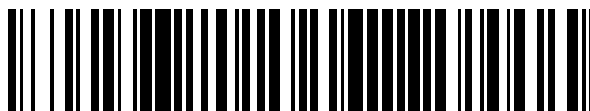


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 552 317**

51 Int. Cl.:

C12N 9/12 (2006.01)
C12N 9/24 (2006.01)
C12N 15/52 (2006.01)
C12N 1/21 (2006.01)
C12P 7/18 (2006.01)
C12P 7/20 (2006.01)
C12P 7/42 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.11.2011 E 11785556 (9)**

97 Fecha y número de publicación de la concesión europea: **19.08.2015 EP 2649179**

54 Título: **Bacterias recombinantes que poseen la capacidad de metabolizar sacarosa**

30 Prioridad:

06.12.2010 US 960646

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

27.11.2015

73 Titular/es:

**E.I. DU PONT DE NEMOURS AND COMPANY
(100.0%)
1007 Market Street
Wilmington, Delaware 19898, US**

72 Inventor/es:

VAN DYK, TINA, K.

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

Observaciones :

Véase nota informativa (Remarks) en el folleto original publicado por la Oficina Europea de Patentes

ES 2 552 317 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

Bacterias recombinantes que poseen la capacidad de metabolizar sacarosa

Campo de la invención

- 5 La invención se refiere a los campos de la microbiología y la biología molecular. Más específicamente, bacterias recombinantes que poseen la capacidad de usar sacarosa como una fuente de carbono y métodos de utilización de dichas bacterias recombinantes para producir productos como glicerol y productos derivados del glicerol.

Antecedentes de la invención

- 10 Muchos microorganismos comerciales útiles usan glucosa como su fuente principal de carbohidratos. Sin embargo, una desventaja del uso de glucosa por microorganismos desarrollados para la producción de productos comercialmente deseables es el alto coste de glucosa. El uso de sacarosa y materias primas mixtas que contienen sacarosa y otros azúcares como fuente de carbohidratos para sistemas microbianos de producción sería más deseable comercialmente porque estos materiales normalmente están inmediatamente disponibles a un bajo coste.

- 15 Un microorganismo de producción puede funcionar más eficientemente cuando puede usar cualquier sacarosa presente en una materia prima mixta. Por lo tanto, cuando un microorganismo de producción no tiene la capacidad de usar eficientemente sacarosa como una fuente principal de carbono no puede operar tan eficientemente. Por ejemplo, las células bacterianas típicamente presentan un uso preferencial de azúcar, siendo la glucosa la más preferida. En medios artificiales que contienen mezclas de azúcares, la glucosa típicamente se metaboliza en su totalidad por delante de otros azúcares. Además, muchas bacterias carecen de la capacidad de usar sacarosa. Por ejemplo, menos de un 50% de las cepas de *Escherichia coli* (*E. Coli*) poseen la capacidad de usar sacarosa. Por
20 consiguiente, cuando un microorganismo de producción no puede usar sacarosa como una fuente de carbohidratos, es deseable modificar genéticamente el microorganismo de modo que pueda usar sacarosa.

- 25 Se han descrito bacterias recombinantes que han sido modificadas genéticamente para usar sacarosa mediante incorporación de genes que utilizan sacarosa. Por ejemplo, Livshits et al. (Patente de EE.UU. Nº 6.960.455) describe la producción de aminoácidos usando cepas de *Escherichia coli* que contienen genes que codifican una vía metabólica para el uso de glucosa. Adicionalmente, Olson et al. (Appl. Microbiol. Biotechnol. 74:1031-1040, 2007) describe cepas de *Escherichia coli* que llevan genes responsables de la degradación de la sacarosa, lo que produce L-tirosina o L-fenilalanina usando sacarosa como fuente de carbono.

- Además, WO 2010/051849 describe la producción de 1,2-propanodiol usando cepas de *Escherichia coli* que contienen genes para el uso de sacarosa.

- 30 Sin embargo, existe una necesidad de cepas bacterianas que sean modificadas genéticamente para usar sacarosa usando nuevos genes de uso de sacarosa y que tengan una mejorada capacidad de usar sacarosa. Adicionalmente, existe una necesidad de cepas bacterianas que sean capaces de producir glicerol y productos derivados del glicerol usando sacarosa como fuente de carbono.

Compendio de la invención

- 35 En una realización, la invención proporciona una bacteria recombinante que comprende en su genoma o en al menos una construcción recombinante:

- (a) una secuencia de nucleótidos heteróloga que codifica un polipéptido que tiene actividad transportadora de sacarosa, teniendo el polipéptido al menos un 95% de identidad de secuencia, basado en el método de alineamiento Clustal V, cuando se compara con una secuencia de aminoácidos según se establece en SEQ ID NO: 24; y
40 (b) una secuencia de nucleótidos heteróloga que codifica un polipéptido que tiene actividad hidrolasa de sacarosa, teniendo el polipéptido al menos un 95% de identidad de secuencia, basado en el método de alineamiento Clustal V, cuando se compara con una secuencia de aminoácidos según se establece en SEQ ID NO: 26;

en donde (a) y (b) están cada una unida operativamente al mismo o diferente promotor en el que además dicha bacteria recombinante es capaz de metabolizar sacarosa.

- 45 En una realización, la bacteria recombinante produce 1,3-propanodiol, glicerol, y/o ácido 3-hidroxipropiónico.

En otra realización, la invención proporciona un proceso de producción de glicerol, 1,3-propanodiol y/o ácido 3-hidroxipropiónico a partir de sacarosa que comprende:

- a) cultivar la bacteria recombinante que produce 1,3-propanodiol, glicerol, y/o ácido 3-hidroxipropiónico, descrita en esta memoria, en presencia de sacarosa; y
50 b) opcionalmente, recuperar el glicerol, 1,3-propanodiol, y/o ácido 3-hidroxipropiónico producidos.

Breve descripción de secuencias

5 Las siguientes secuencias se ajustan a 37 C.F.R. 1.821 1.825 ("Requisitos para Solicitudes de Patentes que Contienen Descripción de Secuencias de Nucleótidos y/o Aminoácidos- las Reglas de Secuencia") y de acuerdo con la Organización Mundial de la Propiedad Intelectual (WIPO) Norma ST.25 (2009) y los requerimientos de listados de secuencias de la EPO y PTC (Reglas 5.2 y 49.5(a bis), y la Sección 208 y el Anexo C de las Instrucciones Administrativas). Los símbolos y formato usados para los datos de las secuencias de nucleótidos y aminoácidos cumplen las reglas representadas en 37 C.F.R. § 1.822.

Tabla A

Compendio de números SEQ ID de genes y proteínas		
Gen	Secuencia codificadora SEQ ID NO:	Proteína codificada SEQ ID NO:
<i>GPD1 de Saccharomyces cerevisiae</i>	1	2
<i>GPD2 de Saccharomyces cerevisiae</i>	3	4
<i>GPP1 de accharomyces cerevisiae</i>	5	6
<i>GPP2 de Saccharomyces cerevisiae</i>	7	8
<i>dhaB1 de Klebsiella pneumoniae</i>	9	10
<i>dhaB2 de Klebsiella pneumoniae</i>	11	12
<i>dhaB3 de Klebsiella pneumoniae</i>	13	14
<i>aldB de Escherichia coli</i>	15	16
<i>aldA de Escherichia coli</i>	17	18
<i>aldH de Escherichia coli</i>	19	20
<i>galP de Escherichia coli</i>	21	22
<i>lac Y de Pseudomonas fluorescens Pf5</i>	23	24
<i>scrB de Pseudomonas fluorescens Pf5</i>	25	26
<i>cscB de Escherichia coli EC3132</i>	27	28
<i>cscA de Escherichia coli EC3132</i>	29	30
<i>scrK de Agrobacterium tumefaciens</i>	31	32
<i>scrK de Streptococcus mutans</i>	33	34
<i>scrK de Escherichia coli</i>	35	36
<i>scrK de Klebsiella pneumoniae</i>	37	38
<i>cscK de Escherichia coli</i>	39	40
<i>cscK de Enterococcus faecalis</i>	41	42
<i>HXK1 de Saccharomyces cerevisiae</i>	43	44
<i>HXK2 de Saccharomyces cerevisiae</i>	45	46
<i>dhaT de Klebsiella pneumoniae</i>	47	48
<i>dhaX de Klebsiella pneumoniae</i>	49	50

SEQ ID NO: 51 es la secuencia de nucleótidos del plásmido pSYCO101.

10 SEQ ID NO: 52 es la secuencia de nucleótidos del plásmido pSYCO103.

SEQ ID NO: 53 es la secuencia de nucleótidos del plásmido pSYCO106.

SEQ ID NO: 54 es la secuencia de nucleótidos del plásmido pSYCO109.

SEQ ID NO: 55 es la secuencia de nucleótidos del plásmido pSYCO400/AGRO.

15 SEQ ID NOs: 56-61 son las secuencias de nucleótidos de los iniciadores usados en los Ejemplos de la presente memoria.

Descripción detallada

La descripción de cada referencia representada en esta memoria se incorpora aquí por referencia en su totalidad.

Como se usa aquí y en las reivindicaciones anexas, las formas singulares “uno/a” y “el/la” incluyen referencia al plural a menos que el contexto indique claramente lo contrario. Así, por ejemplo, la referencia a “una célula” incluye una o más células y equivalentes de la misma conocidas por los expertos en la materia, y así sucesivamente.

En el contexto de esta descripción, se usan un número de términos y abreviaturas. Se proporcionan las siguientes definiciones.

“Marco abierto de lectura” se abrevia como “ORF”

“Reacción en cadena de la polimerasa” se abrevia como “PCR”.

“Colección de cultivos tipo de EE.UU.” se abrevia como “ATCC”.

El término “bacteria recombinante productora de glicerol” se refiere a una bacteria que ha sido modificada genéticamente para ser capaz de producir glicerol y/o productos derivados del glicerol.

El término “polipéptido que tiene actividad transportadora de sacarosa” se refiere a un polipéptido que es capaz de mediar en el transporte de sacarosa dentro de las células microbianas.

El término “polipéptido que tiene actividad fructoquinasa” se refiere a un polipéptido que tiene la capacidad de catalizar la conversión de D-fructosa + ATP en fructosa-fosfato + ADP. Distintivo de la fructoquinasa es EC 2.7.1.4. Las enzimas que tienen alguna capacidad de fosforilar la fructosa, sin importar si esta actividad es su actividad predominante, pueden designarse como fructoquinasas. Abreviaturas usadas para los genes que codifican fructoquinasas y proteínas que tienen actividad fructoquinasa incluyen, por ejemplo, “*Frk*”, “*scrK*”, “*cscK*” “*Fk*”, y “*KHK*”. La fructoquinasa está codificada por el gen *scrK* en *Agrobacterium tumefaciens* y *Streptococcus mutans*; y por el gen *cscK* en ciertas cepas de *Escherichia coli*.

El término “polipéptido que tiene actividad sacarosa hidrolasa” se refiere a un polipéptido que tiene la capacidad de catalizar la hidrólisis de sacarosa para producir glucosa y fructosa. Tales polipéptidos se designan a menudo como “invertasas” o “ β -fructofuranosidasas”.

Los términos “derivados de glicerol” y “productos derivados del glicerol” se usan de manera intercambiable en esta memoria y se refieren a un compuesto que es sintetizado a partir del glicerol o en una vía que incluye glicerol. Ejemplos de tales productos incluyen ácido 3-hidroxipropiónico, metilglioxal, 1,2-propanodiol, y 1,3-propanodiol.

El término “producto microbiano” se refiere a un producto que es producido por microbios, esto es, el resultado de un microorganismo que metaboliza una sustancia. El producto puede ser producido de forma natural por el microorganismo, o el microorganismo puede ser modificado genéticamente para producir el producto.

El término “sistema fosfoenolpiruvato-azúcar fosfotransferasa”, sistema “PTS”, y “PTS” se usan de manera intercambiable en esta memoria y se refieren al sistema de captura de azúcar dependiente de fosfoenolpiruvato.

Los términos “proteína fosfotransportadora HPr” y “PtsH” se refieren a la proteína fosfotransportadora codificada por *ptsH* en *E. Coli*. Los términos “fosfoenolpiruvato-proteína fosfotransferasa” y “PtsI” se refieren a la fosfotransferasa, EC 2.7.3.9, codificada por *ptsI* en *E. Coli*. Los términos “componente glucosa-específico IIA”, y “Crr” se refieren a enzimas designadas como EC 2.7.1.69, codificadas por *crr* en *E.coli*. PtsH, PtsI, y Crr comprenden el sistema PTS.

El término “PTS menos” se refiere a microorganismos que no contienen un sistema PTS en su estado nativo o un microorganismo en el cual el sistema PTS ha sido inactivado a través de la inactivación de un gen PTS.

Los términos “glicerol-3-fosfato deshidrogenasa” y “G3PDH” se refieren a un polipéptido responsable de una actividad enzimática que cataliza la conversión de dihidroxiacetona fosfato (DHAP) en glicerol 3-fosfato (G3P). *In vivo* el G3PDH puede ser dependiente de NAD o NADP. Cuando se hace referencia específicamente a un cofactor específico glicerol-3-fosfato deshidrogenasa, se usarán los términos “glicerol-3-fosfato deshidrogenasa dependiente de NAD” y “glicerol-3-fosfato deshidrogenasa dependiente de NADP”. Como es generalmente el caso, las glicerol-3-fosfato deshidrogenasas dependientes de NAD y dependientes de NADP son capaces de utilizar de manera intercambiable NAD y NADP (por ejemplo por la enzima codificada por *gpsA*), los términos glicerol-3-fosfato deshidrogenasa dependiente de NAD y dependiente de NADP se utilizaran de manera intercambiable. La enzima dependiente de NAD (EC 1.1.1.8) es codificada, por ejemplo, por varios genes incluyendo *GPD1*, también denominado en esta memoria como *DAR1* (secuencia codificadora representada en SEQ ID NO:1; secuencia de proteína codificada representada en SEQ ID NO:2), o *GPD2* (secuencia codificadora representada en SEQ ID NO:3; secuencia de proteína codificada representada en SEQ ID NO:4), o *GPD3*. La enzima dependiente de NADP (EC 1.1.1.94) está codificada, por ejemplo, por *gpsA*.

Los términos “glicerol 3-fosfatasa”, “sn-glicerol 3-fosfatasa”, “D,L-glicerol fosfatasa”, y “G3P fosfatasa” se refieren a un polipéptido que tiene una actividad enzimática que es capaz de catalizar la conversión de glicerol 3-fosfato y agua en glicerol y fosfato inorgánico. La G3P fosfatasa está codificada, por ejemplo, por *GPP1* (secuencia codificadora representada en SEQ ID NO:5; secuencia de proteína codificada representada en SEQ ID NO:6), o *GPP2*

(secuencia codificadora representada en SEQ ID NO:7; secuencia de proteína codificada representada en SEQ ID NO:8).

El término “glicerol deshidratasa” o “enzima deshidratasa” se refiere a un polipéptido que tiene actividad enzimática que es capaz de catalizar la conversión de una molécula de glicerol al producto, 3-hidroxipropionaldehído (3-HPA).

5 Para los fines de la presente invención las enzimas deshidratasas incluyen una glicerol deshidratasa (EC 4.2.1.30) y una diol deshidratasa (EC 4.2.1.28) que tienen sustratos preferidos de glicerol y 1,2-propanodiol, respectivamente. Se han identificado los genes para las enzimas deshidratasas en *Klebsiella pneumoniae*, *Citrobacter freundii*, *Clostridium pasteurianum*, *Salmonella typhimurium*, *Klebsiella oxytoca*, y *Lactobacillus reuteri*, entre otros. En cada caso, la deshidratasa está compuesta por tres subunidades: la grande o subunidad “ α ”, la mediana o subunidad “ β ” y la pequeña o subunidad “ γ ”. Los genes están también descritos en, por ejemplo, Daniel et al. (FEMS Microbiol Rev. 22, 553 (1999)) y Toraya y Mori (J. Biol. Chem. 274, 3372 (1999)). Los genes que codifican la subunidad grande o “ α ” (alfa) de la glicerol deshidratasa incluyen *dhaB1*, (secuencia codificadora representada en SEQ ID NO:9; secuencia de proteína codificada representada en SEQ ID NO:10), *gldA* y *dhaB*; genes que codifican la subunidad mediana o “ β ” (beta) incluyen *dhaB2* (secuencia codificadora representada en SEQ ID NO:11; secuencia de proteína codificada representada en SEQ ID NO:12), *gldB* y *dhaC*; genes que codifican la subunidad “ γ ” (gamma) incluyen *dhaB3* (secuencia codificadora representada en SEQ ID NO:13; secuencia de proteína codificada representada en SEQ ID NO:14), *gldC* y *dhaE*. Otros genes que codifican la subunidad grande o “ α ” de la diol deshidratasa incluyen *pduC* y *pddA*; otros genes que codifican la subunidad mediana o “ β ” incluyen *pduD* y *pddB*; y otros genes que codifican la subunidad pequeña o “ γ ” incluyen *pduE* y *pddC*.

20 Las glicerol y diol deshidratasas están sujetas a un mecanismo basado en la inactivación suicida por glicerol y algunos otros sustratos (Daniel et al. FEMS Microbiol Rev. 22, 553 (1999)). El término “factor de reactivación de deshidratasa” se refiere a aquellas proteínas responsables de reactivar la actividad deshidratasa. Los términos “actividad de reactivación de deshidratasa”, “reactivación de la actividad deshidratasa” y “degeneración de la actividad deshidratasa” se usan de manera intercambiable y se refieren al fenómeno de convertir una deshidratasa incapaz de catálisis de una reacción a una capaz de catálisis de una reacción o al fenómeno de inhibición de la inactivación de una deshidratasa o al fenómeno de extensión de la vida media útil de la enzima deshidratasa *in vivo*. Se han identificado dos proteínas que están implicadas como el factor de reactivación de deshidratasa (mírese, p. ej.; Patente de EE.UU. N° 6.013.494 y referencias en ella; Daniel et al., *supra*; Toraya and Mori, J. Biol. Chem. 274, 3372 (1999); y Tobimatsu et al., J. Bacteriol. 181, 4110 (1999)). Genes que codifican una de las proteínas incluyen, por ejemplo, *orfZ*, *dhaB4*, *gdrA*, *pduG* y *ddrA*. Genes que codifican la segunda de las dos proteínas incluyen, por ejemplo, *orfX*, *orf2b*, *gdrB*, *pduH* y *ddrB*.

Los términos “1,3-propanodiol oxidoreductasa”, “1,3-propanodiol deshidrogenasa” y “DhaT” se usan de manera intercambiable en esta memoria y se refieren a el(los) polipéptido(s) que tienen una actividad enzimática que es capaz de catalizar la interconversión de 3-HPA y 1,3-propanodiol siempre y cuando el(los) gene(s) que codifican tal actividad se encuentra que está físicamente o transcripcionalmente ligado(s) a una enzima deshidratasa en su configuración natural (esto es, tipo silvestre); por ejemplo, el gen se encuentra dentro del regulon *dha* era el caso de *dhaT* de *Klebsiella pneumoniae*. Genes que codifican una 1,3-propanodiol oxidoreductasa incluyen, pero no se limitan a, *dhaT*, de *Klebsiella pneumoniae*, *Citrobacter freundii*, y *Clostridium pasteurianum*. Cada uno de estos genes codifica un polipéptido que pertenece a la familia del tipo III de las alcohol deshidrogenasas, las cuales exhiben un motivo conservado de unión a hierro, y tiene una preferencia por NAD⁺/NADH ligado a la interconversión de 3-HPA y 1,3-propanodiol (Jhonson and Lin, J. Bacteriol. 169, 2050 (1987); Daniel et al., J. Bacteriol. 177, 2151 (1995); y Leurs et al., FEMS Microbiol. Lett. 154, 337 (1997)). Enzimas con propiedades físicas similares se han aislado de *Lactobacillus brevis* y *Lactobacillus buchneri* (Veiga da Dunha y Foster, Appl. Environ. Microbiol. 58, 2005 (1992)).

45 El término “regulon *dha*” se refiere a un conjunto de polinucleótidos asociados o marcos abiertos de lectura que codifican polipéptidos que tienen varias actividades biológicas, incluyendo pero no limitado a una actividad deshidratasa, una actividad de reactivación, y una 1,3-propanodiol oxidoreductasa. Típicamente un regulon *dha* comprende los marcos de lectura *dar*, *orfY*, *dhaT*, *orfX*, *orfW*, *dhaB1*, *dhaB2*, *dhaB3*, y *orfZ* como se describe en la Patente de EE.UU. N° 7.471.558.

50 Los términos “aldehído deshidrogenasa” y “Ald” se refieren a un polipéptido que cataliza la conversión de un aldehído en un ácido carboxílico. Las aldehído deshidrogenasas pueden usar un cofactor redox como NAD, NADP, FAD, o PQQ. Distintivo de aldehído deshidrogenasas es EC 1.2.1.3. (dependiente de NAD); EC 1.2.1.4. (dependiente de NADP); EC 1.2.99.3. (dependiente de PQQ); o EC 1.2.99.7. (dependiente de FAD). Un ejemplo de una aldehído deshidrogenasa dependiente de NADP es AldB (SEQ ID NO:16) codificada por el gen *aldB* de *E. Coli* (secuencia codificadora representada en SEQ ID NO:15). Ejemplos de aldehído deshidrogenasas dependientes de NAD incluyen AldA (SEQ ID NO: 18), codificado por el gen *aldA* de *E. Coli* (secuencia codificadora representada en SEQ ID NO:17); y AldH (SEQ ID NO: 20), codificado por el gen *aldH* de *E. Coli* (secuencia codificadora representada en SEQ ID NO:19).

- Los términos “glucoquinasa” y “Glc” se usan de manera intercambiable en esta memoria y se refieren a una proteína que cataliza la conversión de D-glucosa + ATP en glucosa 6-fosfato + ADP. Distintivo de glucoquinasa es EC 2.7.1.2. Glucoquinasa está codificada por *glk* en *E. Coli*.
- 5 Los términos “fosfoenolpiruvato carboxilasa” y “Ppc” se usan de manera intercambiable en esta memoria y se refieren a una proteína que cataliza la conversión de fosfoenolpiruvato + H₂O + CO₂ en fosfato + ácido oxalacético. Distintivo de fosfoenolpiruvato carboxilasa es EC 4.1.1.31. La fosfoenolpiruvato carboxilasa está codificada por *ppc* en *E. Coli*.
- 10 Los términos “gliceraldehído-3-fosfato deshidrogenasa” y “GapA” se usan de forma intercambiable en esta memoria y se refieren a una proteína que tiene una actividad enzimática capaz de catalizar la conversión de gliceraldehído-3-fosfato + fosfato + NAD⁺ en 3-fosfo-D-glicerol-fosfato + NADH + H⁺. Distintivo de la gliceraldehído-3-fosfato deshidrogenasa es EC 1.2.1.12. La gliceraldehído-3-fosfato deshidrogenasa está codificada por *gapA* en *E. Coli*.
- 15 Los términos “proteína control de respiración aeróbica” y “ArcA” se usan de modo intercambiable en esta memoria y se refieren a una proteína reguladora global. La proteína control de respiración aeróbica está codificada por *ArcA* en *E. Coli*.
- 20 Los términos “metilglixal sintasa” y “MgsA” se usan de forma intercambiable en esta memoria y se refieren a una proteína que tiene una actividad enzimática capaz de catalizar la conversión de dihidroxiacetona fosfato en metilglixal + fosfato. Distintivo de la metilglixal sintasa es EC 4.2.3.3. La metilglixal sintasa está codificada por *mgsA* en *E. Coli*.
- Los términos “fosfogluconato deshidratasa” y “Edd” se usan de forma intercambiable en esta memoria y se refieren a una proteína que tiene una actividad enzimática capaz de catalizar la conversión de 6-fosfo-gluconato en 2-ceto-3-deoxi-6-fosfo-gluconato + H₂O. Distintivo de la fosfogluconato deshidratasa es EC 4.2.1.12. La fosfogluconato deshidratasa está codificada por *edd* en *E. Coli*.
- El término “YciK” se refiere a una supuesta enzima codificada por *yciK* que está traduccionalmente acoplada a *btuR*, el gen que codifica Cob(I)alamina adenosiltransferasa en *E. Coli*.
- 25 El término “cob(I)alamina adenosiltransferasa” se refiere a una enzima capaz de transferir una fracción deoxiadenosilo a partir del ATP a un corrinóide reducido. Distintivo de la cob(I)alamina adenosiltransferasa es EC 2.5.1.17. La cob(I)alamina adenosiltransferasa está codificada por el gen “*btuR*” en *E. Coli*, “*cobA*” en *Salmonella typhimurium*, y “*cobO*” en *Pseudomonas denitrificans*.
- 30 Los términos de “galactosa-protón sym-transportador” y “GalP” se usan de forma intercambiable en esta memoria y se refieren a una proteína que tiene una actividad enzimática capaz de transportar un azúcar y un protón desde el periplasma hacia el citoplasma. La D-glucosa es el sustrato preferido del GalP. El galactosa-protón sym-transportador está codificado por *galP* en *Escherichia Coli* (secuencia codificadora representada en SEQ ID NO: 21, secuencia de proteína codificada representada en SEQ ID NO: 22).
- 35 El término “actividad catalítica no específica” se refiere al(los) polipéptido(s) que tienen una actividad enzimática capaz de catalizar la interconversión de 3-HPA y 1,3-propanodiol y excluye específicamente a 1,3-propanodiol oxidoreductasa(s). Normalmente estas enzimas son alcohol deshidrogenasas. Tales enzimas pueden utilizar cofactores distintos de NAD⁺/NADH, incluyendo pero no limitados a flavinas como FAD o FMN. Se encontró un gen para una alcohol deshidrogenasa no específica (*yqhD*), por ejemplo, que estaba codificado endógenamente y funcionalmente expresado en cepas de *E. Coli* K-12.
- 40 Los términos “promotor GI largo 1.6”, “promotor GI corto/largo 1.2”, y “promotor GI largo 1.5” se refieren a polinucleótidos o fragmentos que contienen un promotor del gen de la glucosa isomerasa de *Streptomyces lividans* como se describe en la Patente de EE.UU. N° 7.132.527. Estos fragmentos promotores incluyen una mutación que disminuye sus actividades comparado con el promotor de la glucosa isomerasa de *Streptomyces lividans* de tipo silvestre.
- 45 Los términos “función” y “función enzimática” se usan de forma intercambiable en esta memoria y se refieren a una actividad catalítica de una enzima de alterar la velocidad a la que ocurre una reacción química específica sin ser ella misma consumida por la reacción. Se entiende que tal actividad puede aplicarse a una reacción en el equilibrio donde la producción de cualquier producto o sustrato puede llevarse a cabo bajo condiciones adecuadas.
- Los términos de “polipéptido” y “proteína” se usan de manera intercambiable en esta memoria.
- 50 Los términos “sustrato de carbono” y “fuente de carbono” se usan de forma intercambiable en esta memoria y se refieren a una fuente de carbono capaz de ser metabolizada por la bacteria recombinante descrita en esta memoria y, particularmente, fuentes de carbono que comprenden sacarosa. La fuente de carbono puede además comprender otros monosacáridos, disacáridos, oligosacáridos; o polisacáridos.

Los términos “célula huésped” y “bacteria huésped” se usan de forma intercambiable en esta memoria y se refieren a una bacteria capaz de recibir genes exógenos o heterólogos y capaz de expresar estos genes para producir un producto genético activo.

- 5 El término “microorganismo de producción” como se usa en esta memoria se refiere a un microorganismo, que incluye, pero no se limita a, aquellos que son recombinantes, usados para crear un producto específico como 1,3-propanodiol, glicerol, ácido 3-hidroxipropiónico, ácidos grasos poliinsaturados, y similares.

10 Como se usa en esta memoria, “ácido nucleico” significa un polinucleótido e incluye un polímero de bases de desoxiribonucleótidos o ribonucleótidos de cadena sencilla o doble. Los ácidos nucleicos también pueden incluir fragmentos y nucleótidos modificados. Así, los términos “polinucleótido”, “secuencia de ácido nucleico”, “secuencia de nucleótidos” o “fragmento de ácido nucleico” se usan de manera intercambiable en esta memoria y se refieren a polímeros de RNA o DNA que es de cadena sencilla o doble, opcionalmente que contiene bases de nucleótidos no naturales o alteradas. Los nucleótidos (usualmente encontrados en su forma 5'-monofosfato) son denominados mediante su sola designación de la letra de la siguiente manera: “A” para adenilato o deoxiadenilato (para RNA o DNA, respectivamente), “C” para citidilato o deoxicitidilato, “G” para guanilato o deoxiguanilato, “U” para uridilato, “T” para deoxitimidilato, “R” para purinas (A o G), “Y” para pirimidinas (C o T), “K” para G o T, “H” para A o C o T, “I” para inosina, y “N” para cualquier nucleótido.

Un polinucleótido puede ser un polímero de RNA o DNA que es de cadena sencilla o doble, que opcionalmente contiene bases de nucleótidos sintéticas, no naturales o alteradas. Un polinucleótido en forma de polímero de DNA puede estar compuesto de una o más segmentos de cDNA, DNA genómico, DNA sintético o mezclas de los mismos.

20 “Gen” se refiere a un fragmento de ácido nucleico que expresa una proteína específica, y que puede referirse a la región codificadora solo o puede incluir secuencias reguladoras que preceden (secuencias 5' no codificadoras) y a continuación (secuencias 3' no codificadoras) de la secuencia codificadora. “Gen nativo” se refieren a un gen tal como se encuentran en la naturaleza con sus propias secuencias reguladoras. “Gen quimérico” se refiere a cualquier gen que no es un gen nativo, que comprende secuencias reguladoras y codificadoras que no se encuentran juntas en la naturaleza. En consecuencia, un gen quimérico comprende secuencias reguladoras y secuencias codificadoras que se derivan de diferentes fuentes, o secuencias reguladoras y secuencias codificadoras derivadas de la misma fuente, pero organizadas de forma diferente a cómo se encuentra en la naturaleza. “Gen endógeno” se refiere a un gen nativo en su localización natural en el genoma de un organismo. Un gen “exógeno” se refiere a un gen que se introduce dentro del organismo huésped mediante transferencia génica. Los genes exógenos pueden comprender genes insertados dentro de un organismo no-nativo, genes introducidos dentro de una nueva localización dentro del huésped nativo, o genes quiméricos.

El término “secuencia de nucleótidos nativa” se refiere a una secuencia de nucleótidos que se encuentra normalmente en el microorganismo huésped.

35 El término “secuencia de nucleótidos no nativa” se refiere a una secuencia de nucleótidos que no se encuentra normalmente en el microorganismo huésped.

El término “polipéptido nativo” se refiere a un polipéptido que se encuentra normalmente en el microorganismo huésped.

El término “polipéptido no nativo” se refiere a un polipéptido que no se encuentra normalmente en el microorganismo huésped.

40 Los términos “encriptación” y “codificación” se usan de modo intercambiable en esta memoria y se refieren al proceso mediante el cual un gen, a través de mecanismos de transcripción y traducción, genera una secuencia de aminoácidos.

El término “secuencia codificadora” se refiere a una secuencia de nucleótidos que codifica una secuencia de aminoácidos específica.

45 Las “secuencias reguladoras adecuadas” se refieren a secuencias de nucleótidos localizadas corriente arriba (secuencias 5' no-codificadoras), dentro de, o corriente abajo (secuencias 3' no-codificadoras) de una secuencia codificadora, y que influencia la transcripción, procesamiento o estabilidad del RNA, o la traducción de la secuencia codificadora asociada. Las secuencias reguladoras pueden incluir promotores, potenciadores, silenciadores, secuencia líder no traducida 5' (p. ej. entre el sitio de comienzo de la transcripción y el codón de iniciación de la traducción), intrones, secuencias de reconocimiento de poliadenilación, sitios de procesamiento de RNA, sitios de unión efectores y estructuras de horquilla.

55 El término “casete de expresión” se refiere a un fragmento de DNA que comprende la secuencia codificadora de un gen seleccionado y las secuencias reguladoras que preceden (secuencias no-codificadoras 5') y que siguen (secuencias no-codificadoras 3') a la secuencia codificadora que se requieren para la expresión del producto del gen seleccionado. Así, un casete de expresión está compuesto típicamente de: 1) una secuencia promotora; 2) una secuencia codificadora (esto es, ORF) y, 3) una región 3' no traducida (p. ej. un terminador) que, en eucariotas,

normalmente contiene un sitio de poliadenilación. El (los) casete(s) de expresión normalmente está incluido dentro de un vector, para facilitar el clonaje y la transformación. Diferentes organismos, incluyendo bacterias, levaduras, y hongos, se pueden transformar con diferentes casetes de expresión siempre y cuando se usen las secuencias reguladoras correctas para cada huésped.

5 “Transformación” se refiere a la transferencia de una molécula de ácido nucleico en un organismo, dando como resultado una herencia genéticamente estable. La molécula de ácido nucleico puede ser un plásmido que se replica autónomamente, por ejemplo, o que puede integrarse en el genoma del organismo huésped. Los organismos huésped transformados con los fragmentos del ácido nucleico se refieren como organismos “recombinantes” o “transformados” o “transformantes”. “Transformación estable” se refiere a la transferencia de una fragmento de ácido
10 nucleico en el genoma de un organismo huésped, que incluye tanto genomas nucleares como organulares, dando como resultado una herencia genéticamente estable. Por el contrario, “transformación transitoria” se refiere a la transferencia de un fragmento de ácido nucleico dentro del núcleo, u orgánulo que contiene DNA, de un huésped que da como resultado la expresión génica sin integración o herencia estable.

15 “Degeneración de codón” se refiere a la naturaleza en el código genético que permite variación de la secuencia de nucleótido sin afectar a la secuencia de aminoácidos de un polipéptido codificado. El artesano experto es muy consciente del “sesgo de codón” exhibido por una célula huésped específica en el uso de codones de nucleótidos para especificar un aminoácido dado. Por lo tanto, cuando se sintetiza un gen para mejorar la expresión en una célula huésped, es deseable para modificar genéticamente el gen tal que su frecuencia de uso de codón se acerque a la frecuencia del uso de codón preferido de la célula huésped.

20 Los términos “sub fragmento que es funcionalmente equivalente” y “sub fragmento funcionalmente equivalente” se usan de manera intercambiable en esta memoria. Estos términos se refieren a una porción o sub secuencia de un fragmento de ácido nucleico aislado en el que se retiene la capacidad de alterar la expresión génica o producir un cierto tipo de fenotipo sin importar si el fragmento o sub fragmento codifica una enzima activa. Los genes quiméricos se pueden designar para su uso en la supresión mediante la ligación de un fragmento de ácido nucleico o sub
25 fragmento del mismo, sin importar si codifica una enzima activa, en la orientación en el sentido o en el antisentido en relación a la secuencia promotora.

El término “dominio conservado” o “motivo” quiere decir un conjunto de aminoácidos conservados en posiciones específicas a lo largo de la secuencia alineada de proteínas evolutivamente relacionadas. Mientras los aminoácidos pueden variar en otras posiciones entre proteínas homólogas, los aminoácidos que están altamente conservados en
30 posiciones específicas indican aminoácidos que son esenciales en la estructura, la estabilidad, o la actividad de una proteína.

El término “substancialmente similar” y “corresponde substancialmente” se usan de manera intercambiable en esta memoria. Se refieren a fragmentos de ácido nucleico donde cambios en uno o más bases de nucleótidos no afectan a la capacidad del fragmento de ácido nucleico para mediar en la expresión génica o de producir un cierto fenotipo.
35 Estos términos se refieren a modificaciones de los fragmentos de ácidos nucleicos de la presente invención como supresión o inserción de uno o más nucleótidos que no alteran substancialmente las propiedades funcionales del fragmento de ácido nucleico resultante en relación al fragmento no modificado inicial. Por lo tanto, se entiende, como los expertos en la técnica apreciarán, que la invención engloba más que las secuencias específicas ilustrativas. Además, el artesano experto identifica que secuencias de ácidos nucleicos substancialmente similares se definen
40 también mediante su capacidad de hibridar (bajo condiciones moderadamente estrictas, p. ej. SSC (citrato sódico estándar) 0,5X, SDS (dodecilsulfato sódico) 0,1%, 60° C) con las secuencias ilustradas en esta memoria, o con cualquier porción de la secuencia de nucleótidos descrita en esta memoria y que son funcionalmente equivalentes a cualquiera de las secuencias de ácidos nucleicos descritas en esta memoria. Las condiciones de astringencia se pueden ajustar a la pantalla para fragmentos moderadamente similares, como secuencias homólogas de organismos alejados, con fragmentos altamente similares, como genes que duplican enzimas funcionales a partir de organismos estrechamente relacionados. Los lavados después de la hibridación determinan las condiciones de astringencia.
45

El término “hibrida selectivamente” incluye la referencia a la hibridación, bajo condiciones de hibridación astringentes, de una secuencia de ácido nucleico con una secuencia diana de ácido nucleico específica a un mayor grado detectable (p. ej. al menos dos veces más de fondo) que su hibridación a secuencias de ácidos nucleico no-diana y a la exclusión substancial de ácidos nucleicos no-diana. Secuencias que hibridan selectivamente son
50 secuencias de dos nucleótidos en donde el complemento de una de las secuencias de nucleótidos tiene típicamente alrededor de 80% de identidad de secuencia, o 90% de identidad de secuencia, hasta e incluyendo 100% de identidad de secuencia (esto es, totalmente complementaria) a la otra secuencia de nucleótidos.

El término “condiciones astringentes” o “condiciones de hibridación astringentes” incluye la referencia a condiciones
55 bajo las que una sonda hibridará selectivamente con su secuencia diana. Las sondas son secuencias de ácido nucleico de cadena sencilla que son complementarias a las secuencias de ácido nucleico a detectar. Las sondas son “hibridables” con la secuencia de ácido nucleico a detectar. Generalmente, una sonda tiene menos de aproximadamente 1.000 nucleótidos de longitud, opcionalmente menor de 500 nucleótidos de longitud.

