Se utilizó un cebador directo ara135 (SEQ ID NO :54) y un cebador inverso ara136 (SEQ ID NO :55) en PCR para sintetizar un fragmento de araE de 1550 pb, que incluye la secuencia codificante de araE (1419 pb) y su 3'UTR (121 bp), agregando un sitio NcoI en el extremo 5' y un sitio EcoRI en el extremo 3' (SEQ ID NO :56). Se utilizó un cebador directo ara137 (SEQ ID NO :57) y un cebador inverso ara138 (SEQ ID NO :58) en PCR para sintetizar un fragmento de araFGH de 3744 pb (SEQ ID NO :59). Este fragmento era idéntico al operón araFGH de *E. coli* pero sin el promotor. Incluía la secuencia codificante de araF, secuencia codificante de araG, secuencia codificante de araH, 3'UTR de araH y regiones intergénicas intactas. Los cebadores agregaron un sitio NcoI 5' y un sitio EcoRI 3'.

El promotor GI de *Actinoplanes missouriensis* (P_{gi}) se seleccionó para dirigir la expresión de araE y araFGH. Es el promotor del gen de xilosa isomerasa y se ha demostrado que funciona en *Z. mobilis* como un promotor constitutivo débil. Para clonar un P_{gi} de *A. missouriensis*, se diseñó un par de cebadores de oligonucleótido. El cebador ara12 (SEQ ID NO :60) era el cebador directo para PCR de P_{gi} que agregaba un sitio SacI y SpeI en el extremo 5' del promotor. El cebador ara13 (SEQ ID NO :61) era el cebador inverso para PCR de P_{gi} que agregaba un sitio NcoI en el extremo 3' del promotor. Estos dos cebadores se utilizaron en una reacción PCR estándar y se utilizó un plásmido que contiene el promotor GI de *Actinoplanes missouriensis* (SEQ ID NO:62) como el ADN plantilla. La reacción PCR produjo un fragmento de ADN de P_{gi} de 201 pb (SEQ ID NO:63) con los sitios SacI y SpeI 5' y un sitio NcoI 3' que se clonó en el vector TOPO Blunt Zero (Invitrogen, Calsbad, CA) siguiendo las instrucciones del fabricante. El plásmido resultante pTP- P_{gi} se propagó en DH5a de *E. coli* y se preparó el ADN de plásmido usando un kit Qiagen DNA Miniprep.

El fragmento SpeI - NcoI P_{gi} de pTP- P_{gi} y el fragmento de PCR de araE NcoI - EcoRI se combinaron en un vector pZB188/aada junto con un marcador de resistencia a cloranfenicol (CM-R; SEQ ID NO:64) creando pARA112 (Figura 6; SEQ ID NO:65). pARA112 contiene un gen quimérico P_{gi} -araE en el vector transportador de *E. coli*/*Zymomonas* derivado por pZB188. El fragmento SpeI - NcoI P_{gi} de pTP- P_{gi} y el fragmento de PCR NcoI - EcoRI araFGH se combinaron en un vector pZB188/aada junto con un marcador de resistencia a cloranfenicol creando pARA113 (Figura 7; SEQ ID NO:66). Los vectores transportadores resultantes se propagaron en DH5a de *E. coli* y se preparó el ADN de plásmido usando un kit Qiagen DNA Miniprep. El gen P_{gi} -araE y operón P_{gi} -araFGH se confirmaron mediante secuenciamiento.

Ejemplo 4

Expresión de los sistemas de transporte de arabinosa de *E. coli* en ZW705-ara354A7 de *Zymomonas*

Se analizaron los efectos de los dos sistemas de transporte de arabinosa de *E. coli* en células de *Zymomonas* que utilizan arabinosa mediante la expresión del gen P_{gi} -araE y operón P_{gi} -araFGH construidos.

1. Transformación de ZW705-ara354A con pARA112 y pARA113.

pARA112 que contiene el gen P_{gi} -araE y pARA113 que contiene el operón P_{gi} -araFGH, ambos preparados en el Ejemplo 3, se transformaron en células de ZW705-ara354A7 (preparadas en los Ejemplos 1 y 2). Las células competentes de la cepa ZW705-ara354A7 se prepararon tal como se describió en el Ejemplo 1. Dado que la transformación de *Z. mobilis* requiere ADN no metilado, pARA112 y pARA113 fueron transformados cada uno en células competentes SCS110 de *E. coli* y se preparó el ADN de plásmido no metilado a partir de un cultivo de 10 mL de una sola colonia usando un kit Qiagen DNA Miniprep. Aproximadamente 500 ng de cada ADN de plásmido se mezcló por separado con 50 μ L de células competentes ZW705-ara354A7 en una cubeta de electroporación VWR 1 MM y se sometió a electroporación en las células a 2,0 KV usando un BT720 Transporater Plus.

Las células transformadas por pARA112 o pARA113 (ZW705-ara354A7-ara112 y ZW705-ara354A7-ara113) se recuperaron en 1 mL de medio MMG5 durante 4 horas a 30°C y luego se cultivaron en placas MMG5-CM120 (MMG5 con 120 mg/L de cloranfenicol y 15 g/L de agar) durante 2 días a 30°C dentro de un frasco anaeróbico con un AnaeroPack. Las colonias individuales se estríaron en una nueva placa MMG5-CM120 y se dejaron crecer en las mismas condiciones que la última etapa. Las estrías crecieron en las placas que contenían cloranfenicol, lo que indica una transformación satisfactoria. 2. Expresión de P_{gi} -araE y P_{gi} -araFGH en las cepas transformadas.

Se seleccionaron varias estirpes de las cepas transformadas de las placas MMG5-CM120 para representar ZW705-ara354A7-ara112 y ZW705-ara354A7-ara113. Se inspeccionó la expresión de P_{gi} -araE o P_{gi} -araFGH mediante el ensayo de crecimiento de 72 horas descrito en el Ejemplo 1. En este ensayo, las células de cada estirpe se cultivaron durante la noche en 2 mL de MRM3G5-CM120 (MRM3G5 con 120 mg/L de cloranfenicol) a 30°C con agitación de 150 rpm. Las células se cosecharon, se lavaron con MRM3A5 y se volvieron a suspender en MRM3A5-CM120 (MRM3A5 que contiene 120 mg/L de cloranfenicol) a una OD₆₀₀ de partida de 0,1. Se cultivaron cuatro mL de la suspensión durante 72 horas a 30°C con agitación de 150 rpm. Al final del crecimiento, se midió la OD₆₀₀ y se analizaron los perfiles de metabolitos usando una columna de exclusión iónica BioRad Aminex HPX-A7H en un sistema de HPLC Agilent 1100 tal como se describió en el Ejemplo 1. Como control, la cepa ZW705-ara354A7 se cultivó y analizó en paralelo con Spec250 reemplazando a CM120. Los resultados para 3 cepas en cada transformación se proporcionan en la Tabla 4.

Tabla 4. Ensayo de crecimiento de 72 horas para ZW705-ara354A7-ara112 y ZW705-ara354A7-ara113 en MRM3A5.

Cepa	Crecimiento (OD600)	Etanol (g/L)	Arabinosa (g/L)
ZW705-ara354A7	3,01	18,57	5,98
ZW705-ara354A7-ara112-1	3,28	19,22	0,43
ZW705-ara354A7-ara112-2	3,33	21,38	0,34
ZW705-ara354A7-ara112-3	3,20	19,65	0,40
ZW705-ara354A7-ara113-5	2,51	16,64	11,95
ZW705-ara354A7-ara113-6	2,12	15,65	15,97
ZW705-ara354A7-ara113-7	2,17	15,32	13,91

En comparación con sus originales, todas las cepas ZW705-ara354A7-ara112 utilizaron más arabinosa durante 72 horas de cultivo, lo que permitió un mayor nivel de crecimiento y producción de etanol. De hecho, estas cepas ZW705-ara354A7-ara112 consumieron casi toda la arabinosa disponible en el medio. Esto indica que araE facilitó la utilización de arabinosa en las cepas diseñadas genéticamente. Por otro lado, parece que la expresión de araFGH tuvo un impacto negativo. Dio como resultado menor utilización de arabinosa, un menor nivel de crecimiento y menor producción de etanol en las cepas ZW705-ara354A7-ara113 durante 72 horas de cultivo.

3. Caracterización de perfiles de crecimiento y metabolitos de la cepa ZW705-ara354A7-ara112.

Dado que las cepas ZW705-ara354A7-ara112 mostraron un metabolismo de arabinosa facilitado, estas cepas se analizaron adicionalmente. La caracterización se realizó siguiendo el procedimiento descrito en el Ejemplo 2.3. Dado que araE se expresó a partir de un vector transportador, el nivel de expresión puede variar entre cepas diferentes. Por lo tanto, se examinaron dos cepas (ZW705-ara354A7-ara112-2 y ZW705-ara354A7-ara112-3) una al lado de la otra. Para caracterizar las cepas en el medio de azúcar único (arabinosa), los cultivos de ZW705-ara354A7-ara112-2 y ZW705-ara354A7-ara112-3 cultivados durante la noche se cosecharon, se lavaron con MRM3A5 y se volvieron a suspender en MRM3A5-CM120 hasta una OD₆₀₀ de partida de 0,1. Veinte mL de las suspensiones se cultivaron a 30°C con agitación de 150 rpm durante un período de tiempo de 96 horas. La OD₆₀₀ se midió a las 0, 6, 12, 24, 48, 72 y 96 horas. En cada punto de tiempo, se analizaron los perfiles de metabolitos usando una columna de exclusión iónica BioRad Aminex HPX-A7H en un sistema de HPLC Agilent 1100. En paralelo, la cepa original de ZW705-ara354A7 se cultivó en 250 mg/L de espectinomicina en lugar de 120 mg/L de cloranfenicol y se analizó como control. Los resultados se muestran en la Figura 8. Estos resultados indican que, sin P_{gi} -araE, ZW705-ara354A7 utilizó arabinosa con una velocidad máxima de 0,93 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 80,4%, a 9,81 g/L. Con expresión de araE, ZW705-ara354A7-ara112-2 y ZW705-ara354A7-ara112-3 utilizaron arabinosa de forma más eficaz, lo que permitió mayores niveles de crecimiento y producción de etanol. Las velocidades máximas del consumo de arabinosa aumentó a 1,18 g/Uhr y 1,28 g/Uhr en las cepas 112-2 y 112-3, respectivamente. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 98%, a 1,02 g/L para ZW705-ara354A7-ara112-2 y en 99,2%, a 0,41 g/L para ZW705-ara354A7-ara112-3. De hecho, ZW705-ara354A7-ara112-2 y ZW705-ara354A7-ara112-3 casi agotaron toda la arabinosa disponible después del cultivo de 72 horas y 48 horas, respectivamente.

Para caracterizar las cepas en un medio que contiene azúcares mezclados, se cultivaron ZW705-ara354A7, ZW705-ara354A7-ara112-2 y ZW705-ara354A7-ara112-3 y analizaron tal como se describió anteriormente pero usando el medio MRM3A2.5X2.5G5. Los resultados se muestran en la Figura 9. Estos resultados muestran que ZW705-ara354A7 agotó de manera eficaz toda la glucosa y xilosa dentro de 24 horas para permitir un fuerte crecimiento y producción de etanol. Su metabolismo de arabinosa fue relativamente más lento e incompleto. La velocidad máxima de consumo de arabinosa fue 0,43 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 62,4%, a 9 g/L. Sin embargo, ZW705-ara354A7-ara112-2 y ZW705-ara354A7-ara112-3 utilizaron arabinosa de forma mucho más eficaz. Las velocidades máximas de consumo de arabinosa

aumentaron a 0,73 g/L/hr y 0,78 g/L/hr, respectivamente. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 90,3%, a 2,33 g/L para ZW705-ara354A7-ara112-2 y en 90,1%, a 2,38 g/L para ZW705-ara354A7-ara112-3. En realidad se había reducido cercano a este nivel dentro de 48 horas en ambas cepas. Por lo tanto, la expresión de araE también facilitó la utilización de arabinosa en el medio de azúcares mezclados, lo que contribuyó a la producción de etanol en la Figura 9. La expresión no tuvo un efecto significativo en el metabolismo de glucosa, pero enlenteció el metabolismo de xilosa de modo que a ambas cepas ZW705-ara354A7-ara112 le tomó 48 horas agotar toda la xilosa en el medio mientras que a la cepa ZW705-ara354A7 le tomó solo 24 horas.

Ejemplo 5

10 Expresión de araE en ZW705-ara354 y ZW801-ara354 de *Zymomonas*

En este ejemplo, se analizan los efectos de la expresión de araE en las cepas ZW705-ara354 y ZW801-ara354 de *Z. mobilis* no adaptadas que utilizan arabinosa.

1. Transformación de ZW705-ara354 y ZW801-ara354 con pARA112.

15 Tal como se describe en el Ejemplo 2, ZW705-ara354 y ZW801-ara354 son cepas de *Z. mobilis* diseñadas genéticamente desarrolladas a partir de ZW705 y ZW801-4 mediante la introducción de P_{gap} -araBAD en el locus de *ldhA*. ZW705-ara354 es la cepa original de ZW705-ara354A7 que no se adaptó en MRM3A5. Se prepararon células competentes de ambas cepas. El ADN no metilado de pARA112 se sometió a electroporación en las células competentes tal como se describió en los ejemplos previos.

20 La ZW705-ara354 (ZW705-ara354-ara112) y ZW801-ara354 ((ZW801-ara354-ara112) transformadas por pARA112 se recuperaron en 1 mL de medio MMG5 durante 4 horas a 30°C y luego se cultivaron en placas MMG5-CM120 durante 2 días a 30°C dentro de un frasco anaeróbico con un AnaeroPack. Las colonias individuales se estrieron en una nueva placa MMG5-CM120 y crecieron en las mismas condiciones que la última etapa. Las estrías crecieron en las placas que contenían cloranfenicol, lo que indica una transformación satisfactoria.

2. Expresión de P_{gi} -araE en las cepas transformadas.

25 Se seleccionaron varias estrías de las cepas transformadas de las placas MMG5-CM120 para representar ZW705-ara354-ara112 y ZW801-ara354-ara112, respectivamente. Se inspeccionó la expresión de P_{gi} -araE mediante el ensayo de crecimiento de 72 horas en MRM3A5. Los detalles del ensayo son los mismos que en los ejemplos anteriores. Como controles, las cepas ZW705-ara354 y ZW801-ara354 se cultivaron y analizaron en paralelo con 250 mg/L de espectrinomicina reemplazando 120 mg/L de cloranfenicol en el medio de cultivo. Los resultados para 3 cepas de cada transformación se proporcionan en la Tabla 5. En comparación con sus cepas originales, todas las cepas ZW705-ara354-ara112 y ZW801-ara354-ara112 utilizaron significativamente más arabinosa durante 72 horas de cultivo, lo que permitió un mayor nivel de crecimiento y producción de etanol. Por lo tanto, araE también facilitó la utilización de arabinosa en ambas cepas ZW705-ara354-ara112 y ZW801-ara354-ara112.

Tabla 5. Ensayo de crecimiento de 72 horas para ZW705-ara354-ara112 y ZW801-ara354-ara112 en MRM3A5

Cepa	Crecimiento (OD600)	Etanol (g/L)	Arabinosa (g/L)
ZW705-ara354	1,15	9,56	27,88
ZW705-ara354-ara112-1	1,56	14,18	17,24
ZW705-ara354-ara112-2	1,67	16,71	10,93
ZW705-ara354-ara112-3	1,47	13,76	19,06
ZW801-ara354	1,39	9,65	27,08
ZW801-ara354-ara112-4	1,95	15,01	15,12
ZW801-ara354-ara112-5	2,07	15,51	12,94
ZW801-ara354-ara112-5	2,29	15,79	13,05

35 3. Caracterización de perfiles de crecimiento y metabolitos de las cepas ZW705-ara354-ara112 y ZW801-ara354-ara112

Las cepas ZW705-ara354-ara112 y ZW801-ara354-ara112 se caracterizaron adicionalmente para determinar sus perfiles de crecimiento y metabolitos durante un período de tiempo de 96 horas. La caracterización se realizó siguiendo el mismo procedimiento descrito en el Ejemplo 4.3. ZW705-ara354-ara112-1 y ZW705-ara354-ara112-2 se examinaron y se compararon con su original ZW705-ara354, mientras ZW801-ara354-ara112-5 y ZW801-ara354-

ara112-6 se examinaron y se compararon con su original ZW801-ara354. Las mediciones y análisis se realizaron a los puntos de tiempo de 0, 6, 12, 24, 48, 72 y 96 horas.

La Figura 10 muestra los resultados obtenidos a partir de las cepas ZW705-ara354 y ZW705-ara354-ara112 cultivadas en MRM3A5. Los resultados muestran que, sin P_{gi} -araE, ZW705-ara354 utilizó poca arabinosa, con una velocidad máxima de 0,25 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo solamente en 38,19%, a 30,22 g/L. Con expresión de araE, ZW705-ara354-ara112-1 y ZW705-ara354-ara112-2 utilizaron arabinosa de forma más eficaz, lo que permitió mayores niveles de crecimiento y producción de etanol. La velocidad máxima de consumo de arabinosa aumentó a 0,46 g/L/hr y 0,48 g/L/hr, respectivamente. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 65,8%, a 16,73 g/L para ZW705-ara354-ara112-1 y en 69,61%, a 14,86 g/L para ZW705-ara354-ara112-2.

La Figura 11 muestra los resultados obtenidos a partir de las cepas ZW705-ara354 y ZW705-ara354-ara112 cultivadas en el medio de azúcares mezclados MRM3A2.5X2.5G5. Los resultados muestran que ZW705-ara354 utilizó de manera eficaz glucosa y xilosa para permitir un fuerte crecimiento y producción de etanol. Su metabolismo de arabinosa fue lento e incompleto. La velocidad máxima de consumo de arabinosa fue 0,29 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 57,32%, a 10,21 g/L. Sin embargo, ZW705-ara354-ara112-1 y ZW705-ara354-ara112-2 utilizaron arabinosa de forma más eficaz. La velocidad máxima de consumo de arabinosa aumentó a 0,32 g/L/hr y 0,35 g/L/hr, respectivamente. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 86,33%, a 3,27 g/L para ZW705-ara354-ara112-1 y en 85,2%, a 3,54 g/L para ZW705-ara354-ara112-2. Estos resultados demostraron que la expresión de araE facilitó la utilización de arabinosa en las cepas ZW705-ara354-ara112 tanto en medio de azúcar único (arabinosa) como en el medio de azúcares mezclados. Por lo tanto, el efecto de araE no requirió un antecedente genético adquirido durante la adaptación de ZW705-ara354A7. De manera similar a los resultados en ZW705-ara354A7-ara112, la expresión de araE enlenteció ligeramente el metabolismo de xilosa en ZW705-ara354-ara112 cultivada en el medio de azúcares mezclados.

La Figura 12 muestra los resultados obtenidos a partir de las cepas ZW801-ara354 y ZW801-ara354-ara112 cultivadas en MRM3A5. Los resultados indican que, sin P_{gi} -araE, ZW801-ara354 utilizó poca arabinosa, con una velocidad máxima de 0,25 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo solamente en 32,99%, a 32,76 g/L. Con expresión de araE, ZW801-ara354-ara112-5 y ZW801-ara354-ara112-6 utilizaron arabinosa de forma más eficaz, lo que permitió mayores niveles de crecimiento y producción de etanol. La velocidad máxima de consumo de arabinosa aumentó a 0,49 g/L/hr y 0,47 g/L/hr, respectivamente. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 69,52%, a 14,90 g/L para ZW801-ara354-ara112-5 y en 65,92%, a 16,66 g/L para ZW801-ara354-ara112-6. La Figura 13 muestra los resultados obtenidos a partir de las cepas ZW801-ara354 y ZW801-ara354-ara112 cultivadas en el medio de azúcares mezclados MRM3A2.5X2.5G5. Esto muestran que ZW801-ara354 utilizó de manera eficaz glucosa y xilosa para permitir un fuerte crecimiento y producción de etanol. Su metabolismo de arabinosa fue lento e incompleto. La velocidad máxima de consumo de arabinosa fue 0,22 g/L/hr. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 45,48%, a 13,04 g/L. Sin embargo, ZW801-ara354-ara112-5 y ZW801-ara354-ara112-6 utilizaron arabinosa de forma más eficaz. La velocidad máxima de consumo de arabinosa aumentó a 0,35 g/L/hr y 0,36 g/L/hr, respectivamente. Al final del período de tiempo, la concentración de arabinosa en el medio se redujo en 89,92%, a 2,41 g/L para ZW801-ara354-ara112-5 y en 88,38%, a 2,78 g/L para ZW801-ara354-ara112-6. Estos resultados demostraron adicionalmente que la expresión de araE facilitó la utilización de arabinosa en las cepas ZW801-ara354-ara112 tanto en medio de azúcar único como en el medio de azúcares mezclados. Por lo tanto, el efecto de araE no se limitó a ZW705-ara354 y las cepas derivadas. De manera similar a ZW705-ara354A7-ara112 y ZW705-ara354-ara112, la expresión de araE enlenteció ligeramente el metabolismo de xilosa en ZW801-ara354-ara112 cultivada en el medio de azúcares mezclados.

Listado de secuencias

<110> E. I. du Pont de Nemours and Company

<120> Zymomonas con utilización de arabinosa mejorada

<130> CL4545PCT

<160> 66

<170> PatentIn versión 3.4

<210> 1

<211> 1416

<212> ADN

<213> Escherichia coli

<400> 1

atggttacta tcaatacggg atctgcttta acgccacggt ctttgcgga tacgcggcgt	60
atgaatatgt ttgtttcggg agctgctgcg gtcgcaggat tgttatttgg tcttgatatc	120
ggcgtaatcg ccggagcgtt gccgttcatt accgatcact ttgtgctgac cagtcgtttg	180
caggaatggg tggtagtag catgatgctc ggtgcagcaa ttggtgcgct gtttaatggt	240
tggctgctgt tccgcctggg gcgtaaatac agcctgatgg cgggggcat cctgtttgta	300
ctcggttcta tagggtccgc ttttgcgacc agcgtagaga tgttaatcgc cgctcgtgtg	360
gtgctgggca ttgctgctgg gatcgctct tacaccgctc ctctgtatct ttctgaaatg	420
gcaagtgaag acgttcgcgg taagatgac agtatgtacc agttgatgg cactcggc	480
atcgtgctgg cgtttttatc cgatacagcg ttcagttata gcggttaactg gcgcgcaatg	540
ttgggggttc ttgctttacc agcagttctg ctgattatc ttgtagtctt cctgccaaat	600
agcccgctt ggctggcgga aaagggcggt catattgagg cggaagaagt attgcgtatg	660
ctgcgcgata cgtcggaaaa agcgcgagaa gaactcaacg aaattcgtga aagcctgaag	720
ttaaaacagg gcggttgggc actgtttaag atcaaccgta acgtccgtcg tgctgtgttt	780
ctcggtatgt tgttgcaggc gatgcagcag tttaccggta tgaacatcat catgtactac	840
gcgcgcgta tcttcaaaat ggcggtctt acgaccacag aacaacagat gattgcgact	900
ctggtcgtag ggctgacctt tatgttcgcc acctttattg cgggtgtttac ggtagataaa	960
gcagggcgta aaccggctct gaaaattggt ttcagcgtga tggcgtagg cactctggtg	1020
ctgggctatt gcctgatgca gtttgataac ggtacggctt ccagtggctt gtcctggctc	1080
tctgttggca tgacgatgat gtgtattgcc ggttatgcga tgagcgccgc gccagtgggtg	1140
tggatcctgt gctctgaaat tcagccgctg aaatgccgcg atttcggtat tacctgttcg	1200
accaccacga actgggtgtc gaatatgatt atcggcgcg ccttcctgac actgcttgat	1260
agcattggcg ctgccggtac gttctggctc tacactgcgc tgaacattgc gtttgtgggc	1320
attactttct ggctcattcc ggaaaccaa aatgtcacgc tggaacatat cgaacgcaa	1380
ctgatggcag gcgagaagtt gagaaatc ggcgtc	1416

<210> 2

<211> 472

<212> PRT

5 <213> Escherichia coli

<400> 2

ES 2 599 579 T3

Met	Val	Thr	Ile	Asn	Thr	Glu	Ser	Ala	Leu	Thr	Pro	Arg	Ser	Leu	Arg	1	5	10	15
Asp	Thr	Arg	Arg	Met	Asn	Met	Phe	Val	Ser	Val	Ala	Ala	Ala	Val	Ala	20	25	30	
Gly	Leu	Leu	Phe	Gly	Leu	Asp	Ile	Gly	Val	Ile	Ala	Gly	Ala	Leu	Pro	35	40	45	
Phe	Ile	Thr	Asp	His	Phe	Val	Leu	Thr	Ser	Arg	Leu	Gln	Glu	Trp	Val	50	55	60	
Val	Ser	Ser	Met	Met	Leu	Gly	Ala	Ala	Ile	Gly	Ala	Leu	Phe	Asn	Gly	65	70	75	80
Trp	Leu	Ser	Phe	Arg	Leu	Gly	Arg	Lys	Tyr	Ser	Leu	Met	Ala	Gly	Ala	85	90	95	
Ile	Leu	Phe	Val	Leu	Gly	Ser	Ile	Gly	Ser	Ala	Phe	Ala	Thr	Ser	Val	100	105	110	
Glu	Met	Leu	Ile	Ala	Ala	Arg	Val	Val	Leu	Gly	Ile	Ala	Val	Gly	Ile	115	120	125	
Ala	Ser	Tyr	Thr	Ala	Pro	Leu	Tyr	Leu	Ser	Glu	Met	Ala	Ser	Glu	Asn	130	135	140	
Val	Arg	Gly	Lys	Met	Ile	Ser	Met	Tyr	Gln	Leu	Met	Val	Thr	Leu	Gly	145	150	155	160
Ile	Val	Leu	Ala	Phe	Leu	Ser	Asp	Thr	Ala	Phe	Ser	Tyr	Ser	Gly	Asn	165	170	175	
Trp	Arg	Ala	Met	Leu	Gly	Val	Leu	Ala	Leu	Pro	Ala	Val	Leu	Leu	Ile	180	185	190	
Ile	Leu	Val	Val	Phe	Leu	Pro	Asn	Ser	Pro	Arg	Trp	Leu	Ala	Glu	Lys	195	200	205	
Gly	Arg	His	Ile	Glu	Ala	Glu	Glu	Val	Leu	Arg	Met	Leu	Arg	Asp	Thr	210	215	220	

Ser Glu Lys Ala Arg Glu Glu Leu Asn Glu Ile Arg Glu Ser Leu Lys
225 230 235 240

Leu Lys Gln Gly Gly Trp Ala Leu Phe Lys Ile Asn Arg Asn Val Arg
245 250 255

Arg Ala Val Phe Leu Gly Met Leu Leu Gln Ala Met Gln Gln Phe Thr
260 265 270

Gly Met Asn Ile Ile Met Tyr Tyr Ala Pro Arg Ile Phe Lys Met Ala
275 280 285

Gly Phe Thr Thr Thr Glu Gln Gln Met Ile Ala Thr Leu Val Val Gly
290 295 300

Leu Thr Phe Met Phe Ala Thr Phe Ile Ala Val Phe Thr Val Asp Lys
305 310 315 320

Ala Gly Arg Lys Pro Ala Leu Lys Ile Gly Phe Ser Val Met Ala Leu
325 330 335

Gly Thr Leu Val Leu Gly Tyr Cys Leu Met Gln Phe Asp Asn Gly Thr
340 345 350

Ala Ser Ser Gly Leu Ser Trp Leu Ser Val Gly Met Thr Met Met Cys
355 360 365

Ile Ala Gly Tyr Ala Met Ser Ala Ala Pro Val Val Trp Ile Leu Cys
370 375 380

Ser Glu Ile Gln Pro Leu Lys Cys Arg Asp Phe Gly Ile Thr Cys Ser
385 390 395 400

Thr Thr Thr Asn Trp Val Ser Asn Met Ile Ile Gly Ala Thr Phe Leu
405 410 415

Thr Leu Leu Asp Ser Ile Gly Ala Ala Gly Thr Phe Trp Leu Tyr Thr
420 425 430

Ala Leu Asn Ile Ala Phe Val Gly Ile Thr Phe Trp Leu Ile Pro Glu
435 440 445

Thr Lys Asn Val Thr Leu Glu His Ile Glu Arg Lys Leu Met Ala Gly
450 455 460

Glu Lys Leu Arg Asn Ile Gly Val
465 470

<210> 3

<211> 1416

<212> ADN

5 <213> Shigella flexneri

<400> 3

```

atgggttacta tcaatacggg atctgcttta acgccacggt ctttgctga tacgcggcgt      60
atgaatatgt ttgtttcggg agctgctgcg gtcgcaggat tggtatttgg tcttgatatc      120
ggcgtaatcg ccggagcgtt gccgttcatt accgatcact ttgtgctgac cagtcgtttg      180
caggaatggg tggttagtag catgatgctc ggcgagcaa ttggtgctgct gtttaatggt      240
tggtctgctg tccgcctggg gcgtaaatac agcctgatgg cgggggcat cctgtttgta      300
ctcggttcta tagggtccgc ttttgcgacc agcgtagaga tgtaategc cgctcgtgtg      360
gtgctgggca ttgctgctgg gatcgctctc tacaccgctc ctctgtatct ttctgaaatg      420
gcaagtgaag acgttcgcgg taagatgctc agtatgtacc agttgatggt cacactcggc      480
atcgtgctgg cgtttttatc cgatacagcg ttcagttata gcggttaactg gcgcgcaatg      540
ttgggggttc ttgctttacc agcagttctg ctgattattc tgggtggtctt cctgccaaat      600
agcccgctgt ggctggcgga aaaggggctt catattgagg cggaagaagt gttgcgtatg      660
ctgcgcgata cgtcggaaaa agcgcgagaa gaactcaacg aaattcgtga aagcctgaag      720
ttaaacagg gcggttgggc actgtttaag atcaaccgta acgtccgtcg tgctgtgttt      780
ctcggtatgt tggtgcaggc gatgcagcag tttaccggtg tgaacatcat catgtactac      840
gcgcgcgcta tcttcaaaat ggccgggcttt acgaccacag aacaacagat gattgcgact      900
ctggtcgtgg gactgacctt tatgttcgcg accttcattg cggctctttac ggtagataaa      960
gcaggtcgta aaccggctct gaaaattggt ttcagcgtga tggcgttagg cactctggtg     1020
ctgggctatt gcctgatgca gtttgataac ggtacggctt ccagtggctt gtcctggctc     1080
tctgttggca tgacgatgat gtgtattgcc ggttatgcga tgagcgccgc gccagtgggtg     1140
tggatcctgt gctctgaaat tcagccgctg aaatgccgcg atttcgggtat tacctgttcg     1200
acgacgacaa actgggtgtc gaatatgatt atcggcgcgg ccttcctgac actgcttgat     1260
agcattggcg ctgccggtac gttctggctc tacactgcgc tgaacattgc gtttgtgggt     1320
attactttct ggctcattcc ggaaaccaa aatgtcacgc tggaacatat cgaacgcaaa     1380
ctgatggcag gcgagaagtt gagaaatatt ggcgtc                                1416