Están bien definidos los métodos de hibridación. Típicamente, la sonda y la muestra se mezclan bajo condiciones que permitirán la hibridación del ácido nucleico. Esto implica poner en contacto la sonda y la muestra en la presencia de una sal inorgánica u orgánica bajo la concentración y condiciones de temperatura apropiadas. Se puede añadir opcionalmente un agente caotrópico. La hibridación de ácidos nucleicos se adapta a una variedad de formatos de ensayo. Uno de los mas apropiados es el formato de ensayo sándwich. Un componente primario de un ensayo tipo sándwich es un soporte sólido. El soporte sólido se ha adsorbido a él o se ha acoplado covalentemente a él una sonda de ácido nucleico inmovilizada que está sin marcar y complementaria a una porción de la secuencia.

Las condiciones astringentes son dependientes de la secuencia y serán diferentes en diferentes circunstancias. Mediante el control de la astringencia de la hibridación y/o condiciones de lavado, se pueden identificar secuencias diana que son 100% complementarias a la sonda (sondeo homólogo). Alternativamente, las condiciones de astringencia se pueden ajustar para permitir alguna discordancia en secuencias de manera que se detectan menores grados de similitud (sondeo heterólogo).

Típicamente, las condiciones astringentes serán aquellas en las que la concentración de sal es menor que aproximadamente ión Na 1,5 M, típicamente alrededor de 0,01 hasta 1,0 M de concentración de ión Na (u otras sales) a un de pH 7,0 hasta 8,3 y la temperatura es al menos alrededor de 30° C para sondas cortas (p. ej., de 10 a 50 nucleótidos) y al menos 60° C para sondas largas (p. ej., mayor de 50 nucleótidos). Se pueden alcanzar condiciones astringentes con la adición de agentes desestabilizantes como formamida. Condiciones ilustrativas de baja astringencia incluyen hibridación con una solución tampón formamida desde el 30 hasta 35%, NaCl 1 M, SDS 1% (dodecilsulfato sódico) a 37° C, y un lavado con SSC desde 1X hasta 2X (SSC 20X = NaCl 3,0 M/citrato trisódico 0,3 M) de 50 hasta 55° C. Condiciones ilustrativas de moderada astringencia incluyen hibridación en formamida desde el 40% hasta el 45%, NaCl 1 M, SDS 1% a 37° C, y un lavado con SSC desde 0,5X hasta 1X desde 55° hasta 60° C. Condiciones ilustrativas de alta astringencia incluyen hibridación en formamida al 50%, NaCl 1 M, SDS 1% a 37° C, y un lavado con SSC desde 0,1X desde 60° hasta 65° C.

La especificidad es típicamente la función de los lavados después de la hibridación, siendo los factores críticos la fuerza iónica y la temperatura de la solución de lavado final. Para híbridos DNA-DNA, el punto de fusión térmico (T_m) se puede aproximar a partir de la ecuación de Meinkoth et al., Anal. Biochem. 138: 267-284 (1984): $T_m = 81,5^\circ \text{C} + 16,6 (\log M) + 0,41 (\% \text{ GC}) - 0,61 (\% \text{ form}) - 500/L$; donde M es la molaridad de los cationes monovalentes, % GC es el porcentaje de nucleótidos de guanosina y citosina en el DNA, % form es el porcentaje de formamida en la solución de hibridación, y L es la longitud del híbrido en pares de bases. La T_m es la temperatura (bajo fuerza iónica y pH definidos) al cual hibrida un 50% de una secuencia diana complementaria con una sonda perfectamente apareada. La T_m se reduce en alrededor de 1° C para cada 1% de desemparejamiento; así, se pueden ajustar T_m , hibridación y/o condiciones de lavado para hibridar con secuencias de la identidad deseada. Por ejemplo, si se buscan secuencias con una identidad $\geq 90\%$, la T_m puede descender 10° C. Generalmente, se seleccionan condiciones de astringencia que sean alrededor de 5° C menor que la T_m para una secuencia determinada y su complementaria a una fuerza iónica y pH definidos. Sin embargo, condiciones de astringencia severas pueden utilizar una hibridación y/o lavados a 1, 2, 3, o 4° C menos que la T_m ; condiciones de astringencia moderadas pueden utilizar una hibridación y/o lavados a 6, 7, 8, 9, o 10° C menos que la T_m ; condiciones de astringencia baja pueden utilizar una hibridación y/o lavados a 11, 12, 13, 14, 15 o 20° C menos que la T_m . Utilizando la ecuación, las composiciones de hibridación y lavado, y T_m deseada, los expertos entenderán que las variaciones en la astringencia de hibridación y/o condiciones de lavado se describen inherentemente. Si el grado de desemparejamiento deseado da lugar a una T_m de menos de 45° C (solución acuosa) o 32° C (solución de formamida) es preferible incrementar la concentración de SSC para que se pueda usar una mayor temperatura. Se encuentra una extensa guía de la hibridación de ácidos nucleicos en Tijssen, Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); y Current Protocols in Molecular Biology, Chapter 2, Ausubel et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). Se pueden aplicar condiciones de hibridación y de lavado durante al menos 10, 30, 60, 90, 120, o 240 minutos.

"Identidad de secuencia" o "identidad" en el contexto de secuencias de ácido nucleico o polipéptido se refiere a bases de ácido nucleico o restos de aminoácidos en dos secuencias que son las mismas cuando se alinean para una correspondencia máxima sobre una ventana de comparación especificada.

Así, el "porcentaje de identidad de secuencia" se refiere al valor determinado mediante comparación de dos secuencias óptimamente alineadas sobre una ventana de comparación, en donde la porción de secuencia de polinucleótido o polipéptido en la ventana de comparación puede comprender adiciones o supresiones (esto es, huecos) en comparación con la secuencia de referencia (que no comprende adiciones o supresiones) para un alineamiento óptimo de las dos secuencias. El porcentaje se calcula determinando el número de posiciones en las cuales aparece la base de ácido nucleico o el resto de aminoácido idénticos en ambas secuencias para dar el número de posiciones coincidentes, dividiendo el número de posiciones coincidentes por el número total de posiciones en la ventana de comparación y multiplicando los resultados por 100 para dar el porcentaje de identidad de secuencia. Ejemplos útiles de porcentajes de identidad de secuencia incluyen, pero no se limitan a, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, o 95%, o cualquier porcentaje entero desde 50% hasta 100%. Estas identidades se pueden determinar usando cualquiera de los programas descritos en esta memoria.

Alineamientos de secuencia y porcentaje de identidad o cálculos similares pueden determinarse usando una variedad de métodos de comparación diseñados para detectar secuencias homólogas incluyendo, pero no limitados a, el programa MegAlign™ de LASERGENE computación bioinformática privada (DNASTAR Inc., Madison, WI). Dentro del contexto de esta aplicación se entiende que donde se usa el software de análisis de secuencia para el análisis, que los resultados del análisis se basarán en los “valores por defecto” del programa de referencia, a menos que se especifique lo contrario. Como se usa aquí los “valores por defecto” significarán cualquier conjunto de valores o parámetros que originalmente se cargan con el software cuando se inicializa la primera vez.

El “método de alineamiento Clustal V” corresponde al método de alineamiento designado como Clustal V (descrito por Higgins and Sharp, CABIOS. 5: 151-153 (1989); Higgins, D.G. et al., Comput. Appl. Biosci. 8: 189-191 (1992)) y establecido en el programa MegAlign™ de LASERGENE computación bioinformática privada (DNASTAR Inc., Madison, WI). Para alineamientos múltiples, los valores por defecto corresponden a PENALIZACIÓN POR HUECO (GAP PENALTY)= 10 y PENALIZACIÓN POR LONGITUD DE HUECO (GAP LENGTH PENALTY) = 10. Parámetros por defecto para alineamientos de pares de bases y cálculo de porcentaje de identidad de secuencias de proteínas usando el método Clustal V son KTUPLE= 1, GAP PENALTY= 3, WINDOW (VENTANA)= 5 y DIAGONALS SALVED (DIAGONALES GUARDADAS) = 5. Para ácidos nucleicos estos parámetros son KTUPLE= 2, GAP PENALTY= 5, WINDOW (VENTANA)= 4 y DIAGONALS SALVED (DIAGONALES GUARDADAS)= 4. Tras el alineamiento de las secuencias usando el programa Clustal V, es posible obtener un “porcentaje de identidad” mediante la visualización de la tabla de las “distancias de secuencia” en el mismo programa.

El “método de alineamiento Clustal W” corresponde al método de alineamiento designado Clustal W (descrito por Higgins and Sharp, *supra*; Higgins, D.G. et al., *supra*) y establecido en el programa MegAlign™ v6.1 de LASERGENE computación bioinformática privada (DNASTAR Inc., Madison, WI). Parámetros por defecto para alineamiento múltiple corresponden a GAP PENALTY= 10, GAP LENGTH PENALTY= 0,2, Delay Divergen Seqs (%)= 30, DNA Transition Weight= 0,5, Protein Weight Matrix= Gonnet Series, DNA Weight Matrix= IUB. Tras el alineamiento de secuencias usando el programa de Clustal W es posible obtener un “porcentaje de identidad” mediante la visualización de la tabla de las “distancias de secuencia” en el mismo programa.

El “método de alineamiento BLASTN” es un algoritmo proporcionado por el National Center for Biotechnology Information (NCBI) para comparar secuencias de nucleótidos usando parámetros por defecto.

Se entiende bien por un experto en la técnica que son útiles muchos niveles de identidad de secuencia para identificar polipéptidos, de otras especies, en donde tales polipéptidos tienen la misma o similar función o actividad. Ejemplos útiles de porcentajes de identidad incluyen, pero no se limitan a, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, o 95%, o cualquier porcentaje entero desde 50% hasta 100%. Ciertamente, puede ser útil cualquier identidad de aminoácidos entera desde 50% hasta 100% en la descripción de la presente invención, como 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%. También, de interés es cualquier complementariedad de longitud completa o parcial de este fragmento de nucleótidos aislado.

Así, la descripción abarcaba más que las secuencias de nucleótidos específicas ejemplo descritas en esta memoria. Por ejemplo, se contemplan las alteraciones en la secuencia génica que refleja la degeneración del código genético. También, es bien conocido en la técnica que es bien común que las alteraciones en un gen que dan como resultado la producción de un aminoácido químicamente equivalente en un sitio dado, no afectan a las propiedades funcionales de la proteína codificada. Se definen las sustituciones para la exposición aquí como intercambios en uno de los cinco grupos siguientes:

1. Restos alifáticos pequeños, no polares o ligeramente polares: Ala, Ser, Thr (Pro, Gly);
2. Restos polares, negativamente cargados y sus amidas: Asp, Asn, Glu, Glen;
3. Restos polares, positivamente cargados : His, Arg, Lys;
4. Restos alifáticos grandes no polares: Met, Leu, Ile, Val (Cys);
5. Restos aromáticos grandes: Phe, Tyr, Trp.

Así, un codón para el aminoácido alanina, un aminoácido hidrofóbico, se puede sustituir por un codón que codifica otros restos menos hidrofóbicos (como glicina) o un resto más hidrofóbico (como valina, leucina, o isoleucina). De modo similar, los cambios que dan como resultado la sustitución de un resto negativamente cargado por otro (como ácido aspártico o ácido glutámico) o un resto positivamente cargados por otro (como lisina o arginina) también se puede esperar que produzca un producto funcionalmente equivalente. En muchos casos, los cambios de nucleótidos que dan como resultado una alteración de las porciones terminales N-terminal y C-terminal de la molécula de proteína tampoco se espera que altere la actividad de la proteína.

Cada una de las modificaciones propuestas está bien dentro de la experiencia rutinaria de la técnica, como es la determinación de la actividad biológica de los productos codificados. Además, el experto en la técnica reconoce que

las secuencias sustancialmente similares también se definen por su capacidad de hibridar bajo condiciones astringentes, como se define anteriormente.

Fragmentos de ácidos nucleicos preferidos sustancialmente similares son aquellos fragmentos de ácidos nucleicos cuyas secuencias de nucleótidos son al menos un 70% idénticas a la secuencia de nucleótidos de los fragmentos de ácidos nucleicos que se citan en esta memoria. Fragmentos de ácidos nucleicos más preferidos son al menos un 90% idénticos a la secuencia de nucleótidos de los fragmentos de ácidos nucleicos que se citan en esta memoria. Los más preferidos son fragmentos de ácidos nucleicos que son al menos un 95% idénticos a la secuencia de nucleótidos de los fragmentos de ácidos nucleicos que se citan en esta memoria.

Una "porción sustancial" de una secuencia de aminoácido o de un nucleótido es aquella porción que comprende suficiente secuencia de aminoácidos de un polipéptido o la secuencia de nucleótidos de un gen para supuestamente identificar ese polipéptido o gen, ya sea mediante evaluación manual de las secuencias por un experto en la técnica, o mediante una comparación de secuencia automatizada por ordenador y identificación usando algoritmos como BLAST, (Basic Local Alignment Search Tool; Altschul, S. F., et al., J. Mol. Biol., 215: 403-410 (1993)). En general, una secuencia de 10 o más aminoácidos contiguos o treinta o más nucleótidos son necesarios para supuestamente identificar un polipéptido o secuencia de ácido nucleico como homóloga a una proteína o gen conocidos. Además, con respecto a secuencias de nucleótidos, sondas de oligonucleótidos específicas del gen que comprenden 20-30 nucleótidos contiguos se pueden usar en métodos de identificación de genes dependientes de secuencia (p. ej. Hibridación de Southern) e aislamiento (p. ej. hibridación *in situ* de colonias bacterianas o placas de bacteriófagos). Además, se pueden usar oligonucleótidos cortos de 12-15 pares de bases como cebadores de amplificación en PCR con el fin de obtener un fragmento de ácido nucleico concreto que comprende los cebadores. En consecuencia, una "porción sustancial" de una secuencia de nucleótidos comprende suficiente secuencia para identificar y/o aislar específicamente un fragmento de ácido nucleico que comprende la secuencia. La presente memoria descriptiva enseña las secuencias de aminoácidos y nucleótidos completas que codifican proteínas concretas. El experto en la materia, teniendo el beneficio de las secuencias como se informa en la presente memoria, puede ahora usar todas o porciones sustanciales de las secuencias descritas para los propósitos conocidos por los expertos en este arte.

El término "complementario" describe la relación entre dos secuencias de bases de nucleótidos que son capaces de apareamiento de bases tipo Watson-Crick cuando se alinean en una orientación antiparalela. Por ejemplo, con respecto al DNA, la adenosina es capaz de aparearse con la timina, y la citosina es capaz de aparearse con la guanina. Por consiguiente, la presente invención puede hacer uso de moléculas de ácido nucleico aisladas que son complementarias a las secuencias completas como se informa en el Listado de Secuencias que acompaña y la especificación al igual que aquellas secuencias de ácidos nucleicos sustancialmente similares.

El término "aislado" se refiere a un polipéptido o secuencia de nucleótidos que se retira de al menos un componente con el cual está naturalmente asociado.

"Promotor" se refiere a una secuencia de DNA capaz de controlar la expresión de una secuencia codificadora o RNA funcional. La secuencia promotora consiste en elementos proximales y más distales corriente arriba, los últimos elementos se refieren a menudo a potenciadores. Por consiguiente, un "potenciador" es una secuencia de DNA que puede estimular la actividad promotora, y puede ser un elemento innato del promotor o un elemento heterólogo insertado para aumentar el nivel o la especificidad de tejido de un promotor. Los promotores se pueden derivar en su totalidad de un gen nativo, o estar compuestos de diferentes elementos derivados de distintos promotores que se encuentran en la naturaleza, o incluso que comprenden segmentos sintéticos de DNA. Se entenderá por los expertos en la técnica que distintos promotores pueden dirigir la expresión de un gen en distintos tipos de tejidos o células, o a distintas etapas de desarrollo, o en respuesta a distintas condiciones ambientales. Se reconoce, además que en la mayoría de los casos los límites exactos de las secuencias reguladoras no se han definido completamente, fragmentos con alguna variación podrían tener una actividad promotora idéntica. Los promotores que causan que un gen se exprese en la mayoría de tipos celulares se denominan comúnmente "promotores constitutivos".

"Secuencias 3' no codificadoras", "terminador de la transcripción" y "secuencias de terminación" se usan de manera intercambiable en esta memoria y se refieren a secuencias de DNA localizadas corriente debajo de una secuencia codificadora, incluyendo secuencias de reconocimiento de poliadenilación y otras secuencias que codifican señales reguladoras capaces de afectar al procesamiento de RNA o la expresión génica. La señal de poliadenilación usualmente se caracteriza por afectar a la adición de extensiones de ácido poliadenílico en el extremo 3' del RNA precursor.

El término "ligado operativamente" se refiere a la asociación de secuencias de ácidos nucleicos a un fragmento único de ácido nucleico de manera que la función de uno se ve afectada por el otro. Por ejemplo, un promotor está operativamente ligado a una secuencia codificadora cuando es capaz de afectar a la expresión de esa secuencia codificadora (esto es, la secuencia codificadora está bajo el control transcripcional del promotor). Las secuencias codificadoras pueden ligarse operativamente a secuencias reguladoras en una orientación sentido o antisentido. En otro ejemplo, las regiones complementarias de RNA pueden estar ligadas operativamente, ya sea directa o indirectamente, 5' al mRNA diana, o 3' al mRNA diana, o dentro del mRNA diana, o una primera región complementaria es 5' y su complemento es 3' para el mRNA diana.

Las técnicas estándar de DNA recombinante y clonación molecular usadas en esta memoria se conocen bien en la técnica y se describen más completamente en Sambrook, J., Fritsch, E.F. and Maniatis, T. Molecular Cloning: A Laboratory Manual; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY (1989). Los métodos de transformación se conocen bien por aquellos expertos en la técnica y se describen más adelante en esta memoria.

5 "PCR" o "reacción en cadena de la polimerasa" es una técnica para la síntesis de grandes cantidades de segmentos específicos DNA y consiste en una serie de ciclos repetitivos (Perkin Elmer Cetus Instruments, Norwalk, CT). Típicamente, el DNA de doble cadena se desnaturaliza con calor, los dos cebadores complementarios al extremo 3' del segmento diana se alinean a baja temperatura y se elongan a temperatura intermedia. Un conjunto de estos pasos consecutivos se conoce como un "ciclo".

10 Un "plásmido" o un "vector" es un elemento extra cromosómico que a menudo lleva genes que no son parte del metabolismo central de la célula, y usualmente en forma de fragmentos circulares de doble cadena. Tales elementos pueden ser secuencias que se replican autónomamente, secuencias que se integran en el genoma, fagos, o secuencias de nucleótidos, lineales o circulares, de DNA o RNA de cadena sencilla o doble, derivados de cualquier fuente, en la que un número de secuencias de nucleótidos se han unido o recombinado en una única construcción
15 que es capaz de introducir un(os) casete(s) de expresión dentro de una célula.

El término "alterado genéticamente" se refiere al proceso de intercambio de material hereditario mediante ingeniería genética, transformación y/o mutación.

El término "recombinante" se refiere a una combinación artificial de dos segmentos de secuencia separados de otro modo, p. ej. mediante síntesis química o por la manipulación de segmentos aislados de ácidos nucleicos mediante técnicas de ingeniería genética. "Recombinante" también incluye la referencia a una célula o vector, que ha sido
20 modificado mediante la introducción de un ácido nucleico heterólogo o una célula derivada de una célula así modificada, pero no abarca la alteración de la célula o vector por eventos naturales que ocurren (p. ej. mutación espontánea, transformación natural, transducción natural, transposición natural) como los que ocurren sin la intervención humana deliberada.

25 Los términos "producto de construcción recombinante", "producto de construcción de expresión", "producto de construcción químico", "producto de construcción", y "producto de construcción de DNA recombinante", se usan de forma intercambiable en esta memoria. Un producto de construcción recombinante comprende una combinación artificial de fragmentos de ácido nucleico, p. ej. secuencias reguladoras y codificadoras que no se encuentran juntas en la naturaleza. Por ejemplo, un producto de construcción recombinante puede comprender secuencias
30 reguladoras y secuencias codificadoras que se derivan de distintas fuentes, o secuencias reguladoras y secuencias codificadoras derivadas de la misma fuente, pero reorganizadas de una manera diferente que la encontrada en la naturaleza. Tal construcción se puede utilizar por sí misma o se puede usar en conjunto con un vector. Si se usa un vector, entonces la elección de un vector depende del método que se utilizará para transformar las células huésped como es bien sabido por los expertos en la técnica. Por ejemplo, se puede usar un vector plasmídico. El experto en
35 la materia es muy consciente de los elementos genéticos que deben estar presentes en el vector para transformar con éxito, seleccionar y propagar las células huésped que comprenden cualquiera de los fragmentos de ácidos nucleicos aislados descritos en esta memoria. El experto en la materia también reconocerá que los eventos independientes de transformación pueden dar como resultado diferentes niveles y patrones de expresión (Jones et al., EMBO J. 4: 2411-2418 (1985); De Almeida et al., Mol. Gen. Genetics 218: 78-86 (1989)), y de esta forma que los eventos múltiples pueden necesitar ser examinados con el fin de obtener líneas que muestran el nivel de expresión y patrón deseados. Tal selección puede realizarse mediante análisis de Southern de DNA, análisis de Northern de la expresión de mRNA, análisis por inmunoelectrotransferencia de la expresión de proteína, o análisis fenotípico, entre
40 otros.

El término "expresión", como se usa aquí, se refiere a la producción de un producto final funcional (p. ej. un mRNA o una proteína [ya sea precursora o madura]).
45

El término "introducido" significa aportar un ácido nucleico (p. ej. producto de construcción de expresión) o proteína dentro de una célula. Introducido incluye la referencia a la incorporación de un ácido nucleico dentro de una célula eucariota o procariota donde el ácido nucleico puede ser incorporado dentro del genoma de la célula, e incluye referencia al suministro transitorio de un ácido nucleico o proteína a la célula. Introducido incluye referencia a
50 métodos de transformación estables o transitorios, así como cruce sexual. Por tanto, "introducido" en el contexto de insertar un fragmento de ácido nucleico (p. ej. un producto de construcción recombinante/ producto de construcción de expresión) dentro de una célula, significa "transfección" o "transformación" o "transducción" e incluye referencia a la incorporación de un fragmento de ácido nucleico dentro de una célula eucariota o procariota donde el fragmento de ácido nucleico se puede incorporar dentro del genoma de la célula (p. ej. cromosoma, plásmido, plasto o DNA mitocondrial), convertido en un replicón autónomo, o expresado transitoriamente (p. ej. mRNA transfectado).
55

El término "homólogo" se refiere a proteínas o polipéptidos de origen evolutivo común con función catalítica similar. La invención puede incluir bacterias que producen proteínas homólogas mediante tecnología recombinante.

Descritas en la presente memoria están las bacterias recombinantes que han sido diseñadas para usar sacarosa usando nuevos genes de utilización de sacarosa de *Pseudomonas fluorescens* (*P. fluorescens*). Específicamente, las bacterias recombinantes descritas en esta memoria comprenden en su genoma o en al menos un producto de construcción recombinante: una secuencia de nucleótidos que codifica un polipéptido que posee actividad transportadora de sacarosa, el polipéptido que tiene al menos una identidad de secuencia del 95%, según un método de alineamiento Clustal V, cuando se compara con una secuencia de aminoácidos según se establece en SEQ ID NO: 24; y una secuencia de nucleótidos que codifica un polipéptido que posee actividad hidrolasa de sacarosa, el polipéptido que tiene al menos una identidad de secuencia del 95%, según un método de alineamiento Clustal V, cuando se compara con una secuencia de aminoácidos según se establece en SEQ ID NO: 26. Las secuencias de nucleótidos están cada una ligadas operativamente al mismo o a diferente promotor. En algunas realizaciones, las bacterias recombinantes son capaces de metabolizar sacarosa para producir glicerol y/o productos derivados del glicerol.

Bacterias huésped adecuadas para usar en la construcción de las bacterias recombinantes descritas en esta memoria incluyen, pero no están limitadas a, organismos del género: *Escherichia*, *Streptococcus*, *Agrobacterium*, *Bacillus*, *Corynebacterium*, *Lacto-bacillus*, *Clostridium*, *Gluconobacter*, *Citrobacter*, *Enterobacter*, *Klebsiella*, *Aerobacter*, *Methylobacter*, *Salmonella*, *Streptomyces*, y *Pseudomonas*.

En una realización, la bacteria huésped se selecciona del género: *Escherichia*, *Klebsiella*, *Citrobacter*, y *Aerobacter*.

En otra realización, la bacteria huésped es *Escherichia coli*.

En algunas realizaciones, la bacteria huésped es PTS menos. En estas realizaciones, la bacteria huésped es PTS menos en su estado nativo, o puede dar PTS menos a través de la inactivación de un gen PTS como se describe debajo.

En producción de microorganismos, es deseable algunas veces desligar el transporte de azúcares y el uso de fosfoenolpiruvato (PEP) para la fosforilación de los azúcares transportados.

El término "regulación a la baja" se refiere a la reducción en, o abolición de, la actividad de la(s) proteína(s) activa(s), en comparación con la actividad de la(s) proteína(s) salvaje(s). Las PTS pueden ser inactivadas (dando como resultado un organismo "PTS menos") mediante regulación a la baja de la expresión de uno o más de los genes endógenos que codifican proteínas requeridas en este tipo de transporte. La regulación a la baja ocurre típicamente cuando uno o más de estos genes tiene una "ruptura", refiriéndose a una inserción, supresión, o mutación dirigida dentro de una porción de ese gen, que da como resultado, ya sea un gen completo *knockout* como el gen de tal manera que se elimina el gen del genoma y no se traduce ninguna proteína, o una proteína que se ha traducido de tal manera que tiene una inserción, supresión, sustitución de aminoácido u otra mutación dirigida. La localización de la interrupción en la proteína puede ser, por ejemplo, dentro de la porción N-terminal de la proteína o dentro de la porción C-terminal de la proteína. La proteína interrumpida tendrá actividad disminuida con respecto a la proteína que no ha sido interrumpida, y puede ser no-funcional. La regulación a la baja que da como resultado una baja o ausente expresión de la proteína, también podría resultar mediante manipulación de las secuencias reguladoras, factores de transcripción y traducción y/o vías de transducción de señales o mediante el uso de sentido, antisentido o tecnología del RNAi, etc.

Los polipéptidos transportadores de sacarosa son polipéptidos que son capaces de mediar en el transporte de sacarosa dentro de las células microbianas. Los péptidos transportadores de sacarosa son conocidos en la técnica, por ejemplo el polipéptido CscB de la cepa salvaje EC3132 de *E. Coli* (establecido en SEQ ID NO:28), codificado por el gen *cscB* (secuencia codificadora representada en SEQ ID NO:27), como se describe por Jahreis et al. (J. Bacteriol. 184: 5307-5316, 2002). El gen *lac Y* de *Pseudomonas fluorescens* Pf5 ATCC® BAA-477 (secuencia codificadora representada en SEQ ID NO: 23) codifica un polipéptido (establecido en SEQ ID NO: 24) que es un 41% idéntico a la conocida proteína transportadora de sacarosa CscB, de la cepa EC3132 de *Escherichia coli*. Sin embargo, no hay actividad conocida demostrada para este polipéptido.

En otra realización, la secuencia de nucleótidos que codifica un polipéptido que tiene actividad transportadora de sacarosa tiene al menos un 95% de identidad de secuencia, según el método de alineamiento BLASTN, cuando se compara con una secuencia de nucleótidos como la representada en SEQ ID NO: 23.

En otra realización, la secuencia de nucleótidos que codifica un polipéptido con actividad transportadora de sacarosa tiene una secuencia de nucleótidos representada en SEQ ID NO: 23.

En otra realización, el polipéptido que tiene actividad transportadora de sacarosa tiene al menos un 95% de identidad de secuencia, según el método de alineamiento Clustal V, cuando se compara con una secuencia de aminoácidos como la representada en SEQ ID NO: 24.

En otra realización, el polipéptido que tiene actividad transportadora de sacarosa tiene una secuencia de aminoácidos representada en SEQ ID NO: 24.

Los polipéptidos que tienen actividad hidrolasa de sacarosa tienen la capacidad de catalizar la hidrólisis de sacarosa a los productos fructosa y glucosa. Son conocidos los polipéptidos que tienen actividad hidrolasa de sacarosa, e incluyen, *CscA* de la cepa salvaje EC3132 de *E. Coli* (establecido en SEQ ID NO: 30), codificado por el gen *cscA* (secuencia codificadora representada en SEQ ID NO: 29), como se describe en Jahreis et al. (J. Bacteriol. 184: 5307-5316, 2002). El gen *scrB* de *Pseudomonas fluorescens* Pf5 ATCC® BAA-477 (secuencia codificadora representada en SEQ ID NO: 25) codifica un polipéptido (establecido en SEQ ID NO: 26) que es un 47% idéntico a la conocida hidrolasa de sacarosa *CscA*, de la cepa EC3132 de *Escherichia coli*. Sin embargo, no hay actividad conocida demostrada para este polipéptido.

En una realización, la secuencia de nucleótidos que codifica un polipéptido que tiene actividad hidrolasa de sacarosa tiene al menos una identidad de secuencia del 95%, según un método de alineamiento BLASTN, cuando se compara con una secuencia de nucleótidos según se establece en SEQ ID NO: 25.

En otra realización, la secuencia de nucleótidos que codifica un polipéptido que tiene actividad hidrolasa de sacarosa tiene una secuencia de nucleótidos representada en SEQ ID NO: 25.

En una realización, el polipéptido que tiene actividad hidrolasa de sacarosa tiene al menos una identidad de secuencia del 95%, según un método de alineamiento Clustal V, cuando se compara con una secuencia de aminoácidos según se establece en SEQ ID NO: 26.

En otra realización, el polipéptido que tiene actividad hidrolasa de sacarosa tiene una secuencia de aminoácidos representada en SEQ ID NO: 26.

Las bacterias recombinantes descritas en esta memoria pueden comprender además en sus genomas o en al menos un producto de construcción recombinante, una secuencia de nucleótidos que codifica un polipéptido que tiene actividad fructoquinasa para permitir a las bacterias utilizar la fructosa producida por la hidrólisis de sacarosa. Los polipéptidos que tienen actividad fructoquinasa incluyen fructoquinasas (designadas EC 2.7.1.4) y varias hexosa quinasas que tienen actividad de fosforilar fructosa (EC 2.7.1.3. y EC 2.7.1.1). La actividad fosforiladora de fructosa puede ser exhibida por hexoquinasas y cetohehexoquinasas. Los genes representativos que codifican polipéptidos de una variedad de microorganismos, que se pueden usar para construir las bacterias recombinantes descritas en esta memoria están enumerados en la Tabla 1. Un experto en la técnica conocerá esas proteínas que son sustancialmente similares a la proteína que es capaz de fosforilar fructosa (como codificadas por los genes enumerados en la Tabla 1) que también se pueden usar.

Tabla 1

Secuencias que Codifican Enzimas con Actividad Fructoquinasa

Fuente	Nombre del gen	Número EC	SEQ ID NO: de Nucleótido	SEQ ID NO: de Proteína
<i>Agrobacterium tumefaciens</i>	<i>scrK</i> (fructoquinasa)	2.7.1.4	31	32
<i>Streptococcus mutans</i>	<i>scrK</i> (fructoquinasa)	2.7.1.4	33	34
<i>Escherichia coli</i>	<i>scrK</i> (fructoquinasa)	2.7.1.4	35	36
<i>Klebsiella pneumoniae</i>	<i>scrK</i> (fructoquinasa)	2.7.1.4	37	38
<i>Escherichia coli</i>	<i>cscK</i> (fructoquinasa)	2.7.1.4	39	40
<i>Enterococcus faecalis</i>	<i>cscK</i> (fructoquinasa)	2.7.1.4	41	42
<i>Saccharomyces cerevisiae</i>	<i>HXK1</i> (hexoquinasa)	2.7.1.1	43	44
<i>Saccharomyces cerevisiae</i>	<i>HXK2</i> (hexoquinasa)	2.7.1.1	45	46

En una realización, el polipéptido que tiene actividad fructoquinasa tiene al menos una identidad de secuencia del 95%, según un método de alineamiento Clustal V, con una secuencia de aminoácidos según se establece en SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, o SEQ ID NO: 46..

En otra realización, el polipéptido que tiene actividad fructoquinasa tiene una secuencia de aminoácidos representada en SEQ ID NO: 40.

Las secuencias codificadoras de los genes que codifican polipéptidos que tienen actividad transportadora de sacarosa y polipéptidos que tienen actividad hidrolasa de sacarosa se pueden usar para aislar secuencias de nucleótidos que codifican polipéptidos homólogos de la misma o de otras especies microbianas. Por ejemplo, se pueden identificar homólogos de los genes usando software de análisis de secuencia, como BLASTN, para buscar bases de datos de secuencias de ácidos nucleicos públicamente disponibles. Además, el aislamiento de genes homólogos usando protocolos dependientes de secuencia se conoce bien en la técnica. Ejemplos de protocolos dependientes de secuencia incluyen, pero no se limitan a, métodos de hibridación de ácidos nucleicos, y métodos de

amplificación de DNA y RNA como se ejemplifica mediante varios usos de tecnologías de amplificación de ácidos nucleicos (p. ej. reacción en cadena de la polimerasa (PCR), Mullis et al., Patente de EE.UU. Nº 4.683.202; reacción en cadena de la ligasa (LCR), Tabor, S. Et al., Proc. Acad. Sci. USA 82, 1074, 1985); o amplificación por desplazamiento de cadena (SDA), Walker, et al., Proc. Natl. Acad. Sci. U.S.A., 89: 392, (1992)). Por ejemplo, la

secuencia de nucleótidos que codifica los polipéptidos descritos arriba se pueden emplear como una sonda de hibridación para la identificación de homólogos.

Uno con experiencia ordinaria en la técnica aprecia que aquellos genes que codifican estos polipéptidos aislados de otras fuentes se pueden usar en las bacterias recombinantes descritas en esta memoria. Además, se pueden hacer variaciones en las secuencias de nucleótidos que codifican los polipéptidos sin afectar a la secuencia de aminoácidos del polipéptido codificado debido a la degeneración de codón, y esas sustituciones de aminoácidos, supresiones o adiciones que producen una proteína sustancialmente similar se pueden incluir en la proteína codificada.

Las secuencias de nucleótidos que codifican los polipéptidos que tienen actividad transportadora de sacarosa y los polipéptidos que tienen actividad hidrolasa de sacarosa se pueden aislar usando PCR (véase, p. ej. Patente de EE.UU. Nº 4.683.202) y cebadores diseñados para unirse a la secuencia deseada, por ejemplo como se describe en el Ejemplo 1 de la presente memoria. Se conocen bien otros métodos de aislamiento de genes por uno experto en la técnica tal como usando cebadores degenerados o hibridación con sonda heteróloga. Las secuencias de nucleótidos se pueden sintetizar químicamente o adquirirse comercialmente de vendedores como DNA2.0 Inc. (Menlo Park, CA).

La expresión de los polipéptidos puede efectuarse usando unos de los muchos métodos conocidos por un experto en la técnica. Por ejemplo, las secuencias de nucleótidos que codifican los polipéptidos descritos más arriba se pueden introducir dentro de la bacteria en al menos un plásmido multicopia, o mediante la integración de una o más copias de las secuencias codificadoras dentro del genoma huésped. Las secuencias de nucleótidos que codifican los polipéptidos se pueden introducir dentro de la bacteria huésped separadamente (p. ej. en plásmidos separados) o en cualquier combinación (p. ej., en un único plásmido, como se describe en los Ejemplos de la presente memoria).

Las regiones codificadoras introducidas que son ya sea en un(os) plásmido(s) o en el genoma, se pueden expresar a partir de al menos un promotor altamente activo. Una región codificadora integrada puede ser introducida ya sea como parte de un gen quimérico que tiene su propio promotor, o ser integrada adyacente a un promotor altamente activo que es endógeno al genoma o en un operón altamente expresado. Promotores adecuados incluyen, pero no se limitan a, *CYC1*, *HIS3*, *GAL1*, *GAL10*, *ADH1*, *PGK*, *PHO5*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO*, y *lac*, *ara*, *tet*, *trp*, *IP_L*, *IP_R*, *T7*, *tac*, y *trc* (útiles para la expresión en *Escherichia coli*) así como los promotores *amy*, *apr*, *npr* y varios promotores de fagos útiles para la expresión en *Bacillus*. El promotor puede también ser el promotor de la glucosa isomerasa de *Streptomyces lividans* o un variante del mismo, descrito por Payne et al. (Patente de EE.UU. Nº 7.132.527).

En una realización, las bacterias recombinantes descritas en esta memoria son capaces de producir glicerol. Se conocen procesos biológicos para la preparación de glicerol usando carbohidratos o azúcares en levaduras y en algunas bacterias, otros hongos, y algas. Ambas, bacterias y levaduras, producen glicerol mediante la conversión de glucosa u otros carbohidratos a través de la vía de la fructosa-1,6-bisfosfato en la glicolisis. En el método de producción de glicerol descrito en esta memoria, las bacterias huésped se pueden usar para que naturalmente produzcan glicerol. Además, las bacterias se pueden modificar genéticamente para la producción de glicerol y derivados de glicerol. La capacidad de producción de glicerol a partir de una variedad de sustratos se puede proporcionar a través de la expresión de las actividades enzimáticas glicerol-3-fosfato deshidrogenasa (G3PDH) y/o glicerol-3-fosfatasa como se describe en la Patente de EE.UU. Nº 7.005.291. Genes que codifican estas proteínas que se pueden usar para expresar las actividades enzimáticas en una bacteria huésped están descritos en la Patente de EE.UU. Nº 7.005.291. Ejemplos adecuados de genes que codifican polipéptidos que tienen actividad glicerol-3-fosfato deshidrogenasa incluyen, pero no se limitan a, *GPD1* de *Saccharomyces cerevisiae* (secuencia codificadora representada en SEQ ID NO: 1, secuencia de proteína codificada representada en SEQ ID NO: 2) y *GPD2* de *Saccharomyces cerevisiae* (secuencia codificadora representada en SEQ ID NO: 3, secuencia de proteína codificada representada en SEQ ID NO: 4). Ejemplos adecuados de genes que codifican polipéptidos que tienen actividad glicerol-3-fosfatasa incluyen, pero no se limitan a, *GPP1* de *Saccharomyces cerevisiae* (secuencia codificadora representada en SEQ ID NO: 5, secuencia de proteína codificada representada en SEQ ID NO: 6) y *GPP2* de *Saccharomyces cerevisiae* (secuencia codificadora representada en SEQ ID NO: 7, secuencia de proteína representada en SEQ ID NO: 8).

El aumento de la producción de glicerol se puede lograr a través de la reducción de la expresión de genes endógenos diana. La regulación a la baja de los genes endógenos que codifican las actividades de glicerol quinasa y de glicerol deshidrogenasa mejora aún más la producción de glicerol como se describe en la Patente de EE.UU. Nº 7.005.291. El aumento de la canalización de carbono a glicerol puede llevarse a cabo mediante la reducción de la expresión del gen endógeno que codifica la glicerolaldehído 3-fosfato deshidrogenasa, como se describe en la Patente EE.UU. Nº 7.371.558. La regulación a la baja puede llevarse a cabo usando cualquier método conocido en la técnica, por ejemplo, los métodos descritos más arriba para la regulación a la baja de genes del sistema PTS.

El glicerol proporciona un sustrato para la producción microbiana de productos útiles. Ejemplos de tales productos, esto es, derivados de glicerol incluyen, pero no se limitan a, ácido 3-hidroxipropiónico, metilglioxal, 1,2-propanodiol, y 1,3-propanodiol.

En otra realización, las bacterias recombinantes descritas en esta memoria son capaces de producir 1,3-propanodiol. El derivado de glicerol 1,3-propanodiol es un monómero que tiene potencial utilidad en la producción de fibras de poliéster y la fabricación de poliuretanos y compuestos cíclicos. El 1,3-propanodiol se puede producir por un único microorganismo mediante la bioconversión de un sustrato de carbono distinto de glicerol o dihidroxiacetona, como se describe en la Patente de EE.UU. N° 5.686.276. En esta bioconversión, el glicerol se produce a partir del sustrato de carbono, como se describió anteriormente. El glicerol es convertido al intermediario 3-hidroxipropionaldehído por la enzima deshidratasa, que puede ser codificada por la bacteria huésped o puede ser introducida dentro del huésped por recombinación. La deshidratasa puede ser glicerol deshidratasa (E.C. 4.2.1.30), diol deshidratasa (E.C. 4.2.1.28) o cualquier otra enzima capaz de catalizar esta conversión. Un ejemplo adecuado de genes que codifican las subunidades "α" (alfa), "β" (beta), y "γ" (gamma) de la glicerol deshidratasa incluyen, pero no se limitan a, *dhaB1* (secuencia codificadora representada en SEQ ID NO: 9), *dhaB2* (secuencia codificadora representada en SEQ ID NO: 11), *dhaB3* (secuencia codificadora representada en SEQ ID NO: 13), respectivamente, de *Klebsiella pneumoniae*. La conversión ulterior de 3-hidroxipropionaldehído en 1,3-propanodiol puede ser catalizada por la 1,3-propanodiol deshidrogenasa (E.C. 1.1.1.202) u otras alcohol deshidrogenasas. Un ejemplo adecuado de un gen que codifica una 1,3-propanodiol deshidrogenasa es *dha T* de *Klebsiella pneumoniae* (secuencia codificadora representada en SEQ ID NO: 47, secuencia de proteína codificada representada en SEQ ID NO: 48).

Las bacterias pueden ser diseñadas de manera recombinante para proporcionar una producción más eficiente de glicerol y del derivado de glicerol 1,3-propanodiol. Por ejemplo, la Patente de EE.UU. N° 7.005.291 describe microorganismos transformados y un método para la producción de glicerol y 1,3-propanodiol con ventajas derivadas de expresar actividades endógenas de uno o ambas, glicerol-3-fosfato deshidrogenasa y glicerol-3-fosfato fosfatasa, mientras que la interrupción de una o ambas de las actividades endógenas al tiempo que se interrumpe una o ambas de las actividades endógenas de glicerol quinasa y glicerol deshidrogenasa.

La Patente de EE.UU. N° 6.013.494 describe un proceso para la producción de 1,3-propanodiol, usando un único microorganismo que comprende glicerol-3-fosfato deshidrogenasas, glicerol-3-fosfato fosfatasa, deshidratasa, y 1,3-propanodiol oxidoreductasa (p.ej. *dhaT*). La Patente de EE.UU. N° 6.136.576 describe un método para la producción de 1,3-propanodiol, que comprende un microorganismo recombinante que comprende además deshidratasa y proteína X (más tarde identificado como un péptido factor de reactivación de la deshidratasa).

La Patente de EE.UU. N° 6.514.733 describe una mejora del proceso donde se produce un aumento significativo en el título (gramos de producto por litro) en virtud de una actividad catalítica no específica (distinguida de 1,3-propanodiol oxidoreductasa codificada por *dha T*) para convertir 3-hidroxipropionaldehído en 1,3-propanodiol. Además, la Patente de EE.UU. N° 7.132.527 describe vectores y plásmidos útiles para la producción de 1,3-propanodiol.

Se puede lograr un aumento en la producción de 1,3-propanodiol mediante modificaciones adicionales a una bacteria huésped, incluyendo regulación a la baja de la expresión de algunos genes diana y regulación al alza, expresión de otros genes diana, como se describe en la Patente de EE.UU. N° 7.371.558. Para la utilización de glucosa como fuente de carbono en un huésped PTS menos se puede aumentar la expresión de la actividad glucoquinasa.

Genes adicionales cuya expresión aumentada o sobre regulada incrementa la producción de 1,3-propanodiol incluyen genes que codifican:

- fosfoenolpiruvato carboxilasa típicamente caracterizada como EC 4.1.1.31
- cob(I)alamina adenosiltransferasa, típicamente caracterizada como EC 2.5.1.17
- actividad catalítica no específica que es suficiente para catalizar la interconversión de 3-HPA y 1,3-propanodiol, y excluye específicamente 1,3-propanodiol oxidoreductasa(s), típicamente estas enzimas son alcohol deshidrogenasas

Genes cuya expresión reducida o sub regulada incrementa la producción de 1,3-propanodiol incluyen genes que codifican:

- proteína de control de la respiración aeróbica
- metilglioxal sintasa
- acetato quinasa
- fosfotransacetilasa

- aldehído deshidrogenasa A
- aldehído deshidrogenasa B
- triosafosfato isomerasa
- fosfogluconato deshidratasa

5 En otra realización, las bacterias recombinantes descritas en esta memoria son capaces de producir ácido 3-hidroxipropiónico. El ácido 3-hidroxipropiónico tiene utilidad en la síntesis de la especialidad y se puede convertir en intermedios comercialmente importantes por la técnica conocida en la industria química, p. ej., ácido acrílico por deshidratación, ácido malónico por oxidación, ésteres por reacciones de esterificación con alcoholes, y 1,3-propanodiol por reducción. El ácido 3-hidroxipropiónico se puede producir biológicamente a partir de una fuente fermentable de carbono por un único microorganismo, como se describe en tramitación con la presente y de propiedad común de EE.UU., solicitud de patente N° 12/815461. En una ruta biosintética representativa, un sustrato de carbono es convertido en 3-hidroxipropionaldehído, como se describe más arriba para la producción de 1,3-propanodiol. El 3-hidroxipropionaldehído es convertido en ácido 3-hidroxipropiónico por una aldehído deshidrogenasa. Ejemplos adecuados de aldehído deshidrogenasas incluyen, pero no se limitan a, AldB (SEQ ID NO: 16), codificada por el gen de *E. Coli aldB* (secuencia codificadora representada en SEQ ID NO: 15); AldA (SEQ ID NO: 18), codificada por el gen de *E. Coli aldA* (secuencia codificadora representada en SEQ ID NO: 17); y AldH (SEQ ID NO: 20), codificada por el gen de *E. Coli aldH* (secuencia codificadora representada en SEQ ID NO: 19).

Muchas de las modificaciones descritas más arriba para mejorar la producción de 1,3-propanodiol por una bacteria recombinante se pueden hacer también para mejorar la producción de ácido 3-hidroxipropiónico. Por ejemplo, la eliminación de la glicerol quinasa previene que el glicerol, formado a partir de G3P por la acción de la G3P fosfatasa, sea reconvertido en G3P a expensas de ATP. También, la eliminación de la glicerol deshidrogenasa (por ejemplo, *gldA*) previene que el glicerol, formado a partir de DHAP por la acción de la glicerol-3-fosfato deshidrogenasa dependiente de NAD, de ser convertido en hidroxiacetona. Las mutaciones se pueden dirigir hacia un gen estructural con el fin de perjudicar o mejorar la actividad de una actividad enzimática o se puede dirigir hacia un gen regulador, incluyendo regiones promotoras y sitios de unión a ribosomas, con el fin de modular el nivel de expresión de una actividad enzimática.

La regulación al alza o regulación a la baja se puede lograr mediante una variedad de métodos que son conocidos por aquellos expertos en la técnica. Se entiende bien que la regulación al alza o regulación a la baja de un gen se refiere a una alteración en el nivel de actividad presente en la célula que se deriva de la proteína codificada por ese gen en relación al nivel de control de la actividad, por ejemplo, por la actividad de la proteína codificada por el gen salvaje (o no alterado) correspondiente.

Genes específicos involucrados en una vía enzimática pueden ser sobre regulados para incrementar la actividad de su(s) función(es) codificada(s). Por ejemplo, se pueden introducir copias adicionales de genes seleccionados dentro de la célula huésped o plásmidos multicopia como pBR322. Tales genes se pueden integrar dentro del cromosoma con secuencias reguladoras apropiadas que dan como resultado una actividad incrementada de sus funciones codificadas. Los genes diana se pueden modificar con el fin de estar bajo el control de promotores no nativos o promotores nativos alterados. Se pueden alterar los promotores endógenos *in vivo* mediante mutación, supresión y/o sustitución.

Alternativamente, puede ser útil reducir o eliminar la expresión de ciertos genes en relación a un nivel de actividad dado. Se conocen los métodos de sub regular (interrupción) genes por aquellos expertos en la materia.

Puede ocurrir regulación a la baja mediante supresión, inserción, o alteración de las regiones codificadoras y/o reguladoras (promotor). Se pueden obtener sub regulaciones específicas mediante mutación aleatoria seguido de chequeo o selección, o, donde se conoce la secuencia del gen, mediante intervención directa por métodos de biología molecular conocidos por aquellos expertos en la técnica. Un método útil en particular, pero no exclusivo, para efectuar regulación a la baja es alterar la fuerza del promotor.

Además, la regulación a la baja de una expresión génica se puede usar para ya sea prevenir la expresión de la proteína de interés o para dar como resultado la expresión de una proteína que es no funcional. Esto se puede lograr por ejemplo, por 1) suprimiendo las regiones codificadoras y/o regiones reguladoras (promotoras), 2) insertando secuencias de ácidos nucleicos exógenas dentro de las regiones codificadoras y/o regiones reguladoras (promotoras), y 3) alterando las regiones codificadoras y/o regiones reguladoras (promotoras) (por ejemplo, haciendo cambios de pares de bases). Se pueden obtener interrupciones específicas mediante mutación aleatoria seguido de chequeo o selección, o, en casos donde se conocen las secuencias génicas, se pueden obtener interrupciones específicas mediante intervención directa usando métodos de biología molecular conocidos por aquellos expertos en la técnica. Un método útil en particular es la supresión de cantidades significativas de regiones codificadoras y/o regiones reguladoras (promotoras).

Se conocen métodos de alterar la expresión de proteína recombinante por aquellos expertos en la técnica, y se tratan en parte en Baneyx, Curr. Opin. Biotechnol. (1999) 10: 411; Ross, et al., J. Bacteriol. (1998) 180: 5375;

deHasefh, et al., J. Bacteriol. (1998) 180: 3019; Smolke and Keasling, Biotechnol. Bioeng. (2002) 80: 762; Swartz, Curr. Opin. Biotech. (2009) 12: 195; y Ma, et al., J. Bacteriol. (2002) 184: 5733.

5 Se pueden construir bacterias recombinantes que contienen cambios en la expresión de genes para metabolizar sacarosa en la producción de productos microbianos incluyendo glicerol y derivados de glicerol, como se describe más arriba, usando técnicas bien conocidas en la técnica, algunas de las cuales están ejemplificadas en los Ejemplos en esta memoria.

10 Se puede lograr la construcción de bacterias recombinantes descritas en esta memoria usando una variedad de vectores y casetes de transformación y expresión adecuados para el clonaje, transformación y expresión de regiones codificadoras que confieren la capacidad de utilizar sacarosa en la producción de glicerol y sus derivados en un microorganismo adecuado. Vectores adecuados son aquellos que son compatibles con el que emplea la bacteria. Los vectores adecuados se pueden derivar, por ejemplo, de una bacteria, un virus (como el bacteriófago T7 o un fago derivado de M-13), un cósmido, una levadura o una planta. Se conocen los protocolos de obtención y utilización de tales vectores por aquellos expertos en la técnica (Sambrook et al., *supra*).

15 Son numerosas y familiares las regiones de control de iniciación, o promotores, que son útiles para conducir la expresión de las regiones codificadoras para la presente invención en la bacteria huésped deseada para aquellos expertos en la técnica. Virtualmente, cualquier promotor capaz de conducir la expresión es adecuado para su uso en esta memoria. Por ejemplo, se puede usar cualquiera de los promotores enumerados más arriba.

Las regiones de control de la terminación se pueden derivar de varios genes nativos a los huéspedes preferidos. Opcionalmente, puede ser innecesario un sitio de terminación; sin embargo, es lo más preferido si está incluido.

20 Para una expresión efectiva de los polipéptidos presentes, las secuencias de nucleótidos que codifican los polipéptidos están ligados operativamente a través de codones de iniciación a regiones control de la expresión seleccionadas de tal manera que la expresión da como resultado la formación del RNA mensajero apropiado.

25 Particularmente útiles son los vectores pSYCO101, pSYCO103, pSYCO106, y pSYCO109, descritos en la Patente de EE.UU. N° 7.371.558, y pSYCO400/AGRO, descrito en la Patente de EE.UU. N° 7.524.660. Los elementos esenciales de estos vectores se derivan del regulón *dha* aislado de *Klebsiella pneumoniae* y de *Saccharomyces cerevisiae*. Cada vector contiene los marcos abiertos de lectura *dhaB1*, *dhaB2*, *dhaB3*, *dhaX* (secuencia codificadora representada en SEQ ID NO: 49; secuencia del polipéptido codificado establecido en SEQ ID NO: 50), *orfX*, *DAR1*, y *GPP2* organizados en tres operones diferentes. Las secuencias de nucleótidos de pSYCO101, pSYCO103, pSYCO106, pSYCO109, y pSYCO400/AGRO están representadas en SEQ ID NO: 51, SEQ ID NO: 52, SEQ ID NO: 30 53, SEQ ID NO: 54, y SEQ ID NO: 55, respectivamente. Las diferencias entre los vectores están ilustradas en el siguiente cuadro [el prefijo "p-" indica un promotor; los marcos abiertos de lectura contenidos entre cada "(" representan la composición de un operón]:

pSYCO101 (SEQ ID NO: 51):

35 p-trc (Dar1_GPP2) en orientación opuesta comparado con los otros dos operones de vía,
p-1.6 largo GI (*dhaB1_dhaB2_dhaB3_dhaX*), y
p-1.6 largo GI (*orfY_orfX_orfW*)

pSYCO103 (SEQ ID NO: 52):

40 p-trc (Dar1_GPP2) en la misma orientación comparado con los otros dos operones de vía,
p-1.5 largo GI (*dhaB1_dhaB2_dhaB3_dhaX*), y
p-1.5 largo GI (*orfY_orfX_orfW*)

pSYCO106 (SEQ ID NO: 53):

p-trc (Dar1_GPP2) en la misma orientación comparado con los otros dos operones de vía,
p-1.6 largo GI (*dhaB1_dhaB2_dhaB3_dhaX*), y
p-1.6 largo GI (*orfY_orfX_orfW*)

45 pSYCO109 (SEQ ID NO: 54):

p-trc (Dar1_GPP2) en la misma orientación comparado con los otros dos operones de vía,
p-1.6 largo GI (*dhaB1_dhaB2_dhaB3_dhaX*), y
p-1.6 largo GI (*orfY_orfX*)

pSYCO400/AGRO (SEQ ID NO: 55):

p-tcr (Dar1_GPP2) en la misma orientación comparado con los otros dos operones de vía,

p-1.6 largo GI (dhaB1_dhaB2_dhaB3_dhaX), y

p-1.6 largo GI (orfY_orfX)

5 p-1.20 corto/largo GI (srcK) orientación opuesta comparado con los otros operones de vía.

Una vez que se construyen los casetes de expresión adecuados, se usan para transformar bacterias huésped apropiadas. Se puede llevar a cabo la introducción del casete que contiene las regiones codificadoras dentro de la bacteria huésped por procedimientos conocidos como mediante transformación (p.ej. usando células permeabilizadas con calcio, o electroporación) o mediante transfección usando un fago virus recombinante (Sambrook et al., *supra*). Los casetes de expresión se pueden mantener en un plásmido estable en una célula huésped. Además, los casetes de expresión se pueden integrar dentro del genoma de la bacteria huésped a través de recombinación homóloga o aleatoria usando vectores y métodos bien conocidos por aquellos expertos en la técnica. Se pueden usar también sistemas de recombinación específica de sitio para la integración genómica de casetes de expresión.

Además de las células ejemplificadas, se pueden usar las células que tienen mutaciones únicas o múltiples designadas específicamente para aumentar la producción de productos microbianos incluyendo glicerol y/o sus derivados. Las células que normalmente desvían una materia prima de carbono en vías no productivas, o que exhiben represión de catabolitos significativa se pueden mutar para evitar estas deficiencias fenotípicas.

Métodos de creación de mutantes son comunes y bien conocidos en la técnica. Se presenta un resumen de algunos métodos en la Patente de EE.UU. Nº 7.371.558. Están bien documentados en la técnica los métodos específicos de creación de mutantes usando radiación o agentes químicos. Véase, por ejemplo, Thomas D. Brock en *Biotechnology: A Textbook of Industrial Microbiology, Second Edition* (1989) Sinauer Associates, Inc., Sunderland, MA., o Deshpande, Mukund V., *Appl. Biochem. Biotechnol.* 36, 227 (1992).

Tras haber ocurrido la mutagénesis, los mutantes que tienen el fenotipo deseado se pueden seleccionar mediante una variedad de métodos. El cheque aleatorio es el más común donde las células mutagenizadas se seleccionan por la capacidad de producir el producto o intermedio deseado. Alternativamente, el aislamiento selectivo de mutantes se puede llevar a cabo mediante crecimiento de una población mutagenizada en medio selectivo donde solamente pueden desarrollarse las colonias resistentes. Los métodos de selección de mutantes están altamente desarrollados y bien conocidos en la técnica de la microbiología industrial. Véase, por ejemplo, Brock, *Supra*; DeMancilha et al., *Food Chem.* 14, 313 (1984).

Los medios de fermentación en la presente invención comprenden sacarosa como sustrato de carbono. También pueden estar presentes otros sustratos de carbono como glucosa y fructosa.

Además del sustrato de carbono, un medio adecuado de fermentación contiene, por ejemplo, minerales adecuados, sales, cofactores, tampones y otros componentes, conocidos por aquellos expertos en la técnica, adecuados para el crecimiento de los cultivos y la promoción de la vía enzimática necesaria para la producción de glicerol y sus derivados, por ejemplo 1,3-propanodiol. Se prestará especial atención a las sales de Co(II) y/o vitamina B₁₂ o precursores de las mismas en la producción de 1,3-propanodiol.

La adenosil-cobalamina (coenzima B₁₂) es un factor importante para la actividad deshidratasa. La síntesis de la coenzima B₁₂ se encuentra en procariotas, algunas de las cuales son capaces de sintetizar el compuesto *de novo*, por ejemplo, *Escherichia blattae*, especies de *Klebsiella*, especies de *Citrobacter*, y especies de *Clostridium*, mientras otras pueden realizar reacciones parciales. *E. Coli*, por ejemplo, no puede fabricar la estructura de anillo de corrina, pero es capaz de catalizar la conversión de cobinamida al corrinoide y puede introducir el grupo 5'-deoxiadenosilo. Así, se sabe en la técnica que se necesita proporcionar un precursor de la coenzima B₁₂, como la vitamina B₁₂, en las fermentaciones de *E. coli*. Se puede añadir continuamente vitamina B₁₂ a las fermentaciones de *E. coli* a una velocidad constante o en etapas como para coincidir con la generación de masa celular, o se puede añadir en adiciones únicas o múltiples de bolos.

Aunque la vitamina B₁₂ se añade a la *E. coli* transformada descrita más arriba en esta memoria, se contempla que otras bacterias, capaces de biosintetizar *de novo* vitamina B₁₂ también serán células productoras adecuadas y la adición de vitamina B₁₂ a estas bacterias será innecesaria.

Típicamente las células bacterianas se crecen de 25 a 40° C en el medio apropiado conteniendo sacarosa. Ejemplos de medios adecuados de crecimiento para su uso en esta memoria son medios comunes comercialmente preparados como caldo Luria Bertani (LB), caldo Sabouraud Dextrosa (SD) o caldo medio de levadura (YM). Se pueden también usar otros medios de crecimiento definidos o sintéticos, y el medio apropiado particular de crecimiento será conocido por alguien experto en la técnica de microbiología o ciencia de la fermentación. El uso de agentes conocidos para modular la represión de catabolitos directamente o indirectamente, p. ej., adenosina 2':3'-

monofosfato cíclica, se puede incorporar en los medios de reacción. De manera similar, el uso de agentes conocidos para modular las actividades enzimáticas (p. ej., metil viológeno) que conduce al aumento de la producción de 1,3-propanodiol se puede usar en conjunción con o como una alternativa a la manipulación genética con cepas productoras de 1,3-propanodiol.

- 5 Los intervalos de pH adecuados para la fermentación están entre pH 5,0 y pH 9,0, donde el pH 6,0 hasta pH 8,0 es típico como condición inicial.

Las reacciones se pueden llevar a cabo bajo condiciones aeróbicas, anóxicas, o anaeróbicas dependiendo de los requerimientos de la bacteria recombinante. Se pueden llevar a cabo fermentaciones de lotes alimentados con alimentación de carbono, por ejemplo, sustrato de carbono, limitado o en exceso.

- 10 La fermentación en lotes es un método comúnmente usado. La fermentación clásica en lote es un sistema cerrado donde la composición del medio se establece al principio de la fermentación y no está sujeta a alteraciones artificiales durante la fermentación. Así, al principio de la fermentación, se inocula el medio con la bacteria deseada y se permite que ocurra la fermentación no añadiendo nada al sistema. Típicamente, sin embargo, la fermentación “en lote” es lote con respecto a la adición de la fuente de carbono, se hacen intentos a menudo en controlar factores como el pH y la concentración de oxígeno. En sistemas en lotes, las composiciones de metabolitos y biomasa del sistema cambian constantemente hasta el tiempo que se para la fermentación. Dentro de los cultivos en lote, las células se moderan a través de una fase estática de latencia hasta una fase logarítmica de alto crecimiento y finalmente hasta una fase estacionaria donde la velocidad de crecimiento se disminuye o se detiene. Las células en fase logarítmica son generalmente responsables de la mayor producción del producto final o intermedio.
- 15
- 20 Una variación del sistema en lote estándar es el sistema Fed-Batch. Los procesos de fermentación Fed_batch son también adecuados para su uso en esta memoria y comprenden un sistema en lote típico con la excepción de que el sustrato se añade en incrementos según progresa la fermentación. Los sistemas Fed-Batch son útiles cuando la represión por catabolito es apta para inhibir el metabolismo de las células y donde es deseable tener cantidades limitadas de sustrato en el medio. Es difícil la medida de la concentración actual de sustrato en sistemas Fed-Batch y por consiguiente se estima en base a los cambios de factores medibles como el pH, oxígeno disuelto y la presión parcial de los gases de deshecho como el CO₂. Las fermentaciones en lote son comunes y bien conocidas en la técnica y se pueden encontrar ejemplos en Brock, *supra*.
- 25

- La fermentación continua es un sistema abierto donde se añade un medio de fermentación definido continuamente a un bioreactor y se elimina simultáneamente una cantidad igual de medio para el procesamiento. La fermentación continua mantiene generalmente los cultivos a una densidad constante alta donde las células están ante todo en fase logarítmica de crecimiento.
- 30

- La fermentación continua permite la modulación de un factor o cualquier número de factores que afectan al crecimiento celular o a la concentración del producto final. Por ejemplo, un método mantendrá un nutriente limitante como la fuente de carbono o el nivel de nitrógeno a una velocidad fija y permitirá moderar todos los otros parámetros. En otros sistemas, se pueden alterar continuamente un número de factores que afectan al crecimiento mientras la concentración celular, medida por la turbidez del medio, se mantiene constante. Los sistemas continuos se esfuerzan por mantener las condiciones de crecimiento en estado estacionario, y así la pérdida de células durante la extracción de medio se debe equilibrar con la velocidad de crecimiento celular en la fermentación. Se conocen bien en la técnica de la microbiología industrial los métodos de modulación de nutrientes y factores de crecimiento para procesos de fermentación continuos así como también técnicas para maximizar la velocidad de formación del producto y se detallan una variedad de métodos por Brock, *supra*.
- 35
- 40

- Se contempla que la presente invención se puede practicar usando procesos en lotes, en fed-batch o continuos y que cualquier modo conocido de fermentación sería adecuado. Además, se contempla que las células pueden ser inmovilizadas en un sustrato como catalizadores de células enteras y sometidas a condiciones de fermentación para la producción de glicerol y derivador de glicerol, como el 1,3-propanodiol.
- 45

- En una realización, se proporciona un proceso para hacer glicerol, 1,3-propanodiol, y/o ácido 3-hidroxipropiónico a partir de sacarosa. El proceso comprende las etapas de cultivar la bacteria, como se describe más arriba, en presencia de sacarosa, y recuperar opcionalmente el glicerol, 1,3-propanodiol, y/o ácido 3-hidroxipropiónico producidos. El producto se puede recuperar usando métodos conocidos en la técnica. Por ejemplo, los sólidos se pueden retirar del medio de fermentación mediante centrifugación, filtración, decantación, o en vivo. Luego, se puede aislar el producto del medio de fermentación, que ha sido tratado para retirar los sólidos como se describe arriba, usando métodos como la destilación, extracción líquido-líquido, o separación basada en membrana.
- 50

Ejemplos

- La presente invención se define además en los siguientes Ejemplos. Se ha de entender que estos Ejemplos, aunque indican realizaciones preferidas de la invención, se dan a modo solamente de ilustración. De la discusión anterior y estos Ejemplos, un experto en la técnica puede determinar las características esenciales en esta invención, y sin apartarse del espíritu y alcance de la misma, puede hacer diversos cambios y modificaciones para adaptarlas a diversos usos y condiciones.
- 55

Métodos Generales

Son bien conocidas en la técnica las técnicas estándar de DNA recombinante y de clonaje molecular descritas en los Ejemplos y se describen por Sambrook, J., Fritsch, E.F. and Maniatis, T. *Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, (1989) (Maniatis) y por T. J. Silhavy, M. L. Bannan, y L.W. Enquist, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. (1984) y por Ausubel, F.M. et al., *Current Protocols in Molecular Biology*, pub. By Greene Publishing Assoc. and Wiley-Interscience (1987).

El significado de las abreviaturas es como sigue: "sec" significa segundo(s), "min" significa minuto(s), "h" significa hora(s), "nm" significa nanómetros, "μL" significa microlitro(s), "mL" significa mililitro(s), "L" significa litro(s), "mM" significa milimolar, "M" significa molar, "g" significa gramo(s), "μg" significa microgramo(s), "bp" significa par(es) de base(s), "kbp" significa kilopar(es) de base(s), "rpm" significa revoluciones por minuto, "ATCC" significa American Type Culture Collection, Manassas, VA.

Medios y condiciones de cultivo

Son bien conocidos en la técnica los materiales métodos adecuados para el mantenimiento y crecimiento de los cultivos celulares. Las técnicas adecuadas para su uso en los siguientes Ejemplos se pueden encontrar como la representada en *Manual of Methods for General Bacteriology* (Phillipp Gerhardt, R.G.E. Murray, Ralph N. Costillow, Eugene W. Nester, Willis A. Wood, Noel R. Krieg and G. Briggs Philips, eds.), American Society for Microbiology, Washington, DC. (1994)) o por Thomas D. Brock en *Biotechnology: A Textbook of Industrial Microbiology*, Second Edition, Sinauer Associates, Inc., Sunderland, MA (1989). Todos los reactivos, enzimas de restricción y materiales descritos para el cultivo y el mantenimiento de las células bacterianas se pueden obtener de Aldrich Chemicals (Milwaukee, WI), BD Diagnostics Systems (Sparks, MD), Life Technologies (Rockville, MD), New England Biolabs (Beverly, MA), o Sigma Chemical Company (St. Louis, MO), a menos que se especifique lo contrario.

El medio LB (Luria Bertani) contiene los siguientes por litro de medio: Bacto-triptona (10 g), extracto Bacto-levadura (5 g), y NaCl (10 g). Se añadieron los suplementos como se describe en los Ejemplos siguientes. Todas las adiciones fueron pre-esterilizadas antes de que fueran añadidas al medio.

Técnicas de Biología Molecular:

Las digestiones con enzimas de restricción, ligaciones, transformaciones, y métodos para electroforesis en geles de agarosa se realizaron como se describe en Sambrook, J., et al., *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press (1989). Las técnicas de las reacciones en cadena de la polimerasa (PCR) fueron encontrados en White, B., *PCR Protocols: Current Methods and Applications*, Volume 15 (1993), Humana Press Inc., New York, NY.

Ejemplo 1

E. coli Recombinante que Contiene Supuestos Genes del Metabolismo de Sacarosa

El propósito de este Ejemplo fue el de construir una cepa de *E. coli* recombinante que contenía supuestos genes del metabolismo de sacarosa de *Pseudomonas fluorescens* Pf5 (ATCC® BAA-477), y que tenía la capacidad de producir 1,3-propanodiol (PDO) a partir de sacarosa.

Los genes de *P. fluorescens* Pf5 *lacY* (SEQ ID NO: 23) y *scrB* (SEQ ID NO: 25), que son genes adyacentes, y la región promotora corriente arriba del gen de Pf5 *lacY* se obtuvieron mediante amplificación por PCR usando una polimerasa de alta fidelidad, Phusion™ Flash (Finnzymes Oy, Espoo, Finland) usando las siguientes condiciones: 98° C/10 sec; 30 ciclos de 98° C/1 sec, 63° C/30 sec, y 72° C/30 sec; y luego 72° C/5 min usando DNA genómico de *P. Fluorescens* Pf5 como molde y los cebadores establecidos en SEQ ID NO: 56 y SEQ ID NO: 57.

Se añadió un nucleótido de adenosina ("A") a los extremos 3' del producto de PCR de 3.063 bp con Taq polimerasa durante 10 min a 72° C. El producto de PCR se clonó en pBAD-TOPO® (Invitrogen, Carlsbad, CA) siguiendo el protocolo del comerciante. Tras la transformación de la cepa de *E. coli* TOP10 (Invitrogen) con selección de resistencia a ampicilina, se aisló el plásmido de DNA. La presencia del DNA insertado se verificó mediante análisis de secuencia de DNA. Un plásmido, llamado "pBAD-TOPO Pf5 cscB&A 3-5", se transformó en una cepa de *E. coli* TTab pSYCO400/AGRO para generar la cepa PDO2241. Un plásmido de control sin inserto de DNA, llamado "pBAD-TOPO ctl 3-8", se transformó también en la cepa de *E. coli* TTab pSYCO400/AGRO para generar la cepa PDO2242.

La cepa de *E. coli* TTab pSYCO400/AGRO, una cepa PTS menos, se construyó como sigue. La cepa TTab se generó por supresión del gen *aldB* de la cepa TT *aldA*, descrita en la Patente de EE.UU. N° 7.371.558 (Ejemplo 17). Brevemente, se hizo una supresión de *aldB* mediante reemplazamiento en primer lugar de la región codificadora de 1,5 kbp de *aldB* en la cepa de *E. coli* MG1655 con el casete FRT-CmR-FRT del plásmido pKD3 (Datsenko and Wanner, Proc. Natl. Acad. Sci. USA 97: 6640-6645, 2000). Se amplificó un casete de reemplazamiento con la pareja de cebadores SEQ ID NO: 58 y SEQ ID NO: 59 usando pKD3 como molde. El cebador SEQ ID NO: 58 contiene 80

bp de homología al extremo 5' terminal de *aldB* y 20 bp de homología con pKD3. El cebador SEQ ID NO: 59 contiene 80 bp de homología al extremo 3' terminal de *aldB* y 20 bp de homología con pKD3. Los productos de PCR se purificaron por gel y se electroporaron en células competentes MG1655/pKD46 (Patente de EE.UU. N° 7.371.558). Las cepas recombinantes se seleccionaron en placas de LB con 12,5 mg/L de cloranfenicol. La supresión del gen *aldB* se confirmó mediante PCR, usando la pareja de cebadores SEQ ID NO: 60 y SEQ ID NO: 61. La cepa salvaje dio un producto de PCR de 1,5 kbp mientras la bacteria recombinante dio un producto de PCR característico de 1,1 kbp. Se preparó un lisado de P1 y se usó para mover la mutación a la cepa TT *aldA* para formar la cepa TT *aldAΔaldB::Cm*. Se chequeó un clon resistente a cloranfenicol mediante PCR genómica con la pareja de cebadores SEQ ID NO: 60 y SEQ ID NO: 61 para asegurarse de que la mutación estaba presente. El marcador de resistencia a cloranfenicol se eliminó usando la recombinasa FLP (Datsenko and Wanner, *supra*) para crear TTab. La cepa TTab se transformó entonces con pSYCO400/AGRO (representada en SEQ ID NO: 55), descrita en la Patente de EE.UU. N°. 7.524.660 (Ejemplo 4), para generar la cepa TTab pSYCO400/AGRO.

Como se describe en las referencias citadas, la cepa TTab es un derivado de la cepa de *E. coli* FM5 (ATCC® N°. 53911) que contiene las siguientes modificaciones:

supresión de los genes *glpK*, *gldA*, *ptsHI*, *crr*, *edd*, *arcA*, *mgsA*, *qor*, *ackA*, *pta*, *aldA* y *aldB*

regulación al alza de los genes *galP*, *glk*, *byuR*, *ppc*, y *yqhD*; y

regulación a la baja del gen *gapA*.

El plásmido pSYCO400/AGRO contiene genes que codifican una vía de producción de glicerol (*DAR1* y *GPP2*) y genes que codifican una glicerol deshidratasa y un factor de reactivación asociado (*dhaB123*, *dhaX*, *orfX*, *orfY*), así como también un gen que codifica la fructoquinasa (*scrK*).

Ejemplo 2

Producción de 1,3-Propanodiol a partir de Sacarosa

El propósito de este Ejemplo fue el de demostrar que una cepa de *E. coli* que comprende los supuestos genes del metabolismo de sacarosa de *Pseudomonas fluorescens* Pf5 (ATCC® BAA-477) era capaz de metabolizar sacarosa y producir 1,3-propanodiol (PDO) y glicerol.

Las cepas de *E. coli* PDO2241 y PDO2242, descritas en el Ejemplo 1, se crecieron toda la noche en L-Broth, Miller's Modification (Teknova, Half Moon Bay, CA) suplementado con 100 mg/l de espectinomicina y 100 mg/L de ampicilina a 33°C. Estas cultivos se usaron para inocular matraces de agitación a una densidad óptica de 0,01 unidades medida a 550 nm en medio mínimo MOPS (Teknova, Half Moon Bay, CA) suplementado con 11 g/L de glucosa. Se añadió vitamina B12 al medio a una concentración de 0,1 mg/L. Los cultivos se incubaron a 34° C con agitación (225 rpm) durante 24 horas, después de tal tiempo se determinaron las concentraciones de sacarosa, glicerol y 1,3-propanodiol (PDO) en el caldo mediante cromatografía líquida de alto rendimiento.

La separación cromatográfica se llevó a cabo usando una columna Aminex HPX-87P (Bio-Rad, Hercules, CA) con una fase móvil isocrática de agua destilada-desionizada a un flujo de 0,5 mL/min y una temperatura de columna de 85° C. Los compuestos eluidos se cuantificaron mediante detección por índice de refracción con una curva estándar preparada con compuestos puros adquiridos comercialmente disueltos a concentraciones conocidas en medio mínimo MOPS. Los tiempos de retención fueron sacarosa a 12,2 min, 1,3-propanodiol a 17,9 min, glicerol a 23,6 min.

La cepa de *E. coli* PDO2241 con los supuestos genes del metabolismo de sacarosa de *P. fluorescens* Pf5 metabolizó toda la sacarosa in el medio (0 g/L residual) y produjo 2,55 g/L de PDO y 2,83 g/L de glicerol. La cepa control PDO2242 fue incapaz de metabolizar sacarosa (10,7 g/L residual) y produjo PDO sólo insignificante (0,06 g/L) y glicerol (0,01 g/L) bajo estas condiciones.

Ejemplo 3

E. coli Recombinante que Contiene los Supuestos Genes del Metabolismo de Sacarosa

El propósito de este Ejemplo fue el de construir una cepa recombinante de *E. coli* que contuviera los genes del metabolismo de sacarosa clonados de *P. fluorescens* Pf5.

El plásmido pBAD-TOPO Pf5 *cscB&A* 3-5, descrito en el Ejemplo 1 se transformó en la cepa de *E. coli* FM5 (ATCC® 53911) para generar la cepa PDO2355. También se transformó un plásmido control sin inserto de DNA, pBAD-TOPO *ctl* 3-8 en la cepa FM5 de *E. coli* para generar la cepa PDO2350.