```

<210> 4

<211> 472

5 <212> PRT

<213> *Shigella flexneri*

<400> 4

```

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

```

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

ES 2 599 579 T3

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

<210> 5

<211> 1416

<212> ADN

5 <213> Shigella boydii

<400> 5

atggttacta tcaatacggg atctgcttta acgccacgtt ctttgcgagg tacgcggcgt	60
atgaatatgt ttgtttcggg agctgctgcg gtcgcaggat tgttatattg tcttgatata	120
ggcgtaatcg ccggagcgtt gccgttcatt accgatcact ttgtgctgac cagtcatttg	180
caggaatggg tggtagtag catgatgctc ggcgagcaa ttggtgcgct gtttaatggt	240

tggctgtcgt tccgcctggg gcgtaaatac agcctgatgg cggggggccat cctgtttgta 300
 ctcggttcta taggggtccgc ttttgcgacc agcgtagaga tgtaaatcgc cgctcgtgtg 360
 gtgctgggca ttgctgtcgg gatcgcgctc tacaccgctc ctctgtatct ttctgaaatg 420
 gcaagtgaaa acgttcgcgg taagatgatc agtatgtacc agttgatggg cacactcggc 480
 atcgtgctgg cgtttttata cgatacagcg ttcagttata gcgtaactg gcgcgcaatg 540
 ttggggggttc ttgctttacc agcagttctg ctgattattc tgggtggtcct cctgccaaat 600
 agcccgcgct ggttggcgga aaaggggct catattgagg cggaagaagt attgcgtatg 660
 ctgcgcgata cgtcggaaaa agcgcgagaa gaactcaacg aaattcgtga aagcctgaag 720
 ttaaaacagg gcggttgggc actgtttaag atcaaccgta acgtccgtcg tgctgtgttt 780
 ctcggtatgt tggtgcaggc gatgcagcag tttaccggta tgaacatcat catgtactac 840
 gcgcgcgta tcttcaaaat ggcgggcttt acgaccacag aacaacagat gattgcgact 900
 ctggtcgtag ggctgacctt tatgttcgcc acctttattg cggtgtttac ggtagataaa 960
 gcagggcgta aaccggctct gaaaattggg ttcagcgtga tggcgtagg cactctggtg 1020
 ctgggctatt gcctgatgca gtttgataac ggtacggctt ccagtggctt gtcctggctc 1080
 tctgttggca tgacgatgat gtgtattgcc ggttatgcga tgagcgccgc gccagtggg 1140
 tggatcctgt gctctgaaat tcagccgctg aaatgccgag atttcggtat tacctgttcg 1200
 accaccacga actgggtgtc gaatatgatt atcggcgcgga ccttcctgac gctgctcgac 1260
 agcattggcg ctgccggtac gttctggctc tacactgcgc tgaacattgc gtttgtgggc 1320
 atcactttct ggctcattcc ggaaacaaa aatgtcacgc tggaacatat cgaacgcaaa 1380
 ctgatggcag gcgagaagtt gagaaatata ggcatac 1416

<210> 6

<211> 472

<212> PRT

5 <213> Shigella boydii

<400> 6

Met	Val	Thr	Ile	Asn	Thr	Glu	Ser	Ala	Leu	Thr	Pro	Arg	Ser	Leu	Arg
1				5					10					15	
Asp	Thr	Arg	Arg	Met	Asn	Met	Phe	Val	Ser	Val	Ala	Ala	Ala	Val	Ala
				20				25					30		
Gly	Leu	Leu	Phe	Gly	Leu	Asp	Ile	Gly	Val	Ile	Ala	Gly	Ala	Leu	Pro
				35			40					45			
Phe	Ile	Thr	Asp	His	Phe	Val	Leu	Thr	Ser	His	Leu	Gln	Glu	Trp	Val
				50			55				60				
Val	Ser	Ser	Met	Met	Leu	Gly	Ala	Ala	Ile	Gly	Ala	Leu	Phe	Asn	Gly

65		70		75		80									
Trp	Leu	Ser	Phe	Arg	Leu	Gly	Arg	Lys	Tyr	Ser	Leu	Met	Ala	Gly	Ala
				85					90					95	
Ile	Leu	Phe	Val	Leu	Gly	Ser	Ile	Gly	Ser	Ala	Phe	Ala	Thr	Ser	Val
			100					105					110		
Glu	Met	Leu	Ile	Ala	Ala	Arg	Val	Val	Leu	Gly	Ile	Ala	Val	Gly	Ile
		115					120					125			
Ala	Ser	Tyr	Thr	Ala	Pro	Leu	Tyr	Leu	Ser	Glu	Met	Ala	Ser	Glu	Asn
	130					135					140				
Val	Arg	Gly	Lys	Met	Ile	Ser	Met	Tyr	Gln	Leu	Met	Val	Thr	Leu	Gly
145					150					155					160
Ile	Val	Leu	Ala	Phe	Leu	Ser	Asp	Thr	Ala	Phe	Ser	Tyr	Ser	Gly	Asn
				165					170					175	
Trp	Arg	Ala	Met	Leu	Gly	Val	Leu	Ala	Leu	Pro	Ala	Val	Leu	Leu	Ile
			180					185					190		
Ile	Leu	Val	Val	Phe	Leu	Pro	Asn	Ser	Pro	Arg	Trp	Leu	Ala	Glu	Lys
		195					200					205			
Gly	Arg	His	Ile	Glu	Ala	Glu	Glu	Val	Leu	Arg	Met	Leu	Arg	Asp	Thr
	210					215					220				
Ser	Glu	Lys	Ala	Arg	Glu	Glu	Leu	Asn	Glu	Ile	Arg	Glu	Ser	Leu	Lys
225					230					235					240
Leu	Lys	Gln	Gly	Gly	Trp	Ala	Leu	Phe	Lys	Ile	Asn	Arg	Asn	Val	Arg
				245					250					255	
Arg	Ala	Val	Phe	Leu	Gly	Met	Leu	Leu	Gln	Ala	Met	Gln	Gln	Phe	Thr
			260					265					270		
Gly	Met	Asn	Ile	Ile	Met	Tyr	Tyr	Ala	Pro	Arg	Ile	Phe	Lys	Met	Ala
		275					280					285			
Gly	Phe	Thr	Thr	Thr	Glu	Gln	Gln	Met	Ile	Ala	Thr	Leu	Val	Val	Gly
	290					295					300				
Leu	Thr	Phe	Met	Phe	Ala	Thr	Phe	Ile	Ala	Val	Phe	Thr	Val	Asp	Lys
305					310					315					320
Ala	Gly	Arg	Lys	Pro	Ala	Leu	Lys	Ile	Gly	Phe	Ser	Val	Met	Ala	Leu
				325					330					335	

Gly Thr Leu Val Leu Gly Tyr Cys Leu Met Gln Phe Asp Asn Gly Thr
340 345 350

Ala Ser Ser Gly Leu Ser Trp Leu Ser Val Gly Met Thr Met Met Cys
355 360 365

Ile Ala Gly Tyr Ala Met Ser Ala Ala Pro Val Val Trp Ile Leu Cys
370 375 380

Ser Glu Ile Gln Pro Leu Lys Cys Arg Asp Phe Gly Ile Thr Cys Ser
385 390 395 400

Thr Thr Thr Asn Trp Val Ser Asn Met Ile Ile Gly Ala Thr Phe Leu
405 410 415

Thr Leu Leu Asp Ser Ile Gly Ala Ala Gly Thr Phe Trp Leu Tyr Thr
420 425 430

Ala Leu Asn Ile Ala Phe Val Gly Ile Thr Phe Trp Leu Ile Pro Glu
435 440 445

Thr Lys Asn Val Thr Leu Glu His Ile Glu Arg Lys Leu Met Ala Gly
450 455 460

Glu Lys Leu Arg Asn Ile Gly Ile
465 470

<210> 7

<211> 1416

<212> ADN

5 <213> Shigella dysenteriae

<400> 7

atggttacta tcaatacggg atctgcttta acgccacgtt ctttgcgtag tacgcggcgt	60
atgaatatgt ttgtttcggg agctgctgcg gtcgcaggat tggtatttgg tcttgatata	120
ggcgtaaatcg ccggagcgtt gccgttcatt accgatcact ttgtgctgac cagtcgtttg	180
caggaatggg tggtagtag catgatgctc ggcgagcaa ttggtgctgt gtttaattgg	240
tggtgctgt tccgcctggg gcgtaaatac agcctgatgg cggggggccat cctgtttgta	300
ctcggttcta tagggccgc ttttgctacc agcgtagaga tgtaaatcgc cgctcgtgtg	360
gtgctgggca ttgctgtcgg gatcgctct tacaccgctc ctctgtatct ttctgaaatg	420
gcaagtgaac acgttcgcgg taagatgata agtatgtacc agttgatgg cacactcgcc	480
atcgtgctgg cgtttttatc cgatacagcg ttcagttata gcggttaactg gcgcgcaatg	540
ttgggggttc ttgctttacc agcagtcctg ctgattattc tgggtgtctt cctgccaaat	600
agcccgcgt ggctggcgga aaaggggcgt catattgagg cggaagaagt gttgcgtatg	660

```

ctgcgcgata cgtcggaaaa agcgcgagaa gaactcaacg aaattcgtga aagcctgaag    720
ttaaacaag gcggttgggc actgtttaag atcaaccgta acgtccgtcg tgctgtgttt    780
ctcggtatgt tggtgcaggc gatgcagcag tttaccggta tgaacatcat catgtactat    840
gcgccgcgta tcttcaaaat ggcgggcttt acgaccacag aacaacagat gattgcgact    900
ctggtcgtgg gactgacctt tatgttcgcg accttcattg cggtctttac ggtagataaa    960
gcaggtcgta aaccggctct gaaaattggg ttcagcgtga tggcgtagg cactctggtg   1020
ctgggctatt gcctgatgca gtttgataac ggtacggctt ccagtggctt gtcctggctc   1080
tctgttggca tgacgatgat gtgtattgcc ggttatgcga tgagcgccgc gccagtgggtg   1140
tggatectgt gctctgaaat tcagccgctg aaatgccacg atttcggtat tacctgttcg   1200
acgacgacaa actgggtgtc gaatatgatt atcggcgcga ccttcctgac actgcttgat   1260
agcattggcg ctgccgtac gttctggctc tacactgcgc tgaacattgc gtttgtgggc   1320
atcactttct ggctcattcc ggaaaccaa aatgtcacgc tggaacatat cgaacgcaa   1380
ctgatggcag gcgagaagtt gagaaatata ggcgtc                                1416

```

<210> 8

<211> 472

<212> PRT

5 <213> Shigella dysenteriae

<400> 8

```

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

Asp Thr Arg Arg Met Asn Met Phe Val Ser Val Ala Ala Ala Val Ala
          20          25          30

Gly Leu Leu Phe Gly Leu Asp Ile Gly Val Ile Ala Gly Ala Leu Pro
          35          40          45

Phe Ile Thr Asp His Phe Val Leu Thr Ser Arg Leu Gln Glu Trp Val
          50          55          60

Val Ser Ser Met Met Leu Gly Ala Ala Ile Gly Ala Leu Phe Asn Gly
65          70          75          80

Trp Leu Ser Phe Arg Leu Gly Arg Lys Tyr Ser Leu Met Ala Gly Ala
          85          90          95

Ile Leu Phe Val Leu Gly Ser Ile Gly Ser Ala Phe Ala Thr Ser Val
          100          105          110

Glu Met Leu Ile Ala Ala Arg Val Val Leu Gly Ile Ala Val Gly Ile
          115          120          125

```

Ala	Ser	Tyr	Thr	Ala	Pro	Leu	Tyr	Leu	Ser	Glu	Met	Ala	Ser	Glu	Asn
130						135					140				
Val	Arg	Gly	Lys	Met	Ile	Ser	Met	Tyr	Gln	Leu	Met	Val	Thr	Leu	Gly
145					150					155					160
Ile	Val	Leu	Ala	Phe	Leu	Ser	Asp	Thr	Ala	Phe	Ser	Tyr	Ser	Gly	Asn
				165					170					175	
Trp	Arg	Ala	Met	Leu	Gly	Val	Leu	Ala	Leu	Pro	Ala	Val	Leu	Leu	Ile
			180					185						190	
Ile	Leu	Val	Val	Phe	Leu	Pro	Asn	Ser	Pro	Arg	Trp	Leu	Ala	Glu	Lys
		195					200					205			
Gly	Arg	His	Ile	Glu	Ala	Glu	Glu	Val	Leu	Arg	Met	Leu	Arg	Asp	Thr
	210					215					220				
Ser	Glu	Lys	Ala	Arg	Glu	Glu	Leu	Asn	Glu	Ile	Arg	Glu	Ser	Leu	Lys
225					230					235					240
Leu	Lys	Gln	Gly	Gly	Trp	Ala	Leu	Phe	Lys	Ile	Asn	Arg	Asn	Val	Arg
			245						250					255	
Arg	Ala	Val	Phe	Leu	Gly	Met	Leu	Leu	Gln	Ala	Met	Gln	Gln	Phe	Thr
			260					265						270	
Gly	Met	Asn	Ile	Ile	Met	Tyr	Tyr	Ala	Pro	Arg	Ile	Phe	Lys	Met	Ala
		275					280					285			
Gly	Phe	Thr	Thr	Thr	Glu	Gln	Gln	Met	Ile	Ala	Thr	Leu	Val	Val	Gly
	290					295					300				
Leu	Thr	Phe	Met	Phe	Ala	Thr	Phe	Ile	Ala	Val	Phe	Thr	Val	Asp	Lys
305					310					315					320
Ala	Gly	Arg	Lys	Pro	Ala	Leu	Lys	Ile	Gly	Phe	Ser	Val	Met	Ala	Leu
				325					330					335	
Gly	Thr	Leu	Val	Leu	Gly	Tyr	Cys	Leu	Met	Gln	Phe	Asp	Asn	Gly	Thr
			340					345					350		
Ala	Ser	Ser	Gly	Leu	Ser	Trp	Leu	Ser	Val	Gly	Met	Thr	Met	Met	Cys
		355					360					365			
Ile	Ala	Gly	Tyr	Ala	Met	Ser	Ala	Ala	Pro	Val	Val	Trp	Ile	Leu	Cys
	370					375					380				

Ser Glu Ile Gln Pro Leu Lys Cys His Asp Phe Gly Ile Thr Cys Ser
385 390 395 400

Thr Thr Thr Asn Trp Val Ser Asn Met Ile Ile Gly Ala Thr Phe Leu
405 410 415

Thr Leu Leu Asp Ser Ile Gly Ala Ala Gly Thr Phe Trp Leu Tyr Thr
420 425 430

Ala Leu Asn Ile Ala Phe Val Gly Ile Thr Phe Trp Leu Ile Pro Glu
435 440 445

Thr Lys Asn Val Thr Leu Glu His Ile Glu Arg Lys Leu Met Ala Gly
450 455 460

Glu Lys Leu Arg Asn Ile Gly Val
465 470

<210> 9

<211> 1416

<212> ADN

5 <213> Salmonella typhimurium

<400> 9

```

atggtctctta ttaatcatga ctctgcttta acgccgcgtt cgcttcgcga cacacgacgt      60
atgaatatgt ttgtttcggg ttctgcagcg gtagcgggac tgttatttgg tctggatatc      120
ggcgttatcg ccggggcgct gccttttatt accgaccatt tcgtactgac cagccggctg      180
caggaatggg tcgtcagcag catgatgctt ggcgcgggcaa ttggcgcat atttaacggc      240
tggttttcat tccggctggg gcgtaagtat agcctgatgg ctggcgcgat tttgttcgtg      300
ctcggtctgc tggggtcggc gtttgcttcc agcgtggaag tattgattgg cgcccgcgtg      360
atactgggcg tagcagtagg gattgcctcc tataccgcgc cgctttatct ctctgaaatg      420
gcaagtgaaa atgttcgcgg caaaatgatc agtatgtatc aactgatggt gacgttaggc      480
attgtgctgg cttttttatc cgatacggca ttcagctaca gcggcaactg gcgcgcgatg      540
ttgggcgtgc tggcgctgcc tgcggtgttg ctcatattc tgggtggtatt cctgccgaat      600
agtccgcgtt ggctggcgca aaaaggctgc catattgaag cggaagaggt gctgcgtatg      660
ctgcgcgata cctcggaaaa agcccgtgat gaactgaatg agattcggga aagcctcaaa      720
ctcaagcagg gaggggtgggc attatttaaa gctaaccgca atgttcgccg cgccgtgttc      780
ctcggtatgc tgctacaggc aatgcagcag ttcaccggca tgaacatcat tatgtactat      840
gcgcgcgca ttttttaaat ggccggcttt accaccacgg aacagcaaat gatcgccacg      900
ctggtggtcg gactgacttt tatgttcgcg acgtttatcg ccgtctttac ggtcgataag      960
gccgggcgta aaccggcggt aaaaatcggg ttcagcgtaa tggcggttagg gacattggtg     1020

```

ttgggctact gcctgatgca gtttgataac ggtacggcat caagcggctc ctcttggtt 1080
 tccgttggga tgacgatgat gtgtatcgcc ggttacgcga tgagcgccgc tccggtggtg 1140
 tggatactgt gttcggaaat ccagccgctg aaatgccgtg attttgcat tacctgttca 1200
 accacgacaa actgggtatc gaacatgac atcggcgcca cattcctgac actgttggac 1260
 agcattggcg cggcaggtac attctggctc tacaccgcgc tgaatatcgc ttttatcggc 1320
 atcactttct ggctgattcc ggaaacaaa aatgtcacc cggagcacat cgaacgcaag 1380
 ctgatggcgg gcgagaagct aagaaatatt ggcgtg 1416

<210> 10

<211> 472

<212> PRT

5 <213> Salmonella typhimurium

<400> 10

Met Val Ser Ile Asn His Asp Ser Ala Leu Thr Pro Arg Ser Leu Arg
 1 5 10 15
 Asp Thr Arg Arg Met Asn Met Phe Val Ser Val Ser Ala Ala Val Ala
 20 25 30
 Gly Leu Leu Phe Gly Leu Asp Ile Gly Val Ile Ala Gly Ala Leu Pro
 35 40 45
 Phe Ile Thr Asp His Phe Val Leu Thr Ser Arg Leu Gln Glu Trp Val
 50 55 60
 Val Ser Ser Met Met Leu Gly Ala Ala Ile Gly Ala Leu Phe Asn Gly
 65 70 75 80
 Trp Leu Ser Phe Arg Leu Gly Arg Lys Tyr Ser Leu Met Ala Gly Ala
 85 90 95
 Ile Leu Phe Val Leu Gly Ser Leu Gly Ser Ala Phe Ala Ser Ser Val
 100 105 110
 Glu Val Leu Ile Gly Ala Arg Val Ile Leu Gly Val Ala Val Gly Ile
 115 120 125
 Ala Ser Tyr Thr Ala Pro Leu Tyr Leu Ser Glu Met Ala Ser Glu Asn
 130 135 140
 Val Arg Gly Lys Met Ile Ser Met Tyr Gln Leu Met Val Thr Leu Gly
 145 150 155 160
 Ile Val Leu Ala Phe Leu Ser Asp Thr Ala Phe Ser Tyr Ser Gly Asn
 165 170 175

Trp Arg Ala Met Leu Gly Val Leu Ala Leu Pro Ala Val Leu Leu Ile
 180 185 190

Ile Leu Val Val Phe Leu Pro Asn Ser Pro Arg Trp Leu Ala Gln Lys
 195 200 205

Gly Arg His Ile Glu Ala Glu Glu Val Leu Arg Met Leu Arg Asp Thr
 210 215 220

Ser Glu Lys Ala Arg Asp Glu Leu Asn Glu Ile Arg Glu Ser Leu Lys
 225 230 235 240

Leu Lys Gln Gly Gly Trp Ala Leu Phe Lys Ala Asn Arg Asn Val Arg
 245 250 255

Arg Ala Val Phe Leu Gly Met Leu Leu Gln Ala Met Gln Gln Phe Thr
 260 265 270

Gly Met Asn Ile Ile Met Tyr Tyr Ala Pro Arg Ile Phe Lys Met Ala
 275 280 285

Gly Phe Thr Thr Thr Glu Gln Gln Met Ile Ala Thr Leu Val Val Gly
 290 295 300

Leu Thr Phe Met Phe Ala Thr Phe Ile Ala Val Phe Thr Val Asp Lys
 305 310 315 320

Ala Gly Arg Lys Pro Ala Leu Lys Ile Gly Phe Ser Val Met Ala Leu
 325 330 335

Gly Thr Leu Val Leu Gly Tyr Cys Leu Met Gln Phe Asp Asn Gly Thr
 340 345 350

Ala Ser Ser Gly Leu Ser Trp Leu Ser Val Gly Met Thr Met Met Cys
 355 360 365

Ile Ala Gly Tyr Ala Met Ser Ala Ala Pro Val Val Trp Ile Leu Cys
 370 375 380

Ser Glu Ile Gln Pro Leu Lys Cys Arg Asp Phe Gly Ile Thr Cys Ser
 385 390 395 400

Thr Thr Thr Asn Trp Val Ser Asn Met Ile Ile Gly Ala Thr Phe Leu
 405 410 415

Thr Leu Leu Asp Ser Ile Gly Ala Ala Gly Thr Phe Trp Leu Tyr Thr
 420 425 430

Ala Leu Asn Ile Ala Phe Ile Gly Ile Thr Phe Trp Leu Ile Pro Glu
435 440 445

Thr Lys Asn Val Thr Leu Glu His Ile Glu Arg Lys Leu Met Ala Gly
450 455 460

Glu Lys Leu Arg Asn Ile Gly Val
465 470

<210> 11

<211> 1431

<212> ADN

5 <213> Salmonella enterica

<400> 11

```
.ttgtggcagg aaaatatggt ctctattaat catgactctg ctttaacgcc gcgttcgctt      60
cgcgacacac gacgtatgaa tatgtttgtt tcggtttctg cagcggtagc gggactgtta      120
tttgggtctgg atatcggcgt tategcgggg gcgctgcctt ttattaccga ccatttcgta      180
ctgaccagcc ggctgcagga atgggtcgtc agcagtatga tgcttggcgc ggcaattggc      240
gcattatttta acggctggct ttcattccgg ctggggcgta agtatagcct gatggctggc      300
gcgattttgt tcgtgctcgg ctcgctgggg tcggcgtttg cttccagcgt ggaagtattg      360
attggcgccc gcgtgatact gggcgtagca gtagggattg cgtcctatac cgcgccgctt      420
tatctctctg aaatggcaag tgaaaatgtt cgcgggcaaaa tgatcagtat gtatcaactg      480
atggtgacgt taggcattgt gctggctttt ttatccgata cggcattcag ctacagcggc      540
aactggcgcg cgatgttggg cgtgctggcg ctgcctgcgg tgttgctcat tattctcgtg      600
gtattcctgc cgaatagtcc gcgttggctg gcgcaaaaag gtcgccatat tgaagcggaa      660
gaggtgctgc gtatgctgcg cgatacctcg gaaaaagccc gtgatgaact gaatgagatt      720
cgggaaagcc tcaaactcaa gcagggcggg tgggcattat ttaaagctaa ccgcaatggt      780
cgccgcgccg tgttcctcgg tatgctgcta caggcaatgc agcagttcac cggcatgaac      840
atcattatgt actatgcgcc gcgcattttt aaaatggccc gctttaccac cacggaacag      900
caaatgatcg ccacgctggt ggtcggactg acctttatgt tcgcgacgtt tatcgcgcgc      960
tttacggctc ataaggccgg gcgtaaaccg gcgttaaaaa tcggtttcag cgtaatggcg     1020
ttagggacat tgggtgttgg ctactgcctg atgcagtttg ataacggtac ggcataaagc     1080
ggtctctcct ggctttccgt tgggatgacg atgatgtgta tcgccgggta cgcgatgagc     1140
gccgctccgg tgggtgtggat actgtgttcg gaaatccagc cgctgaaatg ccgtgatttt     1200
ggcattacct gttcaaccac gacaaactgg gtatcgaaca tgatcatcgg cgcgacattc     1260
ctgacactgt tggacagtat tggcgcgga ggtacattct ggctctacac cgcgctgaat     1320
atcgctttta tcggcatcac tttctggctg attccggaaa ccaaaaatgt caccctggag     1380
catatcgaac gcaagctaata ggcgggagcag aagctaagaa atattggcgt g          1431
```

<210> 12

ES 2 599 579 T3

<211> 477

<212> PRT

<213> Salmonella enterica

<400> 12

```

Met Trp Gln Glu Asn Met Val Ser Ile Asn His Asp Ser Ala Leu Thr
1          5          10          15

Pro Arg Ser Leu Arg Asp Thr Arg Arg Met Asn Met Phe Val Ser Val
          20          25          30

Ser Ala Ala Val Ala Gly Leu Leu Phe Gly Leu Asp Ile Gly Val Ile
          35          40          45

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

Leu Gln Glu Trp Val Val Ser Ser Met Met Leu Gly Ala Ala Ile Gly
65          70          75          80

Ala Leu Phe Asn Gly Trp Leu Ser Phe Arg Leu Gly Arg Lys Tyr Ser
          85          90          95

Leu Met Ala Gly Ala Ile Leu Phe Val Leu Gly Ser Leu Gly Ser Ala
          100          105          110

Phe Ala Ser Ser Val Glu Val Leu Ile Gly Ala Arg Val Ile Leu Gly
          115          120          125

Val Ala Val Gly Ile Ala Ser Tyr Thr Ala Pro Leu Tyr Leu Ser Glu
          130          135          140

Met Ala Ser Glu Asn Val Arg Gly Lys Met Ile Ser Met Tyr Gln Leu
145          150          155          160

Met Val Thr Leu Gly Ile Val Leu Ala Phe Leu Ser Asp Thr Ala Phe
          165          170          175

Ser Tyr Ser Gly Asn Trp Arg Ala Met Leu Gly Val Leu Ala Leu Pro
          180          185          190

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

Trp Leu Ala Gln Lys Gly Arg His Ile Glu Ala Glu Glu Val Leu Arg
210          215          220

```


Met Leu Arg Asp Thr Ser Glu Lys Ala Arg Asp Glu Leu Asn Glu Ile
 225 230 235 240

Arg Glu Ser Leu Lys Leu Lys Gln Gly Gly Trp Ala Leu Phe Lys Ala
 245 250 255

Asn Arg Asn Val Arg Arg Ala Val Phe Leu Gly Met Leu Leu Gln Ala
 260 265 270

Met Gln Gln Phe Thr Gly Met Asn Ile Ile Met Tyr Tyr Ala Pro Arg
 275 280 285

Ile Phe Lys Met Ala Gly Phe Thr Thr Thr Glu Gln Gln Met Ile Ala
 290 295 300

Thr Leu Val Val Gly Leu Thr Phe Met Phe Ala Thr Phe Ile Ala Val
 305 310 315 320

Phe Thr Val Asp Lys Ala Gly Arg Lys Pro Ala Leu Lys Ile Gly Phe
 325 330 335

Ser Val Met Ala Leu Gly Thr Leu Val Leu Gly Tyr Cys Leu Met Gln
 340 345 350

Phe Asp Asn Gly Thr Ala Ser Ser Gly Leu Ser Trp Leu Ser Val Gly
 355 360 365

Met Thr Met Met Cys Ile Ala Gly Tyr Ala Met Ser Ala Ala Pro Val
 370 375 380

Val Trp Ile Leu Cys Ser Glu Ile Gln Pro Leu Lys Cys Arg Asp Phe
 385 390 395 400

Gly Ile Thr Cys Ser Thr Thr Thr Asn Trp Val Ser Asn Met Ile Ile
 405 410 415

Gly Ala Thr Phe Leu Thr Leu Leu Asp Ser Ile Gly Ala Ala Gly Thr
 420 425 430

Phe Trp Leu Tyr Thr Ala Leu Asn Ile Ala Phe Ile Gly Ile Thr Phe
 435 440 445

Trp Leu Ile Pro Glu Thr Lys Asn Val Thr Leu Glu His Ile Glu Arg
 450 455 460

Lys Leu Met Ala Gly Glu Lys Leu Arg Asn Ile Gly Val
 465 470 475

<210> 13

<211> 1419

<212> ADN

5 <213> Klebsiella pneumoniae

<400> 13

```

atgacttcaa tcagtaacga ctctgcatta acgccgcgga cacaacgtga caccgcggcg 60
atgaactggg ttgtttctat cgctgcggcg gtagcggggg tgctctttgg cctggatata 120
ggcgtgatat ccggggcgct gccctttatt accgaccact tcaccttata cagccagctt 180
caggagtggg tggtcagcag tatgatgttg ggggcggcga tcggtgcgct gtttaacggc 240
tggtgtcgt tccgcctcgg ccgtaaatac agcctgatgg cgggggctgt gctctttgtt 300
gccggctcta tcggctccgc ttttgccgcc agcgtggagg tgctgctgat agccgcgctg 360
gtgttggggg tggccgtcgg gatcgcttcc tataccgcgc cggtgtacct ctccgagatg 420
gccagtgaga acgtgcgcgg gaaaatgata agtatgtacc agctgatggg gaccctcggc 480
attgtgctgg cgtttctttc cgatactgcc tttagctaca gcggttaactg gcgcgccatg 540
ttaggcgtgc tggcactgcc ggccgtgatc ctgattatc tggtcgtctt tttgccgaac 600
agcccgcgct ggctggcgga gaaaggacgc catatcgaag cggaagaggg gctgcggatg 660
ctgcgcgata cctcggaaaa ggcgcgcgac gagcttaacg agatccgtga gagcctgaag 720
ctgaagcagg gcggctgggc gttgtttaag gtcaatcgta acgtgcgccg ggccgtgttc 780
cttggcatgc tgctgcaggc gatgcagcag ttcaccggca tgaacatcat catgtactac 840
gcgcgcgcta tctttaaaat ggccggcctt accactaccg aacagcagat gatcgccacc 900
ctggtgggtg gcctgacctt tatgtttgcc acctttattg cgggtgttcac ggtggataaa 960
gcgggtcgta agccggcgct aaaaatcggc tttagcgtga tggcgctggg caccctggtg 1020
ctgggctact gcctgatgca gttcgacaat ggcaccgcct ccagcggctc ctccctggctt 1080
tccgtcggca tgaccatgat gtgtattgcc gggatgcga tgagcgcggc gccgggtggtg 1140
tggatcctct gctccgagat ccagccgctg aaatgccgcg acttcgggtat cacctgctcg 1200
accaccacca actgggtgtc gaacatgata atcggcgcca ccttcttgac gctgcttgac 1260
gcgattggcg ccgcggcac cttctggctc tacacggtgc tcaacgtggc ctttatcggc 1320
gtcaccttct ggctgatccc ggaaaccaag aatgtcacc tcgagcacat tgagcgcaac 1380
ctgatggcgg gcgagaagct gcgcaacatc ggtaaccgt 1419

```

<210> 14

<211> 473

5 <212> PRT

<213> Klebsiella pneumoniae

<400> 14

```

Met Thr Ser Ile Ser Asn Asp Ser Ala Leu Thr Pro Arg Thr Gln Arg
1           5           10           15

```

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

ES 2 599 579 T3

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

<210> 15

<211> 1416

<212> ADN

5 <213> Klebsiella oxytoca

<400> 15

atgaccactc tcagtcacga ctctacaacc atgccgcgta cgcagcgcga taccgggcgc	60
atgaatcagt ttgtctccat tgccgccgcg gtggcagggt tgctgtttgg cctcgatatc	120
ggggtgattg ccggggcgct gccctttatt accgaccatt ttgttttata cagccgcctg	180
caggagtggg tggtgagcag catgatgctg ggagccgccca tcggcgcgtt atttaacggc	240

tggctctctt tccgcctcgg gcgcaaatac agcctgatgg tgggcgcggg gctgttcggt 300
 gccggctccg tgggctccgc gtttgcgacc agcgtcgaaa tgctgctggg ggcaaggatc 360
 gttctcgggg tcgccgtggg gatcgctctt tataccgcgc cgtgtacct gtcggaaatg 420
 gcgagcgaaa acgtgcgcgg caagatgatc agcatgtatc agctgatggg gacgctgggt 480
 atcgtgatgg cgtttctctc cgacaccgcg ttcagctaca gcggcaactg gcgggcatg 540
 cttggcgtag tggcgctgcc ggcggtgggt ctgattattc tggatgatct cctgccgaac 600
 agcccgcgct ggctggcgga aaaagggcgt cacgtggaag cggaagaggg gctgcggatg 660
 ctgcgcgaca cgtcagaaaa agcccgtgac gagctcaacg agatccgcga aagcctgaag 720
 ctgaagcagg gcggctgggc gctgtttaag gtcaaccgca acgtgcggcg ggcggtattc 780
 ctcggcatgc tggtgcaggc gatgcagcag ttaccggta tgaatatcat catgtactac 840
 gcgcgcgca tctttaaaat ggcgggcttc accaccaccg aacagcagat ggtcgcgacc 900
 ctggtgggtg gctgacctt tatgttcgcc acctttatcg ccgtctttac cgtcgataag 960
 gccggacgta agccggcgct gaaaatcggg tttagcgtga tggccatcgg cacgctgggt 1020
 ctgggctact gtctgatgca gtttgataac ggcaccgcct ccagcgggtc ctcctggctg 1080
 tcggtgggga tgaccatgat gtgtatcgcc ggctatgcga tgagcgccgc gccgggtggg 1140
 tggatcctgt gttcggaaat tcagccgctg aagtgcgcgc atttcggcat cacctgctca 1200
 accaccacca actgggtgtc gaacatgatt atcggcgcga ccttcctgac gctgctggac 1260
 gcgatcggcg cggcaggaac cttctggctt tataccgcgc tgaacgtcgc ctttatcggc 1320
 gtgacgttct ggctgatccc ggaaaccaa aacgtcaccc tggagcatat tgaacgcagg 1380
 ctgatgtccg gcgagaagct gcgcaatatc ggcaat 1416

<210> 16

<211> 472

<212> PRT

5 <213> Klebsiella oxytoca

<400> 16

Met Thr Thr Leu Ser His Asp Ser Thr Thr Met Pro Arg Thr Gln Arg
1 5 10 15

Asp Thr Arg Arg Met Asn Gln Phe Val Ser Ile Ala Ala Ala Val Ala
20 25 30

Gly Leu Leu Phe Gly Leu Asp Ile Gly Val Ile Ala Gly Ala Leu Pro
35 40 45

Phe Ile Thr Asp His Phe Val Leu Ser Ser Arg Leu Gln Glu Trp Val
50 55 60

Val Ser Ser Met Met Leu Gly Ala Ala Ile Gly Ala Leu Phe Asn Gly

ES 2 599 579 T3

65		70		75		80									
Trp	Leu	Ser	Phe	Arg	Leu	Gly	Arg	Lys	Tyr	Ser	Leu	Met	Val	Gly	Ala
				85					90					95	
Val	Leu	Phe	Val	Ala	Gly	Ser	Val	Gly	Ser	Ala	Phe	Ala	Thr	Ser	Val
			100					105					110		
Glu	Met	Leu	Leu	Val	Ala	Arg	Ile	Val	Leu	Gly	Val	Ala	Val	Gly	Ile
		115					120					125			
Ala	Ser	Tyr	Thr	Ala	Pro	Leu	Tyr	Leu	Ser	Glu	Met	Ala	Ser	Glu	Asn
	130					135					140				
Val	Arg	Gly	Lys	Met	Ile	Ser	Met	Tyr	Gln	Leu	Met	Val	Thr	Leu	Gly
145					150					155					160
Ile	Val	Met	Ala	Phe	Leu	Ser	Asp	Thr	Ala	Phe	Ser	Tyr	Ser	Gly	Asn
				165					170					175	
Trp	Arg	Ala	Met	Leu	Gly	Val	Leu	Ala	Leu	Pro	Ala	Val	Val	Leu	Ile
			180					185					190		
Ile	Leu	Val	Ile	Phe	Leu	Pro	Asn	Ser	Pro	Arg	Trp	Leu	Ala	Glu	Lys
	195						200					205			
Gly	Arg	His	Val	Glu	Ala	Glu	Glu	Val	Leu	Arg	Met	Leu	Arg	Asp	Thr
	210					215					220				
Ser	Glu	Lys	Ala	Arg	Asp	Glu	Leu	Asn	Glu	Ile	Arg	Glu	Ser	Leu	Lys
225					230					235					240
Leu	Lys	Gln	Gly	Gly	Trp	Ala	Leu	Phe	Lys	Val	Asn	Arg	Asn	Val	Arg
			245						250					255	
Arg	Ala	Val	Phe	Leu	Gly	Met	Leu	Leu	Gln	Ala	Met	Gln	Gln	Phe	Thr
			260					265					270		
Gly	Met	Asn	Ile	Ile	Met	Tyr	Tyr	Ala	Pro	Arg	Ile	Phe	Lys	Met	Ala
		275					280					285			
Gly	Phe	Thr	Thr	Thr	Glu	Gln	Gln	Met	Val	Ala	Thr	Leu	Val	Val	Gly
	290					295					300				
Leu	Thr	Phe	Met	Phe	Ala	Thr	Phe	Ile	Ala	Val	Phe	Thr	Val	Asp	Lys
305					310					315					320
Ala	Gly	Arg	Lys	Pro	Ala	Leu	Lys	Ile	Gly	Phe	Ser	Val	Met	Ala	Ile
				325					330					335	

Gly Thr Leu Val Leu Gly Tyr Cys Leu Met Gln Phe Asp Asn Gly Thr
340 345 350

Ala Ser Ser Gly Leu Ser Trp Leu Ser Val Gly Met Thr Met Met Cys
355 360 365

Ile Ala Gly Tyr Ala Met Ser Ala Ala Pro Val Val Trp Ile Leu Cys
370 375 380

Ser Glu Ile Gln Pro Leu Lys Cys Arg Asp Phe Gly Ile Thr Cys Ser
385 390 395 400

Thr Thr Thr Asn Trp Val Ser Asn Met Ile Ile Gly Ala Thr Phe Leu
405 410 415

Thr Leu Leu Asp Ala Ile Gly Ala Ala Gly Thr Phe Trp Leu Tyr Thr
420 425 430

Ala Leu Asn Val Ala Phe Ile Gly Val Thr Phe Trp Leu Ile Pro Glu
435 440 445

Thr Lys Asn Val Thr Leu Glu His Ile Glu Arg Arg Leu Met Ser Gly
450 455 460

Glu Lys Leu Arg Asn Ile Gly Asn
465 470

<210> 17

<211> 1413

<212> ADN

5 <213> Enterobacter cancerogenus

<400> 17

atgacatctc tcaatgactc taccctcatg cccgcggcgc tgcgcgacac ccgccgcatg 60

aaccagtttg tctccgtcgc ggcgccgcta gcgggtctgc tgtttgggct ggatatcggc 120

gttatcgccg gtgcgtgcc gttatcacc gatcatttca cgtaaagtca tcgcctgcag 180

gagtgggtgg tgagcagcat gatgctgggc gccgcaattg gggcgttgtt caacggctgg 240

ctctcgttcc gcctgggacg aaagtacagc ctgatggctg gggcgatcct gtttgtggcc 300

ggttcactgg ggtcggcggt tgcacaaagc gttgaggtgc tgttgctctc ccgcgtgctg 360

cttggcggtg cggtggggat cgcctcctac accgcgccgc tgtatctctc cgaaatggcg 420

agcgagaacg tgcgcggcaa gatgatcagc atgtatcagc tgatggtgac gctcggcatc 480

gtgctggcgt ttctttccga tacctggttc agctacaccg gtaactggcg cgccatgctc 540

ggcgtgctgg cgttgcccg cctgttgctg atgggtgctg tgattttcct gccgaacagc 600

ccgcgctggc tggcgcaaaa aggccgccac gtcgaggcgg aagaagtgct gcgaatgctg 660

cgtgacacct ctgaaaaagc gcgtgaagag ttgaacgaga tccgcgaaag cctgaagctg 720
 aagcagggcg gctgggcgct gtttaaggct aaccgcaacg tgcgcgcgcg cgtgtttctg 780
 ggaatgctct tgcaggcgat gcagcagttt acgggcatga acatcatcat gtactacgcc 840
 ccgcgcacat ttaaaatggc gggcttcacc acgaccgagc agcagatgat cgccaccctg 900
 gtggtcgggc tgacctttat gttcgccacc tttattgccg tatttaccgt cgataaagcc 960
 ggacgtaaac cggcgctgaa aattggcttt agcgtgatgg cgctcggtac gctgacctc 1020
 ggctactgcc tgatgcagtt tgatcagggc acggcatcga gggggctttc ctggctctcc 1080
 gtcggtatga ccatgatgtg cattgccggt tatgcaatga gcgcgcgcgc ggtggtgtgg 1140
 atcctgtgct ctgaaattca gccgctaaaa tgcgcgcgact ttggtatcac ctgttcacc 1200
 accaccaact ggggtgtcgaa catgattatc ggtgcgacct tctgacgct gctggatgcc 1260
 attggtgcag cgggaacatt ctggctctac acgggtgctga acgtggcggt tattggcgta 1320
 acgttctggc tgatcccaga aaccaaaggg gtgacgctgg agcacattga acgcaagctg 1380
 atggcggggg agaagttaaa aaacataggc gtg 1413

<210> 18

<211> 471

<212> PRT

5 <213> Enterobacter cancerogenus

<400> 18

Met	Thr	Ser	Leu	Asn	Asp	Ser	Thr	Leu	Met	Pro	Ala	Ala	Leu	Arg	Asp
1				5					10					15	
Thr	Arg	Arg	Met	Asn	Gln	Phe	Val	Ser	Val	Ala	Ala	Ala	Val	Ala	Gly
			20					25					30		
Leu	Leu	Phe	Gly	Leu	Asp	Ile	Gly	Val	Ile	Ala	Gly	Ala	Leu	Pro	Phe
		35					40					45			
Ile	Thr	Asp	His	Phe	Thr	Leu	Ser	His	Arg	Leu	Gln	Glu	Trp	Val	Val
	50					55					60				
Ser	Ser	Met	Met	Leu	Gly	Ala	Ala	Ile	Gly	Ala	Leu	Phe	Asn	Gly	Trp
65					70				75					80	
Leu	Ser	Phe	Arg	Leu	Gly	Arg	Lys	Tyr	Ser	Leu	Met	Val	Gly	Ala	Ile
			85					90						95	
Leu	Phe	Val	Ala	Gly	Ser	Leu	Gly	Ser	Ala	Phe	Ala	Thr	Ser	Val	Glu
			100					105						110	
Val	Leu	Leu	Leu	Ser	Arg	Val	Leu	Leu	Gly	Val	Ala	Val	Gly	Ile	Ala
			115				120						125		

Ser Tyr Thr Ala Pro Leu Tyr Leu Ser Glu Met Ala Ser Glu Asn Val
130 135 140

Arg Gly Lys Met Ile Ser Met Tyr Gln Leu Met Val Thr Leu Gly Ile
145 150 155 160

Val Leu Ala Phe Leu Ser Asp Thr Trp Phe Ser Tyr Thr Gly Asn Trp
165 170 175

Arg Ala Met Leu Gly Val Leu Ala Leu Pro Ala Leu Leu Leu Met Val
180 185 190

Leu Val Ile Phe Leu Pro Asn Ser Pro Arg Trp Leu Ala Gln Lys Gly
195 200 205

Arg His Val Glu Ala Glu Glu Val Leu Arg Met Leu Arg Asp Thr Ser
210 215 220

Glu Lys Ala Arg Glu Glu Leu Asn Glu Ile Arg Glu Ser Leu Lys Leu
225 230 235 240

Lys Gln Gly Gly Trp Ala Leu Phe Lys Val Asn Arg Asn Val Arg Arg
245 250 255

Ala Val Phe Leu Gly Met Leu Leu Gln Ala Met Gln Gln Phe Thr Gly
260 265 270

Met Asn Ile Ile Met Tyr Tyr Ala Pro Arg Ile Phe Lys Met Ala Gly
275 280 285

Phe Thr Thr Thr Glu Gln Gln Met Ile Ala Thr Leu Val Val Gly Leu
290 295 300

Thr Phe Met Phe Ala Thr Phe Ile Ala Val Phe Thr Val Asp Lys Ala
305 310 315 320

Gly Arg Lys Pro Ala Leu Lys Ile Gly Phe Ser Val Met Ala Leu Gly
325 330 335

Thr Leu Ile Leu Gly Tyr Cys Leu Met Gln Phe Asp Gln Gly Thr Ala
340 345 350

Ser Ser Gly Leu Ser Trp Leu Ser Val Gly Met Thr Met Met Cys Ile
355 360 365

Ala Gly Tyr Ala Met Ser Ala Ala Pro Val Val Trp Ile Leu Cys Ser
370 375 380

Glu Ile Gln Pro Leu Lys Cys Arg Asp Phe Gly Ile Thr Cys Ser Thr
385 390 395 400

Thr Thr Asn Trp Val Ser Asn Met Ile Ile Gly Ala Thr Phe Leu Thr
405 410 415

Leu Leu Asp Ala Ile Gly Ala Ala Gly Thr Phe Trp Leu Tyr Thr Val
420 425 430

Leu Asn Val Ala Phe Ile Gly Val Thr Phe Trp Leu Ile Pro Glu Thr
435 440 445

Lys Gly Val Thr Leu Glu His Ile Glu Arg Lys Leu Met Ala Gly Glu
450 455 460

Lys Leu Lys Asn Ile Gly Val
465 470

<210> 19

<211> 1392

<212> ADN

5 <213> Bacillus amyloliquefaciens

<400> 19

```
atgaagaatc acccggcacc aattggctca aatgtacctg tcaactcggca gcattccaag      60
tggtttgtca ttctcatctc atgcgcggcc ggactgggag ggcttttgta cggttatgac      120
acggcggtga tttccggcgc tatcggtttc ctgaaagatt tgtaccgctt aagtcctttt      180
atggaagggc tcgtgatttc aagcattatg atcggcgggtg ttttcggcgt cgggatttcc      240
ggatttttga gtgaccgttt cggacggaga aagattttga tggcagcggc gctgttggtt      300
gcggtgtcag cggttgtctc tgcgctttct caaagtgtgt cttccttagt gatcgccaga      360
gtcatcggcg gtctgggaat cggcatgggc tcctcgtttt ctgtcacgta tattaccgaa      420
gccgctccgc cggccatacg cggcagctctg tcttcactgt atcagctgtt tacgatatta      480
gggatataccg gcacttattt tattaacctt gccgtccagc agtccggctc gtatgaatgg      540
ggagtgcaca ccggctggcg gtggatgctc gcttacggca tgattccgtc cgtcatcttt      600
tttatcgtgc tgcttatcgt gccggaaagt ccgcgctggc ttgcgaaagc ggggcgccgg      660
aatgaagccc tcgccgtgct gacgcgcatt aacggcgagc agaccgcgaa agaagaaatc      720
aaacaaatcg aaacgtcttt acaattagaa aaaatgggtt cattgtctca gctgtttaag      780
ccggggctga gaaaagcgct tgtgatcggg attctgctgg ctttattcaa tcaggtcac      840
ggcatgaacg caattacgta ttacgggccg gaaattttca aaatgatggg cttcggacag      900
aatgcggggg ttatcacgac atgcatcgtc ggtgtcgttg aagtgatttt caccattatc      960
gcggttcttt tagtcgataa ggtaggccgg aaaaaactga tgggggtcgg atctgccttt     1020
```

atggcgctgt tcatgatctt aatcgggggca tcctttttatt ttcagctggc gagcgggtccg 1080
gcttttagtcg tcatcatatt gggatttcgtc gccgctttct gcgtatcagt cgggccgatt 1140
acatggatca tgatttcgga aatcttttccg aaccacctcc gcgcacgcgc cgccgggtatt 1200
gcgacgatat tcttatgggg ggcgaactgg gcgatcggcc agttcgtgcc gatgatgatc 1260
agcggggttag ggcttgcgta caccttctgg atattcgccg tcattaatat tctctgtttc 1320
ttgtttgtcg tgacgatctg ccctgagacg aaaaataaat cattagaaga aatagaaaaa 1380
ctctggataa aa 1392

<210> 20

<211> 464

<212> PRT

5 <213> Bacillus amyloliquefaciens

<400> 20

Met Lys Asn His Pro Ala Pro Ile Gly Ser Asn Val Pro Val Thr Arg
1 5 10 15
Gln His Ser Lys Trp Phe Val Ile Leu Ile Ser Cys Ala Ala Gly Leu
20 25 30
Gly Gly Leu Leu Tyr Gly Tyr Asp Thr Ala Val Ile Ser Gly Ala Ile
35 40 45
Gly Phe Leu Lys Asp Leu Tyr Arg Leu Ser Pro Phe Met Glu Gly Leu
50 55 60
Val Ile Ser Ser Ile Met Ile Gly Gly Val Phe Gly Val Gly Ile Ser
65 70 75 80
Gly Phe Leu Ser Asp Arg Phe Gly Arg Arg Lys Ile Leu Met Ala Ala
85 90 95
Ala Leu Leu Phe Ala Val Ser Ala Val Val Ser Ala Leu Ser Gln Ser
100 105 110
Val Ser Ser Leu Val Ile Ala Arg Val Ile Gly Gly Leu Gly Ile Gly
115 120 125
Met Gly Ser Ser Leu Ser Val Thr Tyr Ile Thr Glu Ala Ala Pro Pro
130 135 140
Ala Ile Arg Gly Ser Leu Ser Ser Leu Tyr Gln Leu Phe Thr Ile Leu
145 150 155 160
Gly Ile Ser Gly Thr Tyr Phe Ile Asn Leu Ala Val Gln Gln Ser Gly
165 170 175

Ser Tyr Glu Trp Gly Val His Thr Gly Trp Arg Trp Met Leu Ala Tyr
180 185 190

Gly Met Ile Pro Ser Val Ile Phe Phe Ile Val Leu Leu Ile Val Pro
195 200 205

Glu Ser Pro Arg Trp Leu Ala Lys Ala Gly Arg Arg Asn Glu Ala Leu
210 215 220

Ala Val Leu Thr Arg Ile Asn Gly Glu Gln Thr Ala Lys Glu Glu Ile
225 230 235 240

Lys Gln Ile Glu Thr Ser Leu Gln Leu Glu Lys Met Gly Ser Leu Ser
245 250 255

Gln Leu Phe Lys Pro Gly Leu Arg Lys Ala Leu Val Ile Gly Ile Leu
260 265 270

Leu Ala Leu Phe Asn Gln Val Ile Gly Met Asn Ala Ile Thr Tyr Tyr
275 280 285

Gly Pro Glu Ile Phe Lys Met Met Gly Phe Gly Gln Asn Ala Gly Phe
290 295 300

Ile Thr Thr Cys Ile Val Gly Val Val Glu Val Ile Phe Thr Ile Ile
305 310 315 320

Ala Val Leu Leu Val Asp Lys Val Gly Arg Lys Lys Leu Met Gly Val
325 330 335

Gly Ser Ala Phe Met Ala Leu Phe Met Ile Leu Ile Gly Ala Ser Phe
340 345 350

Tyr Phe Gln Leu Ala Ser Gly Pro Ala Leu Val Val Ile Ile Leu Gly
355 360 365

Phe Val Ala Ala Phe Cys Val Ser Val Gly Pro Ile Thr Trp Ile Met
370 375 380

Ile Ser Glu Ile Phe Pro Asn His Leu Arg Ala Arg Ala Ala Gly Ile
385 390 395 400

Ala Thr Ile Phe Leu Trp Gly Ala Asn Trp Ala Ile Gly Gln Phe Val
405 410 415

Pro Met Met Ile Ser Gly Leu Gly Leu Ala Tyr Thr Phe Trp Ile Phe
420 425 430

Ala Val Ile Asn Ile Leu Cys Phe Leu Phe Val Val Thr Ile Cys Pro
435 440 445

Glu Thr Lys Asn Lys Ser Leu Glu Glu Ile Glu Lys Leu Trp Ile Lys
450 455 460

<210> 21
<211> 1500
<212> ADN
<213> Escherichia coli

5

<400> 21

```
atgacgattt ttgataatta tgaagtgtgg tttgtcattg gcagccagca tctgtatggc      60
ccggaaaccc tgcgtcaggt caccaacat gccgagcacg tcgttaatgc gctgaatacg      120
gaagcgaaac tgccctgcaa actggtgttg aaaccgctgg gcaccacgcc ggatgaaatc      180
accgctatatt gccgcgacgc gaattacgac gatcgttgcg ctggtctggt ggtgtggctg      240
cacacettct ccccgccaa aatgtggatc aacggcctga ccatgctcaa caaacgcttg      300
ctgcaattcc acaccagtt caacgcggcg ctgccgtggg acagtatcga tatggacttt      360
atgaacctga accagactgc acatggcggg cgcgagttcg gcttcattgg cgcgcgtatg      420
cgtcagcaac atgccgtggg taccggtcac tggcaggata aacaagccca tgagcgtatc      480
ggctcctgga tgcgtcaggg ggtctctaaa caggataccc gtcacttgaa agtctgccga      540
tttggcgata acatgcgtga agtggcgggc accgatggcg ataaagtgc cgcacagatc      600
aagttcgggt tctccgtcaa tacctgggcg gttggcgatc tgggtgcaggt ggtgaactcc      660
atcagcgacg gcgatgttaa cgcgctgggc gatgagtacg aaagctgcta caccatgacg      720
cctgccacac aaatccacgg caaaaaacga cagaacgtgc tggaagcggc gcgtattgag      780
ctggggatga agcgtttcct ggaacaaggt ggcttcacac cgttcaccac cacctttgaa      840
gatttgacag gtctgaaaca gcttctggt ctggccgtac agcgtctgat gcagcagggt      900
tacggctttg cgggcgaagg cgactggaaa actgccgccc tgcttcgcat catgaagggtg      960
atgtcaaccg gtctgcaggg cggcacctcc tttatggagg actacaccta tcacttcgag     1020
aaaggtaatg acctggtgct cggctcccat atgctggaag tctgcccgtc gatcgccgca     1080
gaagagaaac cgatcctcga cgttcagcat ctcggtattg gtggttaagg cgatcctgcc     1140
cgcctgatct tcaataacca aaccggccca gcgattgtcg ccagcttgat tgatctcggc     1200
gatcgttacc gtctactggt taactgcac gacacggtga aaacaccgca ctccctgccg     1260
aaactgccgg tggcgaatgc gctgtggaaa gcgcaaccgg atctgccaac tgcttccgaa     1320
gcgtggatcc tcgctggtgg cgcgcacat accgtcttca gccatgcact gaacctcaac     1380
gatatgcgcc aattcgccga gatgcacgac attgaaatca cggtgattga taacgacaca     1440
cgctgccag cgtttaaaga cgcgctgcgc tggaacgaag tgtattacgg atttcgtcgc     1500
```

<210> 22
<211> 500
<212> PRT

10

ES 2 599 579 T3

<213> Escherichia coli

<400> 22

```

Met Thr Ile Phe Asp Asn Tyr Glu Val Trp Phe Val Ile Gly Ser Gln
1          5          10          15

His Leu Tyr Gly Pro Glu Thr Leu Arg Gln Val Thr Gln His Ala Glu
          20          25          30

His Val Val Asn Ala Leu Asn Thr Glu Ala Lys Leu Pro Cys Lys Leu
          35          40          45

Val Leu Lys Pro Leu Gly Thr Thr Pro Asp Glu Ile Thr Ala Ile Cys
50          55          60

Arg Asp Ala Asn Tyr Asp Asp Arg Cys Ala Gly Leu Val Val Trp Leu
65          70          75          80

His Thr Phe Ser Pro Ala Lys Met Trp Ile Asn Gly Leu Thr Met Leu
          85          90          95

Asn Lys Pro Leu Leu Gln Phe His Thr Gln Phe Asn Ala Ala Leu Pro
          100          105          110

Trp Asp Ser Ile Asp Met Asp Phe Met Asn Leu Asn Gln Thr Ala His
115          120          125

Gly Gly Arg Glu Phe Gly Phe Ile Gly Ala Arg Met Arg Gln Gln His
130          135          140

Ala Val Val Thr Gly His Trp Gln Asp Lys Gln Ala His Glu Arg Ile
145          150          155          160

Gly Ser Trp Met Arg Gln Ala Val Ser Lys Gln Asp Thr Arg His Leu
          165          170          175

Lys Val Cys Arg Phe Gly Asp Asn Met Arg Glu Val Ala Val Thr Asp
180          185          190

Gly Asp Lys Val Ala Ala Gln Ile Lys Phe Gly Phe Ser Val Asn Thr
195          200          205

Trp Ala Val Gly Asp Leu Val Gln Val Val Asn Ser Ile Ser Asp Gly
210          215          220

Asp Val Asn Ala Leu Val Asp Glu Tyr Glu Ser Cys Tyr Thr Met Thr
225          230          235          240

```

Pro Ala Thr Gln Ile His Gly Lys Lys Arg Gln Asn Val Leu Glu Ala
245 250 255

Ala Arg Ile Glu Leu Gly Met Lys Arg Phe Leu Glu Gln Gly Gly Phe
260 265 270

His Ala Phe Thr Thr Thr Phe Glu Asp Leu His Gly Leu Lys Gln Leu
275 280 285

Pro Gly Leu Ala Val Gln Arg Leu Met Gln Gln Gly Tyr Gly Phe Ala
290 295 300

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

Met Ser Thr Gly Leu Gln Gly Gly Thr Ser Phe Met Glu Asp Tyr Thr
325 330 335

Tyr His Phe Glu Lys Gly Asn Asp Leu Val Leu Gly Ser His Met Leu
340 345 350

Glu Val Cys Pro Ser Ile Ala Ala Glu Glu Lys Pro Ile Leu Asp Val
355 360 365

Gln His Leu Gly Ile Gly Gly Lys Asp Asp Pro Ala Arg Leu Ile Phe
370 375 380

Asn Thr Gln Thr Gly Pro Ala Ile Val Ala Ser Leu Ile Asp Leu Gly
385 390 395 400

Asp Arg Tyr Arg Leu Leu Val Asn Cys Ile Asp Thr Val Lys Thr Pro
405 410 415

His Ser Leu Pro Lys Leu Pro Val Ala Asn Ala Leu Trp Lys Ala Gln
420 425 430

Pro Asp Leu Pro Thr Ala Ser Glu Ala Trp Ile Leu Ala Gly Gly Ala
435 440 445

His His Thr Val Phe Ser His Ala Leu Asn Leu Asn Asp Met Arg Gln
450 455 460

Phe Ala Glu Met His Asp Ile Glu Ile Thr Val Ile Asp Asn Asp Thr
465 470 475 480

Arg Leu Pro Ala Phe Lys Asp Ala Leu Arg Trp Asn Glu Val Tyr Tyr
485 490 495

Gly Phe Arg Arg
500

<210> 23
<211> 1698

<212> ADN
<213> Escherichia coli

<400> 23

atggcgattg caattggcct cgattttggc agtgattctg tgcgagcttt ggcggtggac	60
tgcgctaccg gtgaagagat cgccaccagc gtagagtggg atccccgttg gcagaaaggg	120
caattttgtg atgccccgaa taaccagttc cgtcatcacc cgcgtgacta cattgagtca	180
atggaagcgg cactgaaaac cgtgcttgca gagcttagcg tcgaacagcg cgcagctgtg	240
gtcgggattg gcggtgacag taccggctcg acgcccgcac cgattgatgc cgacggaaac	300
gtgctggcgc tgcgcccgga gtttgccgaa aacccgaacg cgatgttcgt attgtggaaa	360
gaccacactg cgggtgaaga agcgggaagag attaccctgt tgtgccacgc gccgggcaac	420
gttgactact cccgctacat tgggtgtatt tattccagcg aatgggtctg ggcaaaaatc	480
ctgcatgtga ctgccagga cagcgccgtg gcgcaatctg ccgcatcgtg gattgagctg	540
tgcgactggg tgccagctct gctttccggt accacccgcc cgcaggatat tcgtcgcgga	600
cgttgccagc cggggcataa atctctgtgg cacgaaagct gggggcgccct gccgccagcc	660
agtttctttg atgagctgga cccgatcctc aatcgccatt tgccttcccc gctgttcaact	720
gacacttgga ctgccgatat tccggtgggc accttatgcc cggaatgggc gcagcgtctc	780
ggcctgectg aaagcgtggg gatttccggc ggcgcgtttg actgccatat gggcgagctt	840
ggcgagggcg cacagcctaa cgcactggta aaagttatcg gtacttccac ctgcgacatt	900
ctgattgccg acaaacagag cgttgccgag cgggcagtta aaggatattg cggtcaggtt	960
gatggcagcg tgggtgcctgg atttatcggg ctggaagcag gccaatcggc gtttgggtgat	1020
atctacgcct ggtttggtcg cgtactcggc tggccgctgg aacagcttgc cgcccagcat	1080
ccggaactga aaacgcaaat caacgccagc cagaaacaac tgcttccggc gctgaccgaa	1140
gcatgggcca aaaatccgtc tctggatcac ctgccgggtg tgctcgactg gtttaacggc	1200
cgccgcacac cgaacgctaa ccaacgcctg aaaggggtga ttaccgatct taacctcgct	1260
accgacgctc cgctgctgtt cggcggtttg attgctgcca ccgcctttgg cgcacgcgca	1320
atcatggagt gctttaccga tcaggggatc gccgttaata acgtgatggc actgggcggc	1380
atcgcgcgga aaaaccaggt cattatgcag gcctgctgcg acgtgctgaa tcgcccgtg	1440
caaattgttg cctctgacca gtgctgtgcg ctccggtgcgg cgatttttgc tgccgtcgcc	1500
gcgaaagtgc acgcagacat cccatcagct cagcaaaaaa tggccagtgc ggtagagaaa	1560
accctgcaac cgtgcagcga gcaggcacia cgctttgaac agctttatcg ccgctatcag	1620
caatgggcga tgagcgccga acaacactat cttccaactt ccgccccggc acaggctgcc	1680
5 caggccgttg cgactcta	1698