Las cepas de *E. coli* PDO2355 y PDO2350 se crecieron durante toda la noche en medio LB (Luria Bertani) que contenía 100 µg/mL de ampicilina a 37° C. El siguiente día, se diluyeron estos cultivos a 1:50 en medio LB que contenía ampicilina a 100 µg/mL y se crecieron a 37° C durante 4 horas. Estos cultivos en fase logarítmica se diluyeron 1:100 en pocillos de placas Bioscreen-C (instrumento y placas adquiridas de Growth Curves USA,

- 5 Piscataway, CA) con medio mínimo MOPS (Teknova, Half Moon Bay, CA) conteniendo sacarosa a 10 g/L. Los cultivos se crecieron a 37° C por triplicado con agitación continua. Se monitorizó la densidad óptica. A 20 horas tras la inoculación, la densidad óptica de la PDO2355 era de 0,584+/-0,004, mientras la densidad óptica de la PDO2350 era de 0,065 +/- 0,004. Estos resultados demuestran que la cepa control es incapaz de crecer con sacarosa como única fuente de carbono, mientras que la cepa que expresa *P. fluorescens* Pf5 *lacY* y *scrB* era capaz de crecer con sacarosa como única fuente de carbono.

Listado de Secuencias

- <110> E.I. du Pont de Nemours and Co.
Van Dyk, Tina
- 10 <120> Bacterias recombinantes que poseen la capacidad de metabolizar sacarosa
<130> CL5285 USNA
<160> 61

<170> PatentIn version 3.5
- 15 <210> 1
<211> 1176
<212> DNA
<213> *Saccharomyces cerevisiae*
- <400> 1
atgtctgctg ctgctgatag attaaactta acttccggcc acttgaatgc tggtagaaag 60
agaagttcct cttctgtttc tttgaaggct gccgaaaagc ctttcaaggt tactgtgatt 120
ggatctggta actggggtac tactattgcc aagggtggtg ccgaaaattg taagggatac 180
ccagaagttt tgcctccaat agtacaaatg tgggtgttcg aagaagagat caatggtgaa 240
aaattgactg aaatcataaa tactagacat caaaacgtga aataactgcc tggcatcact 300
ctacccgaca atttggttgc taatccagac ttgattgatt cagtcaagga tgtcgacatc 360
atcgttttca acattccaca tcaatttttg ccccgatatc gtagccaatt gaaaggatc 420
gttgattcac acgtcagagc tatctcctgt cttaaagggtt ttgaagttgg tgctaaagg 480
gtccaattgc taccctctta catcactgag gaactaggta ttcaatgtgg tgctctatct 540
ggtgctaaca ttgccaccga agtcgctcaa gaacactggt ctgaaacaac agttgcttac 600
cacattccaa aggatttcag aggcgagggc aaggacgtcg accataaggt tctaaaggcc 660
ttgttccaca gacctactt ccacgttagt gtcacgaag atgttgctgg tatctccatc 720
tgtggtgctt tgaagaacgt tgttgcttga ggttggtggt tcgtcgaagg tctaggctgg 780
ggtaacaacg cttctgctgc catccaaaga gtcggtttgg gtgagatcat cagattcgg 840
caaagtgttt tccagaatc tagagaagaa acatactacc aagagtctgc tgggtgttgc 900
gatttgatca ccacctgccc tgggtggtaga aacgtcaagg ttgctaggct aatggctact 960
tctggttaagg acgcctggga atgtgaaaag gagttgttga atggccaatc cgctcaaggt 1020
ttaattacct gcaaagaagt tcacgaatgg ttggaaacat gtggctctgt cgaagacttc 1080
ccattatttg aagccgtata ccaaatcgtt tacaacaact acccaatgaa gaacctgccg 1140
gacatgattg aagaattaga tctacatgaa gattag 1176
- 20 <210> 2

ES 2 552 317 T3

<211> 391
<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 2

Met Ser Ala Ala Ala Asp Arg Leu Asn Leu Thr Ser Gly His Leu Asn
1 5 10 15

Ala Gly Arg Lys Arg Ser Ser Ser Ser Val Ser Leu Lys Ala Ala Glu
20 25 30

Lys Pro Phe Lys Val Thr Val Ile Gly Ser Gly Asn Trp Gly Thr Thr
35 40 45

Ile Ala Lys Val Val Ala Glu Asn Cys Lys Gly Tyr Pro Glu Val Phe
50 55 60

Ala Pro Ile Val Gln Met Trp Val Phe Glu Glu Glu Ile Asn Gly Glu
65 70 75 80

Lys Leu Thr Glu Ile Ile Asn Thr Arg His Gln Asn Val Lys Tyr Leu
85 90 95

Pro Gly Ile Thr Leu Pro Asp Asn Leu Val Ala Asn Pro Asp Leu Ile
100 105 110

Asp Ser Val Lys Asp Val Asp Ile Ile Val Phe Asn Ile Pro His Gln
115 120 125

Phe Leu Pro Arg Ile Cys Ser Gln Leu Lys Gly His Val Asp Ser His
130 135 140

Val Arg Ala Ile Ser Cys Leu Lys Gly Phe Glu Val Gly Ala Lys Gly
145 150 155 160

Val Gln Leu Leu Ser Ser Tyr Ile Thr Glu Glu Leu Gly Ile Gln Cys
165 170 175

Gly Ala Leu Ser Gly Ala Asn Ile Ala Thr Glu Val Ala Gln Glu His
180 185 190

Trp Ser Glu Thr Thr Val Ala Tyr His Ile Pro Lys Asp Phe Arg Gly
195 200 205

Glu Gly Lys Asp Val Asp His Lys Val Leu Lys Ala Leu Phe His Arg
210 215 220

Pro Tyr Phe His Val Ser Val Ile Glu Asp Val Ala Gly Ile Ser Ile
225 230 235 240

Cys Gly Ala Leu Lys Asn Val Val Ala Leu Gly Cys Gly Phe Val Glu
245 250 255

Gly Leu Gly Trp Gly Asn Asn Ala Ser Ala Ala Ile Gln Arg Val Gly
260 265 270

Leu Gly Glu Ile Ile Arg Phe Gly Gln Met Phe Phe Pro Glu Ser Arg
275 280 285

Glu Glu Thr Tyr Tyr Gln Glu Ser Ala Gly Val Ala Asp Leu Ile Thr
290 295 300

Thr Cys Ala Gly Gly Arg Asn Val Lys Val Ala Arg Leu Met Ala Thr
305 310 315 320

Ser Gly Lys Asp Ala Trp Glu Cys Glu Lys Glu Leu Leu Asn Gly Gln
325 330 335

Ser Ala Gln Gly Leu Ile Thr Cys Lys Glu Val His Glu Trp Leu Glu
340 345 350

Thr Cys Gly Ser Val Glu Asp Phe Pro Leu Phe Glu Ala Val Tyr Gln
355 360 365

Ile Val Tyr Asn Asn Tyr Pro Met Lys Asn Leu Pro Asp Met Ile Glu
370 375 380

Glu Leu Asp Leu His Glu Asp
385 390

<210> 3

<211> 1323

<212> DNA

5 <213> Saccharomyces cerevisiae

<400> 3

atgcttgctg tcagaagatt aacaagatac acatttcctta agcgaacgca tccggtgtta	60
tatactcgtc gtgcatataa aattttgcct tcaagatcta ctttcctaag aagatcatta	120
ttacaaacac aactgcactc aaagatgact gctcatacta atatcaaaca gcacaaacac	180
tgtcatgagg accatcctat cagaagatcg gactctgccg tgtcaattgt acatttgaaa	240
cgtgcgccct tcaaggttac agtgattggt tctggtaact gggggaccac catcgccaaa	300
gtcattgcgg aaaacacaga attgcattcc catatcttcg agccagaggt gagaatgtgg	360
gtttttgatg aaaagatcgg cgacgaaaat ctgacggata tcataaatac aagacaccag	420
aacgttaaat atctacccaa tattgacctg ccccataatc tagtggccga tcctgatctt	480

ttacactcca tcaaggggtgc tgacatcctt gttttcaaca tccctcatca atttttacca 540
 aacatagtca aacaattgca aggccacgtg gcccctcatg taagggccat ctctgtctca 600
 aaagggttcg agttgggctc caaggggtgtg caattgctat cctcctatgt tactgatgag 660
 ttaggaatcc aatgtggcgc actatctggt gcaaacttgg caccggaagt ggccaaggag 720
 cattggtccg aaaccaccgt ggcttaccaa ctaccaaagg attatcaagg tgatggcaag 780
 gatgtagatc ataagatfff gaaattgctg ttccacagac cttacttcca cgtcaatgtc 840
 atcgatgatg ttgctggtat atccattgcc ggtgccttga agaacgtcgt ggcacttgca 900
 tgtgggtttcg tagaagggtat gggatggggg aacaatgcct ccgcagccat tcaaaggctg 960
 ggttttaggtg aaattatcaa gttcggtaga atgtttttcc cagaatccaa agtcgagacc 1020
 tactatcaag aatccgctgg tgttgcagat ctgatcacca cctgctcagg cggtagaaac 1080
 gtcaagggtt ccacatacat ggccaagacc ggtaagtcag ccttggaagc agaaaaggaa 1140
 ttgcttaacg gtcaatccgc ccaagggata atcacatgca gagaagttca cgagtggcta 1200
 caaacatgtg agttgaccca agaattccca ttattcgagg cagtctacca gatagtctac 1260
 aacaacgtcc gcatggaaga cctaccggag atgattgaag agctagacat cgatgacgaa 1320
 tag 1323

<210> 4

<211> 440

<212> PRT

5 <213> Saccharomyces cerevisiae

<400> 4

Met Leu Ala Val Arg Arg Leu Thr Arg Tyr Thr Phe Leu Lys Arg Thr
 1 5 10 15

His Pro Val Leu Tyr Thr Arg Arg Ala Tyr Lys Ile Leu Pro Ser Arg
 20 25 30

Ser Thr Phe Leu Arg Arg Ser Leu Leu Gln Thr Gln Leu His Ser Lys
 35 40 45

Met Thr Ala His Thr Asn Ile Lys Gln His Lys His Cys His Glu Asp
 50 55 60

His Pro Ile Arg Arg Ser Asp Ser Ala Val Ser Ile Val His Leu Lys
 65 70 75 80

Arg Ala Pro Phe Lys Val Thr Val Ile Gly Ser Gly Asn Trp Gly Thr
 85 90 95

Thr Ile Ala Lys Val Ile Ala Glu Asn Thr Glu Leu His Ser His Ile

ES 2 552 317 T3

100	105	110
Phe Glu Pro Glu Val Arg Met Trp Val Phe Asp Glu Lys Ile Gly Asp 115 120 125		
Glu Asn Leu Thr Asp Ile Ile Asn Thr Arg His Gln Asn Val Lys Tyr 130 135 140		
Leu Pro Asn Ile Asp Leu Pro His Asn Leu Val Ala Asp Pro Asp Leu 145 150 155 160		
Leu His Ser Ile Lys Gly Ala Asp Ile Leu Val Phe Asn Ile Pro His 165 170 175		
Gln Phe Leu Pro Asn Ile Val Lys Gln Leu Gln Gly His Val Ala Pro 180 185 190		
His Val Arg Ala Ile Ser Cys Leu Lys Gly Phe Glu Leu Gly Ser Lys 195 200 205		
Gly Val Gln Leu Leu Ser Ser Tyr Val Thr Asp Glu Leu Gly Ile Gln 210 215 220		
Cys Gly Ala Leu Ser Gly Ala Asn Leu Ala Pro Glu Val Ala Lys Glu 225 230 235 240		
His Trp Ser Glu Thr Thr Val Ala Tyr Gln Leu Pro Lys Asp Tyr Gln 245 250 255		
Gly Asp Gly Lys Asp Val Asp His Lys Ile Leu Lys Leu Leu Phe His 260 265 270		
Arg Pro Tyr Phe His Val Asn Val Ile Asp Asp Val Ala Gly Ile Ser 275 280 285		
Ile Ala Gly Ala Leu Lys Asn Val Val Ala Leu Ala Cys Gly Phe Val 290 295 300		
Glu Gly Met Gly Trp Gly Asn Asn Ala Ser Ala Ala Ile Gln Arg Leu 305 310 315 320		
Gly Leu Gly Glu Ile Ile Lys Phe Gly Arg Met Phe Phe Pro Glu Ser 325 330 335		
Lys Val Glu Thr Tyr Tyr Gln Glu Ser Ala Gly Val Ala Asp Leu Ile 340 345 350		

Thr Thr Cys Ser Gly Gly Arg Asn Val Lys Val Ala Thr Tyr Met Ala
355 360 365

Lys Thr Gly Lys Ser Ala Leu Glu Ala Glu Lys Glu Leu Leu Asn Gly
370 375 380

Gln Ser Ala Gln Gly Ile Ile Thr Cys Arg Glu Val His Glu Trp Leu
385 390 395 400

Gln Thr Cys Glu Leu Thr Gln Glu Phe Pro Leu Phe Glu Ala Val Tyr
405 410 415

Gln Ile Val Tyr Asn Asn Val Arg Met Glu Asp Leu Pro Glu Met Ile
420 425 430

Glu Glu Leu Asp Ile Asp Asp Glu
435 440

<210> 5

<211> 816

<212> DNA

5 <213> *Saccharomyces cerevisiae*

<400> 5

atgaaacggt tcaatgtttt aaaatatatc agaacaacaa aagcaaatat acaaaccatc	60
gcaatgcctt tgaccacaaa acctttatct ttgaaaatca acgccgctct attcgatggt	120
gacggtacca tcatcatctc tcaaccagcc attgctgctt tctggagaga tttcggtaaa	180
gacaagcctt acttcgatgc cgaacacggt attcacatct ctcacggttg gagaacttac	240
gatgccattg ccaagttcgc tccagacttt gctgatgaag aatacgttaa caagctagaa	300
ggtgaaatcc cagaaaagta cggatgaacac tccatcgaag ttccagggtg tgtcaagttg	360
tgtaatgctt tgaacgcctt gccaaaggaa aaatgggctg tcgccacctc tggtaccctg	420
gacatggcca agaaatgggt cgacattttg aagatcaaga gaccagaata cttcatcacc	480
gccaatgatg tcaagcaagg taagcctcac ccagaacatc acttaaaggg tagaaacggt	540
ttgggtttcc caattaatga acaagaccca tccaaatcta aggttggtgt ctttgaagac	600
gcaccagctg gtattgctgc tggtaaggct gctggctgta aaatcggttg tattgctacc	660
actttcgatt tggacttctt gaaggaaaag ggttggtgaca tcattgtcaa gaaccacgaa	720
tctatcagag tcggtgaata caacgctgaa accgatgaag tcgaattgat ctttcatgac	780
tacttatacg ctaaggatga cttggtgaaa tggtaa	816

<210> 6

<211> 271

<212> PRT

10 <213> *Saccharomyces cerevisiae*

<400> 6

ES 2 552 317 T3

Met	Lys	Arg	Phe	Asn	Val	Leu	Lys	Tyr	Ile	Arg	Thr	Thr	Lys	Ala	Asn	1	5	10	15
Ile	Gln	Thr	Ile	Ala	Met	Pro	Leu	Thr	Thr	Lys	Pro	Leu	Ser	Leu	Lys	20	25	30	
Ile	Asn	Ala	Ala	Leu	Phe	Asp	Val	Asp	Gly	Thr	Ile	Ile	Ile	Ser	Gln	35	40	45	
Pro	Ala	Ile	Ala	Ala	Phe	Trp	Arg	Asp	Phe	Gly	Lys	Asp	Lys	Pro	Tyr	50	55	60	
Phe	Asp	Ala	Glu	His	Val	Ile	His	Ile	Ser	His	Gly	Trp	Arg	Thr	Tyr	65	70	75	80
Asp	Ala	Ile	Ala	Lys	Phe	Ala	Pro	Asp	Phe	Ala	Asp	Glu	Glu	Tyr	Val	85	90	95	
Asn	Lys	Leu	Glu	Gly	Glu	Ile	Pro	Glu	Lys	Tyr	Gly	Glu	His	Ser	Ile	100	105	110	
Glu	Val	Pro	Gly	Ala	Val	Lys	Leu	Cys	Asn	Ala	Leu	Asn	Ala	Leu	Pro	115	120	125	
Lys	Glu	Lys	Trp	Ala	Val	Ala	Thr	Ser	Gly	Thr	Arg	Asp	Met	Ala	Lys	130	135	140	
Lys	Trp	Phe	Asp	Ile	Leu	Lys	Ile	Lys	Arg	Pro	Glu	Tyr	Phe	Ile	Thr	145	150	155	160
Ala	Asn	Asp	Val	Lys	Gln	Gly	Lys	Pro	His	Pro	Glu	Pro	Tyr	Leu	Lys	165	170	175	
Gly	Arg	Asn	Gly	Leu	Gly	Phe	Pro	Ile	Asn	Glu	Gln	Asp	Pro	Ser	Lys	180	185	190	
Ser	Lys	Val	Val	Val	Phe	Glu	Asp	Ala	Pro	Ala	Gly	Ile	Ala	Ala	Gly	195	200	205	
Lys	Ala	Ala	Gly	Cys	Lys	Ile	Val	Gly	Ile	Ala	Thr	Thr	Phe	Asp	Leu	210	215	220	
Asp	Phe	Leu	Lys	Glu	Lys	Gly	Cys	Asp	Ile	Ile	Val	Lys	Asn	His	Glu	225	230	235	240

Ser Ile Arg Val Gly Glu Tyr Asn Ala Glu Thr Asp Glu Val Glu Leu
245 250 255

Ile Phe Asp Asp Tyr Leu Tyr Ala Lys Asp Asp Leu Leu Lys Trp
260 265 270

<210> 7
<211> 753
<212> DNA
5 <213> *Saccharomyces cerevisiae*

<400> 7
atgggattga ctactaaacc tctatctttg aaagttaacg ccgctttggt cgacgtcgac 60
ggtaccatta tcatctctca accagccatt gctgcattct ggagggattt cggttaaggac 120
aaaccttatt tcgatgctga acacgttatc caagtctcgc atggttggag aacgtttgat 180
gccattgcta agttcgtctc agactttgcc aatgaagagt atgttaacaa attagaagct 240
gaaattccgg tcaagtacgg tgaaaaatcc attgaagtcc caggtgcagt taagctgtgc 300
aacgctttga acgctctacc aaaagagaaa tgggctgtgg caacttccgg taccctgat 360
atggcacaaa aatgggtcga gcctctggga atcaggagac caaagtactt cattaccgct 420
aatgatgtca aacagggtaa gcctcatcca gaaccatata tgaagggcag gaatggctta 480
ggatatccga tcaatgagca agacccttcc aaatctaagg tagtagtatt tgaagacgct 540
ccagcaggta ttgccgccgg aaaagccgcc ggttgtaaga tcattggtat tgccactact 600
ttcgacttgg acttcctaaa ggaaaaaggc tgtgacatca ttgtcaaaaa ccacgaatcc 660
atcagagttg gcggctacaa tgccgaaaca gacgaagttg aattcatttt tgacgactac 720
ttatatgcta aggacgatct gttgaaatgg taa 753

<210> 8
<211> 250
10 <212> PRT
<213> *Saccharomyces cerevisiae*

<400> 8
Met Gly Leu Thr Thr Lys Pro Leu Ser Leu Lys Val Asn Ala Ala Leu
1 5 10 15
Phe Asp Val Asp Gly Thr Ile Ile Ile Ser Gln Pro Ala Ile Ala Ala
20 25 30
Phe Trp Arg Asp Phe Gly Lys Asp Lys Pro Tyr Phe Asp Ala Glu His
35 40 45
Val Ile Gln Val Ser His Gly Trp Arg Thr Phe Asp Ala Ile Ala Lys
50 55 60

Phe Ala Pro Asp Phe Ala Asn Glu Glu Tyr Val Asn Lys Leu Glu Ala
65 70 75 80

Glu Ile Pro Val Lys Tyr Gly Glu Lys Ser Ile Glu Val Pro Gly Ala
85 90 95

Val Lys Leu Cys Asn Ala Leu Asn Ala Leu Pro Lys Glu Lys Trp Ala
100 105 110

Val Ala Thr Ser Gly Thr Arg Asp Met Ala Gln Lys Trp Phe Glu His
115 120 125

Leu Gly Ile Arg Arg Pro Lys Tyr Phe Ile Thr Ala Asn Asp Val Lys
130 135 140

Gln Gly Lys Pro His Pro Glu Pro Tyr Leu Lys Gly Arg Asn Gly Leu
145 150 155 160

Gly Tyr Pro Ile Asn Glu Gln Asp Pro Ser Lys Ser Lys Val Val Val
165 170 175

Phe Glu Asp Ala Pro Ala Gly Ile Ala Ala Gly Lys Ala Ala Gly Cys
180 185 190

Lys Ile Ile Gly Ile Ala Thr Thr Phe Asp Leu Asp Phe Leu Lys Glu
195 200 205

Lys Gly Cys Asp Ile Ile Val Lys Asn His Glu Ser Ile Arg Val Gly
210 215 220

Gly Tyr Asn Ala Glu Thr Asp Glu Val Glu Phe Ile Phe Asp Asp Tyr
225 230 235 240

Leu Tyr Ala Lys Asp Asp Leu Leu Lys Trp
245 250

<210> 9

<211> 1668

<212> DNA

5 <213> Klebsiella pneumoniae

<220>

<221> CDS

<222> (1)..(1668)

<400> 9

atg aaa aga tca aaa cga ttt gca gta ctg gcc cag cgc ccc gtc aat

48

Met Lys Arg Ser Lys Arg Phe Ala Val Leu Ala Gln Arg Pro Val Asn

1

5

10

15

10

cag gac ggg ctg att ggc gag tgg cct gaa gag ggg ctg atc gcc atg Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met 20 25 30	96
gac agc ccc ttt gac ccg gtc tct tca gta aaa gtg gac aac ggt ctg Asp Ser Pro Phe Asp Pro Val Ser Ser Val Lys Val Asp Asn Gly Leu 35 40 45	144
atc gtc gaa ctg gac ggc aaa cgc cgg gac cag ttt gac atg atc gac Ile Val Glu Leu Asp Gly Lys Arg Arg Asp Gln Phe Asp Met Ile Asp 50 55 60	192
cga ttt atc gcc gat tac gcg atc aac gtt gag cgc aca gag cag gca Arg Phe Ile Ala Asp Tyr Ala Ile Asn Val Glu Arg Thr Glu Gln Ala 65 70 75 80	240
atg cgc ctg gag gcg gtg gaa ata gcc cgt atg ctg gtg gat att cac Met Arg Leu Glu Ala Val Glu Ile Ala Arg Met Leu Val Asp Ile His 85 90 95	288
gtc agc cgg gag gag atc att gcc atc act acc gcc atc acg ccg gcc Val Ser Arg Glu Glu Ile Ile Ala Ile Thr Thr Ala Ile Thr Pro Ala 100 105 110	336
aaa gcg gtc gag gtg atg gcg cag atg aac gtg gtg gag atg atg atg Lys Ala Val Glu Val Met Ala Gln Met Asn Val Val Glu Met Met Met 115 120 125	384
gcg ctg cag aag atg cgt gcc cgc cgg acc ccc tcc aac cag tgc cac Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His 130 135 140	432
gtc acc aat ctc aaa gat aat ccg gtg cag att gcc gct gac gcc gcc Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala 145 150 155 160	480
gag gcc ggg atc cgc ggc ttc tca gaa cag gag acc acg gtc ggt atc Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile 165 170 175	528
gcg cgc tac gcg ccg ttt aac gcc ctg gcg ctg ttg gtc ggt tcg cag Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Val Gly Ser Gln 180 185 190	576
tgc ggc cgc ccc ggc gtg ttg acg cag tgc tcg gtg gaa gag gcc acc Cys Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr 195 200 205	624
gag ctg gag ctg ggc atg cgt ggc tta acc agc tac gcc gag acg gtg Glu Leu Glu Leu Gly Met Arg Gly Leu Thr Ser Tyr Ala Glu Thr Val 210 215 220	672
tcg gtc tac ggc acc gaa gcg gta ttt acc gac ggc gat gat acg ccg Ser Val Tyr Gly Thr Glu Ala Val Phe Thr Asp Gly Asp Asp Thr Pro 225 230 235 240	720
tgg tca aag gcg ttc ctc gcc tcg gcc tac gcc tcc cgc ggg ttg aaa Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys 245 250 255	768
atg cgc tac acc tcc ggc acc gga tcc gaa gcg ctg atg ggc tat tcg Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ser 260 265 270 275 280 285 290 295 300	816

260	265	270	
gag agc aag tcg atg ctc tac ctc gaa tcg cgc tgc atc ttc att act Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr 275 280 285			864
aaa ggc gcc ggg gtt cag gga ctg caa aac ggc gcg gtg agc tgt atc Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile 290 295 300			912
ggc atg acc ggc gct gtg ccg tcg ggc att cgg gcg gtg ctg gcg gaa Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu 305 310 315 320			960
aac ctg atc gcc tct atg ctc gac ctc gaa gtg gcg tcc gcc aac gac Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp 325 330 335			1008
cag act ttc tcc cac tcg gat att cgc cgc acc gcg cgc acc ctg atg Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met 340 345 350			1056
cag atg ctg ccg ggc acc gac ttt att ttc tcc ggc tac agc gcg gtg Gln Met Leu Pro Gly Thr Asp Phe Ile Phe Ser Gly Tyr Ser Ala Val 355 360 365			1104
ccg aac tac gac aac atg ttc gcc ggc tcg aac ttc gat gcg gaa gat Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp 370 375 380			1152
ttt gat gat tac aac atc ctg cag cgt gac ctg atg gtt gac ggc ggc Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly 385 390 395 400			1200
ctg cgt ccg gtg acc gag gcg gaa acc att gcc att cgc cag aaa gcg Leu Arg Pro Val Thr Glu Ala Glu Thr Ile Ala Ile Arg Gln Lys Ala 405 410 415			1248
gcg cgc gcg atc cag gcg gtt ttc cgc gag ctg ggg ctg ccg cca atc Ala Arg Ala Ile Gln Ala Val Phe Arg Glu Leu Gly Leu Pro Pro Ile 420 425 430			1296
gcc gac gag gag gtg gag gcc gcc acc tac gcg cac ggc agc aac gag Ala Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Asn Glu 435 440 445			1344
atg ccg ccg cgt aac gtg gtg gag gat ctg agt gcg gtg gaa gag atg Met Pro Pro Arg Asn Val Val Glu Asp Leu Ser Ala Val Glu Glu Met 450 455 460			1392
atg aag cgc aac atc acc ggc ctc gat att gtc ggc gcg ctg agc cgc Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Gly Ala Leu Ser Arg 465 470 475 480			1440
agc ggc ttt gag gat atc gcc agc aat att ctc aat atg ctg cgc cag Ser Gly Phe Glu Asp Ile Ala Ser Asn Ile Leu Asn Met Leu Arg Gln 485 490 495			1488
cgg gtc acc ggc gat tac ctg cag acc tcg gcc att ctc gat cgg cag Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Gln 500 505 510			1536
ttc gag gtg gtg agt gcg gtc aac gac atc aat gac tat cag ggg ccg			1584

Phe Glu Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro
515 520 525

ggc acc ggc tat cgc atc tct gcc gaa cgc tgg gcg gag atc aaa aat 1632
Gly Thr Gly Tyr Arg Ile Ser Ala Glu Arg Trp Ala Glu Ile Lys Asn
530 535 540

att ccg ggc gtg gtt cag ccc gac acc att gaa taa 1668
Ile Pro Gly Val Val Gln Pro Asp Thr Ile Glu
545 550 555

<210> 10

<211> 555

<212> PRT

5 <213> Klebsiella pneumoniae

<400> 10

Met Lys Arg Ser Lys Arg Phe Ala Val Leu Ala Gln Arg Pro Val Asn
1 5 10 15

Gln Asp Gly Leu Ile Gly Glu Trp Pro Glu Glu Gly Leu Ile Ala Met
20 25 30

Asp Ser Pro Phe Asp Pro Val Ser Ser Val Lys Val Asp Asn Gly Leu
35 40 45

Ile Val Glu Leu Asp Gly Lys Arg Arg Asp Gln Phe Asp Met Ile Asp
50 55 60

Arg Phe Ile Ala Asp Tyr Ala Ile Asn Val Glu Arg Thr Glu Gln Ala
65 70 75 80

Met Arg Leu Glu Ala Val Glu Ile Ala Arg Met Leu Val Asp Ile His
85 90 95

Val Ser Arg Glu Glu Ile Ile Ala Ile Thr Thr Ala Ile Thr Pro Ala
100 105 110

Lys Ala Val Glu Val Met Ala Gln Met Asn Val Val Glu Met Met Met
115 120 125

Ala Leu Gln Lys Met Arg Ala Arg Arg Thr Pro Ser Asn Gln Cys His
130 135 140

Val Thr Asn Leu Lys Asp Asn Pro Val Gln Ile Ala Ala Asp Ala Ala
145 150 155 160

Glu Ala Gly Ile Arg Gly Phe Ser Glu Gln Glu Thr Thr Val Gly Ile
165 170 175

Ala Arg Tyr Ala Pro Phe Asn Ala Leu Ala Leu Leu Val Gly Ser Gln
180 185 190

Cys Gly Arg Pro Gly Val Leu Thr Gln Cys Ser Val Glu Glu Ala Thr
195 200 205

Glu Leu Glu Leu Gly Met Arg Gly Leu Thr Ser Tyr Ala Glu Thr Val
210 215 220

Ser Val Tyr Gly Thr Glu Ala Val Phe Thr Asp Gly Asp Asp Thr Pro
225 230 235 240

Trp Ser Lys Ala Phe Leu Ala Ser Ala Tyr Ala Ser Arg Gly Leu Lys
245 250 255

Met Arg Tyr Thr Ser Gly Thr Gly Ser Glu Ala Leu Met Gly Tyr Ser
260 265 270

Glu Ser Lys Ser Met Leu Tyr Leu Glu Ser Arg Cys Ile Phe Ile Thr
275 280 285

Lys Gly Ala Gly Val Gln Gly Leu Gln Asn Gly Ala Val Ser Cys Ile
290 295 300

Gly Met Thr Gly Ala Val Pro Ser Gly Ile Arg Ala Val Leu Ala Glu
305 310 315 320

Asn Leu Ile Ala Ser Met Leu Asp Leu Glu Val Ala Ser Ala Asn Asp
325 330 335

Gln Thr Phe Ser His Ser Asp Ile Arg Arg Thr Ala Arg Thr Leu Met
340 345 350

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

Pro Asn Tyr Asp Asn Met Phe Ala Gly Ser Asn Phe Asp Ala Glu Asp
370 375 380

Phe Asp Asp Tyr Asn Ile Leu Gln Arg Asp Leu Met Val Asp Gly Gly
385 390 395 400

Leu Arg Pro Val Thr Glu Ala Glu Thr Ile Ala Ile Arg Gln Lys Ala
405 410 415

Ala Arg Ala Ile Gln Ala Val Phe Arg Glu Leu Gly Leu Pro Pro Ile
420 425 430

Ala Asp Glu Glu Val Glu Ala Ala Thr Tyr Ala His Gly Ser Asn Glu
435 440 445

Met Pro Pro Arg Asn Val Val Glu Asp Leu Ser Ala Val Glu Glu Met
450 455 460

Met Lys Arg Asn Ile Thr Gly Leu Asp Ile Val Gly Ala Leu Ser Arg
465 470 475 480

Ser Gly Phe Glu Asp Ile Ala Ser Asn Ile Leu Asn Met Leu Arg Gln
485 490 495

Arg Val Thr Gly Asp Tyr Leu Gln Thr Ser Ala Ile Leu Asp Arg Gln
500 505 510

Phe Glu Val Val Ser Ala Val Asn Asp Ile Asn Asp Tyr Gln Gly Pro
515 520 525

Gly Thr Gly Tyr Arg Ile Ser Ala Glu Arg Trp Ala Glu Ile Lys Asn
530 535 540

Ile Pro Gly Val Val Gln Pro Asp Thr Ile Glu
545 550 555

<210> 11
<211> 585
<212> DNA
5 <213> Klebsiella pneumoniae

<220>
<221> CDS
<222> (1)..(585)

<400> 11
gtg caa cag aca acc caa att cag ccc tct ttt acc ctg aaa acc cgc 48
Val Gln Gln Thr Thr Gln Ile Gln Pro Ser Phe Thr Leu Lys Thr Arg
1 5 10 15
gag ggc ggg gta gct tct gcc gat gaa cgc gcc gat gaa gtg gtg atc 96
Glu Gly Gly Val Ala Ser Ala Asp Glu Arg Ala Asp Glu Val Val Ile
20 25 30
ggc gtc ggc cct gcc ttc gat aaa cac cag cat cac act ctg atc gat 144
Gly Val Gly Pro Ala Phe Asp Lys His Gln His His Thr Leu Ile Asp
35 40 45
atg ccc cat ggc gcg atc ctc aaa gag ctg att gcc ggg gtg gaa gaa 192
Met Pro His Gly Ala Ile Leu Lys Glu Leu Ile Ala Gly Val Glu Glu
50 55 60
gag ggg ctt cac gcc cgg gtg gtg cgc att ctg cgc acg tcc gac gtc 240
Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
65 70 75 80

tcc ttt atg gcc tgg gat gcg gcc aac ctg agc ggc tcg ggg atc ggc 288
Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
85 90 95

atc ggt atc cag tcg aag ggg acc acg gtc atc cat cag cgc gat ctg 336
Ile Gly Ile Gln Ser Lys Gly Thr Val Ile His Gln Arg Asp Leu
100 105 110

ctg ccg ctc agc aac ctg gag ctg ttc tcc cag gcg ccg ctg ctg acg 384
Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
115 120 125

ctg gag acc tac cgg cag att ggc aaa aac gct gcg cgc tat gcg cgc 432
Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
130 135 140

aaa gag tca cct tcg ccg gtg ccg gtg gtg aac gat cag atg gtg ccg 480
Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
145 150 155 160

ccg aaa ttt atg gcc aaa gcc gcg cta ttt cat atc aaa gag acc aaa 528
Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
165 170 175

cat gtg gtg cag gac gcc gag ccc gtc acc ctg cac atc gac tta gta 576
His Val Val Gln Asp Ala Glu Pro Val Thr Leu His Ile Asp Leu Val
180 185 190

agg gag tga 585
Arg Glu

<210> 12
<211> 194
<212> PRT

5 <213> Klebsiella pneumoniae

<400> 12
Val Gln Gln Thr Thr Gln Ile Gln Pro Ser Phe Thr Leu Lys Thr Arg
1 5 10 15

Glu Gly Gly Val Ala Ser Ala Asp Glu Arg Ala Asp Glu Val Val Ile
20 25 30

Gly Val Gly Pro Ala Phe Asp Lys His Gln His His Thr Leu Ile Asp
35 40 45

Met Pro His Gly Ala Ile Leu Lys Glu Leu Ile Ala Gly Val Glu Glu
50 55 60

Glu Gly Leu His Ala Arg Val Val Arg Ile Leu Arg Thr Ser Asp Val
65 70 75 80

Ser Phe Met Ala Trp Asp Ala Ala Asn Leu Ser Gly Ser Gly Ile Gly
85 90 95

Ile Gly Ile Gln Ser Lys Gly Thr Thr Val Ile His Gln Arg Asp Leu
100 105 110

Leu Pro Leu Ser Asn Leu Glu Leu Phe Ser Gln Ala Pro Leu Leu Thr
115 120 125

Leu Glu Thr Tyr Arg Gln Ile Gly Lys Asn Ala Ala Arg Tyr Ala Arg
130 135 140

Lys Glu Ser Pro Ser Pro Val Pro Val Val Asn Asp Gln Met Val Arg
145 150 155 160

Pro Lys Phe Met Ala Lys Ala Ala Leu Phe His Ile Lys Glu Thr Lys
165 170 175

His Val Val Gln Asp Ala Glu Pro Val Thr Leu His Ile Asp Leu Val
180 185 190

Arg Glu

<210> 13
<211> 426
<212> DNA
5 <213> Klebsiella pneumoniae

<220>
<221> CDS
<222> (1)..(426)

<400> 13
atg agc gag aaa acc atg cgc gtg cag gat tat ccg tta gcc acc cgc 48
Met Ser Glu Lys Thr Met Arg Val Gln Asp Tyr Pro Leu Ala Thr Arg
1 5 10 15
tgc ccg gag cat atc ctg acg cct acc ggc aaa cca ttg acc gat att 96
Cys Pro Glu His Ile Leu Thr Pro Thr Gly Lys Pro Leu Thr Asp Ile
20 25 30
acc ctc gag aag gtg ctc tct ggc gag gtg ggc ccg cag gat gtg cgg 144
Thr Leu Glu Lys Val Leu Ser Gly Glu Val Gly Pro Gln Asp Val Arg
35 40 45
atc tcc cgc cag acc ctt gag tac cag gcg cag att gcc gag cag atg 192
Ile Ser Arg Gln Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met
50 55 60
cag cgc cat gcg gtg gcg cgc aat ttc cgc cgc gcg gcg gag ctt atc 240
Gln Arg His Ala Val Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile
65 70 75 80
gcc att cct gac gag cgc att ctg gct atc tat aac gcg ctg cgc ccg 288
Ala Ile Pro Asp Glu Arg Ile Leu Ala Ile Tyr Asn Ala Leu Arg Pro

	85	90	95	
	ttc cgc tcc tcg cag gcg gag ctg ctg gcg atc gcc gac gag ctg gag			336
	Phe Arg Ser Ser Gln Ala Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu			
	100	105	110	
	cac acc tgg cat gcg aca gtg aat gcc gcc ttt gtc cgg gag tcg gcg			384
	His Thr Trp His Ala Thr Val Asn Ala Ala Phe Val Arg Glu Ser Ala			
	115	120	125	
	gaa gtg tat cag cag cgg cat aag ctg cgt aaa gga agc taa			426
	Glu Val Tyr Gln Gln Arg His Lys Leu Arg Lys Gly Ser			
	130	135	140	
	<210> 14			
	<211> 141			
	<212> PRT			
5	<213> Klebsiella pneumoniae			
	<400> 14			
	Met Ser Glu Lys Thr Met Arg Val Gln Asp Tyr Pro Leu Ala Thr Arg			
	1	5	10	15
	Cys Pro Glu His Ile Leu Thr Pro Thr Gly Lys Pro Leu Thr Asp Ile			
	20	25	30	
	Thr Leu Glu Lys Val Leu Ser Gly Glu Val Gly Pro Gln Asp Val Arg			
	35	40	45	
	Ile Ser Arg Gln Thr Leu Glu Tyr Gln Ala Gln Ile Ala Glu Gln Met			
	50	55	60	
	Gln Arg His Ala Val Ala Arg Asn Phe Arg Arg Ala Ala Glu Leu Ile			
	65	70	75	80
	Ala Ile Pro Asp Glu Arg Ile Leu Ala Ile Tyr Asn Ala Leu Arg Pro			
	85	90	95	
	Phe Arg Ser Ser Gln Ala Glu Leu Leu Ala Ile Ala Asp Glu Leu Glu			
	100	105	110	
	His Thr Trp His Ala Thr Val Asn Ala Ala Phe Val Arg Glu Ser Ala			
	115	120	125	
	Glu Val Tyr Gln Gln Arg His Lys Leu Arg Lys Gly Ser			
	130	135	140	
	<210> 15			
	<211> 1539			
10	<212> DNA			
	<213> Escherichia coli			
	<220>			
	<221> CDS			
	<222> (1)..(1539)			
15	<400> 15			

atg acc aat aat ccc cct tca gca cag att aag ccc ggc gag tat ggt Met Thr Asn Asn Pro Pro Ser Ala Gln Ile Lys Pro Gly Glu Tyr Gly 1 5 10 15	48
ttc ccc ctg aag tta aaa gcc cgc tat gac aac ttt att ggc ggc gaa Phe Pro Leu Lys Leu Lys Ala Arg Tyr Asp Asn Phe Ile Gly Gly Glu 20 25 30	96
tgg gta gcc cct gcc gac ggc gag tat tac cag aat ctg acg ccg gtg Trp Val Ala Pro Ala Asp Gly Glu Tyr Tyr Gln Asn Leu Thr Pro Val 35 40 45	144
acc ggg cag ctg ctg tgc gaa gtg gcg tct tcg ggc aaa cga gac atc Thr Gly Gln Leu Leu Cys Glu Val Ala Ser Ser Gly Lys Arg Asp Ile 50 55 60	192
gat ctg gcg ctg gat gct gcg cac aaa gtg aaa gat aaa tgg gcg cac Asp Leu Ala Leu Asp Ala Ala His Lys Val Lys Asp Lys Trp Ala His 65 70 75 80	240
acc tcg gtg cag gat cgt gcg gcg att ctg ttt aag att gcc gat cga Thr Ser Val Gln Asp Arg Ala Ala Ile Leu Phe Lys Ile Ala Asp Arg 85 90 95	288
atg gaa caa aac ctg gag ctg tta gcg aca gct gaa acc tgg gat aac Met Glu Gln Asn Leu Glu Leu Leu Ala Thr Ala Glu Thr Trp Asp Asn 100 105 110	336
ggc aaa ccc att cgc gaa acc agt gct gcg gat gta ccg ctg gcg att Gly Lys Pro Ile Arg Glu Thr Ser Ala Ala Asp Val Pro Leu Ala Ile 115 120 125	384
gac cat ttc cgc tat ttc gcc tcg tgt att cgg gcg cag gaa ggt ggg Asp His Phe Arg Tyr Phe Ala Ser Cys Ile Arg Ala Gln Glu Gly Gly 130 135 140	432
atc agt gaa gtt gat agc gaa acc gtg gcc tat cat ttc cat gaa ccg Ile Ser Glu Val Asp Ser Glu Thr Val Ala Tyr His Phe His Glu Pro 145 150 155 160	480
tta ggc gtg gtg ggg cag att atc ccg tgg aac ttc ccg ctg ctg atg Leu Gly Val Val Gly Gln Ile Ile Pro Trp Asn Phe Pro Leu Leu Met 165 170 175	528
gcg agc tgg aaa atg gct ccc gcg ctg gcg gcg ggc aac tgt gtg gtg Ala Ser Trp Lys Met Ala Pro Ala Leu Ala Ala Gly Asn Cys Val Val 180 185 190	576
ctg aaa ccc gca cgt ctt acc ccg ctt tct gta ctg ctg cta atg gaa Leu Lys Pro Ala Arg Leu Thr Pro Leu Ser Val Leu Leu Leu Met Glu 195 200 205	624
att gtc ggt gat tta ctg ccg ccg ggc gtg gtg aac gtg gtc aat ggc Ile Val Gly Asp Leu Leu Pro Pro Gly Val Val Asn Val Val Asn Gly 210 215 220	672
gca ggt ggg gta att ggc gaa tat ctg gcg acc tcg aaa cgc atc gcc	720

Ala Gly Gly Val Ile Gly Glu Tyr Leu Ala Thr Ser Lys Arg Ile Ala	
225 230 235 240	
aaa gtg gcg ttt acc ggc tca acg gaa gtg ggc caa caa att atg caa	768
Lys Val Ala Phe Thr Gly Ser Thr Glu Val Gly Gln Gln Ile Met Gln	
245 250 255	
tac gca acg caa aac att att ccg gtg acg ctg gag ttg ggc ggt aag	816
Tyr Ala Thr Gln Asn Ile Ile Pro Val Thr Leu Glu Leu Gly Gly Lys	
260 265 270	
tcg cca aat atc ttc ttt gct gat gtg atg gat gaa gaa gat gcc ttt	864
Ser Pro Asn Ile Phe Phe Ala Asp Val Met Asp Glu Glu Asp Ala Phe	
275 280 285	
ttc gat aaa gcg ctg gaa ggc ttt gca ctg ttt gcc ttt aac cag ggc	912
Phe Asp Lys Ala Leu Glu Gly Phe Ala Leu Phe Ala Phe Asn Gln Gly	
290 295 300	
gaa gtt tgc acc tgt ccg agt cgt gct tta gtg cag gaa tct atc tac	960
Glu Val Cys Thr Cys Pro Ser Arg Ala Leu Val Gln Glu Ser Ile Tyr	
305 310 315 320	
gaa cgc ttt atg gaa cgc gcc atc cgc cgt gtc gaa agc att cgt agc	1008
Glu Arg Phe Met Glu Arg Ala Ile Arg Arg Val Glu Ser Ile Arg Ser	
325 330 335	
ggg aac ccg ctc gac agc gtg acg caa atg ggc gcg cag gtt tct cac	1056
Gly Asn Pro Leu Asp Ser Val Thr Gln Met Gly Ala Gln Val Ser His	
340 345 350	
ggg caa ctg gaa acc atc ctc aac tac att gat atc ggt aaa aaa gag	1104
Gly Gln Leu Glu Thr Ile Leu Asn Tyr Ile Asp Ile Gly Lys Lys Glu	
355 360 365	
ggc gct gac gtg ctc aca ggc ggc cgc cgc aag ctg ctg gaa ggt gaa	1152
Gly Ala Asp Val Leu Thr Gly Gly Arg Arg Lys Leu Leu Glu Gly Glu	
370 375 380	
ctg aaa gac ggc tac tac ctc gaa ccg acg att ctg ttt ggt cag aac	1200
Leu Lys Asp Gly Tyr Tyr Leu Glu Pro Thr Ile Leu Phe Gly Gln Asn	
385 390 395 400	
aat atg cgg gtg ttc cag gag gag att ttt ggc ccg gtg ctg gcg gtg	1248
Asn Met Arg Val Phe Gln Glu Glu Ile Phe Gly Pro Val Leu Ala Val	
405 410 415	
acc acc ttc aaa acg atg gaa gaa gcg ctg gag ctg gcg aac gat acg	1296
Thr Thr Phe Lys Thr Met Glu Glu Ala Leu Glu Leu Ala Asn Asp Thr	
420 425 430	
caa tat ggc ctg ggc gcg ggc gtc tgg agc cgc aac ggt aat ctg gcc	1344
Gln Tyr Gly Leu Gly Ala Gly Val Trp Ser Arg Asn Gly Asn Leu Ala	
435 440 445	
tat aag atg ggg cgc ggc ata cag gct ggg cgc gtg tgg acc aac tgt	1392
Tyr Lys Met Gly Arg Gly Ile Gln Ala Gly Arg Val Trp Thr Asn Cys	
450 455 460	
tat cac gct tac ccg gca cat gcg gcg ttt ggt ggc tac aaa caa tca	1440
Tyr His Ala Tyr Pro Ala His Ala Ala Phe Gly Gly Tyr Lys Gln Ser	
465 470 475 480	

ggt atc ggt cgc gaa acc cac aag atg atg ctg gag cat tac cag caa 1488
Gly Ile Gly Arg Glu Thr His Lys Met Met Leu Glu His Tyr Gln Gln
485 490 495

acc aag tgc ctg ctg gtg agc tac tcg gat aaa ccg ttg ggg ctg ttc 1536
Thr Lys Cys Leu Leu Val Ser Tyr Ser Asp Lys Pro Leu Gly Leu Phe
500 505 510

tga 1539

<210> 16
<211> 512
<212> PRT
5 <213> Escherichia coli

<400> 16
Met Thr Asn Asn Pro Pro Ser Ala Gln Ile Lys Pro Gly Glu Tyr Gly
1 5 10 15

Phe Pro Leu Lys Leu Lys Ala Arg Tyr Asp Asn Phe Ile Gly Gly Glu
20 25 30

Trp Val Ala Pro Ala Asp Gly Glu Tyr Tyr Gln Asn Leu Thr Pro Val
35 40 45

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

Asp Leu Ala Leu Asp Ala Ala His Lys Val Lys Asp Lys Trp Ala His
65 70 75 80

Thr Ser Val Gln Asp Arg Ala Ala Ile Leu Phe Lys Ile Ala Asp Arg
85 90 95

Met Glu Gln Asn Leu Glu Leu Leu Ala Thr Ala Glu Thr Trp Asp Asn
100 105 110

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

Asp His Phe Arg Tyr Phe Ala Ser Cys Ile Arg Ala Gln Glu Gly Gly
130 135 140

Ile Ser Glu Val Asp Ser Glu Thr Val Ala Tyr His Phe His Glu Pro
145 150 155 160

Leu Gly Val Val Gly Gln Ile Ile Pro Trp Asn Phe Pro Leu Leu Met
165 170 175

Ala Ser Trp Lys Met Ala Pro Ala Leu Ala Ala Gly Asn Cys Val Val

ES 2 552 317 T3

180	185	190
Leu Lys Pro Ala Arg Leu Thr Pro Leu Ser Val Leu Leu Leu Met Glu		
195	200	205
Ile Val Gly Asp Leu Leu Pro Pro Gly Val Val Asn Val Val Asn Gly		
210	215	220
Ala Gly Gly Val Ile Gly Glu Tyr Leu Ala Thr Ser Lys Arg Ile Ala		
225	230	235 240
Lys Val Ala Phe Thr Gly Ser Thr Glu Val Gly Gln Gln Ile Met Gln		
	245	250 255
Tyr Ala Thr Gln Asn Ile Ile Pro Val Thr Leu Glu Leu Gly Gly Lys		
	260	265 270
Ser Pro Asn Ile Phe Phe Ala Asp Val Met Asp Glu Glu Asp Ala Phe		
	275	280 285
Phe Asp Lys Ala Leu Glu Gly Phe Ala Leu Phe Ala Phe Asn Gln Gly		
	290	295 300
Glu Val Cys Thr Cys Pro Ser Arg Ala Leu Val Gln Glu Ser Ile Tyr		
	310	315 320
Glu Arg Phe Met Glu Arg Ala Ile Arg Arg Val Glu Ser Ile Arg Ser		
	325	330 335
Gly Asn Pro Leu Asp Ser Val Thr Gln Met Gly Ala Gln Val Ser His		
	340	345 350
Gly Gln Leu Glu Thr Ile Leu Asn Tyr Ile Asp Ile Gly Lys Lys Glu		
	355	360 365
Gly Ala Asp Val Leu Thr Gly Gly Arg Arg Lys Leu Leu Glu Gly Glu		
	370	375 380
Leu Lys Asp Gly Tyr Tyr Leu Glu Pro Thr Ile Leu Phe Gly Gln Asn		
	385	390 395 400
Asn Met Arg Val Phe Gln Glu Glu Ile Phe Gly Pro Val Leu Ala Val		
	405	410 415
Thr Thr Phe Lys Thr Met Glu Glu Ala Leu Glu Leu Ala Asn Asp Thr		
	420	425 430

Gln Tyr Gly Leu Gly Ala Gly Val Trp Ser Arg Asn Gly Asn Leu Ala
 435 440 445

Tyr Lys Met Gly Arg Gly Ile Gln Ala Gly Arg Val Trp Thr Asn Cys
 450 455 460

Tyr His Ala Tyr Pro Ala His Ala Ala Phe Gly Gly Tyr Lys Gln Ser
 465 470 475 480

Gly Ile Gly Arg Glu Thr His Lys Met Met Leu Glu His Tyr Gln Gln
 485 490 495

Thr Lys Cys Leu Leu Val Ser Tyr Ser Asp Lys Pro Leu Gly Leu Phe
 500 505 510

<210> 17
 <211> 1440
 <212> DNA
 5 <213> Escherichia coli

<220>
 <221> CDS
 <222> (1)..(1440)

<400> 17
 atg tca gta ccc gtt caa cat cct atg tat atc gat gga cag ttt gtt 48
 Met Ser Val Pro Val Gln His Pro Met Tyr Ile Asp Gly Gln Phe Val
 1 5 10 15
 acc tgg cgt gga gac gca tgg att gat gtg gta aac cct gct aca gag 96
 Thr Trp Arg Gly Asp Ala Trp Ile Asp Val Val Asn Pro Ala Thr Glu
 20 25 30
 gct gtc att tcc cgc ata ccc gat ggt cag gcc gag gat gcc cgt aag 144
 Ala Val Ile Ser Arg Ile Pro Asp Gly Gln Ala Glu Asp Ala Arg Lys
 35 40 45
 gca atc gat gca gca gaa cgt gca caa cca gaa tgg gaa gcg ttg cct 192
 Ala Ile Asp Ala Ala Glu Arg Ala Gln Pro Glu Trp Glu Ala Leu Pro
 50 55 60
 gct att gaa cgc gcc agt tgg ttg cgc aaa atc tcc gcc ggg atc cgc 240
 Ala Ile Glu Arg Ala Ser Trp Leu Arg Lys Ile Ser Ala Gly Ile Arg
 65 70 75 80
 gaa cgc gcc agt gaa atc agt gcg ctg att gtt gaa gaa ggg ggc aag 288
 Glu Arg Ala Ser Glu Ile Ser Ala Leu Ile Val Glu Glu Gly Gly Lys
 85 90 95
 atc cag cag ctg gct gaa gtc gaa gtg gct ttt act gcc gac tat atc 336
 Ile Gln Gln Leu Ala Glu Val Glu Val Ala Phe Thr Ala Asp Tyr Ile
 100 105 110
 gat tac atg gcg gag tgg gca cgg cgt tac gag ggc gag att att caa 384
 Asp Tyr Met Ala Glu Trp Ala Arg Arg Tyr Glu Gly Glu Ile Ile Gln
 115 120 125

agc gat cgt cca gga gaa aat att ctt ttg ttt aaa cgt gcg ctt ggt Ser Asp Arg Pro Gly Glu Asn Ile Leu Leu Phe Lys Arg Ala Leu Gly 130 135 140	432
gtg act acc ggc att ctg ccg tgg aac ttc ccg ttc ttc ctc att gcc Val Thr Thr Gly Ile Leu Pro Trp Asn Phe Pro Phe Phe Leu Ile Ala 145 150 155 160	480
cgc aaa atg gct ccc gct ctt ttg acc ggt aat acc atc gtc att aaa Arg Lys Met Ala Pro Ala Leu Leu Thr Gly Asn Thr Ile Val Ile Lys 165 170 175	528
cct agt gaa ttt acg cca aac aat gcg att gca ttc gcc aaa atc gtc Pro Ser Glu Phe Thr Pro Asn Asn Ala Ile Ala Phe Ala Lys Ile Val 180 185 190	576
gat gaa ata ggc ctt ccg cgc ggc gtg ttt aac ctt gta ctg ggg cgt Asp Glu Ile Gly Leu Pro Arg Gly Val Phe Asn Leu Val Leu Gly Arg 195 200 205	624
ggt gaa acc gtt ggg caa gaa ctg gcg ggt aac cca aag gtc gca atg Gly Glu Thr Val Gly Gln Glu Leu Ala Gly Asn Pro Lys Val Ala Met 210 215 220	672
gtc agt atg aca ggc agc gtc tct gca ggt gag aag atc atg gcg act Val Ser Met Thr Gly Ser Val Ser Ala Gly Glu Lys Ile Met Ala Thr 225 230 235 240	720
gcg gcg aaa aac atc acc aaa gtg tgt ctg gaa ttg ggg ggt aaa gca Ala Ala Lys Asn Ile Thr Lys Val Cys Leu Glu Leu Gly Gly Lys Ala 245 250 255	768
cca gct atc gta atg gac gat gcc gat ctt gaa ctg gca gtc aaa gcc Pro Ala Ile Val Met Asp Asp Ala Asp Leu Glu Leu Ala Val Lys Ala 260 265 270	816
atc gtt gat tca cgc gtc att aat agt ggg caa gtg tgt aac tgt gca Ile Val Asp Ser Arg Val Ile Asn Ser Gly Gln Val Cys Asn Cys Ala 275 280 285	864
gaa cgt gtt tat gta cag aaa ggc att tat gat cag ttc gtc aat cgg Glu Arg Val Tyr Val Gln Lys Gly Ile Tyr Asp Gln Phe Val Asn Arg 290 295 300	912
ctg ggt gaa gcg atg cag gcg gtt caa ttt ggt aac ccc gct gaa cgc Leu Gly Glu Ala Met Gln Ala Val Gln Phe Gly Asn Pro Ala Glu Arg 305 310 315 320	960
aac gac att gcg atg ggg ccg ttg att aac gcc gcg gcg ctg gaa agg Asn Asp Ile Ala Met Gly Pro Leu Ile Asn Ala Ala Ala Leu Glu Arg 325 330 335	1008
gtc gag caa aaa gtg gcg cgc gca gta gaa gaa ggg gcg aga gtg gcg Val Glu Gln Lys Val Ala Arg Ala Val Glu Glu Gly Ala Arg Val Ala 340 345 350	1056
ttc ggt ggc aaa gcg gta gag ggg aaa gga tat tat tat ccg ccg aca Phe Gly Gly Lys Ala Val Glu Gly Lys Gly Tyr Tyr Tyr Pro Pro Thr 355 360 365	1104
ttg ctg ctg gat gtt cgc cag gaa atg tcg att atg cat gag gaa acc Leu Leu Leu Asp Val Arg Gln Glu Met Ser Ile Met His Glu Glu Thr 370 375 380	1152

ttt ggc ccg gtg ctg cca gtt gtc gca ttt gac acg ctg gaa gat gct 1200
Phe Gly Pro Val Leu Pro Val Val Ala Phe Asp Thr Leu Glu Asp Ala
385 390 395 400

atc tca atg gct aat gac agt gat tac ggc ctg acc tca tca atc tat 1248
Ile Ser Met Ala Asn Asp Ser Asp Tyr Gly Leu Thr Ser Ser Ile Tyr
405 410 415

acc caa aat ctg aac gtc gcg atg aaa gcc att aaa ggg ctg aag ttt 1296
Thr Gln Asn Leu Asn Val Ala Met Lys Ala Ile Lys Gly Leu Lys Phe
420 425 430

ggc gaa act tac atc aac cgt gaa aac ttc gaa gct atg caa ggc ttc 1344
Gly Glu Thr Tyr Ile Asn Arg Glu Asn Phe Glu Ala Met Gln Gly Phe
435 440 445

cac gcc gga tgg cgt aaa tcc ggt att ggc ggc gca gat ggt aaa cat 1392
His Ala Gly Trp Arg Lys Ser Gly Ile Gly Gly Ala Asp Gly Lys His
450 455 460

ggc ttg cat gaa tat ctg cag acc cag gtg gtt tat tta cag tct taa 1440
Gly Leu His Glu Tyr Leu Gln Thr Gln Val Val Tyr Leu Gln Ser
465 470 475

<210> 18
<211> 479
<212> PRT
5 <213> Escherichia coli

<400> 18
Met Ser Val Pro Val Gln His Pro Met Tyr Ile Asp Gly Gln Phe Val
1 5 10 15

Thr Trp Arg Gly Asp Ala Trp Ile Asp Val Val Asn Pro Ala Thr Glu
20 25 30

Ala Val Ile Ser Arg Ile Pro Asp Gly Gln Ala Glu Asp Ala Arg Lys
35 40 45

Ala Ile Asp Ala Ala Glu Arg Ala Gln Pro Glu Trp Glu Ala Leu Pro
50 55 60

Ala Ile Glu Arg Ala Ser Trp Leu Arg Lys Ile Ser Ala Gly Ile Arg
65 70 75 80

Glu Arg Ala Ser Glu Ile Ser Ala Leu Ile Val Glu Glu Gly Gly Lys
85 90 95

Ile Gln Gln Leu Ala Glu Val Glu Val Ala Phe Thr Ala Asp Tyr Ile
100 105 110

Asp Tyr Met Ala Glu Trp Ala Arg Arg Tyr Glu Gly Glu Ile Ile Gln
115 120 125

Ser Asp Arg Pro Gly Glu Asn Ile Leu Leu Phe Lys Arg Ala Leu Gly
130 135 140

Val Thr Thr Gly Ile Leu Pro Trp Asn Phe Pro Phe Phe Leu Ile Ala
145 150 155 160

Arg Lys Met Ala Pro Ala Leu Leu Thr Gly Asn Thr Ile Val Ile Lys
165 170 175

Pro Ser Glu Phe Thr Pro Asn Asn Ala Ile Ala Phe Ala Lys Ile Val
180 185 190

Asp Glu Ile Gly Leu Pro Arg Gly Val Phe Asn Leu Val Leu Gly Arg
195 200 205

Gly Glu Thr Val Gly Gln Glu Leu Ala Gly Asn Pro Lys Val Ala Met
210 215 220

Val Ser Met Thr Gly Ser Val Ser Ala Gly Glu Lys Ile Met Ala Thr
225 230 235 240

Ala Ala Lys Asn Ile Thr Lys Val Cys Leu Glu Leu Gly Gly Lys Ala
245 250 255

Pro Ala Ile Val Met Asp Asp Ala Asp Leu Glu Leu Ala Val Lys Ala
260 265 270

Ile Val Asp Ser Arg Val Ile Asn Ser Gly Gln Val Cys Asn Cys Ala
275 280 285

Glu Arg Val Tyr Val Gln Lys Gly Ile Tyr Asp Gln Phe Val Asn Arg
290 295 300

Leu Gly Glu Ala Met Gln Ala Val Gln Phe Gly Asn Pro Ala Glu Arg
305 310 315 320

Asn Asp Ile Ala Met Gly Pro Leu Ile Asn Ala Ala Ala Leu Glu Arg
325 330 335

Val Glu Gln Lys Val Ala Arg Ala Val Glu Glu Gly Ala Arg Val Ala
340 345 350

Phe Gly Gly Lys Ala Val Glu Gly Lys Gly Tyr Tyr Tyr Pro Pro Thr
355 360 365

Leu Leu Leu Asp Val Arg Gln Glu Met Ser Ile Met His Glu Glu Thr

370

375

380

Phe Gly Pro Val Leu Pro Val Val Ala Phe Asp Thr Leu Glu Asp Ala
385 390 395 400

Ile Ser Met Ala Asn Asp Ser Asp Tyr Gly Leu Thr Ser Ser Ile Tyr
405 410 415

Thr Gln Asn Leu Asn Val Ala Met Lys Ala Ile Lys Gly Leu Lys Phe
420 425 430

Gly Glu Thr Tyr Ile Asn Arg Glu Asn Phe Glu Ala Met Gln Gly Phe
435 440 445

His Ala Gly Trp Arg Lys Ser Gly Ile Gly Gly Ala Asp Gly Lys His
450 455 460

Gly Leu His Glu Tyr Leu Gln Thr Gln Val Val Tyr Leu Gln Ser
465 470 475

<210> 19
<211> 1488
<212> DNA
<213> Escherichia coli

<220>
<221> CDS
<222> (1)..(1488)

<400> 19
atg aat ttt cat cat ctg gct tac tgg cag gat aaa gcg tta agt ctc 48
Met Asn Phe His His Leu Ala Tyr Trp Gln Asp Lys Ala Leu Ser Leu
1 5 10 15
gcc att gaa aac cgc tta ttt att aac ggt gaa tat act gct gcg gcg 96
Ala Ile Glu Asn Arg Leu Phe Ile Asn Gly Glu Tyr Thr Ala Ala Ala
20 25 30
gaa aat gaa acc ttt gaa acc gtt gat ccg gtc acc cag gca ccg ctg 144
Glu Asn Glu Thr Phe Glu Thr Val Asp Pro Val Thr Gln Ala Pro Leu
35 40 45
gcg aaa att gcc cgc gcc aag agc gtc gat atc gac cgt gcg atg agc 192
Ala Lys Ile Ala Arg Gly Lys Ser Val Asp Ile Asp Arg Ala Met Ser
50 55 60
gca gca cgc gcc gta ttt gaa cgc gcc gac tgg tca ctc tct tct ccg 240
Ala Ala Arg Gly Val Phe Glu Arg Gly Asp Trp Ser Leu Ser Ser Pro
65 70 75 80
gct aaa cgt aaa gcg gta ctg aat aaa ctc gcc gat tta atg gaa gcc 288
Ala Lys Arg Lys Ala Val Leu Asn Lys Leu Ala Asp Leu Met Glu Ala
85 90 95
cac gcc gaa gag ctg gca ctg ctg gaa act ctc gac acc gcc aaa ccg 336

10

His	Ala	Glu	Glu	Leu	Ala	Leu	Leu	Glu	Thr	Leu	Asp	Thr	Gly	Lys	Pro		
			100					105					110				
att	cgt	cac	agt	ctg	cgt	gat	gat	att	ccc	ggc	gcg	gcg	cgc	gcc	att		384
Ile	Arg	His	Ser	Leu	Arg	Asp	Asp	Ile	Pro	Gly	Ala	Ala	Arg	Ala	Ile		
		115				120					125						
cgc	tgg	tac	gcc	gaa	gcg	atc	gac	aaa	gtg	tat	ggc	gaa	gtg	gcg	acc		432
Arg	Trp	Tyr	Ala	Glu	Ala	Ile	Asp	Lys	Val	Tyr	Gly	Glu	Val	Ala	Thr		
	130					135					140						
acc	agt	agc	cat	gag	ctg	gcg	atg	atc	gtg	cgt	gaa	ccg	gtc	ggc	gtg		480
Thr	Ser	Ser	His	Glu	Leu	Ala	Met	Ile	Val	Arg	Glu	Pro	Val	Gly	Val		
145					150				155						160		
att	gcc	gcc	atc	gtg	ccg	tgg	aac	ttc	ccg	ctg	ttg	ctg	act	tgc	tgg		528
Ile	Ala	Ala	Ile	Val	Pro	Trp	Asn	Phe	Pro	Leu	Leu	Leu	Thr	Cys	Trp		
			165					170						175			
aaa	ctc	ggc	ccg	gcg	ctg	gcg	gcg	gga	aac	agc	gtg	att	cta	aaa	ccg		576
Lys	Leu	Gly	Pro	Ala	Leu	Ala	Ala	Gly	Asn	Ser	Val	Ile	Leu	Lys	Pro		
			180					185					190				
tct	gaa	aaa	tca	ccg	ctc	agt	gcg	att	cgt	ctc	gcg	ggg	ctg	gcg	aaa		624
Ser	Glu	Lys	Ser	Pro	Leu	Ser	Ala	Ile	Arg	Leu	Ala	Gly	Leu	Ala	Lys		
		195					200					205					
gaa	gca	ggc	ttg	ccg	gat	ggt	gtg	ttg	aac	gtg	gtg	acg	ggt	ttt	ggt		672
Glu	Ala	Gly	Leu	Pro	Asp	Gly	Val	Leu	Asn	Val	Val	Thr	Gly	Phe	Gly		
	210					215				220							
cat	gaa	gcc	ggg	cag	gcg	ctg	tcg	cgt	cat	aac	gat	atc	gac	gcc	att		720
His	Glu	Ala	Gly	Gln	Ala	Leu	Ser	Arg	His	Asn	Asp	Ile	Asp	Ala	Ile		
225				230					235					240			
gcc	ttt	acc	ggt	tca	acc	cgt	acc	ggg	aaa	cag	ctg	ctg	aaa	gat	gcg		768
Ala	Phe	Thr	Gly	Ser	Thr	Arg	Thr	Gly	Lys	Gln	Leu	Leu	Lys	Asp	Ala		
			245					250						255			
ggc	gac	agc	aac	atg	aaa	cgc	gtc	tgg	ctg	gaa	gcg	ggc	ggc	aaa	agc		816
Gly	Asp	Ser	Asn	Met	Lys	Arg	Val	Trp	Leu	Glu	Ala	Gly	Gly	Lys	Ser		
			260					265					270				
gcc	aac	atc	gtt	ttc	gct	gac	tgc	ccg	gat	ttg	caa	cag	gcg	gca	agc		864
Ala	Asn	Ile	Val	Phe	Ala	Asp	Cys	Pro	Asp	Leu	Gln	Gln	Ala	Ala	Ser		
		275					280					285					
gcc	acc	gca	gca	ggc	att	ttc	tac	aac	cag	gga	cag	gtg	tgc	atc	gcc		912
Ala	Thr	Ala	Ala	Gly	Ile	Phe	Tyr	Asn	Gln	Gly	Gln	Val	Cys	Ile	Ala		
	290					295				300							
gga	acg	cgc	ctg	ttg	ctg	gaa	gag	agc	atc	gcc	gat	gaa	ttc	tta	gcc		960
Gly	Thr	Arg	Leu	Leu	Leu	Glu	Glu	Ser	Ile	Ala	Asp	Glu	Phe	Leu	Ala		
305				310					315					320			
ctg	tta	aaa	cag	cag	gcg	caa	aac	tgg	cag	ccg	ggc	cat	cca	ctt	gat		1008
Leu	Leu	Lys	Gln	Gln	Ala	Gln	Asn	Trp	Gln	Pro	Gly	His	Pro	Leu	Asp		
			325					330					335				
ccc	gca	acc	acc	atg	ggc	acc	tta	atc	gac	tgc	gcc	cac	gcc	gac	tcg		1056
Pro	Ala	Thr	Thr	Met	Gly	Thr	Leu	Ile	Asp	Cys	Ala	His	Ala	Asp	Ser		
			340					345					350				

gtc cat agc ttt att cgg gaa ggc gaa agc aaa ggg caa ctg ttg ttg 1104
Val His Ser Phe Ile Arg Glu Gly Glu Ser Lys Gly Gln Leu Leu Leu
355 360 365

gat ggc cgt aac gcc ggg ctg gct gcc gcc atc ggc ccg acc atc ttt 1152
Asp Gly Arg Asn Ala Gly Leu Ala Ala Ala Ile Gly Pro Thr Ile Phe
370 375 380

gtg gat gtg gac ccg aat gcg tcc tta agt cgc gaa gag att ttc ggt 1200
Val Asp Val Asp Pro Asn Ala Ser Leu Ser Arg Glu Glu Ile Phe Gly
385 390 395 400

ccg gtg ctg gtg gtc acg cgt ttc aca tca gaa gaa cag gcg cta cag 1248
Pro Val Leu Val Val Thr Arg Phe Thr Ser Glu Glu Gln Ala Leu Gln
405 410 415

ctt gcc aac gac agc cag tac ggc ctt ggc gcg gcg gta tgg acg cgc 1296
Leu Ala Asn Asp Ser Gln Tyr Gly Leu Gly Ala Ala Val Trp Thr Arg
420 425 430

gac ctc tcc cgc gcg cac cgc atg agc cga cgc ctg aaa gcc ggt tcc 1344
Asp Leu Ser Arg Ala His Arg Met Ser Arg Arg Leu Lys Ala Gly Ser
435 440 445

gtc ttc gtc aat aac tac aac gac ggc gat atg acc gtg ccg ttt ggc 1392
Val Phe Val Asn Asn Tyr Asn Asp Gly Asp Met Thr Val Pro Phe Gly
450 455 460

ggc tat aag cag agc ggc aac ggt cgc gac aaa tcc ctg cat gcc ctt 1440
Gly Tyr Lys Gln Ser Gly Asn Gly Arg Asp Lys Ser Leu His Ala Leu
465 470 475 480

gaa aaa ttc act gaa ctg aaa acc atc tgg ata agc ctg gag gcc tga 1488
Glu Lys Phe Thr Glu Leu Lys Thr Ile Trp Ile Ser Leu Glu Ala
485 490 495

<210> 20
<211> 495
<212> PRT
5 <213> Escherichia coli

<400> 20
Met Asn Phe His His Leu Ala Tyr Trp Gln Asp Lys Ala Leu Ser Leu
1 5 10 15

Ala Ile Glu Asn Arg Leu Phe Ile Asn Gly Glu Tyr Thr Ala Ala Ala
20 25 30

Glu Asn Glu Thr Phe Glu Thr Val Asp Pro Val Thr Gln Ala Pro Leu
35 40 45

Ala Lys Ile Ala Arg Gly Lys Ser Val Asp Ile Asp Arg Ala Met Ser
50 55 60

Ala Ala Arg Gly Val Phe Glu Arg Gly Asp Trp Ser Leu Ser Ser Pro
65 70 75 80

Ala Lys Arg Lys Ala Val Leu Asn Lys Leu Ala Asp Leu Met Glu Ala
85 90 95

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

Ile Arg His Ser Leu Arg Asp Asp Ile Pro Gly Ala Ala Arg Ala Ile
115 120 125

Arg Trp Tyr Ala Glu Ala Ile Asp Lys Val Tyr Gly Glu Val Ala Thr
130 135 140

Thr Ser Ser His Glu Leu Ala Met Ile Val Arg Glu Pro Val Gly Val
145 150 155 160

Ile Ala Ala Ile Val Pro Trp Asn Phe Pro Leu Leu Leu Thr Cys Trp
165 170 175

Lys Leu Gly Pro Ala Leu Ala Ala Gly Asn Ser Val Ile Leu Lys Pro
180 185 190

Ser Glu Lys Ser Pro Leu Ser Ala Ile Arg Leu Ala Gly Leu Ala Lys
195 200 205

Glu Ala Gly Leu Pro Asp Gly Val Leu Asn Val Val Thr Gly Phe Gly
210 215 220

His Glu Ala Gly Gln Ala Leu Ser Arg His Asn Asp Ile Asp Ala Ile
225 230 235 240

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

Gly Asp Ser Asn Met Lys Arg Val Trp Leu Glu Ala Gly Gly Lys Ser
260 265 270

Ala Asn Ile Val Phe Ala Asp Cys Pro Asp Leu Gln Gln Ala Ala Ser
275 280 285

Ala Thr Ala Ala Gly Ile Phe Tyr Asn Gln Gly Gln Val Cys Ile Ala
290 295 300

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

Leu Leu Lys Gln Gln Ala Gln Asn Trp Gln Pro Gly His Pro Leu Asp
325 330 335

Pro Ala Thr Thr Met Gly Thr Leu Ile Asp Cys Ala His Ala Asp Ser
340 345 350

Val His Ser Phe Ile Arg Glu Gly Glu Ser Lys Gly Gln Leu Leu Leu
355 360 365

Asp Gly Arg Asn Ala Gly Leu Ala Ala Ala Ile Gly Pro Thr Ile Phe
370 375 380

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

Pro Val Leu Val Val Thr Arg Phe Thr Ser Glu Glu Gln Ala Leu Gln
405 410 415

Leu Ala Asn Asp Ser Gln Tyr Gly Leu Gly Ala Ala Val Trp Thr Arg
420 425 430

Asp Leu Ser Arg Ala His Arg Met Ser Arg Arg Leu Lys Ala Gly Ser
435 440 445

Val Phe Val Asn Asn Tyr Asn Asp Gly Asp Met Thr Val Pro Phe Gly
450 455 460

Gly Tyr Lys Gln Ser Gly Asn Gly Arg Asp Lys Ser Leu His Ala Leu
465 470 475 480

Glu Lys Phe Thr Glu Leu Lys Thr Ile Trp Ile Ser Leu Glu Ala
485 490 495

<210> 21
<211> 1395
<212> DNA
<213> Escherichia coli

<400> 21
atgcctgacg ctaaaaaaca ggggcggtca aacaaggcaa tgacgttttt cgtctgcttc 60
cttgccgctc tggcgggatt actctttggc ctggatatcg gtgtaattgc tggcgactg 120
ccgtttattg cagatgaatt ccagattact tcgcacacgc aagaatgggt cgtaagctcc 180
atgatgttcg gtgcggcagt cgggtgcggtg ggcagcggct ggctctcctt taaactcggg 240
cgcaaaaaga gcctgatgat cggcgcaatt ttgtttgttg ccggttcgct gttctctgcg 300
gctgcgcaa acgttgaagt actgattctt tcccgcgttc tactggggct ggcggtgggt 360
gtggcctctt ataccgcacc gctgtacctc tctgaaattg cgccggaaaa aattcgtggc 420
agtatgatct cgatgtatca gttgatgatc actatcgga tcctcgggtgc ttatctttct 480

gataccgcct tcagctacac cggatgcattg cgctggatgc tgggtgtgat tatcatcccg 540
gcaattttgc tgctgattgg tgtctttctc ctgccagaca gccacgttg gtttgccgcc 600
aaacgccgtt ttgttgatgc cgaacgcgtg ctgctacgcc tgcgtgacac cagcgcggaa 660
gcgaaacgcg aactggatga aatccgtgaa agtttgacag ttaaacagag tggctgggcg 720
ctgtttaaag agaacagcaa cttccgccgc gcggtgttcc ttggcgact gttgcaggta 780
atgcagcaat tcaccgggat gaacgtcatc atgtattacg cgccgaaaat cttcgaactg 840
gcgggttata ccaacactac cgagcaaattg tgggggaccg tgattgtcgg cctgaccaac 900
gtacttgcca cctttatcgc aatcggcctt gttgaccgct ggggacgtaa accaacgcta 960
acgctgggct tcctggatg ggctgctggc atgggcgtac tcggtacaat gatgcatatc 1020
ggtattcact ctccgtcggc gcagtatttc gccatcgcca tgctgctgat gtttattgtc 1080
ggttttgcca tgagtgcggg tccgctgatt tgggtactgt gtcgcaaat tcagccgctg 1140
aaaggccgcg attttgcat cactgctcc actgccacca actggattgc caacatgac 1200
gttggcgcaa cgttcctgac catgctcaac acgctgggta acgccaacac cttctgggtg 1260
tatgcggctc tgaacgtact gtttatcctg ctgacattgt ggctgggtacc ggaaacaaaa 1320
cacgtttcgc tggaacatat tgaacgtaat ctgatgaaag gtcgtaaact gcgcgaaata 1380
ggcgtcacg attaa 1395