<210> 24
<211> 566
<212> PRT
<213> Escherichia coli

<400> 24

Met	Ala	Ile	Ala	Ile	Gly	Leu	Asp	Phe	Gly	Ser	Asp	Ser	Val	Arg	Ala	1	5	10	15
Leu	Ala	Val	Asp	Cys	Ala	Thr	Gly	Glu	Glu	Ile	Ala	Thr	Ser	Val	Glu	20	25	30	
Trp	Tyr	Pro	Arg	Trp	Gln	Lys	Gly	Gln	Phe	Cys	Asp	Ala	Pro	Asn	Asn	35	40	45	
Gln	Phe	Arg	His	His	Pro	Arg	Asp	Tyr	Ile	Glu	Ser	Met	Glu	Ala	Ala	50	55	60	
Leu	Lys	Thr	Val	Leu	Ala	Glu	Leu	Ser	Val	Glu	Gln	Arg	Ala	Ala	Val	65	70	75	80
Val	Gly	Ile	Gly	Val	Asp	Ser	Thr	Gly	Ser	Thr	Pro	Ala	Pro	Ile	Asp	85	90	95	
Ala	Asp	Gly	Asn	Val	Leu	Ala	Leu	Arg	Pro	Glu	Phe	Ala	Glu	Asn	Pro	100	105	110	
Asn	Ala	Met	Phe	Val	Leu	Trp	Lys	Asp	His	Thr	Ala	Val	Glu	Glu	Ala	115	120	125	
Glu	Glu	Ile	Thr	Arg	Leu	Cys	His	Ala	Pro	Gly	Asn	Val	Asp	Tyr	Ser	130	135	140	
Arg	Tyr	Ile	Gly	Gly	Ile	Tyr	Ser	Ser	Glu	Trp	Phe	Trp	Ala	Lys	Ile	145	150	155	160
Leu	His	Val	Thr	Arg	Gln	Asp	Ser	Ala	Val	Ala	Gln	Ser	Ala	Ala	Ser	165	170	175	
Trp	Ile	Glu	Leu	Cys	Asp	Trp	Val	Pro	Ala	Leu	Leu	Ser	Gly	Thr	Thr	180	185	190	
Arg	Pro	Gln	Asp	Ile	Arg	Arg	Gly	Arg	Cys	Ser	Ala	Gly	His	Lys	Ser	195	200	205	
Leu	Trp	His	Glu	Ser	Trp	Gly	Gly	Leu	Pro	Pro	Ala	Ser	Phe	Phe	Asp				

ES 2 599 579 T3

210	215	220
Glu Leu Asp Pro Ile	Leu Asn Arg His	Leu Pro Ser Pro Leu Phe Thr
225	230	235 240
Asp Thr Trp Thr	Ala Asp Ile Pro Val	Gly Thr Leu Cys Pro Glu Trp
	245	250 255
Ala Gln Arg Leu Gly	Leu Pro Glu Ser Val Val	Ile Ser Gly Gly Ala
	260	265 270
Phe Asp Cys His Met	Gly Ala Val Gly Ala Gly	Ala Gln Pro Asn Ala
	275	280 285
Leu Val Lys Val Ile	Gly Thr Ser Thr Cys Asp	Ile Leu Ile Ala Asp
	290	295 300
Lys Gln Ser Val Gly	Glu Arg Ala Val Lys	Gly Ile Cys Gly Gln Val
305	310	315 320
Asp Gly Ser Val Val	Pro Gly Phe Ile Gly	Leu Glu Ala Gly Gln Ser
	325	330 335
Ala Phe Gly Asp Ile	Tyr Ala Trp Phe Gly	Arg Val Leu Gly Trp Pro
	340	345 350
Leu Glu Gln Leu Ala	Ala Gln His Pro Glu	Leu Lys Thr Gln Ile Asn
	355	360 365
Ala Ser Gln Lys Gln	Leu Leu Pro Ala Leu	Thr Glu Ala Trp Ala Lys
	370	375 380
Asn Pro Ser Leu Asp	His Leu Pro Val Val	Leu Asp Trp Phe Asn Gly
385	390	395 400
Arg Arg Thr Pro Asn	Ala Asn Gln Arg Leu	Lys Gly Val Ile Thr Asp
	405	410 415
Leu Asn Leu Ala Thr	Asp Ala Pro Leu Leu	Phe Gly Gly Leu Ile Ala
	420	425 430
Ala Thr Ala Phe Gly	Ala Arg Ala Ile Met	Glu Cys Phe Thr Asp Gln
	435	440 445
Gly Ile Ala Val Asn	Asn Val Met Ala Leu	Gly Gly Ile Ala Arg Lys
	450	455 460
Asn Gln Val Ile Met	Gln Ala Cys Cys Asp	Val Leu Asn Arg Pro Leu
465	470	475 480

Gln Ile Val Ala Ser Asp Gln Cys Cys Ala Leu Gly Ala Ala Ile Phe
485 490 495

Ala Ala Val Ala Ala Lys Val His Ala Asp Ile Pro Ser Ala Gln Gln
500 505 510

Lys Met Ala Ser Ala Val Glu Lys Thr Leu Gln Pro Cys Ser Glu Gln
515 520 525

Ala Gln Arg Phe Glu Gln Leu Tyr Arg Arg Tyr Gln Gln Trp Ala Met
530 535 540

Ser Ala Glu Gln His Tyr Leu Pro Thr Ser Ala Pro Ala Gln Ala Ala
545 550 555 560

Gln Ala Val Ala Thr Leu
565

<210> 25

<211> 693

<212> ADN

5 <213> Escherichia coli

<400> 25

```
atggttagaag atctcaaacg ccaggtatta gaagccaacc tggcgctgcc aaaacacaac      60
ctggtcacgc tcacatgggg caacgtcagc gccgttgatc gcgagcgcg cgtctttgtg      120
atcaaacctt ccggcgctga ttacagcgtc atgaccgctg acgatatggt cgtgggttagc      180
atcgaaaccg gtgaagtggg tgaaggtagc aaaaagccct cctccgacac gccaaactcac      240
cggtgtctct atcaggcatt cccctccatt ggcggcattg tgcatacgca ctcgcgccac      300
gccaccatct gggcgagggc gggtcagtcg attccagcaa ccggcaccac ccacgcccac      360
tatttctacg gcaccattcc ctgtaccgcg aaaatgaccg acgcagaaat caacggcgaa      420
tatgagtggg aaaccggtaa cgtcatcgta gaaacctttg aaaaacaggg tatcgatgca      480
gcgcaaatac ccggcggttct ggtccattcc cacggcccgt ttgcatgggg caaaaatgcc      540
gaagatgagg tgcataacgc catcgtgctg gaagaggctg cttatatggg gatattctgc      600
cgtcagttag cgccgcagtt accggatatg cagcaaacgc tgctggataa acactatctg      660
cgtaagcatg gcgcgaaggc atattacggg cag                                     693
```

<210> 26

<211> 231

10 <212> PRT

<213> Escherichia coli

<400> 26

Met Leu Glu Asp Leu Lys Arg Gln Val Leu Glu Ala Asn Leu Ala Leu

1	5	10	15
Pro Lys His Asn Leu Val Thr Leu Thr Trp Gly Asn Val Ser Ala Val	20	25	30
Asp Arg Glu Arg Gly Val Phe Val Ile Lys Pro Ser Gly Val Asp Tyr	35	40	45
Ser Val Met Thr Ala Asp Asp Met Val Val Val Ser Ile Glu Thr Gly	50	55	60
Glu Val Val Glu Gly Thr Lys Lys Pro Ser Ser Asp Thr Pro Thr His	65	70	75
Arg Leu Leu Tyr Gln Ala Phe Pro Ser Ile Gly Gly Ile Val His Thr	85	90	95
His Ser Arg His Ala Thr Ile Trp Ala Gln Ala Gly Gln Ser Ile Pro	100	105	110
Ala Thr Gly Thr Thr His Ala Asp Tyr Phe Tyr Gly Thr Ile Pro Cys	115	120	125
Thr Arg Lys Met Thr Asp Ala Glu Ile Asn Gly Glu Tyr Glu Trp Glu	130	135	140
Thr Gly Asn Val Ile Val Glu Thr Phe Glu Lys Gln Gly Ile Asp Ala	145	150	155
Ala Gln Met Pro Gly Val Leu Val His Ser His Gly Pro Phe Ala Trp	165	170	175
Gly Lys Asn Ala Glu Asp Ala Val His Asn Ala Ile Val Leu Glu Glu	180	185	190
Val Ala Tyr Met Gly Ile Phe Cys Arg Gln Leu Ala Pro Gln Leu Pro	195	200	205
Asp Met Gln Gln Thr Leu Leu Asp Lys His Tyr Leu Arg Lys His Gly	210	215	220
Ala Lys Ala Tyr Tyr Gly Gln	225	230	

<210> 27

<211> 3226

<212> ADN

5 <213> secuencia artificial

<220>

<223> fragmento de PCR de araA-araB

<400> 27

aaccatggcg attgcaattg gcctcgattt tggcagtgat tctgtgcgag ctttggcggg	60
ggactgcgct accggtgaag agatcgccac cagcgtagag tggatatccc gttggcagaa	120
agggcaattt tgtgatgccc cgaataacca gttccgtcat catccgcgtg actacattga	180
gtcaatggaa gcggcactga aaaccgtgct tgcagagctt agcgtcgaac agcgcgcagc	240
tgtggtcggg attggcggtg acagtaccgg ctcgacgccc gcaccgattg atgccgacgg	300
aaacgtgctg gcgctgcgcc cggagtttgc cgaaaacccg aacgcgatgt tcgtattgtg	360
gaaagaccac actgcggttg aagaagcgga agagattacc cgtttgtgcc acgcgccggg	420
caacgttgac tactcccgt acattggtgg tattttattcc agcgaatggt tctgggcaaa	480
aatcctgcat gtgactcgcc aggacagcgc cgtggcgcaa tctgccgcat cgtggattga	540
gctgtgcgac tgggtgccag ctctgctttc cggtagccacc cgcccgcagg atattcgctg	600
cggacgttgc agcgcggggc ataaatctct gtggcacgaa agctggggcg gcctgccgcc	660
agccagtttc tttgatgagc tggacccgat cctcaatgc catttgctt ccccgtgtt	720
cactgacact tggactgccg atattccggt gggcacctta tgcccgaat gggcgcgagc	780
tctcggcctg cctgaaagcg tggtgatttc cggcggcgcg tttgactgcc atatgggcgc	840
agttggcgca ggcgcacagc ctaacgcact ggtaaaagtt atcggtaact ccacctgcga	900
cattctgatt gccgacaaac agagcgttgg cgagcgggca gttaaaggta tttgcgggtca	960
ggttgatggc agcgtggtgc ctggatttat cggctctggaa gcaggccaat cggcggttgg	1020
tgatatctac gcctggtttg gtgcgctact cggctggccg ctggaacagc ttgccgccca	1080
gcatccggaa ctgaaaacgc aaatcaacgc cagccagaaa caactgcttc cggcgctgac	1140
cgaagcatgg gccaaaaatc cgtctctgga tcacctgccg gtggtgctcg actggtttaa	1200
cggccgccgc acaccgaacg ctaaccaacg cctgaaaggg gtgattaccg atcttaacct	1260
cgctaccgac gctccgctgc tgttcggcgg tttgattgct gccaccgctt ttggcgacg	1320
cgcaatcatg gagtgcctta ccgatcaggg gatcgccgtt aataacgtga tggcactggg	1380
cggcatcgcg cggaaaaacc aggtcattat gcaggcctgc tgcgacgtgc tgaatcgccc	1440
gctgcaaatt gttgcctctg accagtgctg tgcgctcggg gcggcgattt ttgctgccgt	1500
cgccgcgaaa gtgcacgcag acatcccatc agctcagcaa aaaatggcca gtgcggtaga	1560
gaaaaccctg caaccgtgca gcgagcaggc acaacgcttt gaacagcttt atcgccgcta	1620
tcagcaatgg gcgatgagcg ccgaacaaca ctatcttcca acttccgccc cggcacaggc	1680
tgcccaggcc gttgcgactc tataaggaca cgataatgac gatttttgat aattatgaag	1740
tgtggtttgt cattggcagc cagcatctgt atggcccggg aaccctgcgt caggtcaccc	1800
aacatgccga gcacgtcgtt aatgcgctga atacggaagc gaaactgccc tgcaaactgg	1860

tgttgaaacc gctgggcacc acgccggatg aaatcacccg tatttgccgc gacgcgaatt 1920
 acgacgatcg ttgcgctggg ctggtgggtg ggctgcacac cttctccccg gccaaaatgt 1980
 ggatcaacgg cctgaccatg ctcaacaaac cgttgctgca attccacacc cagttcaacg 2040
 cggcgctgcc gtgggacagt atcgatatgg actttatgaa cctgaaccag actgcacatg 2100
 gcggctcgca gttcggcttc attggcgcg cgtatgcgtca gcaacatgcc gtgggtaccg 2160
 gtcactggca ggataaacia gcccatgagc gtatcggtc ctggatgctg caggcggtct 2220
 ctaaacagga taccgctcat ctgaaagtct gccgatttgg cgataacatg cgtgaagtgg 2280
 cggtcaccca tggcgataaa gttgccgcac agatcaagtt cggtttctcc gtcaatacct 2340
 gggcggttgg cgatctgggtg cagggtggta actccatcag cgacggcgat gttaacgcgc 2400
 tggctgatga gtacgaaagc tgctacacca tgacgcctgc cacacaaatc cacggcaaaa 2460
 aacgacagaa cgtgctggaa gcggcgcgta ttgagctggg gatgaagcgt ttcctggaac 2520
 aaggtggctt ccacgcgttc accaccacct ttgaagattt gcacggctctg aaacagcttc 2580
 ctggtctggc cgtacagcgt ctgatgcagc agggttacgg ctttgccggc gaaggcgact 2640
 ggaaaactgc cgccctgctt cgcacatga aggtgatgtc aaccggctctg caggcgggca 2700
 cctcctttat ggaggactac acctatcact tcgagaaagg taatgacctg gtgctcggt 2760
 cccatattgct ggaagtctgc ccgtcgatcg ccgcagaaga gaaaccgatc ctgcacgttc 2820
 agcatctcgg tattggtggg aaggacgatc ctgcccgcct gatcttcaat acccaaaccg 2880
 gccagcgat tgcgccagc ttgattgatc tcggcgatcg ttaccgtcta ctggttaact 2940
 gcatcgacac ggtgaaaaca ccgcactccc tgccgaaact gccggtggcg aatgcgctgt 3000
 ggaaagcgca accggatctg ccaactgctt ccgaagcgtg gatcctcgtt ggtggcgcg 3060
 accatacgt cttcagccat gcactgaacc tcaacgatat gcgccaattc gccgagatgc 3120
 acgacattga aatcacggtg attgataacg acacacgcct gccagcgttt aaagacgcgc 3180
 tgcgctggaa cgaagtgtat tacggatttc gtcgctaagt ctagag 3226

<210> 28

<211> 25

<212> ADN

5 <213> secuencia artificial

<220>

<223> cebador

<400> 28

aaccatggcg attgcaattg gcctc 25

10 <210> 29

<211> 32

<212> ADN

<213> secuencia artificial

<220>

15 <223> cebador

<400> 29

ES 2 599 579 T3

ctctagactt agcgacgaaa tccgtaatac ac 32

<210> 30

<211> 889

<212> ADN

5 <213> secuencia artificial

<220>

<223> fragmento de PCR de araD

<400> 30

```

gtctagagaa ggagtcaca tgtagaaga tctcaaacgc caggtattag aagccaacct 60
ggcgctgcc aaacacaacc tggtcacgct cacatggggc aacgtcagcg cgttgatcg 120
cgagcgcggc gtctttgtga tcaaaccttc cggcgctcgat tacagcgta tgaccgctga 180
cgatatggtc gtggttagca tcgaaaccgg tgaagtgggt gaaggtacga aaaagccctc 240
ctccgacacg ccaactcacc ggctgctcta tcaggcattc ccctccattg gcggcattgt 300
gcatacgcac tcgcgccacg ccaccatctg ggcgcaggcg ggtcagtcga ttccagcaac 360
cggcaccacc cagccgact atttctacgg caccattccc tgtaccgca aaatgaccga 420
cgcagaaatc aacggcgaat atgagtggga aaccggtaac gtcacgtag aaaccttga 480
aaaacagggt atcgatgcag cgcaaatgcc cggcggttctg gtccattccc acggcccgtt 540
tgcatggggc aaaaatgccg aagatgcggt gcataacgcc atcgtgctgg aagaggtcgc 600
ttatatgggg atattctgcc gtcagttagc gccgcagtta ccggatatgc agcaaacgct 660
gctggataaa cactatctgc gtaagcatgg cgcgaggca tattacgggc agtaatgact 720
gtataaaacc acagccaatc aaacgaaacc aggtatact caagcctggt tttttgatgg 780
attttcagcg tggcgaggc aggttttatc ttaaccggac actggcgga caccgccaa 840
gggacagaag tctccttctg gctggcgacg gacaacgggc caagcttgg 889

```

10 <210> 31

<211> 32

<212> ADN

<213> secuencia artificial

<220>

15 <223> cebador

<400> 31

gtctagagaa ggagtcaca tgtagaaga tc 32

<210> 32

<211> 28

20 <212> ADN

<213> secuencia artificial

<220>

<223> cebador

<400> 32

25 ccaagcttgg cccgtgtcc gtcgccag 28

<210> 33

<211> 303

<212> ADN

<213> Zymomonas mobilis

<400> 33

```

tcgatcaaca acccgaaacc tatcgtaatg atgttttgcc cgatcagcct caatcgacaa      60
ttttacgcgt ttcgatcgaa gcagggacga caattggctg ggaacggat actggaataa      120
atggctcttcg ttatgggtatt gatgtttttg gtgcatcggc cccggcggaat gatctatatg      180
ctcatttcgg cttgaccgca gtcggcatca cgaacaaggt gttggccgcg atcgccggta      240
agtcggcagc ttaaaaaata gctatggaat ataatagcta cttataaagt taggagaata      300
aac                                                                 303

```

<210> 34

5 <211> 34

<212> ADN

<213> secuencia artificial

<220>

<223> cebador

10 <400> 34

gggagctcac tagtcgatc aacaaccga atcc 34

<210> 35

<211> 29

<212> ADN

15 <213> secuencia artificial

<220>

<223> cebador

<400> 35

agccatggtt atttcctaa ctattaag 29

20 <210> 36

<211> 323

<212> ADN

<213> secuencia artificial

<220>

25 <223> fragmento de PCR de Pgap

<400> 36

```

gggagctcac tagtcgatc aacaaccga atcctatcgt aatgatgttt tgcccgatca      60
gcctcaatcg acaattttac gcgtttcgat cgaagcaggg acgacaattg gctgggaacg      120
gtatactgga ataaatggtc ttcgttatgg tattgatgtt tttggtgcat cggccccggc      180
gaatgatcta tatgctcatt tcggcttgac cgcagtcggc atcacgaaca aggtggtggc      240
cgcgatcgcc ggtaagtcgg cacgttaaaa aatagctatg gaatataata gctacttaat      300
aagttaggag aataaccatg gct                                                                 323

```

<210> 37

30 <211> 35

<212> ADN

<213> secuencia artificial

<220>

<223> cebador

<400> 37
ctactcattt atcgatggag cacaggatga cgcct 35
<210> 38
<211> 34
5 <212> ADN
<213> secuencia artificial
<220>
<223> cebador
<400> 38
10 catcttacta cgcgttgga ggtcagcaag tgcc 34
<210> 39
<211> 36
<212> ADN
<213> secuencia artificial
15 <220>
<223> oligo de mutagénesis
<400> 39
aagttaggag aataaacatg gcgattgcaa ttggcc 36
<210> 40
20 <211> 36
<212> ADN
<213> secuencia artificial
<220>
<223> oligo de mutagénesis
25 <400> 40
ggccaattgc aatcgccatg ttattctcc taactt 36
<210> 41
<211> 9884
<212> ADN
30 <213> secuencia artificial
<220>
<223> plásmido construido
<400> 41
ctagttcgat caacaacccg aatcctatcg taatgatggt ttgcccgatc agcctcaatc 60
gacaatttta cgcgttttca tcgaagcagg gacgacaatt ggctgggaac ggtatactgg 120

aataaatggt cttcgttatg gtattgatgt ttttgggtgca tcggccccgg cgaatgatct	180
atatgctcat ttcggcttga ccgcagtcgg catcacgaac aagggtgttg ccgcgatcgc	240
cggtaagtcg gcacgttaaa aaatagctat ggaatataat agctacttaa taagttagga	300
gaataaacat ggcgattgca attggcctcg attttggcag tgattctgtg cgagctttgg	360
cgggtggactg cgctaccggt gaagagatcg ccaccagcgt agagtggat ccccggtggc	420
agaaagggca attttgtgat gccccgaata accagttccg tcatcatccg cgtgactaca	480
ttgagtcaat ggaagcggca ctgaaaaccg tgcttgacaga gcttagcgtc gaacagcgcg	540
cagctgtggt cgggattggc gttgacagta ccggctcgac gcccgcacccg attgatgccg	600
acggaaacgt gctggcgctg cggccggagt ttgccgaaaa ccggaacgcg atgttcgtat	660
tgtggaaaga ccacactgcg gttgaagaag cggaagagat tacccgtttg tgccacgcgc	720
cgggcaacgt tgactactcc cgctacattg gtggtattta ttccagcgaa tggttctggg	780
caaaaatcct gcatgtgact cgccaggaca gcgcctggc gcaatctgcc gcatcgtgga	840
ttgagctgtg cgactgggtg ccagctctgc tttccggtac caccgcgccg caggatattc	900
gtcgcggacg ttgcagcgcc gggcataaat ctctgtggca cgaaagctgg ggcggcctgc	960
cgccagccag tttctttgat gagctggacc cgatcctcaa tcgccatttg ccttccccgc	1020
tgttcactga cacttggaact gccgatattc cggtgggcac cttatgcccg gaatgggcgc	1080
agcgtctcgg cctgcctgaa agcgtggtga tttccggcgg cgcgtttgac tgccatatgg	1140
gcgcagttgg cgcaggcgca cagcctaacg cactggtaaa agttatcggg acttccacct	1200
gcgacattct gattgccgac aaacagagcg ttggcgagcg ggcagttaaa ggtatttgcg	1260
gtcaggttga tggcagcgtg gtgcctggat ttatcggctt ggaagcaggc caatcggcgt	1320
ttggtgatat ctacgcctgg tttggtcgcg tactcggtcg gccgctggaa cagcttgccg	1380
cccagcatcc ggaactgaaa acgcaaatca acgccagcca gaaacaactg cttccggcgc	1440
tgaccgaagc atgggcaaaa aatccgtctc tggatcacct gccggtggtg ctcgactggt	1500
ttaacggccg ccgcacaccg aacgctaacc aacgcctgaa aggggtgatt accgatctta	1560
acctcgctac cgacgtcccg ctgctgttcg gcggtttgat tgctgccacc gcctttggcg	1620
cacgcgcaat catggagtgc tttaccgatc aggggatcgc cgtaataaac gtgatggcac	1680
tgggcggcat cgcgcgaaa aaccaggcca ttatgcaggc ctgctgcgac gtgctgaatc	1740
gcccgctgca aattgttgcc tctgaccagt gctgtgcgct cgggtgcggcg atttttgctg	1800
ccgtcgccgc gaaagtgcac gcagacatcc catcagctca gcaaaaaatg gccagtgcgg	1860
tagagaaaac cctgcaaccg tgcagcgagc aggcacaacg ctttgaacag ctttatcgcc	1920
gctatcagca atgggcgatg agcgccgaac aacactatct tccaacttcc gccccggcac	1980
aggctgccca ggccgttgcg actctataag gacacgataa tgacgatttt tgataattat	2040

gaagtgtggt ttgtcattgg cagccagcat ctgtatggcc cggaaaccct gcgtcaggtc 2100
 acccaacatg ccgagcacgt cgtaaatgcg ctgaatacgg aagcgaaact gccctgcaaa 2160
 ctggtgttga aaccgctggg caccacgccg gatgaaatca ccgtattttg ccgcgacgcg 2220
 aattacgacg atcgttgccg tggctcgttg gtgtggctgc acaccttctc cccggccaaa 2280
 atgtggatca acggcctgac catgctcaac aaaccgttgc tgcaattcca caccagttc 2340
 aacgcggcgc tgccgtggga cagtatcgat atggacttta tgaacctgaa ccagactgca 2400
 catggcggtc gcgagttcgg cttcattggc gcgcgtatgc gtcagcaaca tgccgtgggt 2460
 accggtcact ggcaggataa acaagcccat gagcgtatcg gtcctggat gcgtcaggcg 2520
 gtctctaaac aggatacccg tcatctgaaa gtctgccgat ttggcgataa catgcgtgaa 2580
 gtggcgggtca ccgatggcga taaagttgcc gcacagatca agttcgggtt ctccgtcaat 2640
 acctgggcgg ttggcgatct ggtgcaggtg gtgaactcca tcagcgacgg cgatgttaac 2700
 gcgctggctg atgagtacga aagctgctac accatgacgc ctgccacaca aatccacggc 2760
 aaaaaacgac agaacgtgct ggaagcggcg cgtattgagc tggggatgaa gcgtttctctg 2820
 gaacaagggtg gcttcacgc gttcaccacc acctttgaag atttgacgg tctgaaacag 2880
 ctctctggtc tggccgtaca gcgtctgatg cagcagggtt acggctttgc gggcgaaggc 2940
 gactggaaaa ctgccgccct gcttcgcac atgaagggtga tgtcaaccgg tctgcagggc 3000
 ggcacctcct ttatggagga ctacacctat cacttcgaga aaggtaatga cctggtgctc 3060
 ggctcccata tgctggaagt ctgcccgctg atcgccgcag aagagaaacc gatcctcgac 3120
 gttcagcatc tcggtattgg tggttaaggac gatcctgccc gcctgatctt caatacccaa 3180
 accggcccag cgattgtcgc cagcttgatt gatctcggcg atcgttaccg tctactggtt 3240
 aactgcatcg acacggtgaa aacaccgcac tccctgccga aactgccggt ggcgaatgcg 3300
 ctgtggaaag cgcaaccgga tctgccaaact gcttcggaag cgtggatcct cgctggtggc 3360
 gcgcaccata ccgtcttcag ccatgcactg aacctcaacg atatgcgcca attcgccgag 3420
 atgcacgaca ttgaaatcac ggtgattgat aacgacacac gcctgccagc gtttaaagac 3480
 gcgctgcgct ggaacgaagt gtattacgga ttctcgctc aagtctagag aaggagtcaa 3540
 catgttagaa gatctcaaac gccagggtatt agaagccaac ctggcgctgc caaaacacaa 3600
 cctggtcacg ctacatggg gcaacgtcag cgccgttgat cgcgagcgcg gcgtctttgt 3660
 gatcaaacct tccggcgctg attacagcgt catgaccgct gacgatatgg tcgtggttag 3720
 catcgaaacc ggtgaagtgg ttgaaggtag gaaaaagccc tcctccgaca cgccaaactca 3780
 ccggtgctc tatcaggcat tccctccat tggcggcatt gtgcatacgc actcgcgcca 3840
 cgccaccatc tgggcgcagg cgggtcagtc gattccagca accggcacca cccacgccga 3900
 ctatttctac ggcaccattc cctgtaccg caaatgacc gacgcagaaa tcaacggcga 3960
 atatgagtgg gaaaccggtg acgtcatcgt agaaacctt gaaaaacagg gtatcgatgc 4020