<210> 22
<211> 464
<212> PRT
5 <213> Escherichia coli

<400> 22
Met Pro Asp Ala Lys Lys Gln Gly Arg Ser Asn Lys Ala Met Thr Phe
1 5 10 15
Phe Val Cys Phe Leu Ala Ala Leu Ala Gly Leu Leu Phe Gly Leu Asp
20 25 30
Ile Gly Val Ile Ala Gly Ala Leu Pro Phe Ile Ala Asp Glu Phe Gln
35 40 45
Ile Thr Ser His Thr Gln Glu Trp Val Val Ser Ser Met Met Phe Gly
50 55 60
Ala Ala Val Gly Ala Val Gly Ser Gly Trp Leu Ser Phe Lys Leu Gly
65 70 75 80
Arg Lys Lys Ser Leu Met Ile Gly Ala Ile Leu Phe Val Ala Gly Ser
85 90 95

ES 2 552 317 T3

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

340	345	350
Ala Met Leu Leu Met Phe Ile Val Gly Phe Ala Met Ser Ala Gly Pro		
355	360	365
Leu Ile Trp Val Leu Cys Ser Glu Ile Gln Pro Leu Lys Gly Arg Asp		
370	375	380
Phe Gly Ile Thr Cys Ser Thr Ala Thr Asn Trp Ile Ala Asn Met Ile		
385	390	400
Val Gly Ala Thr Phe Leu Thr Met Leu Asn Thr Leu Gly Asn Ala Asn		
405	410	415
Thr Phe Trp Val Tyr Ala Ala Leu Asn Val Leu Phe Ile Leu Leu Thr		
420	425	430
Leu Trp Leu Val Pro Glu Thr Lys His Val Ser Leu Glu His Ile Glu		
435	440	445
Arg Asn Leu Met Lys Gly Arg Lys Leu Arg Glu Ile Gly Ala His Asp		
450	455	460

<210> 23
 <211> 1290
 <212> DNA
 5 <213> Pseudomonas fluorescens

<400> 23

atgcagtttg ccgccaaacg cgagtactgg cttatcagtg gtttggtggt tttcttcttc	60
ttttcgtggt catccagcta ttcattgttt tctatctggc tgcacgagt cattggcttg	120
aatggcacgg aaaccggctt cattttcgcc gccaacgcta ttgcggcgct gctggttcaa	180
cccttctacg gcgccttca agaccgcctc gggctgtcca aaaagcttct ggtgtggatt	240
ggcatcctgc tgtgtgccgc ggccccgttt gcaatttatg tctacgccgg cctgttggcg	300
cagaacgtga tgctcggcgc gttggtcggg gcggcggttc tggcgctggc gatgctggca	360
ggcgttgggg tgatcgagtc gtacaccgag cgcttgctgc ggcacgcagg attcgagttt	420
ggaaccaccc gaatgtgggg gtcgttgggc tgggccagcg cgacgggctt ggtcggcggt	480
gtgttcaaca tcgatcctga cattgcgttt tacatgagca gcctcgccgg catcgtgttt	540
ttgctgatcc tgttccgtct ggacctcgac cggttggccc agccggcagt gcaggcgggc	600
gcggttggtc acccgtgcg cctgaacgat ctctggaagt tgctggcact cccgcgggtc	660
tgggctttca gcctttacct gacgggggta tgccggatct acatgatcta cgagcaacag	720
tttcgggtgt atttctctc gtttttccc acccggagg aggggacctg tgctatggc	780

tacctgaact cgtctcaggt actggtcgag gcggtcctga tgctgcttgc accctgggtg 840
 gtcagccgca caggcgccaa atacgggctg attctggccg gcagcatcat gttcgtgcgc 900
 atccttgggt cggggctggt aacgcaggct tgggccatcg ccgcctgcaa gatgttgac 960
 gccttgaag tgcccatctt gctggtctcg atattcaa acatttcgct caactttgac 1020
 tctoggctgt ccgcctcgat ctacttggtg gggttccagt tcgcccagca actgaccgcc 1080
 atgttgctgt caccgctggt gggctacggc tacgaccatt tcggtttctc cagcgtctac 1140
 gtactgatgg caggcctggt cggcgcttgc ctgctgcttt catggacctt gttgcgcaag 1200
 gaccccgtag gtgacgcctc tcaagtcggg gctggcgatt cacggcagct tcccgccatc 1260
 gcgccatccg cccctcgta tgaaccctag 1290

<210> 24
 <211> 429
 <212> PRT
 5 <213> Pseudomonas fluorescens

<400> 24
 Met Gln Phe Ala Ala Lys Arg Glu Tyr Trp Leu Ile Ser Gly Leu Leu
 1 5 10 15
 Phe Phe Phe Phe Phe Ser Trp Ser Ser Ser Tyr Ser Leu Phe Ser Ile
 20 25 30
 Trp Leu His Arg Val Ile Gly Leu Asn Gly Thr Glu Thr Gly Phe Ile
 35 40 45
 Phe Ala Ala Asn Ala Ile Ala Ala Leu Leu Val Gln Pro Phe Tyr Gly
 50 55 60
 Ala Leu Gln Asp Arg Leu Gly Leu Ser Lys Lys Leu Leu Val Trp Ile
 65 70 75 80
 Gly Ile Leu Leu Cys Ala Ala Ala Pro Phe Ala Ile Tyr Val Tyr Ala
 85 90 95
 Gly Leu Leu Ala Gln Asn Val Met Leu Gly Ala Leu Val Gly Ala Ala
 100 105 110
 Phe Leu Ala Leu Ala Met Leu Ala Gly Val Gly Val Ile Glu Ser Tyr
 115 120 125
 Thr Glu Arg Leu Ser Arg His Ala Gly Phe Glu Phe Gly Thr Thr Arg
 130 135 140
 Met Trp Gly Ser Leu Gly Trp Ala Ser Ala Thr Gly Val Val Gly Val

145		150		155		160
Val Phe Asn Ile	Asp Pro Asp Ile Ala Phe Tyr Met Ser Ser Leu Ala	165		170		175
Gly Ile Val Phe Leu Leu Ile Leu Phe Arg Leu Asp Leu Asp Arg Leu		180		185		190
Ala Gln Pro Ala Val Gln Ala Gly Ala Val Val His Pro Val Arg Leu		195		200		205
Asn Asp Leu Trp Lys Leu Leu Ala Leu Pro Arg Phe Trp Ala Phe Ser		210		215		220
Leu Tyr Leu Thr Gly Val Cys Gly Ile Tyr Met Ile Tyr Glu Gln Gln		225		230		235
Phe Pro Val Tyr Phe Ser Ser Phe Phe Pro Thr Pro Glu Glu Gly Thr		245		250		255
Arg Ala Tyr Gly Tyr Leu Asn Ser Ser Gln Val Leu Val Glu Ala Val		260		265		270
Leu Met Leu Leu Ala Pro Trp Val Val Ser Arg Thr Gly Ala Lys Tyr		275		280		285
Gly Leu Ile Leu Ala Gly Ser Ile Met Phe Val Arg Ile Leu Gly Ser		290		295		300
Gly Leu Val Thr Gln Ala Trp Ala Ile Ala Ala Cys Lys Met Leu His		305		310		315
Ala Leu Glu Val Pro Ile Leu Leu Val Ser Ile Phe Lys Tyr Ile Ser		325		330		335
Leu Asn Phe Asp Ser Arg Leu Ser Ala Ser Ile Tyr Leu Val Gly Phe		340		345		350
Gln Phe Ala Gln Gln Leu Thr Ala Met Leu Leu Ser Pro Leu Val Gly		355		360		365
Tyr Gly Tyr Asp His Phe Gly Phe Ser Ser Val Tyr Val Leu Met Ala		370		375		380
Gly Leu Val Gly Ala Cys Leu Leu Leu Ser Trp Thr Leu Leu Arg Lys		385		390		395
						400

Asp Pro Val Arg Asp Ala Ser Gln Val Gly Ala Gly Asp Ser Arg Gln
405 410 415

Leu Pro Ala Ile Ala Pro Ser Ala Pro Arg Tyr Glu Pro
420 425

<210> 25
<211> 1476
<212> DNA
5 <213> Pseudomonas fluorescens

<400> 25
atgcacgctg cactgttaga gcaagcgcat cgcgctattg aaaaaaact gcctgggcgg 60
ggatgatgtct atgcgctggc ctatcatctt gcgccgcggg tggggtggat gaatgacccg 120
aacggtctgg tttattttgc cggcgagtag catgtgttct accaacatca tccctattcg 180
gctcagtggg ggccgatgca ctggggccat gccaaagagcc gtgacctggt gcactgggag 240
cacctgcccc tcgcgctggc gccgggcgag gcctatgacc gcgacgggtg cttttcaggg 300
tctgcggtgg tcattggacga cgtgttgtac ctgatttaca ccgggcatac ctggctgggt 360
gcgcccgggt acgagcggag cattcgccag gttcagtgcc tggccagcag caccgacggg 420
gttgcggttca gcaagcacgg gccggtgatc gataggcgcc ctgaaccggg catcatgcat 480
tttcgcgacc ccaaggtagt gcggcgagga gagcaatggt ggatggccct gggggcgccg 540
caaggcgacg cccctcagct cctgctctat cgctcaggcg acctgcatca ctggacgtac 600
ctcaggtgag cactgcaagg gcaacgagag tcggacggct atatgtggga gtgtcctgac 660
ctgttcgaac tcgatggctg tgatgtgttt ctctattcgc ctcaaggctt gaaccccagc 720
ggttatgaca actggaacaa gttccagaac agctatcgga tgggcctgct ggaagatcgc 780
ggatacttca gcgagggcgg tgagctgagt gaactggatc atggtcacga tttctatgag 840
gcgcagacct tgctggcgcc agacgggcga cgcctgttgt gggcttggat ggacatgtgg 900
gacagcccga tgccgagtca ggcgcaacac ttgtgcggtg cgctgtcgct acctcgtgaa 960
ctgagccgca atggcgaaac gctacgcagt cggccggccc gcgagttggc agcgctacgc 1020
cagtcgcaac ggacactggc gatcggcgtg gtcgaatccg gcaattgcat actcgtgag 1080
cgagggggcg tgctggaatt cgaactgacc ctggacctgg ctggtagcac ggctgagcgt 1140
ttcgggttgg cgctgcgttg tagtgaggat cggcaagagc ggaccctggt gtacttcgat 1200
gcgatggcgc ggcgtctggt gctggacagg caacactcgg gagcgggggt aagcgggtgcg 1260
cgcagcgtgc cgatagccaa gggccaaatg cagatagcct tgcggatttt ccttgatcga 1320
tcctccattg aggtgtttgt cgatgacgga gcctatagct tgagcagtcg gatctacctt 1380
agccccgaca gcgtggcggg catggcggtt gcggtcaatg gtagcgggtg ttttggccaa 1440
gcgtcgggtc ggcacctggc cgatctgcac ctgtga 1476

<210> 26
<211> 491
10 <212> PRT

ES 2 552 317 T3

<213> Pseudomonas fluorescens

<400> 26

Met His Ala Ala Leu Leu Glu Gln Ala His Arg Ala Ile Glu Lys Lys
1 5 10 15

Leu Pro Gly Arg Gly Asp Val Tyr Arg Leu Ala Tyr His Leu Ala Pro
20 25 30

Pro Val Gly Trp Met Asn Asp Pro Asn Gly Leu Val Tyr Phe Arg Gly
35 40 45

Glu Tyr His Val Phe Tyr Gln His His Pro Tyr Ser Ala Gln Trp Gly
50 55 60

Pro Met His Trp Gly His Ala Lys Ser Arg Asp Leu Val His Trp Glu
65 70 75 80

His Leu Pro Ile Ala Leu Ala Pro Gly Glu Ala Tyr Asp Arg Asp Gly
85 90 95

Cys Phe Ser Gly Ser Ala Val Val Met Asp Asp Val Leu Tyr Leu Ile
100 105 110

Tyr Thr Gly His Thr Trp Leu Gly Ala Pro Gly Asp Glu Arg Ser Ile
115 120 125

Arg Gln Val Gln Cys Leu Ala Ser Ser Thr Asp Gly Val Ala Phe Ser
130 135 140

Lys His Gly Pro Val Ile Asp Arg Ala Pro Glu Pro Gly Ile Met His
145 150 155 160

Phe Arg Asp Pro Lys Val Trp Arg Arg Gly Glu Gln Trp Trp Met Ala
165 170 175

Leu Gly Ala Arg Gln Gly Asp Ala Pro Gln Leu Leu Leu Tyr Arg Ser
180 185 190

Gly Asp Leu His His Trp Thr Tyr Leu Arg Cys Ala Leu Gln Gly Gln
195 200 205

Arg Glu Ser Asp Gly Tyr Met Trp Glu Cys Pro Asp Leu Phe Glu Leu
210 215 220

Asp Gly Cys Asp Val Phe Leu Tyr Ser Pro Gln Gly Leu Asn Pro Ser
225 230 235 240

Gly Tyr Asp Asn Trp Asn Lys Phe Gln Asn Ser Tyr Arg Met Gly Leu
245 250 255

Leu Asp Asp Arg Gly Tyr Phe Ser Glu Gly Gly Glu Leu Arg Glu Leu
260 265 270

Asp His Gly His Asp Phe Tyr Ala Ala Gln Thr Leu Leu Ala Pro Asp
275 280 285

Gly Arg Arg Leu Leu Trp Ala Trp Met Asp Met Trp Asp Ser Pro Met
290 295 300

Pro Ser Gln Ala Gln His Trp Cys Gly Ala Leu Ser Leu Pro Arg Glu
305 310 315 320

Leu Ser Arg Asn Gly Glu Arg Leu Arg Met Arg Pro Ala Arg Glu Leu
325 330 335

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

Ser Gly Asn Cys Ile Leu Ala Glu Arg Gly Ala Leu Leu Glu Phe Glu
355 360 365

Leu Thr Leu Asp Leu Ala Gly Ser Thr Ala Glu Arg Phe Gly Leu Ala
370 375 380

Leu Arg Cys Ser Glu Asp Arg Gln Glu Arg Thr Leu Val Tyr Phe Asp
385 390 395 400

Ala Met Ala Arg Arg Leu Val Leu Asp Arg Gln His Ser Gly Ala Gly
405 410 415

Val Ser Gly Ala Arg Ser Val Pro Ile Ala Lys Gly Gln Met Gln Ile
420 425 430

Ala Leu Arg Ile Phe Leu Asp Arg Ser Ser Ile Glu Val Phe Val Asp
435 440 445

Asp Gly Ala Tyr Ser Leu Ser Ser Arg Ile Tyr Pro Ser Pro Asp Ser
450 455 460

Val Ala Val Met Ala Phe Ala Val Asn Gly Ser Gly Gly Phe Gly Gln
465 470 475 480

Ala Ser Val Trp His Leu Ala Asp Leu His Leu
485 490

ES 2 552 317 T3

<210> 27
 <211> 1248
 <212> DNA
 <213> Escherichia coli

5 <400> 27
 atggcactga atattccatt cagaaatgcg tactatcggt ttgcatccag ttactcattt 60
 ctcttttttta tttcctgggc gctgtggtgg tcgttatacg ctatttggct gaaaggacat 120
 ctaggattaa cagggacgga attaggtaca ctttattcgg tcaaccagtt taccagcatt 180
 ctatttatga tgttctacgg catcggtcag gataaactcg gtctgaagaa accgctcatc 240
 tgggtgatga gtttcattct ggtcctgacc ggaccgttta tgatttacgt ttatgaaccg 300
 ttactgcaaa gcaatttttc tgtaggtcta attctggggg cgctcttttt tggcctgggg 360
 tatctggcgg gatgcggttt gcttgacagc ttcaccgaaa aaatggcgcg aaattttcat 420
 ttcgaatatg gaacagcgcg cgctgggga tcttttggt atgctattgg cgcgttcttt 480
 gccggtatat tttttagtat cagtcccat atcaacttct ggttgggtct gctatttggc 540
 gctgtattta tgatgatcaa catgcgtttt aaagataagg atcaccagtg catagcggcg 600
 gatgcgggag gggtaaaaaa agaggatttt atcgcagttt tcaaggatcg aaacttctgg 660
 gttttcgtca tatttattgt ggggacgtgg tctttctata acatttttga tcaacaactc 720
 tttcctgtct tttatgcagg tttattcgaa tcacacgatg taggaacgcg cctgtatggt 780
 tatctcaact cattccaggt ggtactcgaa gcgctgtgca tggcgattat tcctttcttt 840
 gtgaatcggg tagggccaaa aaatgcatta cttatcggtg ttgtgattat ggcgttgcg 900
 atcctttcct gcgcgttggt cgtaacccc tggattattt cattagtga gctgttacat 960
 gccattgagg ttccactttg tgtcatatcc gtcttcaaat acagcgtggc aaactttgat 1020
 aagcgcctgt cgtcgacgat ctttctgatt ggttttcaaa ttgccagttc gcttgggatt 1080
 gtgctgcttt caacgccgac tgggatactc tttgaccacg caggctacca gacagttttc 1140
 ttcgcaattt cgggtattgt ctgcctgatg ttgctatttg gcattttctt cctgagtaaa 1200
 aaacgcgagc aaatagttat ggaaacgcct gtaccttcag caatatag 1248

<210> 28
 <211> 415
 <212> PRT
 <213> Escherichia coli

10

<400> 28

Met Ala Leu Asn Ile Pro Phe Arg Asn Ala Tyr Tyr Arg Phe Ala Ser
1 5 10 15

Ser Tyr Ser Phe Leu Phe Phe Ile Ser Trp Ser Leu Trp Trp Ser Leu
20 25 30

Tyr Ala Ile Trp Leu Lys Gly His Leu Gly Leu Thr Gly Thr Glu Leu
35 40 45

Gly Thr Leu Tyr Ser Val Asn Gln Phe Thr Ser Ile Leu Phe Met Met
50 55 60

Phe Tyr Gly Ile Val Gln Asp Lys Leu Gly Leu Lys Lys Pro Leu Ile
65 70 75 80

Trp Cys Met Ser Phe Ile Leu Val Leu Thr Gly Pro Phe Met Ile Tyr
85 90 95

Val Tyr Glu Pro Leu Leu Gln Ser Asn Phe Ser Val Gly Leu Ile Leu
100 105 110

Gly Ala Leu Phe Phe Gly Leu Gly Tyr Leu Ala Gly Cys Gly Leu Leu
115 120 125

Asp Ser Phe Thr Glu Lys Met Ala Arg Asn Phe His Phe Glu Tyr Gly
130 135 140

Thr Ala Arg Ala Trp Gly Ser Phe Gly Tyr Ala Ile Gly Ala Phe Phe
145 150 155 160

Ala Gly Ile Phe Phe Ser Ile Ser Pro His Ile Asn Phe Trp Leu Val
165 170 175

Ser Leu Phe Gly Ala Val Phe Met Met Ile Asn Met Arg Phe Lys Asp
180 185 190

Lys Asp His Gln Cys Ile Ala Ala Asp Ala Gly Gly Val Lys Lys Glu
195 200 205

Asp Phe Ile Ala Val Phe Lys Asp Arg Asn Phe Trp Val Phe Val Ile
210 215 220

Phe Ile Val Gly Thr Trp Ser Phe Tyr Asn Ile Phe Asp Gln Gln Leu
225 230 235 240

Phe Pro Val Phe Tyr Ala Gly Leu Phe Glu Ser His Asp Val Gly Thr
245 250 255

Arg Leu Tyr Gly Tyr Leu Asn Ser Phe Gln Val Val Leu Glu Ala Leu
260 265 270

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

Ala Leu Leu Ile Gly Val Val Ile Met Ala Leu Arg Ile Leu Ser Cys
290 295 300

Ala Leu Phe Val Asn Pro Trp Ile Ile Ser Leu Val Lys Leu Leu His
305 310 315 320

Ala Ile Glu Val Pro Leu Cys Val Ile Ser Val Phe Lys Tyr Ser Val
325 330 335

Ala Asn Phe Asp Lys Arg Leu Ser Ser Thr Ile Phe Leu Ile Gly Phe
340 345 350

Gln Ile Ala Ser Ser Leu Gly Ile Val Leu Leu Ser Thr Pro Thr Gly
355 360 365

Ile Leu Phe Asp His Ala Gly Tyr Gln Thr Val Phe Phe Ala Ile Ser
370 375 380

Gly Ile Val Cys Leu Met Leu Leu Phe Gly Ile Phe Phe Leu Ser Lys
385 390 395 400

Lys Arg Glu Gln Ile Val Met Glu Thr Pro Val Pro Ser Ala Ile
405 410 415

<210> 29
<211> 1434
<212> DNA
5 <213> Escherichia coli

<400> 29
atgacgcaat ctcgattgca tgcggcgcaa aacgcactag caaaacttca cgagcgccga 60
ggtaacactt tctatcccca ttttcacctc gcgcctcctg ccgggtggat gaacgatcca 120
aacggcctga tctggtttta cgatcgttat cagcggtttt atcaacatca cccgatgagc 180
gaacactggg ggccaatgca ctggggacat gccaccagcg acgatatgat ccaactggcag 240
catgagccta ttgcgctagc gccaggagac gagaatgaca aagacgggtg tttttcaggt 300
agtgtgtgtg atgacaatgg tgtcctctca cttatctaca ccggacacgt ctggctcgat 360
ggtgcaggta atgacgatgc aattcgcgaa gtacaatgtc tggctaccag tcgggatggg 420
attcatttcg agaaacaggg tgtgatcctc actccaccag aaggcatcat gcacttccgc 480

gatcctaaag tgtggcgtga agccgacaca tgggtggatgg tagtcggggc gaaagaccca 540
ggcaacacgg ggcagatcct gctttatcgc ggcagttcat tgcgtgaatg gactttcgat 600
cgcgtactgg cccacgctga tgcgggtgaa agctatatgt gggaatgtcc ggactttttc 660
agccttggcg atcagcatta tctgatgttt tccccgcagg gaatgaatgc cgagggatac 720
agttatcgaa atcgctttca aagtggcgta ataccggaa tgtggtcgcc aggacgactt 780
tttgacaaat ccggggcattt tactgaactt gataacgggc atgactttta tgcaccacaa 840
agctttgtag cgaaggatgg tcggcgtatt gttatcggct ggatggatat gtgggaatcg 900
ccaatgccct caaaacgtga aggctgggca ggctgcatga cgctggcgcg cgagctatca 960
gagagcaatg gcaaactcct acaacgcccg gtacacgaag ctgagtcggt acgccagcag 1020
catcaatcta tctctccccg cacaatcagc aataaatatg ttttgcagga aaacgcgcaa 1080
gcagttgaga ttcagttgca gtgggcgctg aagaacagtg atgccgaaca ttacggatta 1140
cagctcggcg ctggaatgcg gctgtatatt gataaccaat ctgagcgact tgttttgtgg 1200
cgggtattacc cacacgagaa tttagatggc taccgtagta ttccccctcc gcaggggtgac 1260
atgctcggcc taaggatatt tatcgatata tcatccgtgg aagtatttat taacgacggg 1320
gaggcgggtga tgagtggcg aatatatccg cagccagaag aacgggaact gtcgctctat 1380
gcctcccacg gagtggctgt gctgcaacat ggagcactct ggcaactggg ttaa 1434

<210> 30
<211> 477
<212> PRT
5 <213> Escherichia coli

<400> 30
Met Thr Gln Ser Arg Leu His Ala Ala Gln Asn Ala Leu Ala Lys Leu
1 5 10 15
His Glu Arg Arg Gly Asn Thr Phe Tyr Pro His Phe His Leu Ala Pro
20 25 30
Pro Ala Gly Trp Met Asn Asp Pro Asn Gly Leu Ile Trp Phe Asn Asp
35 40 45
Arg Tyr His Ala Phe Tyr Gln His His Pro Met Ser Glu His Trp Gly
50 55 60
Pro Met His Trp Gly His Ala Thr Ser Asp Asp Met Ile His Trp Gln
65 70 75 80
His Glu Pro Ile Ala Leu Ala Pro Gly Asp Glu Asn Asp Lys Asp Gly
85 90 95

Cys Phe Ser Gly Ser Ala Val Asp Asp Asn Gly Val Leu Ser Leu Ile
 100 105 110
 Tyr Thr Gly His Val Trp Leu Asp Gly Ala Gly Asn Asp Asp Ala Ile
 115 120 125
 Arg Glu Val Gln Cys Leu Ala Thr Ser Arg Asp Gly Ile His Phe Glu
 130 135 140
 Lys Gln Gly Val Ile Leu Thr Pro Pro Glu Gly Ile Met His Phe Arg
 145 150 155 160
 Asp Pro Lys Val Trp Arg Glu Ala Asp Thr Trp Trp Met Val Val Gly
 165 170 175
 Ala Lys Asp Pro Gly Asn Thr Gly Gln Ile Leu Leu Tyr Arg Gly Ser
 180 185 190
 Ser Leu Arg Glu Trp Thr Phe Asp Arg Val Leu Ala His Ala Asp Ala
 195 200 205
 Gly Glu Ser Tyr Met Trp Glu Cys Pro Asp Phe Phe Ser Leu Gly Asp
 210 215 220
 Gln His Tyr Leu Met Phe Ser Pro Gln Gly Met Asn Ala Glu Gly Tyr
 225 230 235 240
 Ser Tyr Arg Asn Arg Phe Gln Ser Gly Val Ile Pro Gly Met Trp Ser
 245 250 255
 Pro Gly Arg Leu Phe Ala Gln Ser Gly His Phe Thr Glu Leu Asp Asn
 260 265 270
 Gly His Asp Phe Tyr Ala Pro Gln Ser Phe Val Ala Lys Asp Gly Arg
 275 280 285
 Arg Ile Val Ile Gly Trp Met Asp Met Trp Glu Ser Pro Met Pro Ser
 290 295 300
 Lys Arg Glu Gly Trp Ala Gly Cys Met Thr Leu Ala Arg Glu Leu Ser
 305 310 315 320
 Glu Ser Asn Gly Lys Leu Leu Gln Arg Pro Val His Glu Ala Glu Ser
 325 330 335
 Leu Arg Gln Gln His Gln Ser Ile Ser Pro Arg Thr Ile Ser Asn Lys
 340 345 350

Tyr Val Leu Gln Glu Asn Ala Gln Ala Val Glu Ile Gln Leu Gln Trp
355 360 365

Ala Leu Lys Asn Ser Asp Ala Glu His Tyr Gly Leu Gln Leu Gly Ala
370 375 380

Gly Met Arg Leu Tyr Ile Asp Asn Gln Ser Glu Arg Leu Val Leu Trp
385 390 395 400

Arg Tyr Tyr Pro His Glu Asn Leu Asp Gly Tyr Arg Ser Ile Pro Leu
405 410 415

Pro Gln Gly Asp Met Leu Ala Leu Arg Ile Phe Ile Asp Thr Ser Ser
420 425 430

Val Glu Val Phe Ile Asn Asp Gly Glu Ala Val Met Ser Ser Arg Ile
435 440 445

Tyr Pro Gln Pro Glu Glu Arg Glu Leu Ser Leu Tyr Ala Ser His Gly
450 455 460

Val Ala Val Leu Gln His Gly Ala Leu Trp Gln Leu Gly
465 470 475

<210> 31

<211> 927

<212> DNA

5 <213> Agrobacterium tumefaciens

<400> 31

atgatcctgt gttgtggtga agccctgata gacatgctgc cccggcagac gacgctgggt	60
gaggcgggct ttgcccctta cgcaggcgga gcggtcttca acacggcaat tgcgctgggg	120
cgtcttggeg tcccttcagc cttttttacc ggtctttccg acgacatgat gggcgatata	180
ctgcgggaga ccctgcgggc cagcaagggt gatttcagct attgcgccac cctgtcgcgc	240
cccaccacca ttgcgttcgt taagctggtt gatggccatg cgacctacgc tttttacgac	300
gagaacaccg ccggccgat gatcaccgag gccgaacttc cggccttggg agcggattgc	360
gaagcgctgc atttcggcgc catcagcctt attccogaac cctgcggcag cacctatgag	420
gcgctgatga cgcgcgagca tgagaccgc gtcactctgc tcgatccgaa cattcgtccc	480
ggcttcatcc agaacaagca gtcgcacatg gcccgcatcc gccgcatggc ggcgatgtct	540
gacatcgtca agttctcgga tgaggacctg gcgtggttcg gtctggaagg cgacgaggac	600
acgcttgccc gccactggct gcaccacggt gcaaaactcg tcgttgtcac ccgtggcgcc	660
aagggtgccg tgggttacag cgccaatctc aagggtggaag tggcctccga gcgcgtcgaa	720

gtgggtcgata cggtcggcgc cggcgatacg ttccgatgccg gcattcttgc ttccgtgaaa 780
 atgcagggcc tgcagaccaa agcgcaggtg gcttcgctga gcgaagagca gatcagaaaa 840
 gctttggcgc ttggcgcgaa agccgctgcg gtcactgtct cgcgggctgg cgcaaaccg 900
 cctttcgcgc atgaaatcgg tttgtga 927

<210> 32
 <211> 308
 <212> PRT
 5 <213> Agrobacterium tumefaciens

<400> 32
 Met Ile Leu Cys Cys Gly Glu Ala Leu Ile Asp Met Leu Pro Arg Gln
 1 5 10 15

Thr Thr Leu Gly Glu Ala Gly Phe Ala Pro Tyr Ala Gly Gly Ala Val
 20 25 30

Phe Asn Thr Ala Ile Ala Leu Gly Arg Leu Gly Val Pro Ser Ala Phe
 35 40 45

Phe Thr Gly Leu Ser Asp Asp Met Met Gly Asp Ile Leu Arg Glu Thr
 50 55 60

Leu Arg Ala Ser Lys Val Asp Phe Ser Tyr Cys Ala Thr Leu Ser Arg
 65 70 75 80

Pro Thr Thr Ile Ala Phe Val Lys Leu Val Asp Gly His Ala Thr Tyr
 85 90 95

Ala Phe Tyr Asp Glu Asn Thr Ala Gly Arg Met Ile Thr Glu Ala Glu
 100 105 110

Leu Pro Ala Leu Gly Ala Asp Cys Glu Ala Leu His Phe Gly Ala Ile
 115 120 125

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

Arg Glu His Glu Thr Arg Val Ile Ser Leu Asp Pro Asn Ile Arg Pro
 145 150 155 160

Gly Phe Ile Gln Asn Lys Gln Ser His Met Ala Arg Ile Arg Arg Met
 165 170 175

Ala Ala Met Ser Asp Ile Val Lys Phe Ser Asp Glu Asp Leu Ala Trp
 180 185 190

Phe Gly Leu Glu Gly Asp Glu Asp Thr Leu Ala Arg His Trp Leu His
195 200 205

His Gly Ala Lys Leu Val Val Val Thr Arg Gly Ala Lys Gly Ala Val
210 215 220

Gly Tyr Ser Ala Asn Leu Lys Val Glu Val Ala Ser Glu Arg Val Glu
225 230 235 240

Val Val Asp Thr Val Gly Ala Gly Asp Thr Phe Asp Ala Gly Ile Leu
245 250 255

Ala Ser Leu Lys Met Gln Gly Leu Leu Thr Lys Ala Gln Val Ala Ser
260 265 270

Leu Ser Glu Glu Gln Ile Arg Lys Ala Leu Ala Leu Gly Ala Lys Ala
275 280 285

Ala Ala Val Thr Val Ser Arg Ala Gly Ala Asn Pro Pro Phe Ala His
290 295 300

Glu Ile Gly Leu
305

- <210> 33
<211> 1404
<212> DNA
5 <213> Streptococcus mutans

<400> 33
cagctgatta tgcgctcagtt gaaaccctcg cttcttcagg aactgttgct gtaggtgata 60
gcttacttga agttaaaaa taagaaatat tadcagaaag accgtaaggt ctttttgact 120
gcttaaaaga ttcagtaaca atagtattaa agccttttg gctaactaata cttgaaattt 180
agcaaattat gatataatgt taagtagtcc ttaagggtag attaagggtta ttcaaattcca 240
aaaattgatt tggttaagtt agtaaaatat aagaggttta ttatgtctaa attatatggc 300
agcatcgaag ctggcggaac aaaatttgct tgtgctgtag gtgatgaaaa ttttcaaatt 360
ttagaaaaag ttcagttccc aacaacaaca ctttatgaaa caatagaaaa aacagttgct 420
ttctttaaaa aatttgaagc tgatttagcc agtggtgcca ttggttcttt tggccctatt 480
gatattgatc aaaattcaga cacttatggt tacattactt caacaccaa gccaaactgg 540
gctaacgttg attttgcgg ctttaatttct aaagatttta aaattccatt ttactttacg 600
acagatgtta attcttctgc ttatggggaa acaattgctc gttcaaattgt taaaagtctg 660
gtttattata ctattggaac aggcattgga gcaggggcta ttcaaaatgg cgaattcatt 720

ggcggatatgg gacatacggg agctggacac gtttacatgg ctccgcatcc caatgatgtt 780
catcatgggtt ttgtaggcac ctgtcctttc cataaaggct gtttagaagg acttgcagcg 840
ggtcctagct tagaggctcg tactgggtatt cgtgggtgagt taattgagca aaactcagaa 900
gtttgggata ttcaggcata ctacattgct caggcgggcta ttcaagcgac tgtcctttat 960
cgtccgcaag tcattgtatt tggcggaggc gttatggcac aagaacatat gctcaatcgg 1020
gttcgtgaaa aatttacttc acttttgaat gactatcttc cagttccaga tgttaaagat 1080
tatattgtga caccagctgt tgcagaaaaat gggttcagcaa cattgggaaa tctcgcttta 1140
gctaaaaaga tagcagcgcg ttaattaaat atgaattgga agattaaagc accttctaata 1200
attcaatatt aaactgttag aatttacgtg aacgaaattt tcattttatg aggataatga 1260
agtgaatata attactcttg atttcctctg aaactagata gtggatatatt gaaaaacaga 1320
aaggagaaca ctatggaagg acctttgttt ttacaatcac aaatgcataa aaaaatctgg 1380
ggcggcaatc ggctcagaaa agaa 1404

<210> 34
<211> 293
<212> PRT
5 <213> Streptococcus mutans

<400> 34
Met Ser Lys Leu Tyr Gly Ser Ile Glu Ala Gly Gly Thr Lys Phe Val
1 5 10 15
Cys Ala Val Gly Asp Glu Asn Phe Gln Ile Leu Glu Lys Val Gln Phe
20 25 30
Pro Thr Thr Thr Pro Tyr Glu Thr Ile Glu Lys Thr Val Ala Phe Phe
35 40 45
Lys Lys Phe Glu Ala Asp Leu Ala Ser Val Ala Ile Gly Ser Phe Gly
50 55 60
Pro Ile Asp Ile Asp Gln Asn Ser Asp Thr Tyr Gly Tyr Ile Thr Ser
65 70 75 80
Thr Pro Lys Pro Asn Trp Ala Asn Val Asp Phe Val Gly Leu Ile Ser
85 90 95
Lys Asp Phe Lys Ile Pro Phe Tyr Phe Thr Thr Asp Val Asn Ser Ser
100 105 110
Ala Tyr Gly Glu Thr Ile Ala Arg Ser Asn Val Lys Ser Leu Val Tyr
115 120 125

Tyr Thr Ile Gly Thr Gly Ile Gly Ala Gly Ala Ile Gln Asn Gly Glu
130 135 140

Phe Ile Gly Gly Met Gly His Thr Glu Ala Gly His Val Tyr Met Ala
145 150 155 160

Pro His Pro Asn Asp Val His His Gly Phe Val Gly Thr Cys Pro Phe
165 170 175

His Lys Gly Cys Leu Glu Gly Leu Ala Ala Gly Pro Ser Leu Glu Ala
180 185 190

Arg Thr Gly Ile Arg Gly Glu Leu Ile Glu Gln Asn Ser Glu Val Trp
195 200 205

Asp Ile Gln Ala Tyr Tyr Ile Ala Gln Ala Ala Ile Gln Ala Thr Val
210 215 220

Leu Tyr Arg Pro Gln Val Ile Val Phe Gly Gly Gly Val Met Ala Gln
225 230 235 240

Glu His Met Leu Asn Arg Val Arg Glu Lys Phe Thr Ser Leu Leu Asn
245 250 255

Asp Tyr Leu Pro Val Pro Asp Val Lys Asp Tyr Ile Val Thr Pro Ala
260 265 270

Val Ala Glu Asn Gly Ser Ala Thr Leu Gly Asn Leu Ala Leu Ala Lys
275 280 285

Lys Ile Ala Ala Arg
290

<210> 35
<211> 915
<212> DNA
5 <213> Escherichia coli

<400> 35
atgtcagcca aagtatgggt tttaggggat gcggtcgtag atctcttgcc agaatcagac 60
gggcgcctac tgccttgccc tggcggcgcg ccagctaacc ttgcggtggg aatcgccaga 120
ttaggcggaa caagtgggtt tataggtcgg gtgggggatg atccttttgg tgcgttaatg 180
caaagaacgc tgctaactga gggagtcgat atcacgtatc tgaagcaaga tgaatggcac 240
cggacatcca cggtgcttgt cgatctgaac gatcaagggg aacgttcatt tacgtttatg 300
gtccgccccca gtgccgatct ttttttagag acgacagact tgccctgctg gcgacatggc 360

gaatggttac atctctgttc aattgcgttg tctgcgagc cttcgcgtac cagcgcattt 420
actgcgatga cggcgatccg gcatgccgga gggtttgtca gcttcgatcc taatattcgt 480
gaagatctat ggcaagacga gcatttgctc cgcttggtgtt tgcggcaggc gctacaactg 540
gcggtatgtcg tcaagctctc ggaagaagaa tggcgactta tcagtggaaa aacacagaac 600
gatcaggata tatgcgccct ggcaaaagag tatgagatcg ccatgctgtt ggtgactaaa 660
gggtgcagaag gggtggtggt ctgttatcga ggacaagttc accattttgc tggaatgtct 720
gtgaattgtg tcgatagcac gggggcgga gatgcgttcg ttgccgggtt actcacaggt 780
ctgtcctcta cgggattatc tacagatgag agagaaatgc gacgaattat cgatctcgct 840
caacgttgcg gagcgttgc agtaacggcg aaaggggcaa tgacagcgct gccatgtcga 900
caagaactgg aatag 915

<210> 36
<211> 304
<212> PRT
5 <213> Escherichia coli

<400> 36
Met Ser Ala Lys Val Trp Val Leu Gly Asp Ala Val Val Asp Leu Leu
1 5 10 15
Pro Glu Ser Asp Gly Arg Leu Leu Pro Cys Pro Gly Gly Ala Pro Ala
20 25 30
Asn Val Ala Val Gly Ile Ala Arg Leu Gly Gly Thr Ser Gly Phe Ile
35 40 45
Gly Arg Val Gly Asp Asp Pro Phe Gly Ala Leu Met Gln Arg Thr Leu
50 55 60
Leu Thr Glu Gly Val Asp Ile Thr Tyr Leu Lys Gln Asp Glu Trp His
65 70 75 80
Arg Thr Ser Thr Val Leu Val Asp Leu Asn Asp Gln Gly Glu Arg Ser
85 90 95
Phe Thr Phe Met Val Arg Pro Ser Ala Asp Leu Phe Leu Glu Thr Thr
100 105 110
Asp Leu Pro Cys Trp Arg His Gly Glu Trp Leu His Leu Cys Ser Ile
115 120 125
Ala Leu Ser Ala Glu Pro Ser Arg Thr Ser Ala Phe Thr Ala Met Thr
130 135 140

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

Glu Asp Leu Trp Gln Asp Glu His Leu Leu Arg Leu Cys Leu Arg Gln
165 170 175

Ala Leu Gln Leu Ala Asp Val Val Lys Leu Ser Glu Glu Glu Trp Arg
180 185 190

Leu Ile Ser Gly Lys Thr Gln Asn Asp Gln Asp Ile Cys Ala Leu Ala
195 200 205

Lys Glu Tyr Glu Ile Ala Met Leu Leu Val Thr Lys Gly Ala Glu Gly
210 215 220

Val Val Val Cys Tyr Arg Gly Gln Val His His Phe Ala Gly Met Ser
225 230 235 240

Val Asn Cys Val Asp Ser Thr Gly Ala Gly Asp Ala Phe Val Ala Gly
245 250 255

Leu Leu Thr Gly Leu Ser Ser Thr Gly Leu Ser Thr Asp Glu Arg Glu
260 265 270

Met Arg Arg Ile Ile Asp Leu Ala Gln Arg Cys Gly Ala Leu Ala Val
275 280 285

Thr Ala Lys Gly Ala Met Thr Ala Leu Pro Cys Arg Gln Glu Leu Glu
290 295 300

- <210> 37
<211> 924
<212> DNA
5 <213> *Klebsiella pneumoniae*

<400> 37
atgaatggaa aaatctgggt actcggcgat gcggtcgtcg atctcctgcc cgatggagag 60
ggccgcctgc tgcaatgcc cggcggcgcg ccggccaacg tggcggtcgg cgtggcgcg 120
ctcggcggtg acagcgggtt tatcggccgc gtcggcgacg atcccttcgg ccgttttatg 180
cgtcacaccc tggcgagga gcaagtggat gtgaactata tgcgcctcga tgcggcgag 240
cgcacctcca cggtggtggt cgatctcgat agccacgggg agcgcacctt tacctttatg 300
gtccgtccga gcgcgcacct gttccttcag cccgaggatc tcccgcggtt tgccgccggt 360
cagtggctgc acgtctgctc catcgctctc agcgcggagc cgagccgcag caccgacattc 420
gcggcgatgg aggcgataaa gcgcgcgggg ggctatgtca gcttcgaccc caatatccgc 480

agcgacctgt ggcaggatcc gcaggacctt cgcgactgtc tcgaccgggc gctggccctc 540
 gccgacgcca taaaactttc ggaagaggag ctggcggtta tcagcggcag cgacgacatc 600
 gtcagcggca ccgcccggct gaacgcccgc ttccagccga cgctactgct ggtgaccag 660
 ggtaaagcgg ggggccaggc cgccctgcgc gggcaggtta gccacttccc tgcccgcccg 720
 gtggtggcgc tcgataccac cggcgccggc gatgcctttg tcgcccgggt actcgccggc 780
 ctcgccgcc acggtatccc ggacaacctc gcagccctgg ctcccgacct cgcgctggcg 840
 caaacctgcg gcgccctggc caccaccgcc aaaggcgcca tgaccgccct gccctacagg 900
 gacgatcttc agcgtcgcgt gtga 924

<210> 38
 <211> 307
 <212> PRT
 5 <213> Klebsiella pneumoniae

<400> 38
 Met Asn Gly Lys Ile Trp Val Leu Gly Asp Ala Val Val Asp Leu Leu
 1 5 10 15
 Pro Asp Gly Glu Gly Arg Leu Leu Gln Cys Pro Gly Gly Ala Pro Ala
 20 25 30
 Asn Val Ala Val Gly Val Ala Arg Leu Gly Gly Asp Ser Gly Phe Ile
 35 40 45
 Gly Arg Val Gly Asp Asp Pro Phe Gly Arg Phe Met Arg His Thr Leu
 50 55 60
 Ala Gln Glu Gln Val Asp Val Asn Tyr Met Arg Leu Asp Ala Ala Gln
 65 70 75 80
 Arg Thr Ser Thr Val Val Val Asp Leu Asp Ser His Gly Glu Arg Thr
 85 90 95
 Phe Thr Phe Met Val Arg Pro Ser Ala Asp Leu Phe Leu Gln Pro Glu
 100 105 110
 Asp Leu Pro Pro Phe Ala Ala Gly Gln Trp Leu His Val Cys Ser Ile
 115 120 125
 Ala Leu Ser Ala Glu Pro Ser Arg Ser Thr Thr Phe Ala Ala Met Glu
 130 135 140
 Ala Ile Lys Arg Ala Gly Gly Tyr Val Ser Phe Asp Pro Asn Ile Arg
 145 150 155 160

Ser Asp Leu Trp Gln Asp Pro Gln Asp Leu Arg Asp Cys Leu Asp Arg
165 170 175

Ala Leu Ala Leu Ala Asp Ala Ile Lys Leu Ser Glu Glu Glu Leu Ala
180 185 190

Phe Ile Ser Gly Ser Asp Asp Ile Val Ser Gly Thr Ala Arg Leu Asn
195 200 205

Ala Arg Phe Gln Pro Thr Leu Leu Leu Val Thr Gln Gly Lys Ala Gly
210 215 220

Val Gln Ala Ala Leu Arg Gly Gln Val Ser His Phe Pro Ala Arg Pro
225 230 235 240

Val Val Ala Val Asp Thr Thr Gly Ala Gly Asp Ala Phe Val Ala Gly
245 250 255

Leu Leu Ala Gly Leu Ala Ala His Gly Ile Pro Asp Asn Leu Ala Ala
260 265 270

Leu Ala Pro Asp Leu Ala Leu Ala Gln Thr Cys Gly Ala Leu Ala Thr
275 280 285

Thr Ala Lys Gly Ala Met Thr Ala Leu Pro Tyr Arg Asp Asp Leu Gln
290 295 300

Arg Ser Leu
305

<210> 39
<211> 915
<212> DNA
<213> Escherichia coli

<400> 39
atgtcagcca aagtatgggt tttaggggat gcggtcgtag atctcttgcc agaatacagac 60
gggcgcctac tgccttgccc tggcggcgcg ccagctaacg ttgcggtggg aatcgccaga 120
ttaggcggaa caagtgggtt tataggtcgg gtgggggatg atccttttgg tgcgttaatg 180
caaagaacgc tgctaactga gggagtcgat atcacgtatc tgaagcaaga tgaatggcac 240
cggacatcca cggtgcttgt cgatctgaac gatcaagggg aacgttcatt tacgtttatg 300
gtccgcccc gtgccgatct ttttttagag aegacagact tgccctgctg gcgacatggc 360
gaatggttac atctctgttc aattgcgttg tctgccgagc cttcgcgtac cagcgcat 420
actgcgatga cggcgatccg gcatgccgga ggttttgtca gcttcgatcc taatatcgt 480

gaagatctat ggcaagacga gcatttgctc cgcttgtgtt tgcggcaggc gctacaactg 540
 gcggatgtcg tcaagctctc ggaagaagaa tggcgactta tcagtggaaa aacacagaac 600
 gatcaggata tatgcgccct ggcaaaagag tatgagatcg ccatgctggt ggtgactaaa 660
 ggtgcagaag ggggtggtgt ctgttatcga ggacaagttc accattttgc tggaatgtct 720
 gtgaattgtg tcgatagcac gggggcggga gatgcgttcg ttgccggggt actcacaggt 780
 ctgtcctcta cgggattatc tacagatgag agagaaatgc gacgaattat cgatctcgct 840
 caacgttgcg gagcgcttgc agtaacggcg aaaggggcaa tgacagcgct gccatgtcga 900
 caagaactgg aatag 915

<210> 40
 <211> 304
 <212> PRT
 5 <213> Escherichia coli

<400> 40
 Met Ser Ala Lys Val Trp Val Leu Gly Asp Ala Val Val Asp Leu Leu
 1 5 10 15
 Pro Glu Ser Asp Gly Arg Leu Leu Pro Cys Pro Gly Gly Ala Pro Ala
 20 25 30
 Asn Val Ala Val Gly Ile Ala Arg Leu Gly Gly Thr Ser Gly Phe Ile
 35 40 45
 Gly Arg Val Gly Asp Asp Pro Phe Gly Ala Leu Met Gln Arg Thr Leu
 50 55 60
 Leu Thr Glu Gly Val Asp Ile Thr Tyr Leu Lys Gln Asp Glu Trp His
 65 70 75 80
 Arg Thr Ser Thr Val Leu Val Asp Leu Asn Asp Gln Gly Glu Arg Ser
 85 90 95
 Phe Thr Phe Met Val Arg Pro Ser Ala Asp Leu Phe Leu Glu Thr Thr
 100 105 110
 Asp Leu Pro Cys Trp Arg His Gly Glu Trp Leu His Leu Cys Ser Ile
 115 120 125
 Ala Leu Ser Ala Glu Pro Ser Arg Thr Ser Ala Phe Thr Ala Met Thr
 130 135 140
 Ala Ile Arg His Ala Gly Gly Phe Val Ser Phe Asp Pro Asn Ile Arg
 145 150 155 160

Glu Asp Leu Trp Gln Asp Glu His Leu Leu Arg Leu Cys Leu Arg Gln
165 170 175

Ala Leu Gln Leu Ala Asp Val Val Lys Leu Ser Glu Glu Glu Trp Arg
180 185 190

Leu Ile Ser Gly Lys Thr Gln Asn Asp Gln Asp Ile Cys Ala Leu Ala
195 200 205

Lys Glu Tyr Glu Ile Ala Met Leu Leu Val Thr Lys Gly Ala Glu Gly
210 215 220

Val Val Val Cys Tyr Arg Gly Gln Val His His Phe Ala Gly Met Ser
225 230 235 240

Val Asn Cys Val Asp Ser Thr Gly Ala Gly Asp Ala Phe Val Ala Gly
245 250 255

Leu Leu Thr Gly Leu Ser Ser Thr Gly Leu Ser Thr Asp Glu Arg Glu
260 265 270

Met Arg Arg Ile Ile Asp Leu Ala Gln Arg Cys Gly Ala Leu Ala Val
275 280 285

Thr Ala Lys Gly Ala Met Thr Ala Leu Pro Cys Arg Gln Glu Leu Glu
290 295 300

<210> 41
<211> 879
<212> DNA
5 <213> Enterococcus faecalis

<400> 41
atgacagaaa aacttttagg aagtatcgaa gccggtggca caaaatttgt atgtggcggt 60
gggacagatg atttgaccat cgtagaacgt gtcagttttc ccacaacaac cccagaagaa 120
acaatgaaaa aagtaataga atttttccaa caatatcctt taaaagcgat tgggattggt 180
tcatttggtc cgattgatat tcacgttgat tctcctacgt atggttatat cacttctaca 240
ccaaaattag cttggcgtaa ctttgacttg ttaggaacta tgaaacaaca ttttgatgtg 300
ccaatggctt ggacaacgga tgtgaatgct gcggcatatg gtgagtatgt tgctggaaat 360
gggcaacata catctagtgtg tgtatattat acaattggaa ctggtgttgg cgctggagcg 420
attcaaaacg gtgagtttat tgaaggcttt agccaccag aaatggggca tgcgttagtt 480
cgtcgtcatc ctgaagatac gtatgcagga aattgtcctt atcatggaga ttgtttagaa 540
gggattgcag caggaccagc agttgaaggt cgttctggtg aaaaaggaca tttattggaa 600

gaggatcata aaacttgga attagaagct tattatttag cgcaagcggc gtacaatacg 660
actttattat tagcgccaga agtgatcatt ttaggtggcg gcgtcatgaa acaacgtcat 720
ttgatgccga aagtctgga aaaatttgct gaattagtca atggatatgt ggaaacaccg 780
cctttagaaa aatacttggt gacgcctctt ttagaagata atccaggaac aatcgggtgc 840
tttgccttgg caaaaaaagc tttaatggct caaaaataa 879

<210> 42
<211> 292
<212> PRT
5 <213> Enterococcus faecalis

<400> 42
Met Thr Glu Lys Leu Leu Gly Ser Ile Glu Ala Gly Gly Thr Lys Phe
1 5 10 15

Val Cys Gly Val Gly Thr Asp Asp Leu Thr Ile Val Glu Arg Val Ser
20 25 30

Phe Pro Thr Thr Thr Pro Glu Glu Thr Met Lys Lys Val Ile Glu Phe
35 40 45

Phe Gln Gln Tyr Pro Leu Lys Ala Ile Gly Ile Gly Ser Phe Gly Pro
50 55 60

Ile Asp Ile His Val Asp Ser Pro Thr Tyr Gly Tyr Ile Thr Ser Thr
65 70 75 80

Pro Lys Leu Ala Trp Arg Asn Phe Asp Leu Leu Gly Thr Met Lys Gln
85 90 95

His Phe Asp Val Pro Met Ala Trp Thr Thr Asp Val Asn Ala Ala Ala
100 105 110

Tyr Gly Glu Tyr Val Ala Gly Asn Gly Gln His Thr Ser Ser Cys Val
115 120 125

Tyr Tyr Thr Ile Gly Thr Gly Val Gly Ala Gly Ala Ile Gln Asn Gly
130 135 140

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

Arg Arg His Pro Glu Asp Thr Tyr Ala Gly Asn Cys Pro Tyr His Gly
165 170 175

Asp Cys Leu Glu Gly Ile Ala Ala Gly Pro Ala Val Glu Gly Arg Ser
180 185 190

Gly Lys Lys Gly His Leu Leu Glu Glu Asp His Lys Thr Trp Glu Leu
195 200 205

Glu Ala Tyr Tyr Leu Ala Gln Ala Ala Tyr Asn Thr Thr Leu Leu Leu
210 215 220

Ala Pro Glu Val Ile Ile Leu Gly Gly Gly Val Met Lys Gln Arg His
225 230 235 240

Leu Met Pro Lys Val Arg Glu Lys Phe Ala Glu Leu Val Asn Gly Tyr
245 250 255

Val Glu Thr Pro Pro Leu Glu Lys Tyr Leu Val Thr Pro Leu Leu Glu
260 265 270

Asp Asn Pro Gly Thr Ile Gly Cys Phe Ala Leu Ala Lys Lys Ala Leu
275 280 285

Met Ala Gln Lys
290

- <210> 43
<211> 1458
<212> DNA
5 <213> *Saccharomyces cerevisiae*

<400> 43
ttaagcgcca atgataccaa gagacttacc ttcggaatt cttttttcgg acaatgcagc 60
aataacagca gcacctgcac ctgaaccatc ctcagctgga acaatcgtaa ttggatcttt 120
gcttgcgta ccagtccatc catagatata tctcaaacc ttagcggcgg cttccttgaa 180
acctgggtat ttgttataga cagaaccgtc agcggcaatg tgaccagtct tgtaacctct 240
cttttgcaa atagcggcaa taccacaaac agctaatacta gcagctctgg taccgatcaa 300
ttcacaaagt cttctaataca acttacgttc tggcagagtg gtcttgacac caaagtcctt 360
ttggaagatg tcatcagtat cttccaagtt ttcaaagtga tcatcctcga ttcttgctgg 420
gtaggaggtg tccatgatgt atggttgttt caacttgctt agatcttgat cttcaacat 480
caagcccttc tcgtttaatt caagtaaacac tagacgcaac aattcaccca agtagtaacc 540
ggaggtcatc ttttcaaaag cttgttgacc aggtcttgga gattgttcgt cgacagcaac 600
atcgtacttg gttcttgga agaccaaag ttcatatcg aaggaaccat attcacaatt 660
gatagccatt ggagagttac ttggaatata gtctgctaatt ttgcctcca acttttcgat 720
atcggaaca acatcataga aagcacggtt gacaccagta ccgaaaatca cacccatctt 780

```

agtctctggg tcagtgtagt atgaggcaat taaagtacca acagtatcat taatcaatgc      840
tacaatttca ataggcaact ctctcttgga aatttcgttt tgtagcaatg ggacgacatc      900
gtggccttcg acatttgga taccgaaacc cttgggtccat ctttgcaaaa taccttcgtt      960
aatcttggtt tgggaagctg ggtacgagaa ggtgaaacct aatggtaagg tgccttggt      1020
gttttagcaat tcttgctcga ccataaagtc cttcaaagag tcggcaataa aggaccataa      1080
ctcctcttgg tgcttagtgg ttctcatgtc atgtggtagt ttatacttgg attgagtgg      1140
gtcaaaggta tggttaccgc tcaacttgac caacacgact cttaagttag taccacccaa      1200
atcaatggcc aaatagttac cagattcttt acctgttggg aattccatga cccaaccggg      1260
aatcattgga atgttacctc ccttctttgt caaaccttta ttcaattcgt cgataaagt      1320
cttaacaacc tttctcaagg tctcgtgtc aactgtaaac atatcttcca actgatgaat      1380
ttcatccatc aattccttgg gcacatcagc catggaacc tttctagcct gtggtttctt      1440
tggacctaaa tgaaccat                                          1458

```

<210> 44
 <211> 485
 <212> PRT
 5 <213> *Saccharomyces cerevisiae*

```

<400> 44
Met Val His Leu Gly Pro Lys Lys Pro Gln Ala Arg Lys Gly Ser Met
1          5          10          15

Ala Asp Val Pro Lys Glu Leu Met Asp Glu Ile His Gln Leu Glu Asp
20          25          30

Met Phe Thr Val Asp Ser Glu Thr Leu Arg Lys Val Val Lys His Phe
35          40          45

Ile Asp Glu Leu Asn Lys Gly Leu Thr Lys Lys Gly Gly Asn Ile Pro
50          55          60

Met Ile Pro Gly Trp Val Met Glu Phe Pro Thr Gly Lys Glu Ser Gly
65          70          75          80

Asn Tyr Leu Ala Ile Asp Leu Gly Gly Thr Asn Leu Arg Val Val Leu
85          90          95

Val Lys Leu Ser Gly Asn His Thr Phe Asp Thr Thr Gln Ser Lys Tyr
100          105          110

Lys Leu Pro His Asp Met Arg Thr Thr Lys His Gln Glu Glu Leu Trp
115          120          125

```

Ser Phe Ile Ala Asp Ser Leu Lys Asp Phe Met Val Glu Gln Glu Leu
130 135 140

Leu Asn Thr Lys Asp Thr Leu Pro Leu Gly Phe Thr Phe Ser Tyr Pro
145 150 155 160

Ala Ser Gln Asn Lys Ile Asn Glu Gly Ile Leu Gln Arg Trp Thr Lys
165 170 175

Gly Phe Asp Ile Pro Asn Val Glu Gly His Asp Val Val Pro Leu Leu
180 185 190

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

Ile Asn Asp Thr Val Gly Thr Leu Ile Ala Ser Tyr Tyr Thr Asp Pro
210 215 220

Glu Thr Lys Met Gly Val Ile Phe Gly Thr Gly Val Asn Gly Ala Phe
225 230 235 240

Tyr Asp Val Val Ser Asp Ile Glu Lys Leu Glu Gly Lys Leu Ala Asp
245 250 255

Asp Ile Pro Ser Asn Ser Pro Met Ala Ile Asn Cys Glu Tyr Gly Ser
260 265 270

Phe Asp Asn Glu His Leu Val Leu Pro Arg Thr Lys Tyr Asp Val Ala
275 280 285

Val Asp Glu Gln Ser Pro Arg Pro Gly Gln Gln Ala Phe Glu Lys Met
290 295 300

Thr Ser Gly Tyr Tyr Leu Gly Glu Leu Leu Arg Leu Val Leu Leu Glu
305 310 315 320

Leu Asn Glu Lys Gly Leu Met Leu Lys Asp Gln Asp Leu Ser Lys Leu
325 330 335

Lys Gln Pro Tyr Ile Met Asp Thr Ser Tyr Pro Ala Arg Ile Glu Asp
340 345 350

Asp Pro Phe Glu Asn Leu Glu Asp Thr Asp Asp Ile Phe Gln Lys Asp
355 360 365

Phe Gly Val Lys Thr Thr Leu Pro Glu Arg Lys Leu Ile Arg Arg Leu

370 375 380

Cys Glu Leu Ile Gly Thr Arg Ala Ala Arg Leu Ala Val Cys Gly Ile
385 390 395 400

Ala Ala Ile Cys Gln Lys Arg Gly Tyr Lys Thr Gly His Ile Ala Ala
405 410 415

Asp Gly Ser Val Tyr Asn Lys Tyr Pro Gly Phe Lys Glu Ala Ala Ala
420 425 430

Lys Gly Leu Arg Asp Ile Tyr Gly Trp Thr Gly Asp Ala Ser Lys Asp
435 440 445

Pro Ile Thr Ile Val Pro Ala Glu Asp Gly Ser Gly Ala Gly Ala Ala
450 455 460

Val Ile Ala Ala Leu Ser Glu Lys Arg Ile Ala Glu Gly Lys Ser Leu
465 470 475 480

Gly Ile Ile Gly Ala
485

<210> 45
<211> 1461
<212> DNA
5 <213> *Saccharomyces cerevisiae*

<400> 45
atggttcatt taggtccaaa aaaaccacaa gccagaaagg gttccatggc cgatgtgcca 60
aaggaattga tgcaacaaat tgagaatttt gaaaaaattt tcactgttcc aactgaaact 120
ttacaagccg ttaccaagca cttcatttcc gaattggaaa agggtttgtc caagaagggt 180
ggtaacattc caatgattcc aggttgggtt atggatttcc caactggtaa ggaatccggt 240
gatttcttgg ccattgattt ggggtgtacc aacttgagag ttgtcttagt caagttgggc 300
ggtgaccgtg cctttgacac cactcaatct aagtacagat taccagatgc tatgagaact 360
actcaaaatc cagacgaatt gtgggaattt attgccgact ctttgaaagc ttttattgat 420
gagcaattcc cacaaggtat ctctgagcca attccattgg gtttcacctt ttctttccca 480
gcttctcaaa acaaaatcaa tgaaggtatc ttgcaaagat ggactaaagg ttttgatatt 540
ccaaacattg aaaaccacga tgttgttcca atgttgcaaa agcaaatcac taagaggaat 600
atcccaattg aagttgttgc tttgataaac gacactaccg gtacttttgt tgcttcttac 660
tacttgacc cagaaactaa gatgggtgtt atcttcggta ctggtgtcaa tgggtgcttac 720
tacgatgttt gttccgatat cgaaaagcta caaggaaaac tatctgatga cattccacca 780

tctgctccaa tggccatcaa ctgtgaatac gggttccttcg ataatgaaca tgtcgttttg 840
ccaagaacta aatacgatat caccattgat gaagaatctc caagaccagg ccaacaaacc 900
tttgaaaaaa tgtcttctgg ttactactta ggtgaaattt tgcgtttggc cttgatggac 960
atgtacaaac aaggtttcat cttcaagaac caagacttgt ctaagttcga caagcctttc 1020
gtcatggaca cttcttacct agccagaatc gaggaagatc cattcgagaa cctagaagat 1080
accgatgact tgttccaaaa tgagttcggg atcaacacta ctggttcaaga acgtaaattg 1140
atcagacggt tatctgaatt gattggtgct agagctgcta gattgtccgt ttgtggtatt 1200
gctgctatct gtcaaaagag aggttacaag accggtcaca tcgctgcaga cggttccggt 1260
tacaacagat acccaggttt caaagaaaag gctgccaatg ctttgaagga catttacggc 1320
tggaactcaa cctcactaga cgactaccca atcaagattg ttcctgctga agatggttcc 1380
ggtgctggtg ccgctgttat tgctgctttg gcccaaaaaa gaattgctga aggtaagtcc 1440
gttgggtatca tcggtgctta a 1461

<210> 46
<211> 486
<212> PRT
5 <213> *Saccharomyces cerevisiae*

<400> 46
Met Val His Leu Gly Pro Lys Lys Pro Gln Ala Arg Lys Gly Ser Met
1 5 10 15
Ala Asp Val Pro Lys Glu Leu Met Gln Gln Ile Glu Asn Phe Glu Lys
20 25 30
Ile Phe Thr Val Pro Thr Glu Thr Leu Gln Ala Val Thr Lys His Phe
35 40 45
Ile Ser Glu Leu Glu Lys Gly Leu Ser Lys Lys Gly Gly Asn Ile Pro
50 55 60
Met Ile Pro Gly Trp Val Met Asp Phe Pro Thr Gly Lys Glu Ser Gly
65 70 75 80
Asp Phe Leu Ala Ile Asp Leu Gly Gly Thr Asn Leu Arg Val Val Leu
85 90 95
Val Lys Leu Gly Gly Asp Arg Thr Phe Asp Thr Thr Gln Ser Lys Tyr
100 105 110
Arg Leu Pro Asp Ala Met Arg Thr Thr Gln Asn Pro Asp Glu Leu Trp
115 120 125

Glu Phe Ile Ala Asp Ser Leu Lys Ala Phe Ile Asp Glu Gln Phe Pro
130 135 140

Gln Gly Ile Ser Glu Pro Ile Pro Leu Gly Phe Thr Phe Ser Phe Pro
145 150 155 160

Ala Ser Gln Asn Lys Ile Asn Glu Gly Ile Leu Gln Arg Trp Thr Lys
165 170 175

Gly Phe Asp Ile Pro Asn Ile Glu Asn His Asp Val Val Pro Met Leu
180 185 190

Gln Lys Gln Ile Thr Lys Arg Asn Ile Pro Ile Glu Val Val Ala Leu
195 200 205

Ile Asn Asp Thr Thr Gly Thr Leu Val Ala Ser Tyr Tyr Thr Asp Pro
210 215 220

Glu Thr Lys Met Gly Val Ile Phe Gly Thr Gly Val Asn Gly Ala Tyr
225 230 235 240

Tyr Asp Val Cys Ser Asp Ile Glu Lys Leu Gln Gly Lys Leu Ser Asp
245 250 255

Asp Ile Pro Pro Ser Ala Pro Met Ala Ile Asn Cys Glu Tyr Gly Ser
260 265 270

Phe Asp Asn Glu His Val Val Leu Pro Arg Thr Lys Tyr Asp Ile Thr
275 280 285

Ile Asp Glu Glu Ser Pro Arg Pro Gly Gln Gln Thr Phe Glu Lys Met
290 295 300

Ser Ser Gly Tyr Tyr Leu Gly Glu Ile Leu Arg Leu Ala Leu Met Asp
305 310 315 320

Met Tyr Lys Gln Gly Phe Ile Phe Lys Asn Gln Asp Leu Ser Lys Phe
325 330 335

Asp Lys Pro Phe Val Met Asp Thr Ser Tyr Pro Ala Arg Ile Glu Glu
340 345 350

Asp Pro Phe Glu Asn Leu Glu Asp Thr Asp Asp Leu Phe Gln Asn Glu
355 360 365

Phe Gly Ile Asn Thr Thr Val Gln Glu Arg Lys Leu Ile Arg Arg Leu
370 375 380

Ser Glu Leu Ile Gly Ala Arg Ala Ala Arg Leu Ser Val Cys Gly Ile
385 390 395 400

Ala Ala Ile Cys Gln Lys Arg Gly Tyr Lys Thr Gly His Ile Ala Ala
405 410 415

Asp Gly Ser Val Tyr Asn Arg Tyr Pro Gly Phe Lys Glu Lys Ala Ala
420 425 430

Asn Ala Leu Lys Asp Ile Tyr Gly Trp Thr Gln Thr Ser Leu Asp Asp
435 440 445

Tyr Pro Ile Lys Ile Val Pro Ala Glu Asp Gly Ser Gly Ala Gly Ala
450 455 460

Ala Val Ile Ala Ala Leu Ala Gln Lys Arg Ile Ala Glu Gly Lys Ser
465 470 475 480

Val Gly Ile Ile Gly Ala
485

<210> 47

<211> 1164

<212> DNA

5 <213> Klebsiella pneumoniae

<400> 47

tcagaatgcc tggcggaaaa tcgcggaat ctctgctcg ttgcctttac gcgggttcga	60
gaacgcattg cagtctttca gagccatctc cgccatgtag gggaagtcgg cctcttttac	120
tcccagatcg cgcagatgct gcggaatacc gatatccatc gacagacgcg tgatagcggc	180
gatggctttt tccgcgcgt cgagagtga cagtcgggtg atattttcgc ccatcagttc	240
agcgatatcg gcaattttct ccgggttggc gatcaggttg tagcgggccca catgcggcag	300
caggacagcg ttggccacgc cgtgcggcat gtcgtacagg ccgcccagct ggtgcgccat	360
ggcgtgcacg tagccgaggt tggcgttatt gaaagccatc ccggccagca gagaggcata	420
ggccatgttt tccgcgcct gcagattgct gccgagggcc acggcctggc gcaggttgcg	480
ggcgatgagg cggatcgct gcatggcggc ggcgtccgtc accgggttag cgtctttgga	540
gatataggcc tctacggcgt gggtcagggc atccatcccg gtcgcccgcg tcagggcggc	600
cggtttaccg atcatcagca gcggatcggt gatagagacc gacggcaggt tgcgccagct	660
gacgatcaca aacttcactt tggtttcggt gttggtcagg acgcagtggc gggtgacctc	720
gctggcgggtg ccggcgggtg tattgaccgc gacgataggc ggcagcgggt tggtcagggt	780
ctcgattccg gcatactggt acagatcgcc ctcatgggtg gcggcgatgc cgatgccttt	840

gccgcaatcg tgcgggctgc cgccgccac ggtgacgatg atgtcgact gttcgcggcg 900
 aaacacggcg aggcgctgc gcacgttggg gtctttcggg ttcggctcga cgccgtcaaa 960
 gatcgccacc tcgatcccg cctcccgcag ataatgcagg gttttgtcca ctgcgccatc 1020
 ttttaattgcc cgcaggcctt tgtcggtgac cagcagggt tttttcccc ccagcagctg 1080
 gcagcgttcg ccgactacgg aaatggcgtt ggggccaaaa aagttaacgt ttggcaccag 1140
 ataatcaaac atacgatagc tcat 1164

<210> 48
 <211> 387
 <212> PRT

5 <213> Klebsiella pneumoniae

<400> 48
 Met Ser Tyr Arg Met Phe Asp Tyr Leu Val Pro Asn Val Asn Phe Phe
 1 5 10 15
 Gly Pro Asn Ala Ile Ser Val Val Gly Glu Arg Cys Gln Leu Leu Gly
 20 25 30
 Gly Lys Lys Ala Leu Leu Val Thr Asp Lys Gly Leu Arg Ala Ile Lys
 35 40 45
 Asp Gly Ala Val Asp Lys Thr Leu His Tyr Leu Arg Glu Ala Gly Ile
 50 55 60
 Glu Val Ala Ile Phe Asp Gly Val Glu Pro Asn Pro Lys Asp Thr Asn
 65 70 75 80
 Val Arg Asp Gly Leu Ala Val Phe Arg Arg Glu Gln Cys Asp Ile Ile
 85 90 95
 Val Thr Val Gly Gly Gly Ser Pro His Asp Cys Gly Lys Gly Ile Gly
 100 105 110
 Ile Ala Ala Thr His Glu Gly Asp Leu Tyr Gln Tyr Ala Gly Ile Glu
 115 120 125
 Thr Leu Thr Asn Pro Leu Pro Pro Ile Val Ala Val Asn Thr Thr Ala
 130 135 140
 Gly Thr Ala Ser Glu Val Thr Arg His Cys Val Leu Thr Asn Thr Glu
 145 150 155 160
 Thr Lys Val Lys Phe Val Ile Val Ser Trp Arg Asn Leu Pro Ser Val
 165 170 175

Ser Ile Asn Asp Pro Leu Leu Met Ile Gly Lys Pro Ala Ala Leu Thr
180 185 190

Ala Ala Thr Gly Met Asp Ala Leu Thr His Ala Val Glu Ala Tyr Ile
195 200 205

Ser Lys Asp Ala Asn Pro Val Thr Asp Ala Ala Ala Met Gln Ala Ile
210 215 220

Arg Leu Ile Ala Arg Asn Leu Arg Gln Ala Val Ala Leu Gly Ser Asn
225 230 235 240

Leu Gln Ala Arg Glu Asn Met Ala Tyr Ala Ser Leu Leu Ala Gly Met
245 250 255

Ala Phe Asn Asn Ala Asn Leu Gly Tyr Val His Ala Met Ala His Gln
260 265 270

Leu Gly Gly Leu Tyr Asp Met Pro His Gly Val Ala Asn Ala Val Leu
275 280 285

Leu Pro His Val Ala Arg Tyr Asn Leu Ile Ala Asn Pro Glu Lys Phe
290 295 300

Ala Asp Ile Ala Glu Leu Met Gly Glu Asn Ile Thr Gly Leu Ser Thr
305 310 315 320

Leu Asp Ala Ala Glu Lys Ala Ile Ala Ala Ile Thr Arg Leu Ser Met
325 330 335

Asp Ile Gly Ile Pro Gln His Leu Arg Asp Leu Gly Val Lys Glu Ala
340 345 350

Asp Phe Pro Tyr Met Ala Glu Met Ala Leu Lys Asp Gly Asn Ala Phe
355 360 365

Ser Asn Pro Arg Lys Gly Asn Glu Gln Glu Ile Ala Ala Ile Phe Arg
370 375 380

Gln Ala Phe
385

<210> 49
<211> 1824
<212> DNA
5 <213> Klebsiella pneumoniae

<220>
<221> CDS
<222> (1)..(1824)
<400> 49

atg ccg tta ata gcc ggg att gat atc ggc aac gcc acc acc gag gtg Met Pro Leu Ile Ala Gly Ile Asp Ile Gly Asn Ala Thr Thr Glu Val 1 5 10 15	48
gcg ctg gcg tcc gac tac ccg cag gcg agg gcg ttt gtt gcc agc ggg Ala Leu Ala Ser Asp Tyr Pro Gln Ala Arg Ala Phe Val Ala Ser Gly 20 25 30	96
atc gtc gcg acg acg ggc atg aaa ggg acg cgg gac aat atc gcc ggg Ile Val Ala Thr Thr Gly Met Lys Gly Thr Arg Asp Asn Ile Ala Gly 35 40 45	144
acc ctc gcc gcg ctg gag cag gcc ctg gcg aaa aca ccg tgg tcg atg Thr Leu Ala Ala Leu Glu Gln Ala Leu Ala Lys Thr Pro Trp Ser Met 50 55 60	192
agc gat gtc tct cgc atc tat ctt aac gaa gcc gcg ccg gtg att ggc Ser Asp Val Ser Arg Ile Tyr Leu Asn Glu Ala Ala Pro Val Ile Gly 65 70 75 80	240
gat gtg gcg atg gag acc atc acc gag acc att atc acc gaa tcg acc Asp Val Ala Met Glu Thr Ile Thr Glu Thr Ile Ile Thr Glu Ser Thr 85 90 95	288
atg atc ggt cat aac ccg cag acg ccg ggc ggg gtg ggc gtt ggc gtg Met Ile Gly His Asn Pro Gln Thr Pro Gly Gly Val Gly Val Gly Val 100 105 110	336
ggg acg act atc gcc ctc ggg cgg ctg gcg acg ctg ccg gcg gcg cag Gly Thr Thr Ile Ala Leu Gly Arg Leu Ala Thr Leu Pro Ala Ala Gln 115 120 125	384
tat gcc gag ggg tgg atc gta ctg att gac gac gcc gtc gat ttc ctt Tyr Ala Glu Gly Trp Ile Val Leu Ile Asp Asp Ala Val Asp Phe Leu 130 135 140	432
gac gcc gtg tgg tgg ctc aat gag gcg ctc gac cgg ggg atc aac gtg Asp Ala Val Trp Trp Leu Asn Glu Ala Leu Asp Arg Gly Ile Asn Val 145 150 155 160	480
gtg gcg gcg atc ctc aaa aag gac gac ggc gtg ctg gtg aac aac cgc Val Ala Ala Ile Leu Lys Lys Asp Asp Gly Val Leu Val Asn Asn Arg 165 170 175	528
ctg cgt aaa acc ctg ccg gtg gtg gat gaa gtg acg ctg ctg gag cag Leu Arg Lys Thr Leu Pro Val Val Asp Glu Val Thr Leu Leu Glu Gln 180 185 190	576
gtc ccc gag ggg gta atg gcg gcg gtg gaa gtg gcc gcg ccg ggc cag Val Pro Glu Gly Val Met Ala Ala Val Glu Val Ala Ala Pro Gly Gln 195 200 205	624
gtg gtg cgg atc ctg tcg aat ccc tac ggg atc gcc acc ttc ttc ggg Val Val Arg Ile Leu Ser Asn Pro Tyr Gly Ile Ala Thr Phe Phe Gly 210 215 220	672
cta agc ccg gaa gag acc cag gcc atc gtc ccc atc gcc cgc gcc ctg	720