agcgcaaatg cccggcggttc tgggtccattc ccacggccccg tttgcatggg gcaaaaatgc 4080
 cgaagatgcg gtgcataacg ccatcgtgct ggaagaggte gcttatatgg ggatattctg 4140
 ccgtcagtta ggcgcgcagt taccggatat gcagcaaacg ctgctggata aacactatct 4200
 gcgtaagcat ggcgcgaagg catattacgg gcagtaatga ctgtataaaa ccacagccaa 4260
 tcaaacgaaa ccaggctata ctcaagcctg gttttttgat ggattttcag cgtggcgag 4320
 gcagggtttta tcttaaccgg aactggcg gacaccccg aagggacaga agtctccttc 4380
 tggctggcga cggacaacgg gccaaagctt gaagggcgaa ttctgcagat atccatcaca 4440
 ctggcgcccg ctaattccgg atgagcattc atcaggcggg caagaatgtg aataaaggcc 4500
 ggataaaact tgtgcttatt tttctttacg gtctttaaaa aggccgtaat atccagctga 4560
 acggtctggt tataggtaca ttgagcaact gactgaaatg cctcaaaatg ttctttacga 4620
 tgccattggg atatatcaac ggtggtatat ccagtgattt tttctccat tttagcttcc 4680
 ttagctcctg aaaatctcga taactcaaaa aatacgcccc gtagtgatct tatttcatta 4740
 tgggtgaaagt tggaacctct tacgtgccga tcaacgtctc attttcgcca aaagttggcc 4800
 cagggttcc cggtatcaac agggacacca ggatttattt attctgcgaa gtgatcttcc 4860
 gtcacaggta tttattcggc gcaaagtgcg tcgggtgatg ctgccaaactt actgatttag 4920
 tgtatgatgg tgtttttgag gtgctccagt ggcttctgtt tctatcagct gtccctcctg 4980
 ttcagctact gacggggtgg tgcgtaacgg caaaagcacc gccggacatc agcgctagcg 5040
 gagtgtatac tggcttacta tgttggcact gatgagggtg tcagtgaagt gcttcatgtg 5100
 gcaggagaaa aaaggctgca ccggtgcgtc agcagaatat gtgatacagg atatattccg 5160
 cttcctcgct cactgactcg ctacgctcgg tcgttcgact gcggcgagcg gaaatggctt 5220
 acgaacgggg cggagatttc ctggaagatg ccaggaagat acttaacagg gaagtgagag 5280
 ggccgcggca aagccgtttt tccataggct ccgccccct gacaagcatc acgaaatctg 5340
 acgctcaaat cagtgggtgg gaaacccgac aggactataa agataccagg cgtttcccc 5400
 tggcggctcc ctggtgcgt ctctgttcc tgccttcgg tttaccggtg tcattccgct 5460
 gttatggcgg cgtttgtctc attccacgcc tgacactcag ttccgggtag gcagttcgct 5520
 ccaagctgga ctgtatgcac gaacccccg ttcagtccga ccgctgcgcc ttatccggtg 5580
 actatcgtct tgagtccaac ccggaaagac atgcaaaagc accactggca gcagccactg 5640
 gtaattgatt tagaggagtt agtcttgaag tcatgcgccg gttaaggcta aactgaaagg 5700
 acaagttttg gtgactgcgc tcctccaagc cagttacctc ggttcaaaga gttggtagct 5760
 cagagaacct tcgaaaaacc gccctgcaag gcggtttttt cgttttcaga gcaagagatt 5820
 acgcgagac caaacgatc tcaagaagat catcttatta atcagataaa atatttctag 5880
 atttcagtgc aatttatctc ttcaaatgta gcacctgaag tcagcccat acgatataag 5940

ttgtaatttct catgtttgac agcttatcat cgatggagca caggatgacg cctaacaatt	6000
cattcaagcc gacaccgctt cgcggcgcgcg ctttaattcag gagttaaaca tcatgaggga	6060
agcggtgatc gccgaagtat cgactcaact atcagaggta gttggcgctca tcgagcgcca	6120
tctcgaaccg acgttgctgg ccgtacattt gtacggctcc gcagtggatg gcggcctgaa	6180
gccacacagt gatattgatt tgctggttac ggtgactgta aggcttgatg aaacaacgcg	6240
gcgagctttg atcaacgacc ttttggaaac ttcggttcc cctggagaga gcgagattct	6300
ccgcgctgta gaagtcacca ttgttggtgca cgacgacatc attccgtggc gttatccagc	6360
taagcgcgaa ctgcaatttg gagaatggca gcgcaatgac attccttgag gtatcttcga	6420
gccagccacg atcgacattg atctggctat cttgctgaca aaagcaagag aacatagcgt	6480
tgccttggtgta ggtccagcgg cggaggaact ctttgatccg gttcctgaac aggatctatt	6540
tgaggcgcta aatgaaacct taacgctatg gaactcgccg cccgactggg ctggcgatga	6600
gcgaaatgta gtgcttacgt tgtcccgcat ttggtacagc gcagtaaccg gcaaaatcgc	6660
gccgaaggat gtcgctgccg actgggcaat ggagcgcctg ccggcccagt atcagcccgt	6720
catacttgaa gctaggcagg cttatcttgg acaagaagat cgcttggcct cgcgcgcaga	6780
tcagttggaa gaatttggtc actacgtgaa aggcgagatc accaaggtag tcggcaaata	6840
atgtctaaca attcgttcaa gccgacgccg cttcgcggcg cggcttaact caagcgttag	6900
agagctgggg aagactatgc gcgatctgtt gaaggtggtt ctaagcctcg tacttgcat	6960
ggcatcgggg caggcacttg ctgacctgcc aacgcgcctt tgtagtcttg gcctgttgtg	7020
tgcatgagca aatcaatggc accacccccct cctttttgag ctgaatggtc ataaaattta	7080
taattatcta tcgtaattcg gaatctatgt tcagggtctc gccattgctt tttgtctgct	7140
gggtcaagtt ccatgcctaa ggtttttaag acatcagaaa gaggtattgc acgcatgcta	7200
tcagcttttc ttctagctaa tgacagggct tcctctgctc tatctgctcg tttttttct	7260
tccacatata tcgccgcttt gtcagccagc ggctgtatta cggaaagtgc cgatttttgg	7320
gcttttaggc gttctttttc tgcccatctt tccttatattg taaaaattga ggggtgggatg	7380
ggtgcctgaa tcttgggatc tagctgtaaa gttttgttga tatttccgta atgtctttgg	7440
actctttgat gcgttgcttt tgaacctttt acgcctcttg ccagccctag aggetccata	7500
gaagccgcat aatccgtctg gagggcagaa agggcttttc gaccatcaaa ccatctcgat	7560
gcgtttaaac ggcctgtatc ggggtctcta ggcaccataa agccggttaa gtgggggtgtt	7620
gtttcatcag catgtagctg aagagataca aggttgtttt ctccaaagggt ttgttccgcc	7680
cattgctggg tgattgtttt ccagtgttcg agtttttcag gagtggcctg ttttgaccat	7740
tctggagaca taccaaagaa cagttctatg gcctgcacac cgttttttct aagaggettt	7800
cccgtttctt tctgaatttt attcagcata gatttaacat ctgctgatgg gtcagtagag	7860
cctttgagta tttcgtttag ttctttttcta tctgggtcag cgttttgtgt ttcgcggcct	7920

cgcgatcatat gcaggetcgc ggctttaatc gtgccaactg ttttatgttt ttcaaaccta	7980
aagattgcat agttcggcat gttttaactg ctttaatttg agaaaagacc agaggaaata	8040
atccagccta tatttccttc cctagtagcg aactggaatt gtttttccga aggaaaaaag	8100
caattccgta gtgagtactg aatttattct gattcgtctt gcttttgag cgtcttttg	8160
cgttctataa ctgttgtgaa agctacgcgg tcgccattga aaacgaaatt aggattaata	8220
aaataccatc cttggcgaa atgctttgca atgattttag ctttttctaa ttcggctaga	8280
cctcttgcaa aggtagcttg agatagtgc agtttttttt cttgtgcgtt aagaaagtcc	8340
tctaaaacga atttgtctaa agggacgagg tctttgctga tgctttgtc ttgaagtac	8400
caaaccagaa cgctgaaagc ttttattcca gcggctccta gttcaaaagt tagcgcgata	8460
ttggtgctaa ataattttac aaattcttca ctatcaacac gtctgtaagt cgtcacatga	8520
gtgccttgca tctcaccagt ggcttgattg accagaatgt tatcatctcg tcctaatacg	8580
gataactgaa ccctctgact ttttaactggc acaaccatac ctctgatgaa aggattctcg	8640
tcatatctga ttggctgctt tctcaatttt gtgcgcatat ttgataaacc tttaatcaaa	8700
aaaaccacat tttttgatta tacctattca tcgaatgagg caaggctctat caattttacc	8760
cctttttttg atagacgggt taatcaatat tgatagacc ctccacagat tctgaaaatc	8820
gacttcccta ttttagggat attttcacga ttccctttct tagttcttcc tagtggggaa	8880
attcgttgaa tctgcctcg gaaaaacat gagaaagctg ttggttatat acacgggcaa	8940
agccacccta tttttagcta ctggggaaag agataaggca gggattttgt aaaattaaaa	9000
ccggattttt cgctttacgg tttgtttagg cgcaactgtc tttttaagac cgcgtttaac	9060
catcaaaaga tcgttccaat cttttccgtg tatcatctgt tctttagggtg ggagccagtt	9120
ttcaactttt tttgttgaa acgcggcttt aatcgctccg actaatagcg atgctgctct	9180
ttgtcctaca gcatcccaat cataggcaat atggacagaa gatgcctttt caacgatttt	9240
tccgagagtt ttagtaagag acgttcttac gccgctggtg ctttaataatt ttacgccagc	9300
tttaattttt tctgggctta aaaagccgac tactgaaatc gcgtctatcg cactttcagc	9360
gatataaaga tcatactttt cgtcattttt tacattgatg ctgccagtaa aatgggcttc	9420
gcgactgctt cccaaggcta accctttaaa accactgctt gttccgcgta attctgcgcc	9480
ctgaagtgta tctttatcgt catacatcaa gaaggctaca ttaccgcat catctgttcg	9540
gatagagtca ggaatattgt taaatgatat tctcgggca gcgttggtc ctggccacgg	9600
gtgcgcatga tcgtgctcct gtcgttgagg acccgctag gctggcggg ttgccttact	9660
ggttagcaga atgaatcacc gatacgcgag cgaacgtgaa gcgactgctg ctgcaaacg	9720
tctgcgacct gagcaacaac atgaatggtc ttcggtttcc gtgtttcgta aagtctggaa	9780
acgcggaagt cccctacgtg ctgctgaagt tgcccgaac agagagtga accaaccggt	9840
gataccacga tactatgact gagagtcaac gccatgggag ctca	9884

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

 <220>
 5 <223> cebador

 <400> 42

 atgggagctc gttttctat ccccatcacc tcgg 34

 <210> 43
 <211> 35
 10 <212> ADN
 <213> secuencia artificial

 <220>
 <223> cebador

 <400> 43

 15 atcgactagt gggtcataat atgggcaaag acgct 35

 <210> 44
 <211> 895
 <212> ADN
 <213> secuencia artificial

 20 <220>
 <223> fragmento de PCR LDH-L

 <400> 44

 atgggagctc gttttctat ccccatcacc tcgggtttgt tgacaaaaaa aggtggccac 60
 taaattggct ttccgcaccg atgggatgat ttttattctt tgctattctt cgctctttgc 120
 ccaattcatt aaaagcggaa atcatcacca aagatagaag acgcagcctt caccatttca 180
 gattgccctt ctcgggcatt ttctgctgct agaatcctct taaaaatatt aaattccact 240
 ctattggtaa tatgtttccc tctttaggga acaaataaag cccttctttg ttctataaaa 300
 gttagcttac cgattttaca aaaaataata ccgcttcatt caatcggtaa tacatatctt 360
 ttttcttcaa aaaacttttc aagaggggtgt ctatgcgcgt cgcaatatc agttccaaaa 420
 actatgacca tcattctatt gaaaaagaaa atgaacatta tggccatgac cttgtttttc 480
 tgaatgagcg gcttaccaa gagacagcag aaaaagccaa agacgcagaa gctgtttgta 540
 tctttgtgaa tgacgaagcc aatgccgaag tgctggaaat ttggcaggc ttaggcatca 600
 agttgggtgc tcttcgttgc gccggttata acaatgtcga tctcgatgcg gccaaaaagc 660
 tgaatatcaa ggttgtgcgc gtgcctgcct attcgcccta ttcggttgcc gaatatgcag 720
 tagggatgtt gctcaccctg aatcggcaaa tttcacgcgg tttgaagcgg gttcgggaaa 780
 ataacttctc cttggaaggt ttgattggcc ttgatgtgca tgacaaaaca gtcggcatta 840
 tcggtgttgg tcatatcggg agcgtctttg cccatattat gaccactag tcgat 895

 <210> 45
 25 <211> 33
 <212> ADN
 <213> secuencia artificial

<220>

<223> cebador

<400> 45

gcgaattcat ggttttggtg ccaatgttat cgc 33

5 <210> 46

<211> 35

<212> ADN

<213> secuencia artificial

<220>

10 <223> cebador

<400> 46

ttaggcggcc gcgcggctga catacatctt gcgaa 35

<210> 47

<211> 1169

15 <212> ADN

<213> secuencia artificial

<220>

<223> fragmento de PCR LDH-R

<400> 47

gcgaattcat ggttttggtg ccaatgttat cgcctataaa ccgcattccag accccgaatt 60

ggcgaaaaag gtcgggttcc gcttcacctc tctcgatgaa gtgatcgaga ccagcgacat 120

catttcgctt cactgtccgc tcacgccaga aaatcatcac atgattaatg aagaaacact 180

ggcaagggca aaaaaaggct ttacctcgt caataccagt cgcggcggtt tggttgatac 240

caaggcgggtg attaaatcgc tgaaagccaa acatctcggc ggttatgcgg cggatgttta 300

cgaagaggag gggcctttat tcttcgaaaa tcacgctgac gatattatcg aagatgatat 360

tctcgaaagg ttgatcgctt tcccgaatgt ggttttcacg ggacatcagg cctttttgac 420

gaaagaggcc ttatcaaaca ttgctcacag tattctacaa gatatcagcg atgccgaagc 480

tggaaaagaa atgccggatg cgcttggtta gtagacaagc gacaattaac cttttgaaga 540

tcataatgat caaatttttg ggttaattcg gtagttatgg cataggctat tacgcgctaa 600

ttgatatcaa aaaaaagcat agccggacat cataccggct atgtttttta ttaggaaaaa 660

atttcctttc accttgctta gccatcgccg cattatttaa tcaatatgcc gagtttttct 720

tgaaatccct atcttacacc aaggccaaca agggaatcat ccatactcgg tgtcctatcc 780

tatgactttt taaattttct ccaaatttac taaaatcacg ccattctcagc ggctgctatt 840

20 ttcaaaaagc gcctctcaaa accgcttttt cctgctcaaa tatcggtacc caaaattccc 900

tcaaaaaagg cagggtatctt ttacaaaat cgcccctaata atctctcaat ccgctgcctt 960

gttcatatgt ttttgcaaat gattttttatt aaactttttt aggcgtatctt ttatcaagaa 1020

aatttaaata atcacatttt tattatttta gatttaagta ttgatacaag tgatatctat 1080

aaatgttttt ataactttct ggatcgtaat cggctggcaa tcgtttttccc tatattcgca 1140

agatgtatgt cagccgcgcg gccgcctaa 1169

<210> 48
 <211> 1098
 <212> ADN
 <213> secuencia artificial

5 <220>
 <223> fragmento de PCR LoxPw-aadA-LoxPw

<400> 48

```

ataacttcgt ataatgtatg ctatacgaag ttatgcggcc gcagcacagg atgacgccta      60
acaattcatt caagccgaca ccgcttcgcg gcgcgggetta attcaggagt taaacatcat      120
gaggggaagcg gtgatcgccg aagtatcgac tcaactatca gaggtagttg gcgtcatcga      180
gcgccatctc gaaccgacgt tgctggccgt acatttgtag ggctccgcag tggatggcgg      240
cctgaagcca cacagtgata ttgatttgct ggttacgggtg actgtaaggc ttgatgaaac      300
aacgcggcga gctttgatca acgacctttt ggaaacttcg gcttcccctg gagagagcga      360
gattctccgc gctgtagaag tcaccattgt tgtgcacgac gacatcattc cgtggcggtta      420
tccagctaag cgcgaactgc aatttggaaga atggcagcgc aatgacattc ttgcagggtat      480
cttcgagcca gccacgatcg acattgatct ggctatcttg ctgacaaaag caagagaaca      540
tagcgttgcc ttggtaggtc cagcggcgga ggaactcttt gatccgggtc ctgaacagga      600
tctatttgag gcgctaaatg aaaccttaac gctatggaac tcgccgcccg actgggctgg      660
cgatgagcga aatgtagtgc ttacgttgct ccgcatttgg tacagcgag taaccggcaa      720
aatcgcgccg aaggatgtcg ctgccgactg ggcaatggag cgcctgccgg cccagtatca      780
gcccgtcata cttgaagcta ggcaggctta tcttggaaca gaagatcgct tggcctcgcg      840
cgcatatcag ttggaagaat ttgttacta cgtgaaaggc gagatcacca aggtagtcgg      900
caaataatgt ctaacaattc gttcaagccg acgccgcttc gcggcgcggc ttaactcaag      960
cgtagagag ctggggaaga ctatgcgcga tctgttgaag gtggttctaa gcctcgtagt      1020
tgcgatggca tcggggcagg cacttgctga cctgccttaa ttaaataact tcgtataatg      1080
tatgctatac gaagttat                                     1098
    
```

<210> 49
 <211> 10441
 <212> ADN
 <213> secuencia artificial

<220>
 <223> plásmido construido

<400> 49

ctagttcgat caacaacccg aatcctatcg taatgatgtt ttgcccgatc agcctcaatc 60
 gacaatttta cgcgtttcga tcgaagcagg gacgacaatt ggctgggaac ggtatactgg 120
 aataaatggt cttcggttatg gtattgatgt ttttgggtgca tcggccccgg cgaatgatct 180
 atatgctcat ttcggttga ccgcagtcgg catcacgaac aagggtgttg ccgcgatcgc 240
 cggtaagtcg gcacgttaaa aatagctat ggaatataat agctacttaa taagttagga 300
 gaataaacat ggcgattgca attggcctcg attttggcag tgattctgtg cgagctttgg 360
 cgggtggactg cgctaccggt gaagagatcg ccaccagcgt agagtggat ccccggtggc 420
 agaaagggca attttgtgat gccccgaata accagttccg tcatcatccg cgtgactaca 480
 ttgagtcaat ggaagcggca ctgaaaacccg tgcttgca gaagcgtc gaacagcgcg 540
 cagctgtggt cgggattggc gttgacagta ccggtcgcac gccgcaccg attgatgccg 600
 acggaaacgt gctggcgctg cgcgcggagt ttgccgaaaa cccgaacgcg atgttcgtat 660
 tgtggaaaga ccacactgcg gttgaagaag cggaagagat taccggttg tgccacgcgc 720
 cgggcaacgt tgactactcc cgctacattg gtggtattta ttccagcgaa tggttctggg 780
 caaaaatcct gcatgtgact cgccaggaca gcgccgtggc gcaatctgcc gcatcgtgga 840
 ttgagctgtg cgactgggtg ccagctctgc tttccggtac caccgcgcg caggatattc 900
 gtcgcggacg ttgcagcgcc gggcataaat ctctgtggca cgaaagctgg ggcggcctgc 960
 cgccagccag tttctttgat gagctggacc cgatcctcaa tcgccatttg ccttccccgc 1020
 tgttcaactga cacttggaact gccgatattc cgggtggcac cttatgcccg gaatgggcgc 1080
 agcgtctcgg cctgcctgaa agcgtggtga tttccggcgg cgcgtttgac tgccatatgg 1140
 gcgcagttgg cgcaggcgca cagcctaacg cactggtaaa agttatcggg acttccacct 1200
 gcgacattct gattgccgac aaacagagcg ttggcgagcg ggcagttaaa ggtatttgcg 1260
 gtcaggttga tggcagcgtg gtgcctggat ttatcggtct ggaagcaggc caatcggcgt 1320
 ttggtgatat ctacgcctgg tttggtcgcg tactcggtcg gccgctggaa cagcttgccg 1380
 cccagcatcc ggaactgaaa acgcaaatac acgccagcca gaaacaactg cttccggcgc 1440
 tgaccgaagc atggggccaaa aatccgtctc tggatcacct gccggtggtg ctcgactggt 1500
 ttaacggccg ccgcacaccg aacgctaacc aacgcctgaa aggggtgatt accgatctta 1560
 acctcgctac cgacgctccg ctgctgttcg gcggtttgat tgetgccacc gcctttggcg 1620
 cacgcgcaat catggagtgc tttaccgatc aggggatcgc cgtaataaac gtgatggcac 1680
 tgggcggcat cgcgcggaaa aaccaggtca ttatgcaggc ctgctgcgac gtgctgaatc 1740
 gcccgctgca aattgttgcc tctgaccagt gctgtgcgct cgggtgcggcg atttttgctg 1800
 ccgtcgccgc gaaagtgcac gcagacatcc catcagctca gcaaaaaatg gccagtgcgg 1860

tagagaaaac cctgcaaccg tgcagcgagc aggcacaacg ctttgaacag ctttatcgcc	1920
gctatcagca atgggcgatg agcgccgaac aacactatct tccaacttcc gccccggcac	1980
aggctgcccga ggccgttgcg actctataag gacacgataa tgacgatttt tgataattat	2040
gaagtgtggt ttgtcattgg cagccagcat ctgtatggcc cggaaacctt gcgtcaggtc	2100
acccaacatg ccgagcacgt cgttaatgcg ctgaatacgg aagcgaaact gccctgcaaa	2160
ctggtgttga aaccgctggg caccacgccg gatgaaatca ccgctatttg ccgcgacgcg	2220
aattacgacg atcggtgcgc tggctcgttg gtgtggtgc acaccttctc cccggccaaa	2280
atgtggatca acggcctgac catgctcaac aaaccgttgc tgcaattcca caccagttc	2340
aacgcggcgc tgccgtggga cagtatcgat atggacttta tgaacctgaa ccagactgca	2400
catggcggtc gcgagttcgg cttcattggc gcgcgtatgc gtcagcaaca tgccgtggtt	2460
accggtcact ggcaggataa acaagcccat gagcgtatcg gtcctggat gcgtcaggcg	2520
gtctctaaac aggatacccg tcatctgaaa gtctgccgat ttggcgataa catgcgtgaa	2580
gtggcggtca ccgatggcga taaagttgcc gcacagatca agttcggttt ctccgtcaat	2640
acctggggcg ttggcgatct ggtgcaggtg gtgaactcca tcagcgacgg cgatgttaac	2700
gcgctggctg atgagtacga aagctgctac accatgacgc ctgccacaca aatccacggc	2760
aaaaaacgac agaacgtgct ggaagcggcg cgtattgagc tggggatgaa gcgtttcctg	2820
gaacaagggtg gcttcacgc gttcaccacc acctttgaag atttgacgg tctgaaacag	2880
cttctcgttc tggccgtaca gcgtctgatg cagcagggtt acggctttgc gggcggaaggc	2940
gactggaaaa ctgccgccct gttcgcac atgaagggtga tgtcaaccgg tctgcagggc	3000
ggcacctcct ttatggagga ctacacctat cacttcgaga aaggtaatga cctggtgctc	3060
ggctcccata tgctggaagt ctgcccgctg atcgccgcag aagagaaacc gatcctcgac	3120
gttcagcatc tcggtattgg tggttaaggac gatcctgccc gcctgatctt caatacccaa	3180
accggcccag cgattgtcgc cagcttgatt gatctcggcg atcgttaccg tctactggtt	3240
aactgcatcg acacggtgaa aacaccgcac tccctgccga aactgccggt ggcgaatgcg	3300
ctgtggaaaag cgcaaccgga tctgccaaact gcttcggaag cgtggatcct cgctggtggc	3360
gcgcaccata ccgtcttcag ccatgcactg aacctcaacg atatgcgcca attcgccgag	3420
atgcacgaca ttgaaatcac ggtgattgat aacgacacac gcctgccagc gtttaaagac	3480
gcgctgcgct ggaacgaagt gtattacgga tttcgtcgct aagtctagag aaggagtcaa	3540
catgttagaa gatctcaaac gccagggtatt agaagccaac ctggcgctgc caaaacacaa	3600
cctggtcacg ctcacatggg gcaacgtcag cgccgttgat cgcgagcgcg gcgtctttgt	3660
gatcaaacct tccggcgctg attacagcgt catgaccgct gacgatatgg tcgtggttag	3720
catcgaaacc ggtgaagtgg ttgaaggtag gaaaaagccc tcctccgaca cgccaactca	3780
ccggctgctc tatcaggcat tcccctccat tggcggcatt gtgcatacgc actcgcgcca	3840

cgccaccatc	tgggcgagc	cggtcagtc	gattccagca	accggcacca	cccacgccga	3900
ctattttctac	ggcaccattc	cctgtacccg	caaaatgacc	gacgcagaaa	tcaacggcga	3960
atatgagtgg	gaaaccggta	acgtcatcgt	agaaaccttt	gaaaaacagg	gtatcgatgc	4020
agcgcaaatg	cccggcggtc	tgggccattc	ccacggcccc	tttgcattgg	gcaaaaatgc	4080
cgaagatgcg	gtgcataacg	ccatcgtgct	ggaagaggtc	gcttatatgg	ggatattctg	4140
ccgtcagtta	gcgcgcagtc	taccggatat	gcagcaaacg	ctgctggata	aacactatct	4200
gcgtaagcat	ggcgcggaag	catattacgg	gcagtaatga	ctgtataaaa	ccacagccaa	4260
tcaaacgaaa	ccaggctata	ctcaagcctg	gttttttgat	ggattttcag	cgtggcgag	4320
gcaggtttta	tcttaacccg	acactggcgg	gacaccccg	aaggacaga	agtctccttc	4380
tggctggcga	cggacaacgg	gccaagcttg	gaaggcgaa	ttcgcgatcg	cataacttcg	4440
tataatgtat	gctatacgaa	gttatgcggc	cgcagcacag	gatgacgcct	aacaattcat	4500
tcaagccgac	accgcttcgc	ggcgcggtt	aattcaggag	ttaaactca	tgagggaagc	4560
ggtgatcgcc	gaagtatcga	ctcaactatc	agaggtagtt	ggcgtcatcg	agcgccatct	4620
cgaaccgacg	ttgctggccg	tacatttgta	cggctccgca	gtggatggcg	gcctgaagcc	4680
acacagtgat	attgatttgc	tggttacggc	gactgtaagg	cttgatgaaa	caacgcggcg	4740
agctttgatc	aacgaccttt	tggaaacttc	ggcttccctt	ggagagagcg	agattctccg	4800
cgctgtagaa	gtcaccattg	ttgtgcacga	cgacatcatt	ccgtggcggt	atccagctaa	4860
gcgcgaactg	caatttgag	aatggcagcg	caatgacatt	cttgcaggta	tcttcgagcc	4920
agccacgatc	gacattgatc	tggctatctt	gctgacaaaa	gcaagagaac	atagcggtgc	4980
cttggtaggt	ccagcggcgg	aggaactctt	tgatccgggt	cctgaacagg	atctatttga	5040
ggcgctaaat	gaaaccttaa	cgctatggaa	ctcgccgccc	gactgggctg	gcgatgagcg	5100
aaatgtagtg	cttacgttgt	cccgcatttg	gtacagcgca	gtaaccggca	aaatcgcgcc	5160
gaaggatgtc	gctgccgact	gggcaatgga	gcgcctgccg	gccagtatc	agcccgatc	5220
acttgaagct	aggcaggctt	atcttgga	agaagatcgc	ttggcctcgc	gcgcagatca	5280
gttggaaaga	tttgttcact	acgtgaaagg	cgagatcacc	aaggtagtcg	gcaaataatg	5340
tctaacaatt	cggtcaagcc	gacgccgctt	cgcggcgcgg	cttaactcaa	gcgttagaga	5400
gctggggaag	actatgcgcg	atctgttgaa	ggtggttcta	agcctcgta	ttgcgatggc	5460
atcggggcag	gcacttgctg	acctgcctta	attaaataac	ttcgtataat	gtatgctata	5520
cgaagttagt	gccggccaat	tcatggtttt	ggtgccaatg	ttatcgccct	taaaccgcat	5580
ccagaccccg	aattggcgaa	aaaggctcgt	ttccgcttca	cctctctcga	tgaagtgatc	5640
gagaccagcg	acatcatttc	gcttcaactg	ccgctcacgc	cagaaaatca	tcacatgatt	5700
aatgaagaaa	cactggcaag	ggcaaaaaaa	ggcttttacc	tcgtcaatac	cagtcgcggc	5760

ggcttggttg ataccaaggc ggtgattaaa tcgctgaaag ccaaacatct cggcggttat 5820
 ggcgcggatg ttacgaaga ggaggggcct ttattcttcg aaaatcacgc tgacgatatt 5880
 atcgaagatg atattctcga aagggtgatc gctttcccg atgtggtttt cacgggacat 5940
 caggcctttt tgacgaaaga ggccttatca aacattgctc acagtattct acaagatata 6000
 agcgatgccg aagctggaaa agaaatgccg gatgcgcttg tttagtagac aagcgacaat 6060
 taaccttttg aagatcataa tgatcaaatt tttgggttaa ttcggtagtt atggcatagg 6120
 ctattacgcg ctaattgata tcaaaaaaaaa gcatagccgg acatcatacc ggctatgttt 6180
 ttatttagga aaaaatttcc tttcaccttg cttagccatc gccgcattat ttaatcaata 6240
 tgccgagttt ttcttgaaat ccctatctta caccaaggcc aacaaggga tcatccatac 6300
 tcggtgtcct atcctatgac tttttaaat ttctccaaat ttactaaaat cagccatct 6360
 cagcggtgc tattttcaaa aagcgctct caaaaccgct tttcctgct caaatatcgg 6420
 atcccaaat tccctcaaaa aaggcagggt attttttaca aaatcgcccc taatatctct 6480
 caatccgctg ccttgttcat atgtttttgc aaatgatttt tattaaactt ttttaggcgt 6540
 atttttatca agaaaattta aataatcaca tttttattat tttagattta agtattgata 6600
 caagtatat ctataaatgt ttttataact ttctggatcg taatcggctg gcaatcgttt 6660
 tccctatatt cgcaagatgt atgtcagccg cgcggccgct ggtaccaat tcgccctata 6720
 gtgagtcgta ttacgcgcgc tactggccg tcgttttaca acgtcgtgac tgggaaaacc 6780
 ctggcgttac ccaacttaat cgccttgacg cacatcccc tttcgccagc tggcgtaata 6840
 gcgaagaggc cgcacccgat cgccttccc aacagttgcg cagcctgaat ggcgaatggg 6900
 acgcgccctg tagcggcgca ttaagcgcg cgggtgtggt ggttacgcgc agcgtgaccg 6960
 ctacacttgc cagcgcccta gcgccgctc ctttcgcttt cttcccttc tttctcgcca 7020
 cgttcgccgg ctttccccgt caagctctaa atcgggggct ccctttaggg ttccgattta 7080
 gtgctttacg gcacctcgac ccaaaaaaac ttgattaggg tgatggttca cgtagtgggc 7140
 catcgccctg atagacggtt tttcgccctt tgacggttga gtccacgttc ttaaatagt 7200
 gactcttggt ccaaactgga acaacactca accctatctc ggtctattct tttgatttat 7260
 aagggtttt gccgatttcg gcctattggt taaaaaatga gctgatttaa caaaaattta 7320
 acgcgaattt taacaaaata ttaacgctta caatttaggt ggcaactttc ggggaaatgt 7380
 gcgcggaacc cctatttggt tatttttcta aatacattca aatatgtatc cgctcatgag 7440
 acaataaacc tgataaatgc ttcaataata ttgaaaagg aagagtatga gtattcaaca 7500
 tttccgtgtc gcccttattc ctttttttgc ggcattttgc cttcctgttt ttgctcacc 7560
 agaaacgctg gtgaaagtaa aagatgctga agatcagttg ggtgcacgag tgggttacat 7620
 cgaactggat ctcaacagcg gtaagatcct tgagagtttt cgcgccgaag aacgttttcc 7680
 aatgatgagc acttttaag ttctgctatg tggcgcggtta ttatcccgt ttagcgcgg 7740

gcaagagcaa	ctcggtcgcc	gcatacacta	ttctcagaat	gacttggttg	agtactcacc	7800
agtcacagaa	aagcatctta	cggatggcat	gacagtaaga	gaattatgca	gtgctgccat	7860
aaccatgagt	gataaactg	cggccaactt	acttctgaca	acgatcggag	gaccgaagga	7920
gctaaccgct	tttttgaca	acatggggga	tcatgtaact	cgccttgatc	gttgggaacc	7980
ggagctgaat	gaagccatac	caaacgacga	gcgtgacacc	acgatgcctg	tagcaatggc	8040
aacaacgttg	cgcaaactat	taactggcga	actacttact	ctagcttccc	ggcaacaatt	8100
aatagactgg	atggaggcgg	ataaagttgc	aggaccactt	ctgcgctcgg	cccttcgggc	8160
tggttggttt	attgctgata	aatctggagc	cggtgagcgt	gggtctcggc	gtatcattgc	8220
agcactgggg	ccagatggta	agccctcccg	tatcgtagtt	atctacacga	cggggagtc	8280
ggcaactatg	gatgaacgaa	atagacagat	cgctgagata	ggtgcctcac	tgattaagca	8340
ttggtaactg	tcagaccaag	tttactcata	tatactttag	attgatttaa	aacttcattt	8400
ttaattttaa	aggatctagg	tgaagatcct	ttttgataat	ctcatgacca	aatccctta	8460
acgtgagttt	tcgttccact	gagcgtcaga	ccccgtagaa	aagatcaaag	gatcttcttg	8520
agatcccttt	tttctgcgcg	taatctgctg	cttgcaaaca	aaaaaaccac	cgctaccagc	8580
ggtggtttgt	ttgccggatc	aagagctacc	aactcttttt	ccgaaggtaa	ctggcttcag	8640
cagagcgcag	ataccaaata	ctgtccttct	agtgtagccg	tagttaggcc	accacttcaa	8700
gaactctgta	gcaccgccta	catacctcgc	tctgctaata	ctgttaccag	tggctgctgc	8760
cagtggcgat	aagtcgtgtc	ttaccggggt	ggactcaaga	cgatagttac	cggataaggc	8820
gcagcggctg	ggctgaacgg	ggggttcgtg	cacacagccc	agcttgagag	gaacgacct	8880
caccgaactg	agatacctac	agcgtgagct	atgagaaagc	gccacgcttc	ccgaaggag	8940
aaaggcggac	aggatccgg	taagcggcag	ggtcgggaaca	ggagagcgca	cgagggagct	9000
tccaggggga	aacgcctggt	atctttatag	tcctgtcggg	tttcgccacc	tctgacttga	9060
gcgtcgat	ttgtgatgct	cgtcaggggg	gcggagccta	tggaaaaacg	ccagcaacgc	9120
ggccttttta	cggttcctgg	ccttttgctg	gccttttgct	cacatgttct	ttcctgcgtt	9180
atccctgat	tctgtggata	accgtattac	cgcctttgag	tgagctgata	ccgctcgccg	9240
cagccgaacg	accgagcgca	gcgagtcagt	gagcgaggaa	gcggaagagc	gcccaatacg	9300
caaaccgcct	ctccccgcgc	gttggccgat	tcattaatgc	agctggcacg	acaggtttcc	9360
cgactggaaa	gcgggagctg	agcgcaacgc	aattaatgtg	agttagctca	ctcattaggc	9420
acccagggct	ttacacttta	tgcttccggc	tcgtatgttg	tgtggaattg	tgagcggata	9480
acaatttcac	acaggaaaca	gctatgacca	tgattacgcc	aagcgcgcaa	ttaaccctca	9540
ctaaagggaa	caaaagctgg	agctcggttt	tctatcccca	tcacctcgg	tttggtgaca	9600
aaaaaaggtg	gccactaaat	tggctttccg	caccgatggg	atgattttta	ttctttgcta	9660

ttcttctgctc tttgcccaat tcattaaaag cggaaatcat caccaaagat agaagacgca 9720
gccttcacca tttcagattg cccttctcgg gcattttctg ctgctagaat cctcttaaaa 9780
atattaaatt ccactctatt ggtaatatgt ttccctcttt aggggaacaaa taaagccctt 9840
ctttgttcta taaaagttag cttaccgatt ttacaaaaaa taataccgct tcattcaatc 9900
ggtaatacat atcttttttc ttcaaaaaac ttttcaagag ggtgtctatg cgcgtcgcaa 9960
tattcagttc caaaaactat gaccatcatt ctattgaaaa agaaaatgaa cattatggcc 10020
atgaccttgt ttttctgaat gagcggctta ccaaagagac agcagaaaaa gccaaagacg 10080
cagaagctgt ttgtatcttt gtgaatgacg aagccaatgc cgaagtgtg gaaattttgg 10140
caggcttagg catcaagttg gttgctcttc gttgcgccgg ttataacaat gtcgatctcg 10200
atgcccgaata aaagctgaat atcaaggttg tgcgcgtgcc tgcctattcg ccctattcgg 10260
ttgccgaata tgcagtaggg atgttgctca ccctgaatcg gcaaatttca cgcgggttga 10320
agcgggttcg ggaaaataac ttctccttgg aaggtttgat tggccttgat gtgcatgaca 10380
aaacagtcgg cattatcggg gttggtcata tcgggagcgt ctttgcccat attatgaccc 10440
a 10441

<210> 50

<211> 20

<212> ADN

5 <213> secuencia artificial

<220>

<223> cebador

<400> 50

gccttgggct tttaaagcct 20

10 <210> 51

<211> 22

<212> ADN

<213> secuencia artificial

<220>

15 <223> cebador

<400> 51

tcaatccacg atgcggcaga tt 22

<210> 52

<211> 20

20 <212> ADN

<213> secuencia artificial

<220>

<223> cebador

<400> 52

25 ccagtatcag cccgtcatatc

20

<210> 53

<211> 26

<212> ADN

<213> secuencia artificial

<220>
<223> cebador

<400> 53

tctcggagag atagagggtca gtcgac 26

5 <210> 54
<211> 27
<212> ADN
<213> secuencia artificial

<220>
10 <223> cebador

<400> 54

aaccatgggt actatcaata cggaatc 27

<210> 55
<211> 27
15 <212> ADN
<213> secuencia artificial

<220>
<223> cebador

<400> 55

20 ttgaattcct gatgtgtgtt accgcaa 27

<210> 56
<211> 1550
<212> ADN
<213> secuencia artificial

25 <220>
<223> fragmento de PCR de araE

<400> 56

aaccatgggt actatcaata cggaatctgc tttaacgccca cggttctttgc gggatacgcg 60

gcgtatgaat atgtttgttt cggtagctgc tgcggtcgca ggattgttat ttggtcttga 120

tatcggcgta atcgccggag cgttgccgtt cattaccgat cactttgtgc tgaccagtcg 180

tttgcaggaa tgggtgggta gtagcatgat gctcgggtgca gcaattggtg cgctgtttaa 240

tgggttggctg tcgttccgcc tggggcgtaa atacagcctg atggcggggg ccacacctgtt 300

tgtactcggt tctatagggc ccgcttttgc gaccagcgta gagatgttaa tcgccgctcg 360

tgtggtgctg ggcattgctg tcgggatcgc gtcttacacc gtcctctctgt atctttctga 420

aatggcaagt gaaaacgttc gcggtaagat gatcagtatg taccagttga tggtcacact 480

cggcacgtg ctggcggttt tatccgatac agcgttcagt tatagcggtg actggcgcg 540

aatgttgggg gttcttgctt taccagcagt tctgctgatt attctggtag tcttcctgcc 600

aaatagcccg cgctggctgg cggaaaaggg gcgtcatatt gaggcggaag aagtattgcg 660
 tatgctgccc gatacgtcgg aaaaagcgcg agaagaactc aacgaaattc gtgaaagcct 720
 gaagttaaaa cagggcggtt gggcactgtt taagatcaac cgtaacgtcc gtcgtgctgt 780
 gtttctcggg atgttggtgc aggcgatgca gcagtttacc ggtatgaaca tcatcatgta 840
 ctacgcgccc cgtatcttca aaatggcggg ctttacgacc acagaacaac agatgattgc 900
 gactctggtc gtagggctga cctttatgtt cgccaccttt attgcggtgt ttacggtaga 960
 taaagcaggg cgtaaaccgg ctctgaaaat tggtttcagc gtgatggcgt taggcactct 1020
 ggtgctgggc tattgcctga tgcagtttga taacggtagc gcttccagtg gcttgctcctg 1080
 gctctctgtt ggcattgacga tgatgtgtat tgccggttat gcgatgagcg ccgcgccagt 1140
 ggtgtggatc ctgtgctctg aaattcagcc gctgaaatgc cgcgatttcg gtattacctg 1200
 ttcgaccacc acgaactggg tgcgaatat gattatcggc gcgaccttcc tgacactgct 1260
 tgatagcatt ggcgtgccc gtacgttctg gctctacact gcgctgaaca ttgcgtttgt 1320
 gggcattact ttctggctca ttccggaaac caaaaatgtc acgctggaac atatcgaacg 1380
 caaactgatg gcaggcgaga agttgagaaa tatcggcgtc tgatttcacg ggccggatgt 1440
 gctgtacatc cggccctttt ttctgtaata gagattgggc acttggccgt tgaggcggtt 1500
 gtctcgttcc ttattcagcc ttgttgcggg aacacacatc aggaattcaa 1550

<210> 57

<211> 32

<212> ADN

5 <213> secuencia artificial

<220>

<223> cebador

<400> 57

aaccatggcg cacaatttta ctaaagccct gg 32

10 <210> 58

<211> 30

<212> ADN

<213> secuencia artificial

<220>

15 <223> cebador

<400> 58

ccgaattcct tctctttct tattgtgtg 30

<210> 59

<211> 3744

20 <212> ADN

<213> secuencia artificial

<220>

<223> fragmento de PCR de araFGH

<400> 59

aaccatggcg cacaaattta ctaaagccct ggcagccatt ggtctggcag ccgttatgtc	60
acaatccgct atggcggaga acctgaagct cggttttctg gtgaagcaac cggaagagcc	120
gtggttccag accgaatgga agtttgccga taaagccggg aaggatttag ggtttgaggt	180
tattaagatt gccgtgccgg atggcgaaaa aacattgaac gcgatcgaca gcctggctgc	240
cagtggcgca aaaggtttcg ttatttgtac tccggacccc aaactcggct ctgccatcgt	300
cgcgaaagcg cgtggctacg atatgaaagt cattgccgtg gatgaccagt ttgttaacgc	360
caaaggtaag ccaatggata ccgttccgct ggtgatgatg gggcgacta aaattggcga	420
acgtcagggc caggaactgt ataaagagat gcagaaacgt ggctgggatg tcaaagaaag	480
cgcggtgatg gcgattaccg ccaacgaact ggataccgcc cgcgcgcgta ctacgggatc	540
tatggatgcg ctgaaagcgg ccggattccc ggaaaaacaa atttatcagg tacctaccaa	600
atetaacgac atccccgggg catttgacgc tgccaactca atgctgggtc aacatccgga	660
agttaaacad tggctgatcg tcggtatgaa cgacagcacc gtgctgggcg gcgtacgcgc	720
gacggaaggt cagggtctta aagcgccga tatcatcggc attggcatta acggtgtgga	780
tgcggtgagc gaactgtcta aagcacaggc aaccggcttc tacggttccc tgctgccaaag	840
cccgacgta catggctata aatccagcga aatgctttac aactgggtag caaaagacgt	900
tgaaccgcca aaatttaccg aagttaccga cgtggtactg atcacgcgtg acaacttta	960
agaagaactg gagaaaaaag gtttaggcgg taagtaattt gccgaaaaa tccccctctg	1020
catgatgcag agggggtgtg aacgaccagt gattcacgga gacgttatgc aacagtctac	1080
cccgtatctc tcatttcgcg gcacggtaa aacgtttccc ggcgttaagg cgctgacgga	1140
tattagtttt gactgctatg ccggtcaggt tcatgcgttg atgggtgaaa atggcgcagg	1200
aaaatcaact ctcttaaaaa tcctcagcgg caactatgcg ccaaccacgg gttctgtagt	1260
gattaatggg caggaaatgt cttttccga cacgaccgca gcacttaacg cgggcgtggc	1320
gattatttac caggaactgc atctcgtgcc ggaaatgacc gtcgcggaaa acatctatct	1380
cggccagctg ccgcataaag gcggcattgt gaatcgctca ttgctgaatt atgagggggg	1440
tttacaactt aaacatcttg gtatggatat tgacccggac acgccgtga aatatctctc	1500
cattggtcag tggcagatgg ttgaaatcgc caaagcgctg gcgcgtaacg ccaaaattat	1560
cgcctttgat gagccaacca gctccctctc tgcccgtgaa atcgacaatc ttttcgcgct	1620
tattcgtgaa ctgcgaaaag aggggcgggt aatcttatac gtttctcacc gtatggaaga	1680
aatatttgcc ctacagcatg ccattactgt ctttaaagat ggacgttatg tcaaaacctt	1740
taccgatatg cagcaggttg accacgacgc gctggtgcag gcgatggctg ggcgcgacat	1800
tggcgatatc tacggctggc aaccgcgtag ttatggcgag gagcgcctac gtcttgatgc	1860
tgtgaaagca ccaggcgtgc gtacgccaat aagtctggcg gttcgcagtg gtgaaattgt	1920

tgggctgttt ggtctggtag gggcggggcg tagcgaatta atgaaaggca tgtttggcgg	1980
gacgcaaate accgccggtc aggtttatat cgaccaacag ccgatcgata ttcgtaaacc	2040
gagccacgcc attgccgcag gcatgatgct ctgcccggaa gatcgcaaag cggaaggcat	2100
tattcccgtg cactccgttc gcgacaatat caacatcagt gccagacgta aacatgtgct	2160
cggcggttgt gtaatcaaca acggttggga agaaaacaat gccgatcacc acattcgttc	2220
gctcaacate aaaacgccgg gcgcggagca actgatcatg aatctctcag gcggaaatca	2280
gcaaaaagcc attctgggcc gctggttate ggaagagatg aaggtcattt tgctggatga	2340
acctacgcgc ggcattgatg ttggcgctaa gcacgaaata tataacgtaa tttatgcgct	2400
ggcggcgagc ggcgtggcgg tgctgtttgc ctccagcgac ttacctgaag tcctcggcgt	2460
tgccgaccgg attgtggtga tgcgggaagg tgaaatcgcc ggtgaattgt tacacgagca	2520
ggcagatgag cgtcaggcac tgagccttgc gatgcctaaa gtcagccagg ctgttgccctg	2580
agtaaggaga gtatgatgtc ttctgtttct acatcgggggt ctggcgacac taagtcgtca	2640
ttcagcttcg ggcgtatctg ggatcagtac ggcattgctg tggtgtttgc ggtgctcttt	2700
atcgctctg ccatTTTTgt cccaaatttt gccaccttca ttaatatgaa agggttgggc	2760
ctggcaattt ccatgtcggg gatggtggct tgtggcatgt tgttctgcct cgcttccggt	2820
gactttgacc tttctgtcgc ctccgtaatt gcctgtgcgg gtgtcaccac ggcggtggtt	2880
attaacctga ctgaaagcct gtggattggc gtggcagcgg ggttgttget gggcgttctc	2940
tgtggcctgg tcaatggctt tgttatcgcc aaactgaaaa taaatgctct gatcacgaca	3000
ttggcaacga tgcagattgt tcgaggtctg gcgtacatca tttcagacgg taaagcggtc	3060
ggtatcgaag atgaaagctt ctttgcctt ggttacgcca actggttcgg tctgcctgcg	3120
ccaatctggc tcaccgtcgc gtgtctgatt atctttggtt tgctgctgaa taaaaccacc	3180
tttggctgta acaccctggc gattggcggg aacgaagagg ccgcgcgtct ggcgggtgta	3240
ccggttggtc gcacaaaaat tattatcttt gttctctcag gcctggatc agcgatagcc	3300
ggaattattc tggcttcacg tatgaccagt gggcagccaa tgacgtcgat tggttatgag	3360
ctgattgtta tctccgcctg cgttttaggt ggcgtttctc tgaaagggtg catcggaata	3420
atctcatatg tgggtggcggg tatcttaatt ttaggcaccg tggaaaacgc catgaacctg	3480
cttaatatTT ctccttctgc gcagtacgtg gtctgcggct taatcctgct ggcagcgggtg	3540
atcttcgacc gttacaagca aaaagcgaaa cgcaactgtct gatgcttttt tctgcaacaa	3600
tttagcggtt tttccaccca tagccaaccg ccataacggg tggctgttct tcggtgcaaa	3660
tggcgacccc cgtcacactg tctatactta catgtctgta aagcgcgttc tgcgcaacac	3720
aataagaaaa gagaaggaat tcgg	3744

<210> 60

<211> 27

<212> ADN

5 <213> secuencia artificial

<220>
 <223> cebador
 <400> 60
 gggagctcac tagtcgatct gtgctgt 27

5 <210> 61
 <211> 23
 <212> ADN
 <213> secuencia artificial

10 <220>
 <223> cebador
 <400> 61
 agccatggtt acctccggga aac 23

15 <210> 62
 <211> 181
 <212> ADN
 <213> Actinoplanes missouriensis
 <400> 62
 cgatctgtgc tgtttgccac ggtatgcagc accagcgcga gattatgggc tcgcacgctc 60
 gactgtcgga cgggggcact ggaacgagaa gtcaggcgag ccgtcacgcc cttgacaatg 120
 ccacatcctg agcaaataat tcaaccacta aacaaatcaa ccgcgtttcc cggaggtaac 180
 c 181

20 <210> 63
 <211> 201
 <212> ADN
 <213> secuencia artificial
 <220>
 <223> fragmento de PCR Pgi

25 <400> 63
 gggagctcac tagtcgatct gtgctgtttg ccacgggatg cagcaccagc gcgagattat 60
 gggctcgcaac gctcgactgt cggacggggg cactggaacg agaagtcagg cgagccgtca 120
 cggccttgac aatgccacat cctgagcaaa taattcaacc actaaacaaa tcaaccgcgt 180
 ttcccggagg taaccatggc t 201

30 <210> 64
 <211> 911
 <212> ADN
 <213> secuencia artificial
 <220>
 <223> marcador de resistencia a cloranfenicol
 <400> 64
 gtgacggaag atcacttcgc agaataaata aatcctggtg tccctgttga taccgggaag 60

ccctgggcca acttttggcg aaaatgagac gttgatcggc acgtaagagg ttccaacttt 120
caccataatg aaataagatc actaccgggc gtattttttg agttatcgag attttcagga 180
gctaaggaag ctaaaatgga gaaaaaaatc actggatata ccaccgttga tatatcccaa 240
tggtatcgta aagaacattt tgaggcattt cagtcagttg ctcaatgtac ctataaccag 300
accgttcagc tggatattac ggccttttta aagaccgtaa agaaaaataa gcacaagttt 360
tatecggcct ttattcacat tcttgcccgc ctgatgaatg ctcatccgga attccgtatg 420
gcaatgaaag acggtgagct ggtgatatgg gatagtgttc acccttggtta caccgttttc 480
catgagcaaa ctgaaacgtt ttcacgcgtc tggagtgaat accacgacga tttccggcag 540
tttctacaca tatattcgca agatgtggcg tgttacgggtg aaaacctggc ctatttccct 600
aaagggttta ttgagaatat gtttttcgtc tcagccaatc cctgggtgag tttcaccagt 660
tttgatttaa acgtggccaa tatggacaac ttcttcgccc ccgttttcac catgggcaaa 720
tattatacgc aaggcgacaa ggtgctgatg ccgctggcga ttcagggttca tcatgccgtt 780
tgtgatggct tccatgtcgg cagaatgctt aatgaattac aacagtactg cgatgagtgg 840
cagggcgggg cgtaattttt ttaaggcagt tattggtgcc cttaaaccgcc tggttgctac 900
gcctgaataa g 911

<210> 65

<211> 7224

<212> ADN

5 <213> secuencia artificial

<220>

<223> plásmido construido

<400> 65

ggcttactat gttggcactg atgaggggtgt cagtgaagtg cttcatgtgg caggagaaaa 60
aaggctgcac cgggtgcgtca gcagaatatg tgatacagga tatattccgc ttcctcgtc 120
actgactcgc tacgctcggc cgttcgactg cggcgagcgg aaatggctta cgaacggggc 180
ggagatttcc tggaagatgc caggaagata cttaacaggg aagtgagagg gccgcggcaa 240
agccgttttt ccataggctc cgccccctg acaagcatca cgaaatctga cgctcaaatc 300
agtgggtggc aaacccgaca ggactataaa gataccaggc gtttccccct ggcggtccc 360
tcgtgcgtc tectgttct gcctttcggc ttaccgggtg cattccgctg ttatggccgc 420
gtttgtctca ttccacgcct gacactcagt tccgggtagg cagttcgtc caagctggac 480
tgtatgcacg aacccccgt tcagtccgac cgctgcgcct tatccggtaa ctatcgtctt 540
gagtccaacc cggaaagaca tgcaaaagca cactggcag cagccactgg taattgattt 600
agaggagtta gtcttgaagt catgcgccgg ttaaggctaa actgaaagga caagttttgg 660
tgactgcgct cctccaagcc agttacctcg gttcaaagag ttggtagctc agagaacctt 720

cgaaaaaccg ccctgcaagg cgggttttttc gttttcagag caagagatta cgcgcagacc 780
 aaaacgatct caagaagatc atcttattaa tcagataaaa tattttctaga tttcagtgca 840
 atttatctct tcaaatgtag cacctgaagt cagccccata cgatataagt tgtaattctc 900
 atgtttgaca gcttatcatc gatggagcac aggatgacgc ctaacaattc attcaagccg 960
 acaccgcttc gcggcgcggc ttaattcagg agttaaacat catgagggaa gcggtgatcg 1020
 ccgaagtatc gactcaacta tcagaggtag ttggcgatc cgagcgccat ctccaaccga 1080
 cgttgctggc cgtacatttg tacggctccg cagtggatgg cggcctgaag ccacacagtg 1140
 atattgattt gctggttacg gtgactgtaa ggcttgatga aacaacgcgg cgagctttga 1200
 tcaacgacct tttggaaact tcggcttccc ctggagagag cgagattctc cgcgctgtag 1260
 aagtcaccat tgttgtgcac gacgacatca ttccgtggcg ttatccagct aagcgcgaac 1320
 tgcaatttgg agaatggcag cgcaatgaca ttcttgacgg tatcttcgag ccagccacga 1380
 tcgacattga tctggctatc ttgctgacaa aagcaagaga acatagcgtt gccttggttag 1440
 gtccagcggc ggaggaactc tttgatccgg ttccctgaaca ggatctattt gaggcgctaa 1500
 atgaaacctt aacgctatgg aactcgccgc ccgactgggc tggcgatgag cgaaatgtag 1560
 tgcttacgtt gtccgcatt tggtagcgc cagtaaccgg caaaatcgcg ccgaaggatg 1620
 tcgctgccga ctgggcaatg gagcgctgc cggcccagta tcagcccgtc atacttgaag 1680
 ctaggcaggc ttatcttgga caagaagatc gcttggcctc gcgcgcagat cagttggaag 1740
 aatttgttca ctacgtgaaa ggcgagatca ccaaggtagt cggcaaataa tgtctaacaa 1800
 ttctgtcaag ccgacgcgc ttcgcggcgc ggcttaactc aagcgtaga gagctgggga 1860
 agactatgcg cgatctgttg aaggtaggtc taagcctcgt acttgcgatg gcatcggggc 1920
 aggcacttgc tgacctgcca acgcgccttt gtagtcttgg cctgttgtgt gcatgagcaa 1980
 atcaatggca ccacccctc ctttttgagc tgaatggtca taaaatttat aattatctat 2040
 cgtaattcgg aatctatgtt cagggtctcg ccattgcttt ttgtctgctg ggtcaagttc 2100
 catgcctaag gtttttaaga catcagaaaag aggtattgca cgcagctat cagcttttct 2160
 tctagctaat gacagggtt cctctgctct atctgctcgt tttttttctt ccacatatct 2220
 cgccgctttg tcagccagcg gctgtattac ggaaagtgcc gatttttggg ctttttaggcg 2280
 ttctttttct gccattctt ccttatttgt aaaaattgag ggtgggatgg gtgcctgaat 2340
 cttgggatct agctgtaaag ttttggtgat atttccgtaa tgtctttgga ctctttgatg 2400
 cgttgctttt gaacctttta cgcctctggc cagccctaga ggctccatag aagccgcata 2460
 atccgtctgg agggcagaaa gggcttttcg accatcaaac catctcgatg cgtttaaacg 2520
 gcctgtatcg gggctcttag gcaccataaa gccgggttaag tggggtgttg tttcatcagc 2580
 atgtagctga agagatacaa ggttggtttt tccaaagggt tgttccgcc attgctgggt 2640
 gattgttttc cagtgttcga gtttttcagg agtggcctgt tttgaccatt ctggagacat 2700

accaaagaac agttctatgg cctgcacacc gttttttcta agaggctttc cegtttcttt	2760
ctgaatttta ttcagcatag atttaacatc tgctgatggg tcagtagagc ctttgagtat	2820
ttcgtttagt tcttttctat ctgggtcagc gttttgtgtt tcgcggcctc gcgtcatatg	2880
caggctcgcg gctttaatcg tgccaactgt tttatgtttt tcaaacctaa agattgcata	2940
gttcggcatg ttttaactgc ttttaattga gaaaagacca gaggaataa tccagcctat	3000
atttctttcc ctagtagcga actggaattg tttttccgaa ggaaaaaagc aattccgtag	3060
tgagtactga atttattctg attcgtcttg cttttggagc gtctttttgc gttctataac	3120
tgttgtgaaa gctacgcggt cgccattgaa aacgaaatta ggattaataa aataccatcc	3180
ttggcgaaaca tgctttgcaa tgatttttagc tttttctaata tcggctagac ctcttgcaaa	3240
ggtagcttga gatagtgcc a gttttttttc ttgtgcgtta agaaagtcct ctaaaacgaa	3300
tttgctctaaa gggacgaggt ctttgctgat gcctttgtct tgaagtatcc aaaccagaac	3360
gctgaaagct tttattccag cggctcctag ttcaaaaagtt agcgcgatat tgggtgctaaa	3420
taattttaca aattcttcac tatcaacacg tctgtaagtc gtcacatgag tgccttgcat	3480
ctcaccagtg gcttgattga ccagaatgtt atcatctcgt cctaatacgag ataactgaac	3540
cctctgactt ttaactggca caaccatacc ttcgatgaaa ggattctcgt catatctgat	3600
tggctgcttt ctcaattttg tcgccatatt tgataaacct ttaatcaaaa aaaccacatt	3660
ttttgattat acctattcat cgaatgaggc aaggctctatc aattttaccc ctttttttga	3720
tagacggttt aatcaatatt gatagacccc ttcacagatt ctgaaaatcg acttccttat	3780
tttagggata ttttcacgat tccctttctt agttcttcct agtggggaaa ttcgttgaat	3840
cctgcctcgg aaaaaccatg agaaagctgt tggttatata cacgggcaaa gccaccctat	3900
ttttagctac tggggaaaga gataaggcag ggtatttgta aaattaaaac cggatttttc	3960
gctttacggg ttgttttaggc gcaactgtct ttttaagacc gcgtttaacc atcaaaagat	4020
cgttccaatc ttttcctgtt atcatctgtt ctttaggtgg gagccagttt tcaacttttt	4080
ttgttgga aa cgcggttta atcgctccga ctaatagcga tgctgctctt tgtcctacag	4140
catcccaatc ataggcaata tggacagaag atgccttttc aacgattttt cggagagttt	4200
tagtaagaga cgttcttacg ccgctggtgc ttaataattt tacgccagct ttaatttttt	4260
ctgggcttaa aaagccgact actgaaatcg cgtctatcgc actttcagcg atataaagat	4320
catacttttc gtcatttttt acattgatgc tgccagtaaa atgggcttcg cgactgcttc	4380
ccaaggctaa ccctttaaaa ccaactgcttg ttccgcgtaa ttctgcgcc tgaagtgtat	4440
ctttatcgtc atacatcaag aaggctacat taccgcgac atctgttcgg atagagtcag	4500
gaatattgtt aatgatatt cctcggttag tcgatctgtg ctgtttgcca cggtatgcag	4560
caccagcgcg agattatggg ctgcacgct cgactgtcgg acgggggcac tggaacgaga	4620

agtcaggcga gccgtcacgc ccttgacaat gccacatcct gagcaaataa ttcaaccact 4680
 aaacaaatca accgcgtttc ccggaggtaa ccatgggttac tatcaatacg gaatctgctt 4740
 taacgccacg ttctttgcgg gatacgcggc gtatgaatat gtttgtttcg gtagctgctg 4800
 cggtcgcagg attgttattt ggtcttgata tcggcgtaat cgccggagcg ttgccgttca 4860
 ttaccgatca ctttgtgctg accagtcgtt tgcaggaatg ggtgggttagt agcatgatgc 4920
 tcggtgcagc aattgggtgc ctgtttaatg gttggctgtc gttccgcctg gggcgtaaata 4980
 acagcctgat ggccggggcc atcctgtttg tactcggttc tatagggtcc gcttttgcca 5040
 ccagcgtaga gatgttaate gccgctcgtg tgggtgctggg cattgctgtc gggatcgctg 5100
 cttacaccgc tcctctgtat ctttctgaaa tggcaagtga aaacggtcgc ggtaagatga 5160
 tcagtatgta ccagttgatg gtcacactcg gcatcggtgt ggcgttttta tccgatacag 5220
 cgttcagtta tagcggtaac tggcgcgcaa tgttgggggt tcttgcttta ccagcagttc 5280
 tgctgattat tctggtagtc ttcctgccaa atagcccgcg ctggctggcg gaaaaggggc 5340
 gtcattattga ggcggaagaa gtattgcgtg tgctgcgcga tacgtcggaa aaagcgcgag 5400
 aagaactcaa cgaaattcgt gaaagcctga agttaaaca gggcggttgg gcactgttta 5460
 agatcaaccg taacgtccgt cgtgctgtgt ttctcggtat gttgttgagc gcgatgcagc 5520
 agtttaccgg tatgaacatc atcatgtact acgcgccgcg tatcttcaaa atggcgggct 5580
 ttacgaccac agaacaacag atgattgcga ctctggctgt agggctgacc tttatgttcg 5640
 ccacctttat tgcggtgttt acggtagata aagcagggcg taaaccggct ctgaaaattg 5700
 gtttcagcgt gatggcgtaa ggcactctgg tgctgggcta ttgcctgatg cagtttgata 5760
 acggtaacgg ttccagtggc ttgtcctggc tctctgttgg catgacgatg atgtgtattg 5820
 ccggttatgc gatgagcgcc gcgccagtgg tgtggatcct gtgctctgaa attcagccgc 5880
 tgaaatgccg cgatttcggt attacctgtt cgaccaccac gaactgggtg tcgaatatga 5940
 ttatcgggcg gaccttcctg aactgcttg atagcattgg cgctgccggg acgttctggc 6000
 tctacactgc gctgaacatt gcgtttgtgg gcattacttt ctggctcatt ccggaaacca 6060
 aaaatgtcac gctggaacat atcgaacgca aactgatggc aggcgagaag ttgagaaata 6120
 tcggcgctctg atttcacggg ccggatgtgc tgtacatccg gccctttttt cgtaataaga 6180
 gattgggcac ttggccgttg aggcgtttgt ctcgttcctt attcagcctt gttgcggtaa 6240
 cacacatcag gaattctgca gatatccatc aactggcgcg ccgcgtgacg gaagatcact 6300
 tcgcagaata aataaatcct ggtgtccctg ttgataccgg gaagccctgg gccaaactttt 6360
 ggcgaaaatg agacgttgat cggcacgtaa gaggttccaa ctttcacat aatgaaataa 6420
 gatcactacc gggcgatttt tttgagttat cgagattttc aggagctaag gaagctaaaa 6480
 tggagaaaaa aatcactgga tataccaccg ttgatataat ccaatggcat cgtaaagaac 6540
 attttgaggc atttcagtca gttgctcaat gtacctataa ccagaccgtt cagctggata 6600

ttacggcctt	tttaaagacc	gtaaagaaaa	ataagcacia	gttttatccg	gcctttattc	6660
acattcttgc	ccgcctgatg	aatgctcatc	cggaattccg	tatggcaatg	aaagacggtg	6720
agctggtgat	atgggatagt	gttcaccctt	gttacaccgt	tttccatgag	caaactgaaa	6780
cgttttcatc	gctctggagt	gaataccacg	acgatttccg	gcagtttcta	cacatatatt	6840
cgcaagatgt	ggcgtgttac	ggtgaaaacc	tggcctattt	ccctaaaggg	tttattgaga	6900
atatgttttt	cgtctcagcc	aatccctggg	tgagtttcac	cagttttgat	ttaaacgtgg	6960
ccaatatgga	caacttcttc	gcccccgttt	tcaccatggg	caaataattat	acgcaaggcg	7020
acaaggtgct	gatgccgctg	gcgattcagg	ttcatcatgc	cgtttgtgat	ggcttccatg	7080
tccgcagaat	gcttaatgaa	ttacaacagt	actgcgatga	gtggcagggc	ggggcgtaat	7140
ttttttaagg	cagttattgg	tgcccttaaa	cgcctgggtg	ctacgcctga	ataagttaat	7200
taatgcgcta	gcggagtgtg	tact				7224

<210> 66

<211> 9418

<212> ADN

5 <213> secuencia artificial

<220>

<223> plásmido construido

<400> 66

ctagtcgac	tgtgctgttt	gccacgggat	gcagcaccag	cgcgagatta	tgggctcgca	60
cgctcgactg	tccgacgggg	gcactggaac	gagaagtcag	gcgagccgtc	acgcccttga	120
caatgccaca	tcctgagcaa	ataattcaac	cactaaacaa	atcaaccgcg	tttcccggag	180
gtaaccatgg	cgcacaaatt	tactaaagcc	ctggcagcca	ttgggtctggc	agccgttatg	240
tcacaatccg	ctatggcgga	gaacctgaag	ctcgggtttc	tgggtgaagca	accggaagag	300
ccgtgggttc	agaccgaatg	gaagtttgcc	gataaagccg	ggaaggattt	agggtttgag	360
gttattaaga	ttgccgtgcc	ggatggcgaa	aaaacattga	acgcgatega	cagcctggct	420
gccagtggcg	caaaagggtt	cgttatttgt	actccggacc	ccaaactcgg	ctctgccatc	480
gtcgcgaaag	cgcgtggcta	cgatatgaaa	gtcattgccg	tggatgacca	gtttgttaac	540
gccaaaggta	agccaatgga	taccgttccg	ctggtgatga	tggcggcgac	taaaattggc	600
gaacgtcagg	gccaggaact	gtataaagag	atgcagaaac	gtggctggga	tgtcaaagaa	660
agcgcggtga	tggcgattac	cgccaacgaa	ctggataccg	cccgcgcgcg	tactacggga	720
tctatggatg	cgctgaaagc	ggccggattc	ccgaaaaaac	aaatttatca	ggtacctacc	780
aaatctaacg	acatcccggg	ggcatttgac	gctgccaaact	caatgctggt	tcaacatccg	840
gaagttaaac	attggctgat	cgtcggtatg	aacgacagca	ccgtgctggg	cggcgtagcg	900
gcgacggaag	gtcagggcct	taaagcggcc	gatatcatcg	gcattggcat	taacggtgtg	960

gatgcggtga gcgaactgtc taaagcacag gcaaccggct tctacggttc cctgctgcc 1020
agcccgacg tacatggcta taaatccagc gaaatgcttt acaactgggt agcaaaagac 1080
gttgaaccgc caaaatttac cgaagttacc gacgtggtac tgatcacgcg tgacaacttt 1140
aaagaagaac tggagaaaaa aggttttaggc ggtaagtaat ttgccggaaa aattccccctc 1200
tgcgatgatgc agaggggggtg tgaacgacca gtgattcacg gagacgttat gcaacagtct 1260
accccgatatc totcatttcg cggcatcggg aaaacgtttc ccggcggttaa ggcgctgacg 1320
gatattagtt ttgactgcta tgcgggtcag gttcatgcgt tgatgggtga aaatggcgca 1380
ggaaaatcaa ctctcttaaa aatcctcagc ggcaactatg cgccaaccac gggttctgta 1440
gtgattaatg ggcaggaaat gtccttttcc gacacgaccg cagcacttaa cgcggggcgtg 1500
gcgattatct accaggaact gcatctcgtg ccggaaatga ccgtcgcgga aaacatctat 1560
ctcgccacgc tgcgcataa aggcggcatt gtgaatcgtc cattgctgaa ttatgaggcg 1620
ggttttacaac ttaaacatct tggatggat attgaccggg acacgcgcgt gaaatatctc 1680
tccattgggc agtggcagat ggttgaaatc gccaaagcgc tggcgcgtaa cgccaaaatt 1740
atcgcccttg atgagccaac cagctccctc tctgcccgtg aaatcgacaa tcttttccgc 1800
gttattcgtg aactgcgaaa agagggggcg gtaatcttat acgtttctca ccgtatggaa 1860
gaaatatttg ccctcagcga tgccattact gtcttttaaag atggacgtta tgtcaaaacc 1920
tttaccgata tgcagcaggt tgaccacgac gcgctgggtc aggcgatggg cgggcgcgac 1980
attggcgata tctacggctg gcaaccgcgt agttatggcg aggagcgct acgtcttgat 2040
gctgtgaaag caccaggcgt gcgtacgcca ataagctctg cggttcgcag tgggtgaaatt 2100
gttgggctgt ttggctctgt aggggcgggg cgtagcgaat taatgaaagg catgtttggc 2160
gggacgcaaa tcaccgccg tccaggtttat atcgaccaac agccgatcga tattcgtaaa 2220
ccgagccacg ccattgccgc aggcgatgat ctctgcccgg aagatcgcaa agcgggaaggc 2280
attattcccg tgcactccgt tcgcgacaat atcaacatca gtgccagacg taaacatgtg 2340
ctcggcgggt gtgtaatcaa caacggttg gaagaaaaca atgccgatca ccacattcgt 2400
tcgctcaaca tcaaaacgcc gggcgcgagg caactgatca tgaatctctc aggcggaaat 2460
cagcaaaaag ccattctggg ccgctgggta tcggaagaga tgaaggcat tttgctggat 2520
gaacctacgc gcggcattga tgttggcgct aagcacgaaa tatataacgt aatttatgcg 2580
ctggcgggcg agggcggtgg ggtgctgttt gcctccagcg acttacctga agtcctcggc 2640
gttgccgacc ggattgtggg gatgcgggaa ggtgaaatcg ccggtgaatt gttacacgag 2700
caggcagatg agcgtcaggc actgagcctt gcgatgccta aagtcagcca ggctgttgcc 2760
tgagtaagga gagtatgatg tcttctgttt ctacatcggg gtctggcgca cctaagtcgt 2820
cattcagctt cgggcgtatc tgggatcagt acggcatgct ggtgggtgtt gcgggtgctct 2880
ttatcgctg tgcattttt gtcccaaaatt ttgccacctt cattaatatg aaagggttg 2940

gcctggcaat ttccatgtcg gggatgggtgg cttgtggcat gttgttctgc ctcgcttccg 3000
 gtgactttga cctttctgtc gcctccgtaa ttgcctgtgc ggggtgcacc acggcggtgg 3060
 ttattaacct gactgaaagc ctgtggattg gcgtggcagc ggggttggtg ctgggcgttc 3120
 tctgtggcct ggtcaatggc tttgttatcg ccaaactgaa aataaatgct ctgatcacga 3180
 cattggcaac gatgcagatt gttcgaggtc tggcgtacat catttcagac ggtaaagcgg 3240
 tcggtatcga agatgaaagc ttctttgccc ttggttacgc caactgggtc ggtctgcctg 3300
 cgccaatctg gctcaccgtc gcgtgtctga ttatctttgg tttgctgctg aataaaacca 3360
 cctttggctg taacaccctg gcgattggcg ggaacgaaga ggccgcgcgt ctggcggtg 3420
 taccggttgt tcgcaccaa attattatct ttgttctctc aggcctggta tcagcgatag 3480
 ccggaattat tctggcttca cgtatgacca gtgggcagcc aatgacgtcg attggttatg 3540
 agctgattgt tatctccgcc tgcgttttag gtggcgtttc tctgaaaggt ggcacgga 3600
 aaatctcata tgtgggtggcg ggtatcttaa ttttaggcac cgtggaaaac gccatgaacc 3660
 tgcttaatat ttctccttcc gcgcagtacg tggttcgcgg cttaatcctg ctggcagcgg 3720
 tgatcttcga ccgttacaag caaaaagcga aacgcactgt ctgatgcttt tttctgcaac 3780
 aatttagcgt ttttccac catagccaac cgccataacg gttggctgtt cttcgttgca 3840
 aatggcgacc cccgtcacac tgtctatact tacatgtctg taaagcgcgt tctgcgcaac 3900
 acaataagaa aagagaagga attctgcaga tatccatcac actggcgggc gcgtgacgga 3960
 agatcacttc gcagaataaa taaatcctgg tgtccctgtt gataccggga agccctgggc 4020
 caacttttgg cgaaaatgag acgttgatcg gcacgtaaga ggttccaact ttcaccataa 4080
 tgaaataaga tcaactaccg gcgtattttt tgagttatcg agattttcag gagctaagga 4140
 agctaaaatg gagaaaaaaa tcaactggata taccaccgtt gatatatccc aatggcatcg 4200
 taaagaacat tttgaggcat ttcagtcagt tgcctaatgt acctataacc agaccgttca 4260
 gctggatatt acggcctttt taaagaccgt aaagaaaaat aagcacaagt tttatccggc 4320
 ctttattcac attcttgcgc gcctgatgaa tgctcatccg gaattccgta tggcaatgaa 4380
 agacggtgag ctggtgatat gggatagtggt tcacccttgt tacaccgttt tccatgagca 4440
 aactgaaacg ttttcatcgc tctggagtga ataccacgac gatttccggc agtttctaca 4500
 catatattcg caagatgtgg cgtgttacgg tgaaaacctg gcctatttcc ctaaagggtt 4560
 tattgagaat atgtttttcg tctcagccaa tccctgggtg agtttcacca gttttgattt 4620
 aaacgtggcc aatatggaca acttcttcgc ccccgttttc accatgggca aatattatac 4680
 gcaaggcgac aaggtgctga tgccgctggc gattcaggtt catcatgccg tttgtgatgg 4740
 cttccatgtc ggcagaatgc ttaatgaatt acaacagtac tgcatgagt ggcagggcgg 4800
 ggcgtaattt ttttaaggca gttattgggt cccttaaacg cctggttgct acgcctgaat 4860

aagttaatta atgcgctagc ggagtgtata ctggcttact atgttggcac tgatgagggt 4920
gtcagtgaag tgcttcatgt ggcaggagaa aaaaggctgc accggtgcgt cagcagaata 4980
tgtgatacag gatataattcc gcttcctcgc tcaactgactc gctacgctcg gtcgttcgac 5040
tgcggcgagc ggaaatggct tacgaacggg gcggagattt cctggaagat gccaggaaga 5100
tacttaacag ggaagtgaga gggccgcggc aaagccgttt ttccataggc tccgcccccc 5160
tgacaagcat cacgaaatct gacgctcaaa tcagtgggtgg cgaaaccoga caggactata 5220
aagataccag gcgtttcccc ctggcggtc cctcgtgcgc tctcctgttc ctgcctttcg 5280
gtttaccggt gtcattccgc tgttatggcc gcgtttgtct cattccacgc ctgacactca 5340
gttcgggta ggcagttcgc tccaagctgg actgtatgca cgaaccccc gttcagtcgc 5400
accgctgcgc cttatccggt aactatcgtc ttgagtccaa cccggaaaga catgcaaaag 5460
caccactggc agcagccact ggtaattgat ttagaggagt tagtcttgaa gtcattgcgc 5520
ggttaaggct aaactgaaag gacaagtttt ggtgactgcg ctctccaaag ccagttacct 5580
cggttcaaag agttggtagc tcagagaacc ttcgaaaaac cgccctgcaa ggcggttttt 5640
tcgttttcag agcaagagat tacgcgcaga caaaacgat ctcaagaaga tcatcttatt 5700
aatcagataa aatattttcta gatttcagtg caatttatct cttcaaagt agcacctgaa 5760
gtcagcccca tacgatataa gttgtaattc tcatgtttga cagcttatca tcgatggagc 5820
acaggatgac gcctaacaat tcattcaagc cgacaccgct tcgcggcgcg gcttaattca 5880
ggagttaaac atcatgaggg aagcggtagt cgccgaagta tcgactcaac tatcagaggt 5940
agttggcgtc atcgagcgcc atctcgaacc gacgttgctg gccgtacatt tgtacggctc 6000
cgcagtggtt ggcggcctga agccacacag tgatattgat ttgctggtta cggtgactgt 6060
aaggcttgat gaaacaacgc ggcgagcttt gatcaacgac cttttggaaa cttcggcttc 6120
ccctggagag agcgagattc tccgcgctgt agaagtcacc attgttgtgc acgacgacat 6180
cattccgtgg cgttatccag ctaagcgcga actgcaattt ggagaatggc agcgcaatga 6240
cattcttgca ggtatcttcg agccagccac gatcgacatt gatctggcta tcttgctgac 6300
aaaagcaaga gaacatagcg ttgccttggg aggtccagcg gcggaggaac tctttgatcc 6360
ggttcctgaa caggatctat ttgaggcgct aaatgaaacc ttaacgctat ggaactcgcc 6420
gcccgactgg gctggcgatg agcgaaatgt agtgcttacg ttgtcccga tttggtacag 6480
cgcagtaacc ggcaaaatcg cgccgaagga tgtcgttgcc gactgggcaa tggagcgect 6540
gccggcccag tatcagcccc tcatacttga agctaggcag gcttatcttg gacaagaaga 6600
tcgcttgccc tcgcgcgcag atcagttgga agaatttgtt cactacgtga aaggcgagat 6660
caccaaggta gtcggcaaat aatgtctaac aattcgttca agccgacgcc gcttcgcggc 6720
gcggcttaac tcaagcgta gagagctggg gaagactatg cgcgatctgt tgaaggtggg 6780
tctaagcctc gtacttgca tggcatcggg gcaggcactt gctgacctgc caacgcgcct 6840

ttgtagtctt	ggcctgttgt	gtgcatgagc	aaatcaatgg	caccaccccc	tcctttttga	6900
gctgaatggg	cataaaattt	ataattatct	atcgtaattc	ggaatctatg	ttcaggggtct	6960
cgccattgct	ttttgtctgc	tgggtcaagt	tccatgccta	aggtttttaa	gacatcagaa	7020
agagggtattg	cacgcatgct	atcagctttt	cttctagcta	atgacagggc	ttcctctgct	7080
ctatctgctc	gttttttttc	ttccacatat	ctcgccgctt	tgtcagccag	cggctgtatt	7140
acggaaagtg	cgcatttttg	ggcttttagg	cgttcttttt	ctgcccattc	ttccttattt	7200
gtaaaaattg	aggggtggat	gggtgcctga	atcttgggat	ctagctgtaa	agttttggtg	7260
atatttccgt	aatgtctttg	gactctttga	tgcgttgctt	ttgaaccttt	tacgcctctg	7320
gccagcccta	gaggctccat	agaagccgca	taatccgtct	ggagggcaga	aagggttttt	7380
cgaccatcaa	accatctcga	tgcgtttaaa	cggcctgtat	cggggctctc	aggcaccata	7440
aagccgggta	agtgggggtg	tgtttcatca	gcatgtagct	gaagagatac	aagggtgttt	7500
tctccaaagg	tttgttccgc	ccattgctgg	gtgattgttt	tccagtgttc	gagtttttca	7560
ggagtggcct	gttttgacca	ttctggagac	ataccaaaga	acagttctat	ggcctgcaca	7620
ccgttttttc	taagaggctt	tcccgtttct	ttctgaattt	tattcagcat	agatttaaca	7680
tctgctgatg	ggtcagtaga	gcctttgagt	atttcgttta	gttcttttct	atctgggtca	7740
gcgttttggt	tttcgcggcc	tcgctcata	tgcaggctcg	cggctttaat	cgtgccaaact	7800
gttttatggt	tttcaaacct	aaagattgca	tagttcggca	tgttttaact	gctttaattt	7860
gagaaaagac	cagaggaaat	aatccagcct	atatttcttt	ccctagtagc	gaactggaat	7920
tgtttttccg	aaggaaaaaa	gcaattccgt	agtgagtact	gaatttattc	tgattcgtct	7980
tgccttttga	gcgtcttttt	gcgttctata	actggtgtga	aagctacgcg	gtcgccattg	8040
aaaacgaaat	taggattaat	aaaataccat	ccttggcgaa	catgctttgc	aatgatttta	8100
gctttttcta	attcggctag	acctcttgca	aaggtagctt	gagatagtgc	cagttttttt	8160
tcttgtgcgt	taagaaagtc	ctctaaaacg	aatttgtcta	aagggacgag	gtctttgctg	8220
atgcctttgt	cttgaagtat	ccaaaccaga	acgctgaaag	cttttattcc	agcggctcct	8280
agttcaaaag	ttagcgcgat	attggtgcta	aataatttta	caaattcttc	actatcaaca	8340
cgtctgtaag	tcgtcacatg	agtgccttgc	atctcaccag	tggcttgatt	gaccagaatg	8400
ttatcatctc	gtcctaateg	agataactga	accctctgac	ttttaactgg	cacaaccata	8460
ccttcgatga	aaggattctc	gtcatatctg	attggctgct	ttctcaattt	tgtcgccata	8520
tttgataaac	ctttaatcaa	aaaaaccaca	ttttttgatt	atacctattc	atcgaatgag	8580
gcaagggtcta	tcaattttac	cccttttttt	gatagacggg	ttaatcaata	ttgatagacc	8640
ccttcacaga	ttctgaaaat	cgacttccct	attttagggg	tattttcacg	attccctttc	8700
ttagttcttc	ctagtgggga	aattcgttga	atcctgcctc	ggaaaaacca	tgagaaagct	8760

gttggttata tacacgggca aagccaccct attttttagct actggggaaa gagataaggc	8820
agggatatttg taaaattaaa accggatttt tcgcttttacg gtttgtttag gcgcaactgt	8880
ctttttaaga ccgcgtttaa ccatcaaaag atcgttccaa tcttttccgt gtatcatctg	8940
ttcttttaggt gggagccagt tttcaacttt tttgttgga aacgcggctt taatcgctcc	9000
gactaatagc gatgctgctc tttgtcctac agcatcccaa tcataggcaa tatggacaga	9060
agatgccttt tcaacgattt ttcggagagt tttagtaaga gacgttctta cgccgctggg	9120
gcttaataat tttacgccag ctttaatttt ttctgggctt aaaaagccga ctactgaaat	9180
cggtctatc gcactttcag cgatataaag atcatacttt tcgtcatttt ttacattgat	9240
gctgccagta aaatgggctt cgcgactgct tccaaggct aaccctttaa aaccactgct	9300
tgttccgcgt aattctgcgc cctgaagtgt atctttatcg tcatacatca agaaggctac	9360
attaccgcga tcatctgttc ggatagagtc aggaatattg ttaaatgata ttcctcgg	9418

REIVINDICACIONES

1. Un microorganismo recombinante del género *Zymomonas* o *Zymobacter* que utiliza arabinosa para producir etanol, dicho microorganismo comprende al menos un gen heterólogo que codifica un simportador de protones de arabinosa, en donde dicho simportador es expresado por dicho microorganismo.
- 5 2. El microorganismo recombinante de la reivindicación 1 en donde el simportador de protones de arabinosa está codificado por la región codificante de un gen araE.
3. El microorganismo recombinante de la reivindicación 1, en donde la utilización de arabinosa se mejora en al menos aproximadamente 10% en comparación con un microorganismo original en donde dicho microorganismo original carece del al menos un gen heterólogo que codifica un simportador de protones de arabinosa.
- 10 4. El microorganismo recombinante de la reivindicación 1, en donde la cepa utiliza adicionalmente xilosa para producir etanol.
5. Un proceso para generar un microorganismo recombinante del género *Zymomonas* o *Zymobacter* que tiene mayor utilización de arabinosa que comprende:
 - a) proporcionar una cepa recombinante de *Zymomonas* o *Zymobacter* que utilice arabinosa para producir etanol en condiciones adecuadas; y
 - 15 b) introducir al menos un gen heterólogo que codifique un simportador de protones de arabinosa en la cepa de (a) para la expresión de dicho simportador.
6. El proceso según la reivindicación 5, que comprende además adaptar la cepa ya sea antes o después de la etapa (b), o tanto antes como después de la etapa (b), mediante cultivo en serie en medio que contiene arabinosa como la única fuente de carbono mediante el cual se produce una cepa adaptada y en donde dicha cepa ha mejorado adicionalmente la utilización de arabinosa en comparación con la cepa sin adaptación.
- 20 7. El proceso según la reivindicación 6, en donde la cepa adaptada utiliza de manera adicional xilosa y glucosa para la producción de etanol en medio de azúcares mezclados que comprende arabinosa, xilosa y glucosa.
8. Un proceso para producir etanol que comprende:
 - a) proporcionar una cepa recombinante de *Zymomonas* o *Zymobacter* que utiliza arabinosa para producir etanol, dicha cepa comprende al menos un gen heterólogo que codifica un simportador de protones de arabinosa, en donde dicho simportador es expresado por dicha cepa; y
 - 25 b) cultivar la cepa de (a) en un medio que comprende arabinosa a través del cual la arabinosa es convertida en etanol.
- 30 9. El proceso según la reivindicación 8 en donde el simportador de protones de arabinosa está codificado por la región codificante de un gen araE.
10. El proceso según la reivindicación 8, en donde la utilización de arabinosa se mejora en al menos aproximadamente 10% en comparación con un microorganismo original en donde dicho microorganismo original carece de un gen heterólogo que codifica un simportador de protones de arabinosa.
- 35 11. El proceso según la reivindicación 8 en donde la cepa de (a) además es capaz de utilizar xilosa y glucosa para producir etanol.
12. El proceso según la reivindicación 8 en donde la cepa de (a) se ha adaptado mediante cultivo en serie en medio que contiene arabinosa como la única fuente de carbono mediante el cual se produce una cepa adaptada para arabinosa en donde dicha cepa adaptada para arabinosa ha aumentado la producción de etanol en comparación con la cepa de (a) que no ha sido adaptada.
- 40 13. El proceso según la reivindicación 8 en donde la conversión de arabinosa en etanol se ve aumentada con respecto a la conversión de arabinosa en etanol mediante una cepa original recombinante sin al menos un gen heterólogo que codifica un simportador de protones de arabinosa.
14. El proceso de la reivindicación 8 en donde el medio comprende ya sea una mezcla de azúcares que comprende arabinosa o arabinosa como azúcar único.
- 45 15. Un método para mejorar la utilización de arabinosa mediante un microorganismo que utiliza arabinosa que comprende:
 - (a) proporcionar un microorganismo que utiliza arabinosa, donde dicho microorganismo se selecciona del grupo que consiste en una cepa recombinante de *Zymomonas* o *Zymobacter* que utiliza arabinosa para producir etanol;

(b) introducir en el genoma de dicho microorganismo al menos un gen heterólogo que codifica un simportador de protones de arabinosa, donde dicho simportador es expresado por dicho microorganismo; y

- 5 (c) poner en contacto el microorganismo de (b) con un medio que comprende arabinosa, donde dicho microorganismo metaboliza dicha arabinosa con una mayor velocidad en comparación con dicho microorganismo que no presenta el simportador de protones de arabinosa.

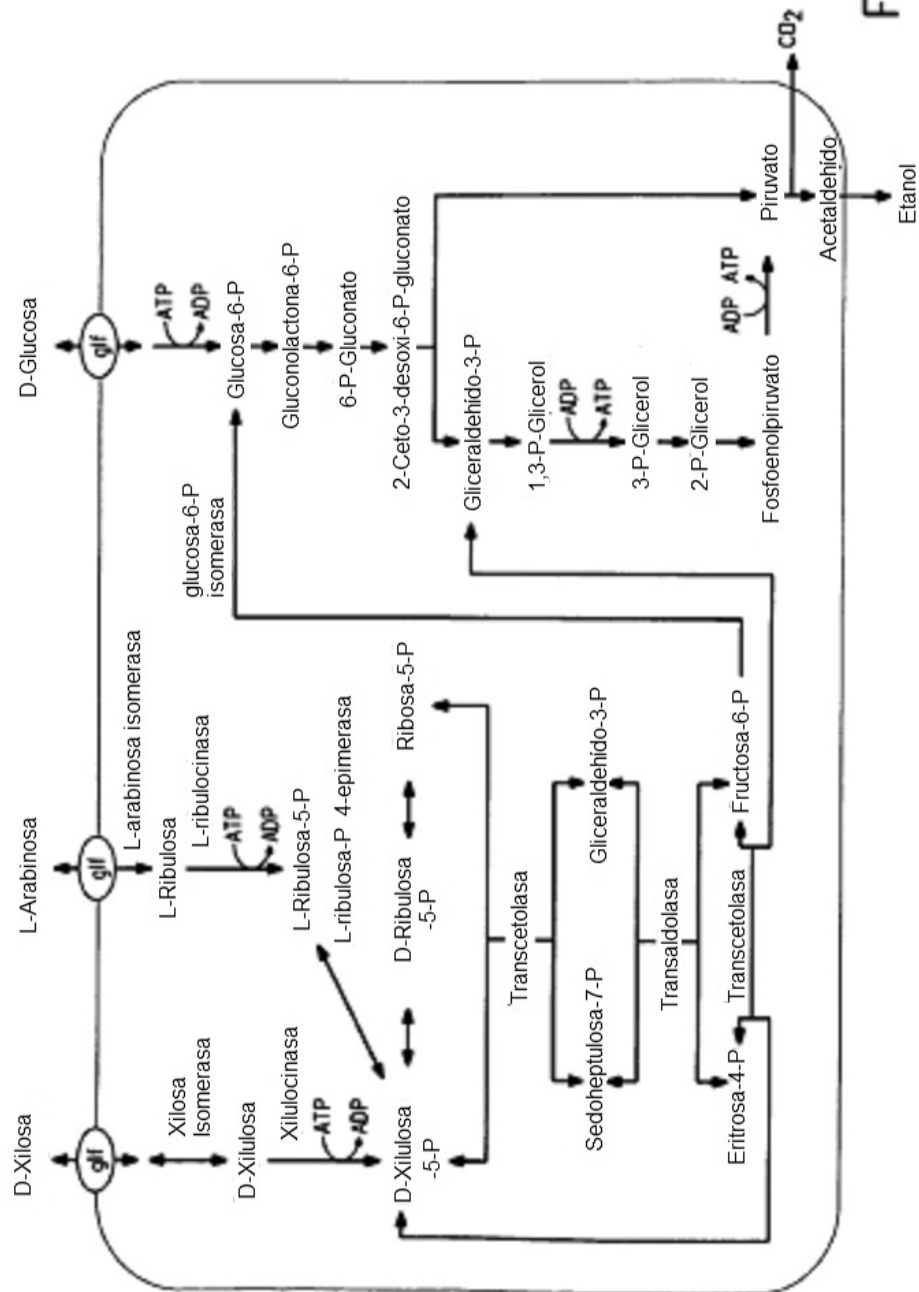


FIG. 1

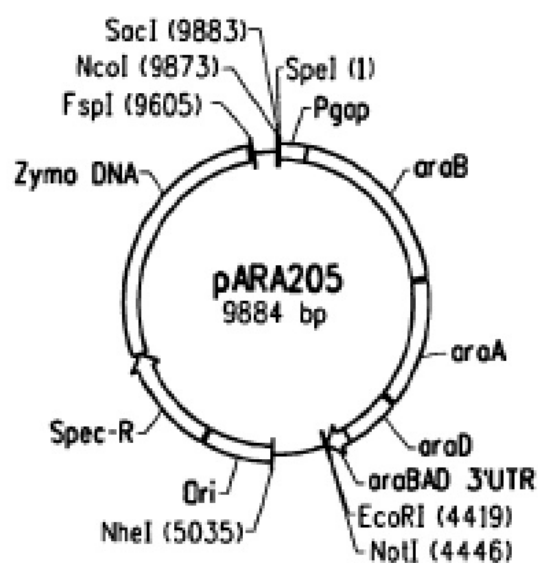


FIG. 2

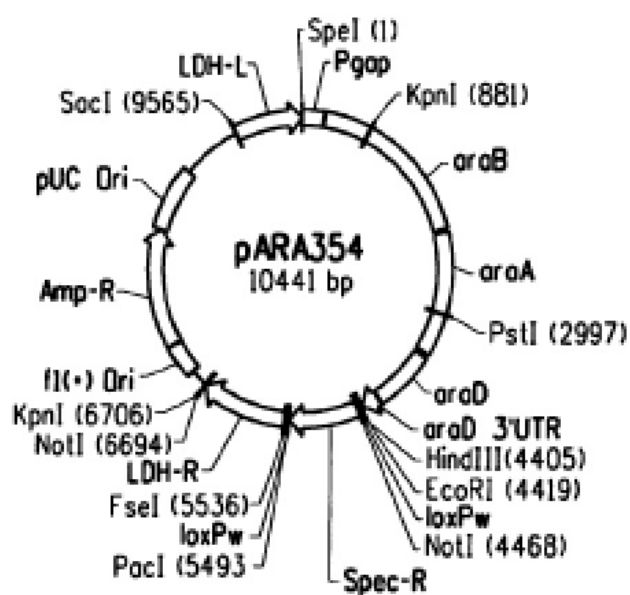


FIG. 3

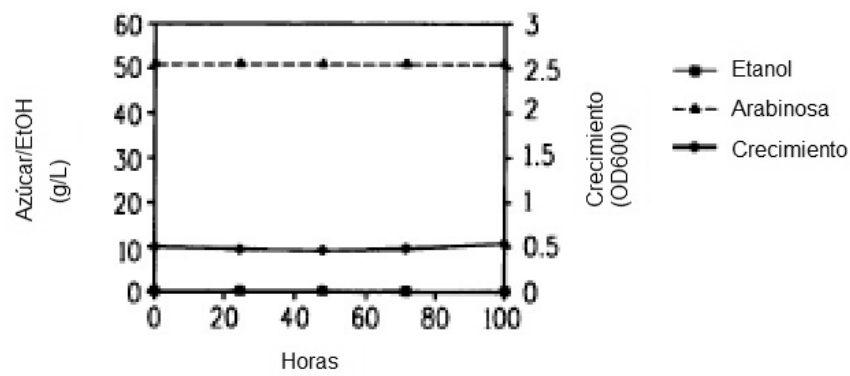


FIG. 4A

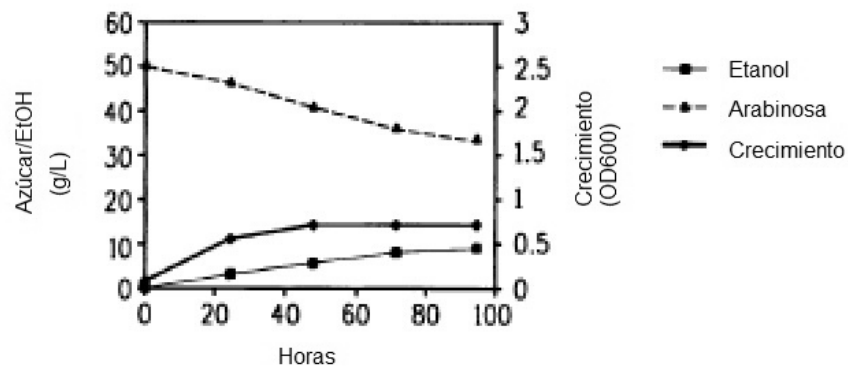


FIG. 4B

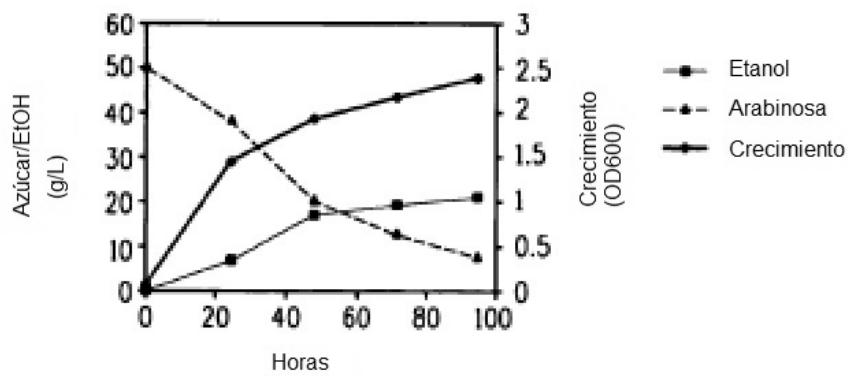
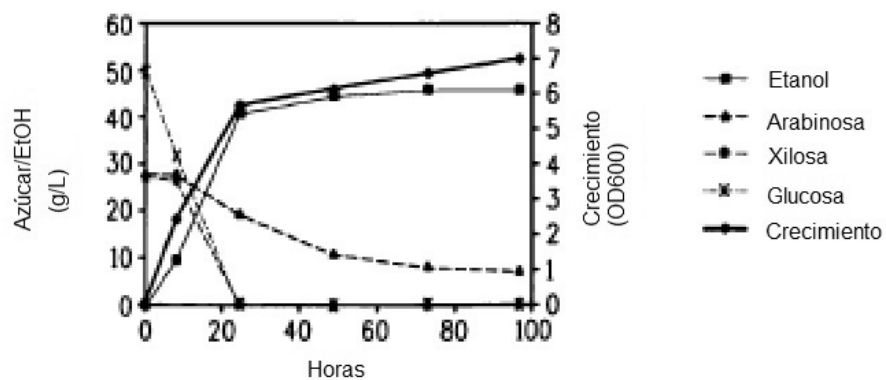
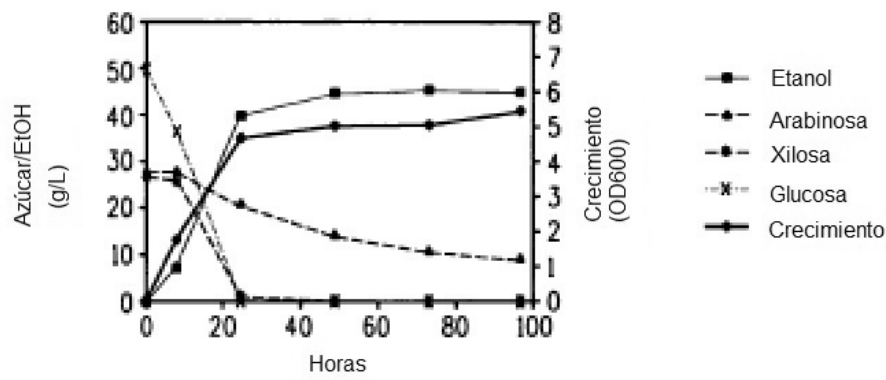
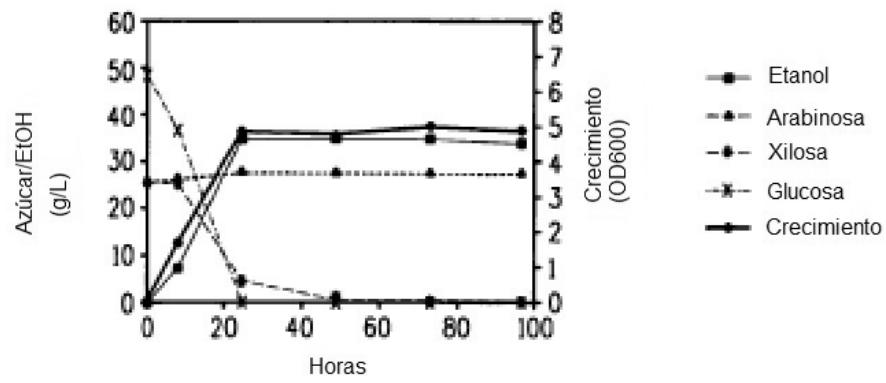


FIG. 4C



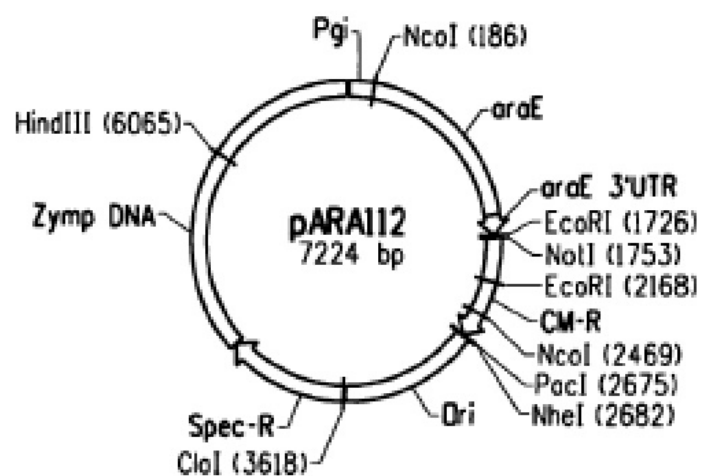


FIG. 6

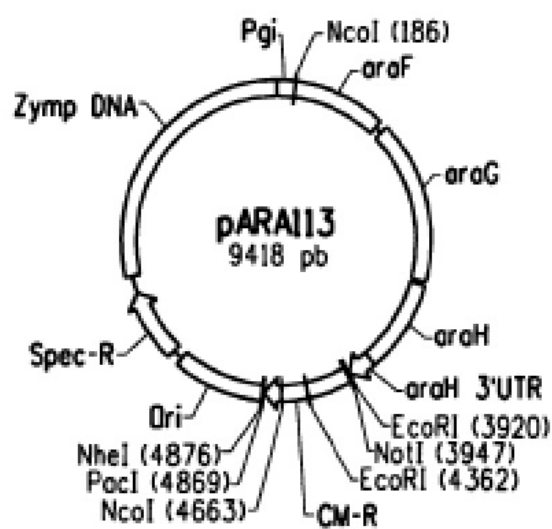


FIG. 7

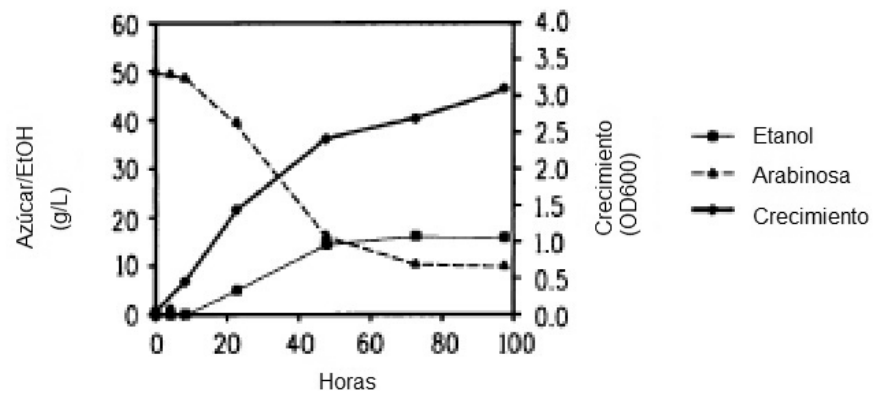


FIG. 8A

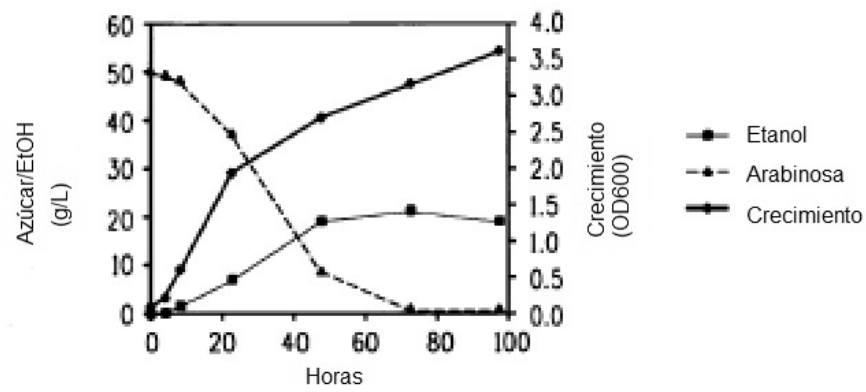


FIG. 8B

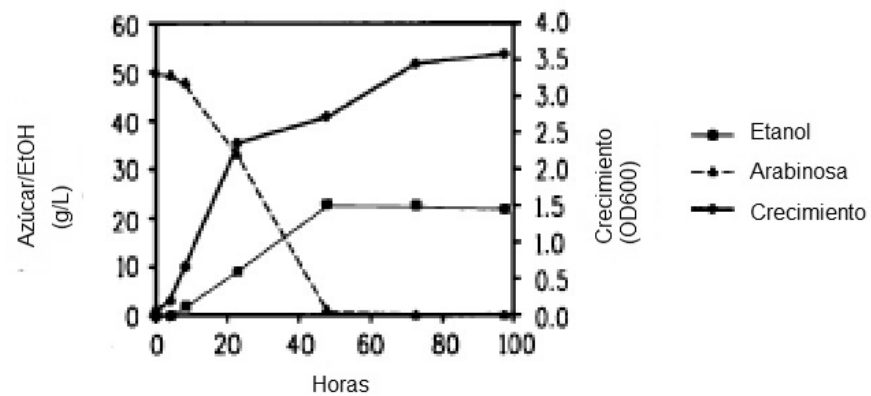


FIG. 8C

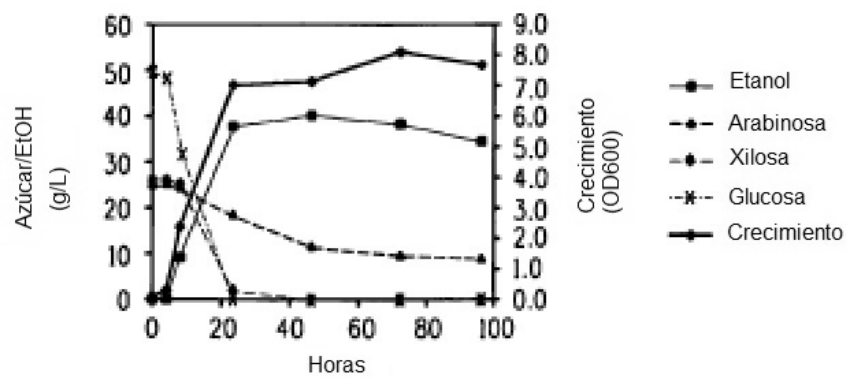


FIG. 9A

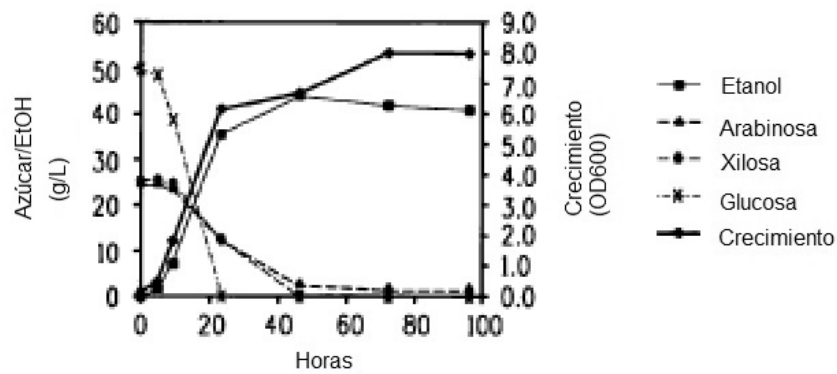


FIG. 9B

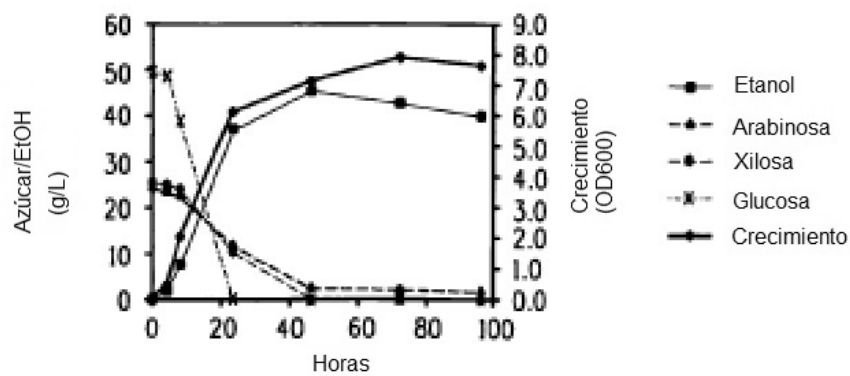


FIG. 9C

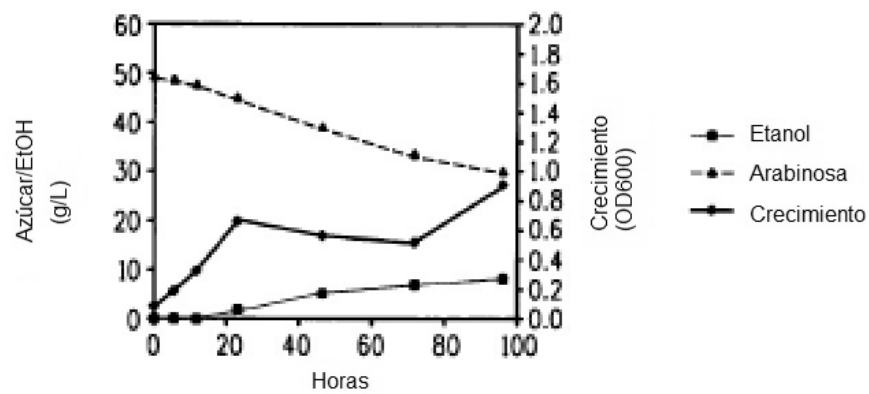


FIG. 10A

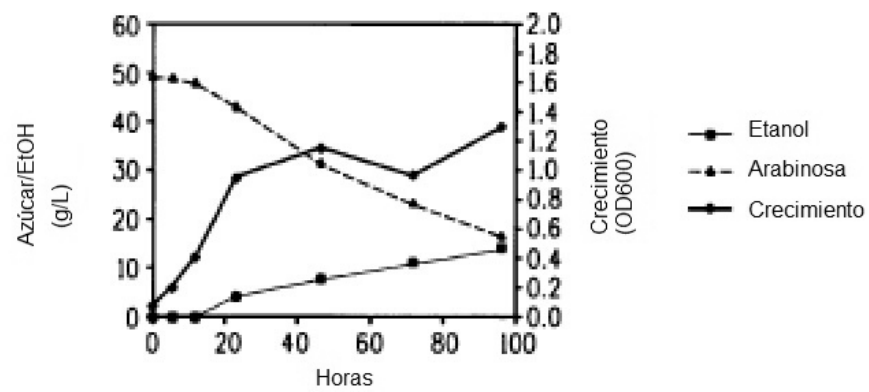


FIG. 10B

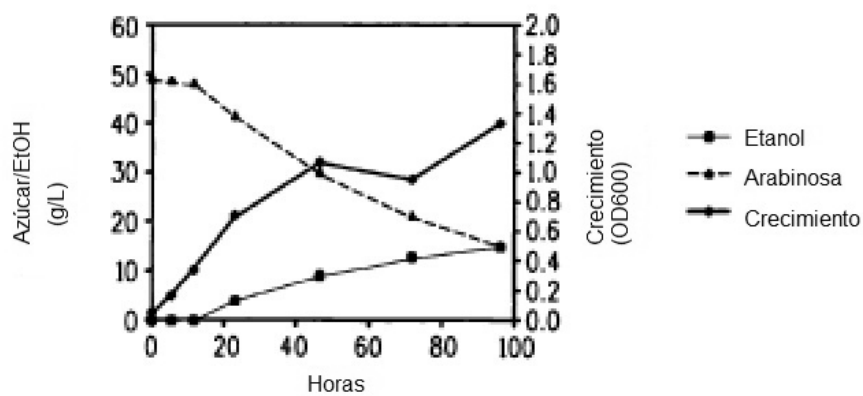


FIG. 10C

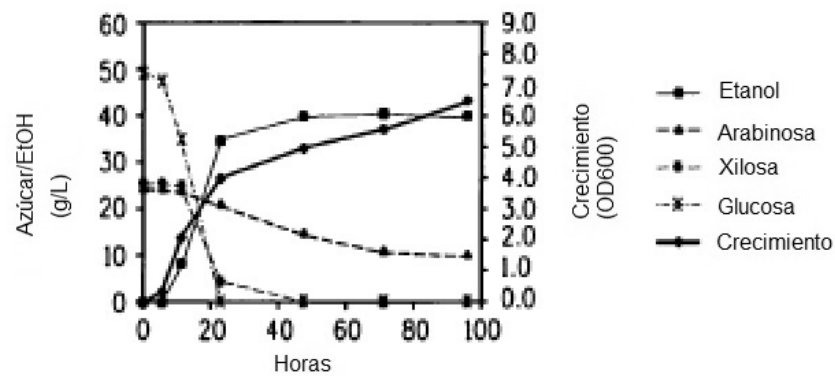


FIG. 11A

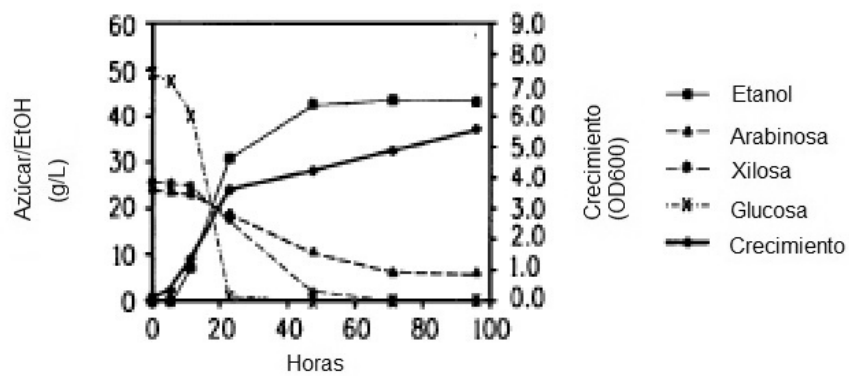


FIG. 11B

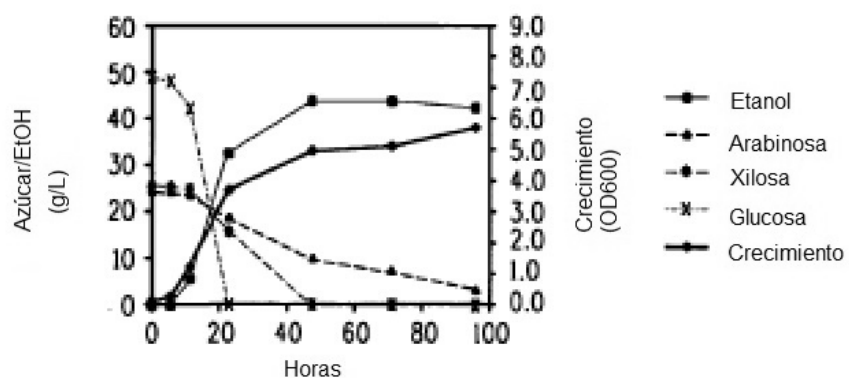


FIG. 11C

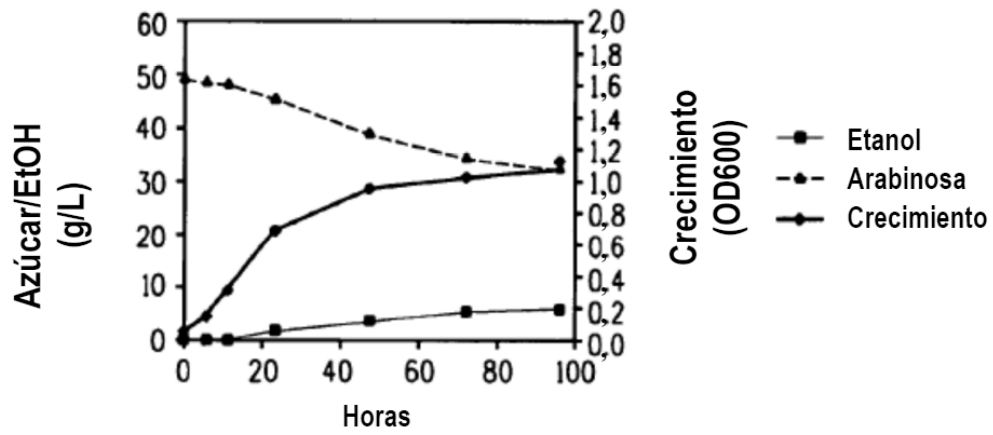


FIG. 12A

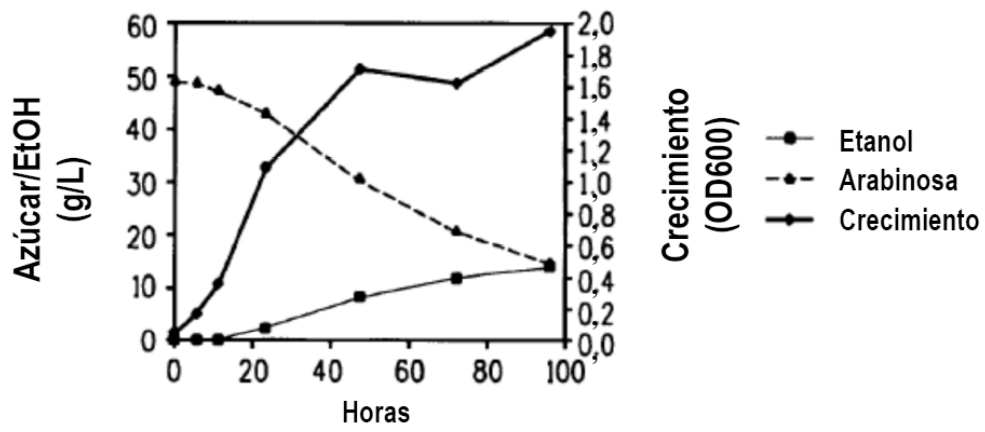


FIG. 12B

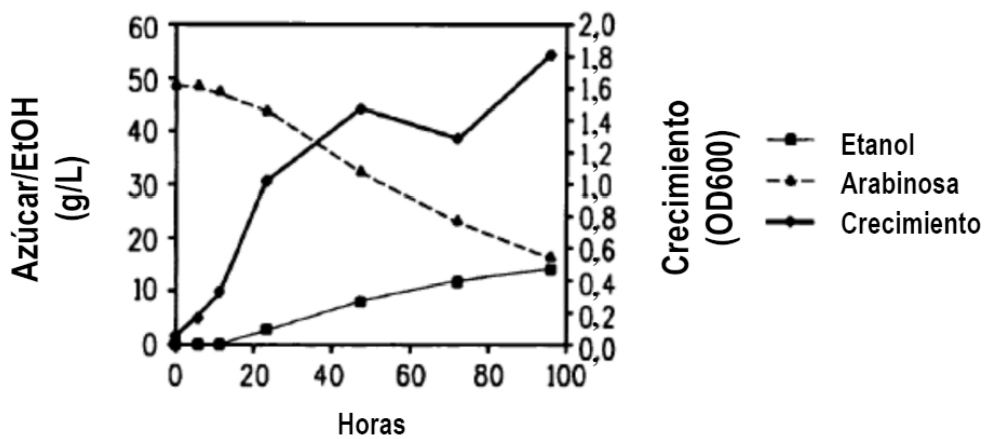


FIG. 12C

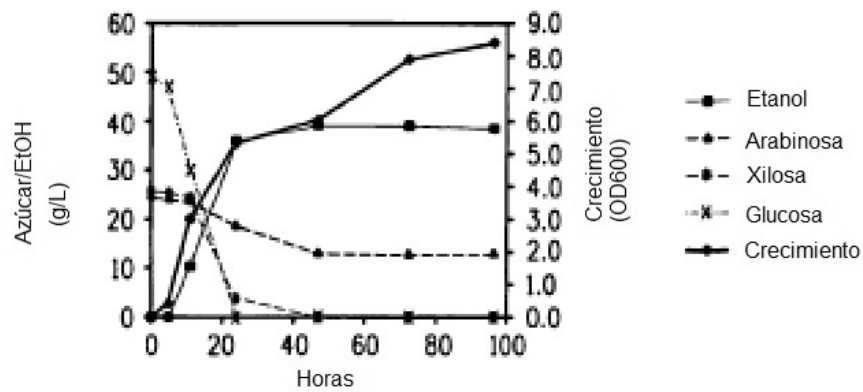


FIG. 13A

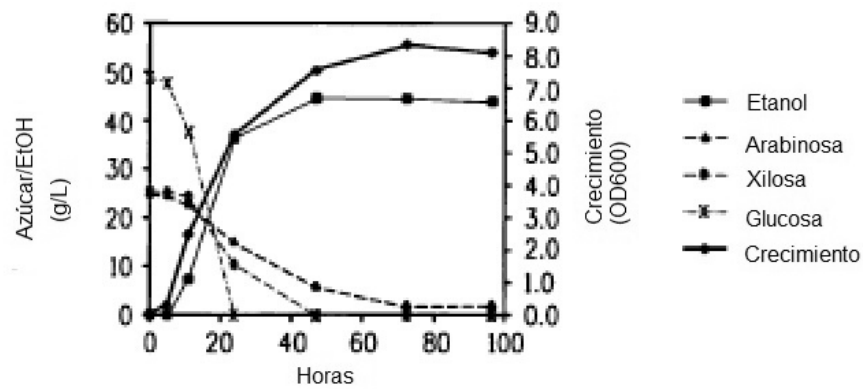


FIG. 13B

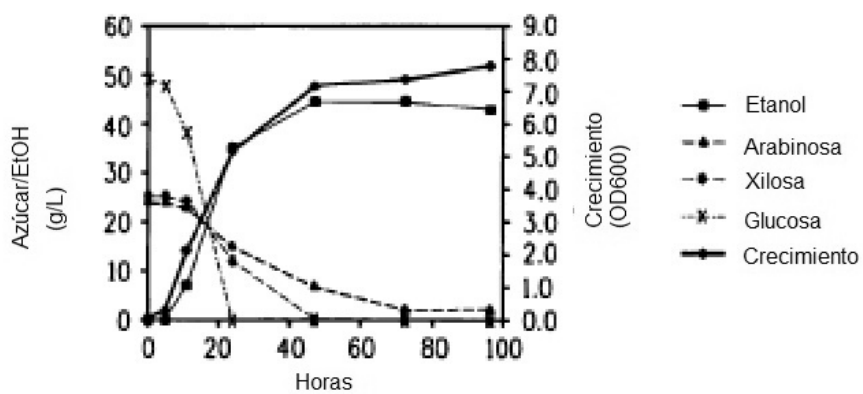


FIG. 13C