Leu	Ser	Pro	Glu	Glu	Thr	Gln	Ala	Ile	Val	Pro	Ile	Ala	Arg	Ala	Leu	
225					230					235					240	
att	ggc	aac	cgt	tcc	gcg	gtg	gtg	ctc	aag	acc	ccg	cag	ggg	gat	gtg	768
Ile	Gly	Asn	Arg	Ser	Ala	Val	Val	Leu	Lys	Thr	Pro	Gln	Gly	Asp	Val	
			245					250					255			
cag	tcg	cgg	gtg	atc	ccg	gcg	ggc	aac	ctc	tac	att	agc	ggc	gaa	aag	816
Gln	Ser	Arg	Val	Ile	Pro	Ala	Gly	Asn	Leu	Tyr	Ile	Ser	Gly	Glu	Lys	
			260					265					270			
cgc	cgc	gga	gag	gcc	gat	gtc	gcc	gag	ggc	gcg	gaa	gcc	atc	atg	cag	864
Arg	Arg	Gly	Glu	Ala	Asp	Val	Ala	Glu	Gly	Ala	Glu	Ala	Ile	Met	Gln	
		275					280					285				
gcg	atg	agc	gcc	tgc	gct	ccg	gta	cgc	gac	atc	cgc	ggc	gaa	ccg	ggc	912
Ala	Met	Ser	Ala	Cys	Ala	Pro	Val	Arg	Asp	Ile	Arg	Gly	Glu	Pro	Gly	
	290					295					300					
acc	cac	gcc	ggc	ggc	atg	ctt	gag	cgg	gtg	cgc	aag	gta	atg	gcg	tcc	960
Thr	His	Ala	Gly	Gly	Met	Leu	Glu	Arg	Val	Arg	Lys	Val	Met	Ala	Ser	
305					310					315					320	
ctg	acc	ggc	cat	gag	atg	agc	gcg	ata	tac	atc	cag	gat	ctg	ctg	gcg	1008
Leu	Thr	Gly	His	Glu	Met	Ser	Ala	Ile	Tyr	Ile	Gln	Asp	Leu	Leu	Ala	
			325						330					335		
gtg	gat	acg	ttt	att	ccg	cgc	aag	gtg	cag	ggc	ggg	atg	gcc	ggc	gag	1056
Val	Asp	Thr	Phe	Ile	Pro	Arg	Lys	Val	Gln	Gly	Gly	Met	Ala	Gly	Glu	
			340					345					350			
tgc	gcc	atg	gag	aat	gcc	gtc	ggg	atg	gcg	gcg	atg	gtg	aaa	gcg	gat	1104
Cys	Ala	Met	Glu	Asn	Ala	Val	Gly	Met	Ala	Ala	Met	Val	Lys	Ala	Asp	
		355					360					365				
cgt	ctg	caa	atg	cag	gtt	atc	gcc	cgc	gaa	ctg	agc	gcc	cga	ctg	cag	1152
Arg	Leu	Gln	Met	Gln	Val	Ile	Ala	Arg	Glu	Leu	Ser	Ala	Arg	Leu	Gln	
	370					375					380					
acc	gag	gtg	gtg	gtg	ggc	ggc	gtg	gag	gcc	aac	atg	gcc	atc	gcc	ggg	1200
Thr	Glu	Val	Val	Val	Gly	Gly	Val	Glu	Ala	Asn	Met	Ala	Ile	Ala	Gly	
385					390					395					400	
gcg	tta	acc	act	ccc	ggc	tgt	gcg	gcg	ccg	ctg	gcg	atc	ctc	gac	ctc	1248
Ala	Leu	Thr	Thr	Pro	Gly	Cys	Ala	Ala	Pro	Leu	Ala	Ile	Leu	Asp	Leu	
				405					410					415		
ggc	gcc	ggc	tcg	acg	gat	gcg	gcg	atc	gtc	aac	gcg	gag	ggg	cag	ata	1296
Gly	Ala	Gly	Ser	Thr	Asp	Ala	Ala	Ile	Val	Asn	Ala	Glu	Gly	Gln	Ile	
			420					425					430			
acg	gcg	gtc	cat	ctc	gcc	ggg	gcg	ggg	aat	atg	gtc	agc	ctg	ttg	att	1344
Thr	Ala	Val	His	Leu	Ala	Gly	Ala	Gly	Asn	Met	Val	Ser	Leu	Leu	Ile	
			435				440					445				
aaa	acc	gag	ctg	ggc	ctc	gag	gat	ctt	tcg	ctg	gcg	gaa	gcg	ata	aaa	1392
Lys	Thr	Glu	Leu	Gly	Leu	Glu	Asp	Leu	Ser	Leu	Ala	Glu	Ala	Ile	Lys	
		450				455					460					
aaa	tac	ccg	ctg	gcc	aaa	gtg	gaa	agc	ctg	ttc	agt	att	cgt	cac	gag	1440
Lys	Tyr	Pro	Leu	Ala	Lys	Val	Glu	Ser	Leu	Phe	Ser	Ile	Arg	His	Glu	
465					470					475					480	

aat ggc gcg gtg gag ttc ttt cgg gaa gcc ctc agc ccg gcg gtg ttc 1488
Asn Gly Ala Val Glu Phe Phe Arg Glu Ala Leu Ser Pro Ala Val Phe
485 490 495

gcc aaa gtg gtg tac atc aag gag ggc gaa ctg gtg ccg atc gat aac 1536
Ala Lys Val Val Tyr Ile Lys Glu Gly Glu Leu Val Pro Ile Asp Asn
500 505 510

gcc agc ccg ctg gaa aaa att cgt ctc gtg cgc cgg cag gcg aaa gag 1584
Ala Ser Pro Leu Glu Lys Ile Arg Leu Val Arg Arg Gln Ala Lys Glu
515 520 525

aaa gtg ttt gtc acc aac tgc ctg cgc gcg ctg cgc cag gtc tca ccc 1632
Lys Val Phe Val Thr Asn Cys Leu Arg Ala Leu Arg Gln Val Ser Pro
530 535 540

ggc ggt tcc att cgc gat atc gcc ttt gtg gtg ctg gtg ggc ggc tca 1680
Gly Gly Ser Ile Arg Asp Ile Ala Phe Val Val Leu Val Gly Gly Ser
545 550 555 560

tcg ctg gac ttt gag atc ccg cag ctt atc acg gaa gcc ttg tcg cac 1728
Ser Leu Asp Phe Glu Ile Pro Gln Leu Ile Thr Glu Ala Leu Ser His
565 570 575

tat ggc gtg gtc gcc ggg cag ggc aat att cgg gga aca gaa ggg ccg 1776
Tyr Gly Val Val Ala Gly Gln Gly Asn Ile Arg Gly Thr Glu Gly Pro
580 585 590

cgc aat gcg gtc gcc acc ggg ctg cta ctg gcc ggt cag gcg aat taa 1824
Arg Asn Ala Val Ala Thr Gly Leu Leu Ala Gly Gln Ala Asn
595 600 605

<210> 50

<211> 607

<212> PRT

5 <213> Klebsiella pneumoniae

<400> 50

Met Pro Leu Ile Ala Gly Ile Asp Ile Gly Asn Ala Thr Thr Glu Val
1 5 10 15

Ala Leu Ala Ser Asp Tyr Pro Gln Ala Arg Ala Phe Val Ala Ser Gly
20 25 30

Ile Val Ala Thr Thr Gly Met Lys Gly Thr Arg Asp Asn Ile Ala Gly
35 40 45

Thr Leu Ala Ala Leu Glu Gln Ala Leu Ala Lys Thr Pro Trp Ser Met
50 55 60

Ser Asp Val Ser Arg Ile Tyr Leu Asn Glu Ala Ala Pro Val Ile Gly
65 70 75 80

Asp Val Ala Met Glu Thr Ile Thr Glu Thr Ile Ile Thr Glu Ser Thr
85 90 95

Met Ile Gly His Asn Pro Gln Thr Pro Gly Gly Val Gly Val Gly Val
100 105 110

Gly Thr Thr Ile Ala Leu Gly Arg Leu Ala Thr Leu Pro Ala Ala Gln
115 120 125

Tyr Ala Glu Gly Trp Ile Val Leu Ile Asp Asp Ala Val Asp Phe Leu
130 135 140

Asp Ala Val Trp Trp Leu Asn Glu Ala Leu Asp Arg Gly Ile Asn Val
145 150 155 160

Val Ala Ala Ile Leu Lys Lys Asp Asp Gly Val Leu Val Asn Asn Arg
165 170 175

Leu Arg Lys Thr Leu Pro Val Val Asp Glu Val Thr Leu Leu Glu Gln
180 185 190

Val Pro Glu Gly Val Met Ala Ala Val Glu Val Ala Ala Pro Gly Gln
195 200 205

Val Val Arg Ile Leu Ser Asn Pro Tyr Gly Ile Ala Thr Phe Phe Gly
210 215 220

Leu Ser Pro Glu Glu Thr Gln Ala Ile Val Pro Ile Ala Arg Ala Leu
225 230 235 240

Ile Gly Asn Arg Ser Ala Val Val Leu Lys Thr Pro Gln Gly Asp Val
245 250 255

Gln Ser Arg Val Ile Pro Ala Gly Asn Leu Tyr Ile Ser Gly Glu Lys
260 265 270

Arg Arg Gly Glu Ala Asp Val Ala Glu Gly Ala Glu Ala Ile Met Gln
275 280 285

Ala Met Ser Ala Cys Ala Pro Val Arg Asp Ile Arg Gly Glu Pro Gly
290 295 300

Thr His Ala Gly Gly Met Leu Glu Arg Val Arg Lys Val Met Ala Ser
305 310 315 320

Leu Thr Gly His Glu Met Ser Ala Ile Tyr Ile Gln Asp Leu Leu Ala
325 330 335

Val Asp Thr Phe Ile Pro Arg Lys Val Gln Gly Gly Met Ala Gly Glu
340 345 350

Cys Ala Met Glu Asn Ala Val Gly Met Ala Ala Met Val Lys Ala Asp
355 360 365

Arg Leu Gln Met Gln Val Ile Ala Arg Glu Leu Ser Ala Arg Leu Gln
370 375 380

Thr Glu Val Val Val Gly Gly Val Glu Ala Asn Met Ala Ile Ala Gly
385 390 395 400

Ala Leu Thr Thr Pro Gly Cys Ala Ala Pro Leu Ala Ile Leu Asp Leu
405 410 415

Gly Ala Gly Ser Thr Asp Ala Ala Ile Val Asn Ala Glu Gly Gln Ile
420 425 430

Thr Ala Val His Leu Ala Gly Ala Gly Asn Met Val Ser Leu Leu Ile
435 440 445

Lys Thr Glu Leu Gly Leu Glu Asp Leu Ser Leu Ala Glu Ala Ile Lys
450 455 460

Lys Tyr Pro Leu Ala Lys Val Glu Ser Leu Phe Ser Ile Arg His Glu
465 470 475 480

Asn Gly Ala Val Glu Phe Phe Arg Glu Ala Leu Ser Pro Ala Val Phe
485 490 495

Ala Lys Val Val Tyr Ile Lys Glu Gly Glu Leu Val Pro Ile Asp Asn
500 505 510

Ala Ser Pro Leu Glu Lys Ile Arg Leu Val Arg Arg Gln Ala Lys Glu
515 520 525

Lys Val Phe Val Thr Asn Cys Leu Arg Ala Leu Arg Gln Val Ser Pro
530 535 540

Gly Gly Ser Ile Arg Asp Ile Ala Phe Val Val Leu Val Gly Gly Ser
545 550 555 560

Ser Leu Asp Phe Glu Ile Pro Gln Leu Ile Thr Glu Ala Leu Ser His
565 570 575

Tyr Gly Val Val Ala Gly Gln Gly Asn Ile Arg Gly Thr Glu Gly Pro
580 585 590

Arg Asn Ala Val Ala Thr Gly Leu Leu Leu Ala Gly Gln Ala Asn
595 600 605

<211> 13669
<212> DNA
<213> Secuencia Artificial

<220>

5 <223> Plásmido

<400> 51

tagtaaagcc ctcgctagat tttaatgcgg atgttgcgat tacttcgcca actattgcga	60
taacaagaaa aagccagcct ttcattgatat atctcccaat ttgtgtaggg ottattatgc	120
acgcttaaaa ataataaaag cagacttgac ctgatagttt ggctgtgagc aattatgtgc	180
ttagtgcata taacgcttga gtttaagccgc gccgcgaagc ggctgcggct tgaacgaatt	240
gttagacatt atttgccgac taccttggtg atctcgcctt tcacgtagtg gacaaattct	300
tccaactgat ctgcgcgcga ggccaagcga tcttcttctt gtccaagata agcctgtcta	360
gcttcaagta tgacgggctg atactgggac gccagggcgt ccattgcccc gtcggcagcg	420
acatccttcg gcgcgatttt gccgggtact gcgctgtacc aaatgcggga caacgtaagc	480
actacatttc gctcatcgcc agcccagtcg gccggcgagt tccatagcgt taaggtttca	540
tttagcgcct caaatagata ctgttcagga accggatcaa agagttcctc cgccgctgga	600
cctaccaagg caacgctatg ttctcttgct ttgtgcagca agatagccag atcaatgtcg	660
atcgtggctg gctcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc	720
agttcgcgct tagctggata acgccacgga atgatgtcgt cgtgcacaac aatgggtgact	780
tctacagcgc ggagaatctc gctctctcca ggggaagccg aagtttcaa aaggtcgttg	840
atcaaagctc gccgcgttgt ttcattcaagc cttacgggtca ccgtaaccag caaatcaata	900
tcaactgttg gcttcaggcc gccatccact gcggagccgt acaaatgtac ggccagcaac	960
gtcggttcga gatggcgctc gatgacgcca actacctctg atagttgagt cgataacttcg	1020
gcgatcaccg cttccctcat gatgtttaac tttgttttag gccgactgcc ctgctgcgta	1080
acatcgttgc tgctccataa catcaaactc cgacccacgg cgtaacgcgc ttgctgcttg	1140
gatgcccag gcatagactg taccctcaaaa aaacagtcct aacaagccat gaaaaccgcc	1200
actgcgccgt taccaccgct gcgttcggtc aaggttctgg accagttgcg tgagcgcata	1260
cgctacttgc attacagctt acgaaccgaa caggcttatg tccactgggt tcgtgccttc	1320
atccgtttcc acggtgtgcg tcacccggca accttgggca gcagcgaagt cgaggcattt	1380
ctgtcctggc tggcgaacga gcgcaagggt tcgggtctcca cgcacgtca ggcattggcg	1440
gccttgctgt tcttctacgg caaggtgctg tgcacggatc tgccttggt tcaggagatc	1500
ggaagacctc ggccgtcgcg gcgcttgccg gtggtgctga ccccgatga agtggttcgc	1560

atcctcgggtt ttctggaagg cgagcatcgt ttgttcgccc agcttctgta tggaacgggc	1620
atgcggatca gtgagggttt gcaactgcgg gtcaaggatc tggatttcga tcacggcacg	1680
atcatcgtgc gggagggcaa gggctccaag gatcgggcct tgatgttacc cgagagcttg	1740
gcaccagacc tgcgcgagca ggggaattaa ttcccacggg ttttgctgcc cgcaaacggg	1800
ctgttctggt gttgctagtt tggtatcaga atcgcagatc cggcttcagc cggtttgccg	1860
gctgaaagcg ctatttcttc cagaattgcc atgatttttt cccacaggga ggcgtcactg	1920
gctcccggtg tgcgggcagc tttgattcga taagcagcat cgcctgtttc aggctgtcta	1980
tgtgtgactg ttgagctgta acaagtgtgc tcagggtgtc aatttcatgt tctagttgct	2040
ttgttttact ggtttcacct gttctattag gtgttacatg ctgttcatct gttacattgt	2100
cgatctgttc atggtgaaca gctttgaatg caccaaaaac tcgtaaaagc tctgatgtat	2160
ctatcttttt tacaccgttt tcatctgtgc atatggacag ttttcccttt gatatgtaac	2220
ggtgaacagt tgttctactt ttgtttgtta gtcttgatgc ttcactgata gatacaagag	2280
ccataagaac ctcagatcct tccgtattta gccagtatgt tctctagtgt ggttcgtgtg	2340
ttttgcgtga gccatgagaa cgaaccattg agatcatact tactttgcat gtcactcaaa	2400
aattttgcct caaaactggt gagctgaatt tttgcagtta aagcatcgtg tagtgttttt	2460
cttagtccgt tatgtaggta ggaatctgat gtaatggttg ttggtatttt gtcaccattc	2520
attttttatct ggttgttctc aagttcgggt acgagatcca tttgtctatc tagttcaact	2580
tggaaaatca acgtatcagt cgggcggcct cgcttatcaa ccaccaatth catattgctg	2640
taagtgttta aatctttact tattggtttc aaaaccatt ggtaagcct tttaaactca	2700
tggtagtatt tttcaagcat taacatgaac tttaattcat caaggctaatt ctctatattt	2760
gccttgtagg ttttcttttg tggttagttct tttaataacc actcataaat cctcatagag	2820
tatttgtttt caaaagactt aacatgttcc agattatatt ttatgaatth ttttaactgg	2880
aaaagataag gcaatatctc ttcactaaaa actaattcta atttttcgct tgagaacttg	2940
gcatagtttg tccactggaa aatctcaaag cctttaacca aaggattcct gatttccaca	3000
gttctcgtca tcagctctct ggttgcttta gctaatacac cataagcatt ttccctactg	3060
atgttcatca tctgagcgta ttggttataa gtgaacgata ccgtccgttc tttccttgta	3120
gggttttcaa tcgtgggggt gagtagtgcc acacagcata aaattagctt ggtttcatgc	3180
tccgttaagt catagcgact aatcgctagt tcatthgctt tgaaaacaac taattcagac	3240
atacatctca attggtctag gtgattthaa tcactatacc aattgagatg ggctagtcaa	3300
tgataattac tagtcccttt cctthgagtt gtgggtatct gtaaatctg ctagacctth	3360
gctggaaaac ttgtaaattc tgctagaccc tctgtaaatt ccgctagacc tttgtgtgtt	3420

ttttttgttt atattcaagt ggttataatt tatagaataa agaaagaata aaaaaagata	3480
aaaagaatag atcccagccc tgtgtataac tcactacttt agtcagttcc gcagtattac	3540
aaaaggatgt cgcaaacgct gtttgctcct ctacaaaaca gaccttaaaa ccctaaaggc	3600
ttaagtagca ccctcgcaag ctogggcaaa tcgctgaata ttccttttgt ctccgaccat	3660
caggcacctg agtcgctgtc tttttcgtga cattcagttc gctgcgctca cggctctggc	3720
agtgaatggg ggtaaattggc actacaggcg ccttttatgg attcatgcaa ggaaactacc	3780
cataatacaa gaaaagcccc tcacgggctt ctcagggcgt tttatggcgg gtctgctatg	3840
tggtgctatc tgactttttg ctgttcagca gttcctgccc tctgattttc cagtctgacc	3900
acttcggatt atcccgtgac aggtcattca gactggctaa tgcacccagt aaggcagcgg	3960
tatcatcaac aggcttacct gtcttactgt cgggaattca tttaaatagt caaaagcctc	4020
cgaccggagg cttttgactg ctaggcgatc tgtgctgttt gccacggtat gcagcaccag	4080
cgcgagatta tgggctcgca cgctcgactg tcggacgggg gcactggaac gagaagtcag	4140
gcgagccgtc acgcccttga caatgccaca tcctgagcaa ataattcaac cactaaacaa	4200
atcaaccgcg tttcccgag gtaaccaagc ttgcgggaga gaatgatgaa caagagccaa	4260
caagttcaga caatcacctt gcccgccgcc cagcaaatgg cggcggcggg ggaaaaaaa	4320
gccactgaga tcaacgtggc ggtggtgttt tccgtagtgt accgcggagg caacacgctg	4380
cttatccagc ggatggacga ggccttcgtc tccagctgcg atatttccct gaataaagcc	4440
tggagcgcct gcagcctgaa gcaaggtacc catgaaatta cgtcagcggg ccagccagga	4500
caatctctgt acggtctgca gctaaccaac caacagcgaa ttattatttt tggcggcggc	4560
ctgccagtta tttttaatga gcaggttaatt ggcgcgctcg gcgttagcgg cggtagcggtc	4620
gagcaggatc aattattagc ccagtgcgcc ctggattgtt tttccgcatt ataacctgaa	4680
gcgagaaggc atattatgag ctatcgtatg ttccgccagg cattctgagt gttaacgagg	4740
ggaccgtcat gtcgctttca ccgccaggcg tacgcctgtt ttacgatccg cgcgggcacc	4800
atgccggcgc catcaatgag ctgtgctggg ggctggagga gcaggggggc ccctgccaga	4860
ccataacctt tgacggaggc ggtgacgccg ctgcgctggg cgccttggcg gccagaagct	4920
cgcctctgcg ggtgggtatc gggctcagcg cgtccggcga gatagccctc actcatgccc	4980
agctgccggc ggacgcgccg ctggctaccg gacacgtcac cgatagcgac gatcaactgc	5040
gtacgctcgg cgccaacgcc gggcagctgg ttaaagtcct gccgttaagt gagagaaact	5100
gaatgtatcg tatctatacc cgcaccgggg ataaaggcac caccgccctg tacggcggca	5160
gccgcacga gaaagaccat attcgcgtcg aggcctacgg caccgtcgat gaactgatat	5220
cccagctggg cgtctgctac gccacgacct gcgacgccgg gctgcgggaa agcctgcacc	5280
atattcagca gacgctgttc gtgctggggg ctgaactggc cagcgatgcg cggggcctga	5340

ccgcctgag ccagacgac gccgaagagg agatcaccgc cctggagcgg cttatcgacc	5400
gcaatatggc cgagagcggc ccgttaaaac agttcgtgat cccggggagg aatctcgcc	5460
ctgcccagct gcacgtggcg cgcaccagc cccgtcggt cgaacgcctg ctgacggcca	5520
tggaccgcgc gcatccgctg cgcgacgcgc tcaaacgcta cagcaatcgc ctgtcggatg	5580
ccctgttctc catggcgga atcgaagaga ctaggcctga tgcttgctg tgaactggcc	5640
tagcaaacac agaaaaaagc ccgcacctga cagtgcgggc ttttttttc ctaggcgatc	5700
tgtgctgttt gccacggtat gcagcaccag cgcgagatta tgggctcgca cgctcgactg	5760
tcggacgggg gcaactggaac gagaagtcag gcgagccgct acgcccttga caatgccaca	5820
tcctgagcaa ataattcaac cactaaacaa atcaaccgcg tttcccgagg gtaaccaagc	5880
ttcacctttt gagccgatga acaatgaaaa gatcaaacg atttgcagta ctggcccagc	5940
gccccgtcaa tcaggacggg ctgattggcg agtggcctga agaggggctg atcgccatgg	6000
acagcccctt tgacccggtc tcttcagtaa aagtggacaa cggctcgatc gtcgaactgg	6060
acggcaaacg ccgggaccag tttgacatga tcgaccgatt tatcgccgat tacgcgatca	6120
acgttgagcg cacagagcag gcaatgcgc tggaggcggg ggaaatagcc cgtatgctgg	6180
tggatattca cgtcagccgg gaggagatca ttgccatcac taccgccatc acgccggcca	6240
aagcggctga ggtgatggcg cagatgaacg tggtggagat gatgatggcg ctgcagaaga	6300
tgctgccccg ccggaccccc tccaaccagt gccacgtcac caatctcaa gataatccgg	6360
tgcagattgc cgtgacgcc gccgaggccg ggatccgcgg cttctcagaa caggagacca	6420
cggtcggtat cgcgcgtac gcgccgttta acgccctggc gctgttggtc ggttcgcagt	6480
gcggccgccc cggcgtgttg acgcagtgt cggtggaaga ggccaccgag ctggagctgg	6540
gcatgcgtgg cttaaccagc tacgccgaga cgggtgcgg ctacggcacc gaagcggat	6600
ttaccgacgg cgatgatacg ccgtgggtcaa aggcgctcct cgcctcgcc tacgcctccc	6660
gcgggttgaa aatgcgctac acctccggca ccggatccga agcgtgatg ggctattcgg	6720
agagcaagtc gatgctctac ctogaatcgc gctgcatctt cattactaaa ggcgcgggg	6780
ttcagggaact gcaaacggc gcggtgagct gtatcgccat gaccggcgct gtgccgtcgg	6840
gcattcgggc ggtgctggcg gaaaacctga tcgcctctat gctcgacctc gaagtggcgt	6900
ccgccaacga ccagactttc tccactcgg atattcgccg caccgcgcgc accctgatgc	6960
agatgctgcc gggcaccgac tttattttct ccggctacag cgcggtgccg aactacgaca	7020
acatgttcgc cggctcgaac ttcatgacgg aagattttga tgattacaac atcctgcagc	7080
gtgacctgat ggttgacggc ggcctgcgtc cggtgaccga ggcggaacc attgccattc	7140
gccagaaagc ggcgcggggc atccaggcgg ttttcgcga gctggggctg ccgccaatcg	7200

ccgacgagga ggtggaggcc gccacctacg cgcacggcag caacgagatg ccgccgcgta	7260
acgtggtgga ggatctgagt gcggtggaag agatgatgaa gcgcaacatc accggcctcg	7320
atattgtcgg cgcgctgagc cgcagcggct ttgaggatat cgccagcaat attctcaata	7380
tgctgcgcca gcgggtcacc ggcgattacc tgcagacctc ggccattctc gatcggcagt	7440
tcgaggtggt gagtgcggtc aacgacatca atgactatca ggggccgggc accggctatc	7500
gcatctctgc cgaacgctgg gcggagatca aaaatattcc gggcgtggtt cagccccaga	7560
ccattgaata aggcgggtatt cctgtgcaac agacaaccca aattcagccc tcttttacct	7620
tgaaaacccg cgagggcggg gtagcttctg ccgatgaacg cgccgatgaa gtggtgatcg	7680
gcgtcggccc tgccttcgat aaacaccagc atcacactct gatcgatatg ccccatggcg	7740
cgatcctcaa agagctgatt gccggggtgg aagaagaggg gcttcacgcc cggtggtgc	7800
gcattctgcg cacgtccgac gtctccttta tggcctggga tgcggccaac ctgagcggct	7860
cggggatcgg catcgggtatc cagtogaagg ggaccacggt catccatcag cgcgatctgc	7920
tgccgctcag caacctggag ctgttctccc aggcgccgct gctgacgctg gagacctacc	7980
ggcagattgg caaaaacgct gcgcgctatg cgcgcaaaga gtcaccttcg ccggtgccgg	8040
tggtgaacga tcagatggtg cggcogaaat ttatggccaa agccgcgcta tttcatatca	8100
aagagaccaa acatgtggtg caggacgccg agcccgtcac cctgcacatc gacttagtaa	8160
gggagtgacc atgagcgaga aaaccatgcg cgtgcaggat tatccgttag ccacccgctg	8220
cccgagcat atcctgacgc ctaccggcaa accattgacc gatattacct tcgagaaggt	8280
gctctctggc gaggtgggccc cgcaggatgt gcggatctcc cgccagacct ttgagtacca	8340
ggcgagatt gccgagcaga tgcagcgcca tgcggtggcg cgcaatttcc gccgcgcggc	8400
ggagcttacc gccattcctg acgagcgcgt tctggctatc tataacgcgc tgcgccggtt	8460
ccgctcctcg caggcggagc tgctggcgat gcgcgacgag ctggagcaca cctggcatgc	8520
gacagtgaat gccgcctttg tccgggagtc ggcggaagtg tatcagcagc ggcataagct	8580
gcgtaaagga agctaagcgg aggtcagcat gccgttaata gccgggattg atatcggcaa	8640
cgccaccacc gaggtggcgc tggcgtccga ctaccgcgag gcgagggcgt ttgttgccag	8700
cgggatcgtc gcgacgacgg gcatgaaagg gacgcgggac aatatcgccg ggaccctcgc	8760
cgcgctggag caggccctgg cgaaaacacc gtggtcgatg agcgatgtct ctgcgatcta	8820
tcttaacgaa gccgcgccgg tgattggcga tgtggcgatg gagaccatca ccgagaccat	8880
tatcaccgaa tcgaccatga tcggtcataa cccgcagacg ccgggcgggg tgggcgttg	8940
cgtggggacg actatcgccc tcgggcggct ggcgacgctg ccggcgggcg agtatgccga	9000
ggggtggatc gtactgattg acgacgccgt cgatttcctt gacgccgtgt ggtggctcaa	9060
tgaggcgctc gaccggggga tcaacgtggt ggcggcgatc ctcaaaaagg acgacggcgt	9120

gctggtgaac aaccgcctgc gtaaaaccct gccggtggtg gatgaagtga cgctgctgga	9180
gcaggtcccc gagggggtaa tggcggcggt ggaagtggcc gcgccgggcc aggtggtgcg	9240
gacctgtcgc aatccctacg ggatcgccac cttcttcggg ctaagcccg aagagacca	9300
ggccatcgtc cccatcgccc gcgccctgat tggcaaccgt tccgcggtgg tgctcaagac	9360
cccgcagggg gatgtgcagt cgcgggtgat cccggcgggc aacctctaca ttagcggcga	9420
aaagcgccgc ggagaggccg atgtcgccga gggcgcgga gccatcatgc aggcgatgag	9480
cgcctgcgct ccggtacgcg acatccgcg cgaaccgggc acccacgcg gcggcatgct	9540
tgagcgggtg cgcaaggtaa tggcgtccct gaccggccat gagatgagcg cgatatacat	9600
ccaggatctg ctggcggtgg atacgtttat tccgcgcaag gtgcagggcg ggatggccgg	9660
cgagtgcgcc atggagaatg ccgtcgggat ggcggcgatg gtgaaagcgg atcgtctgca	9720
aatgcaggtt atcgccgcg aactgagcgc ccgactgcag accgaggtgg tggggggcg	9780
cgtggaggcc aadatggcca tcgccggggc gttaaccact cccggtctgt cggcgccgct	9840
ggcgatcctc gacctcggcg ccggctcgac ggatgcggcg atcgtcaacg cggaggggca	9900
gataacggcg gtccatctcg ccggggcggg gaatatggtc agcctgttga ttaaaaccga	9960
gctgggcctc gaggatcttt cgctggcgga agcgataaaa aaataccgc tggccaaagt	10020
ggaaagcctg ttcagtattc gtcacgagaa tggcgcggtg gagttctttc gggaagccct	10080
cagcccggcg gtgttcgcca aagtgggtga catcaaggag ggcgaactgg tgccgatcga	10140
taagccagc ccgctgga aaattcgtct cgtgcgcgg caggcgaag agaaagtgtt	10200
tgtcaccaac tgctgcgcg cgctgcgcca ggtctcacc gcggttcca ttcgcgatat	10260
cgcctttgtg gtgctggtgg gcggctcatc gctggacttt gagatcccg agcttatcac	10320
ggaagccttg tcgcactatg gcgtggtcgc cgggcagggc aatattcggg gaacagaagg	10380
gccgcgcaat gcggtcgcca ccgggtgct actggcgggt caggcgaatt aaacgggcgc	10440
tcgcgccagc ctctaggtac aaataaaaaa ggcacgtcag atgacgtgcc ttttttcttg	10500
tctagagtac tggcgaaagg gggatgtgct gcaaggcgat taagttgggt aacgccagg	10560
ttttccagc cagcagcttg taaaacgacg gccagtgaat tcgagctcgg taccggggc	10620
ggccgcgcta gcgcccgatc cagctggagt ttgtagaaac gcaaaaaggc catccgtcag	10680
gatggccttc tgcttaattt gatgcctggc agtttatggc ggcgctcctg cccgccaccc	10740
tccgggcccgt tgcttcgcaa cgttcaaadc cgtcccgcg ggatttgtcc tactcaggag	10800
agcgttcacc gacaaacaac agataaaacg aaaggcccag tctttcgact gagcctttcg	10860
ttttatttga tgctggcag ttccctactc tcgcatgggg agaccccaca ctaccatcg	10920
cgtacggcg tttcacttct gagttcggca tggggtcagg tgggaccacc gcgctactgc	10980

cgccaggcaa attctgtttt atcagaccgc ttctgcgttc tgatttaatc tgtatcaggg	11040
tgaaaatctt ctctcatccg ccaaaacagc caagcttgca tgcttcagc ccgggttacc	11100
atttcaacag atcgtcctta gcatataagt agtcgtcaaa aatgaattca acttcgtctg	11160
tttcggcatt gtagccgcca actctgatgg attcgtgggt tttgacaatg atgtcacagc	11220
ctttttcctt taggaagtcc aagtcgaaag tagtggcaat accaatgatc ttacaaccgg	11280
cggtttttcc ggccgcaata cctgctggag cgtcttcaaa tactactacc ttagatttgg	11340
aagggtcttg ctcatatgatc ggatataccta agccattcct gcccttcaga tatgggtctg	11400
gatgaggctt accctgtttg acatcattag cggtaatgaa gtactttgggt ctctgatctc	11460
ccagatgctc gaaccatttt tgtgccatat cacgggtacc ggaagttgcc acagcccatt	11520
tctcttttgg tagagcgttc aaagcgttgc acagcttaac tgcacctggg acttcaatgg	11580
atttttcacc gtacttgacc ggaatttcag cttctaattt gttaacatac tcttcattgg	11640
caaagtctgg agcgaactta gcaatggcat caaacgttct ccaaccatgc gagacttgg	11700
taacgtgttc agcatcgaaa taaggtttgt ccttaccgaa atccctccag aatgcagcaa	11760
tggctgggtg agagatgata atggtaccgt cgacgtcgaa caaagcggcg ttaactttca	11820
aagatagagg ttagtagtc aatccataa ttctagtctg tttcctggat ccaataaatc	11880
taatcttcat gtagatctaa ttcttcaatc atgtccggca ggttcttcat tgggtagtgt	11940
ttgtaaacga tttgtatac ggcttcaaat aatgggaagt cttegacaga gccacatgtt	12000
tccaaccatt cgtgaacttc tttgcaggta attaaacctt gagcggattg gccattcaac	12060
aactcctttt cacattccca ggcttcctta ccagaagtag ccattagcct agcaaccttg	12120
acgtttctac caccagcgca ggtggtgatc aaatcagcaa caccagcaga ctcttggtag	12180
tatgtttctt ctctagattc tgggaaaaac atttgaccga atctgatgat ctcacccaaa	12240
ccgactcttt ggatggcagc agaagcgttg ttacccagc ctagaccttc gacgaaacca	12300
caacctaagg caacaacgtt cttcaaagca ccacagatgg agataccagc aacatcttcg	12360
atgacactaa cgtggaagta aggtctgtgg aacaaggcct ttagaacctt atggtcgacg	12420
tccttgccct cgcctctgaa atcctttgga atgtggttaag caactgttgt ttcagaccag	12480
tgttcttgag cgacttcggt ggcaatgtta gcaccagata gagcaccaca ttgaatacct	12540
agttcctcag tgatgtaaga ggatagcaat tggacacctt tagcaccaac ttcaaaaccc	12600
tttagacagg agatagctct gacgtgtgaa tcaacatgac ctttcaattg gctacagata	12660
cggggcaaaa attgatgtgg aatgttgaaa acgatgatgt cgacatcctt gactgaatca	12720
atcaagtctg gattagcaac caaattgtcg ggtagagtga tgccaggcaa gtatttcacg	12780
ttttgatgtc tagtatttat gatttcagtc aatttttcac cattgatctc ttcttcgaac	12840
accacattt gtactattgg agcgaaaact tctgggtatc ccttacaatt ttcggaacc	12900

accttggcaa tagtagtacc ccagttacca gatccaatca cagtaacctt gaaaggcttt	12960
tccggcagcct tcaaagaaac agaagaggaa cttctctttc taccagcatt caagtggccg	13020
gaagttaagt ttaatctatc agcagcagca gccatggaat tgtcctcctt actagtcag	13080
gtctgtttcc tgtgtgaaat tggtatccgc tcacaattcc acacattata cgagccggat	13140
gattaattgt caacagctca ttccagaata ttgcccagaa ccgttatgat gtcggcgcaa	13200
aaaacattat ccagaacggg agtgcgccct gagcgacacg aattatgcag tgatttacga	13260
cctgcacagc cataccacag cttccgatgg ctgcctgacg ccagaagcat tggcgcacgc	13320
tagccagtac atttaaatgg taccctctag tcaaggcctt aagtgcgtcg tattacggac	13380
tggccgtcgt ttacaacgt cgtgactggg aaaaccctgg cgttacccaa cttaatcgcc	13440
ttgcagcaca tcccccttc gccagctggc gtaatagcga agaggcccg accgatcgcc	13500
cttcccaaca gttgcgcagc ctgaatggcg aatggcgccct gatgcggtat tttctcetta	13560
cgcatctgtg cggatatttc caccgcatat ggtgcactct cagtacaatc tgctctgatg	13620
ccgcatagtt aagccagccc cgacacccgc caacacccgc tgacgagct	13669

<210> 52
 <211> 13543
 <212> DNA
 <213> Secuencia Artificial

5

<220>
 <223> Plásmido

<400> 52	
tagtaaaagcc ctgcgtagat tttaatgagg atgttgcgat tacttcgcca actattgcga	60
taacaagaaa aagccagcct ttcagatgat atctcccaat ttgtgtaggg cttattatgc	120
acgcttaaaa ataataaaag cagacttgac ctgatagttt ggctgtgagc aattatgtgc	180
ttagtgcata taacgcttga gtaagccgc gccgcgaagc ggcgtcggct tgaacgaatt	240
gtagacatt atttgccgac taccttggtg atctgcctt tcacgtagtg gacaaattct	300
tccaactgat ctgcgcgcga ggccaagoga tcttcttctt gtccaagata agcctgtcta	360
gcttcaagta tgacgggctg atactgggoc ggcaggcgct ccattgcca gtcggcagcg	420
acatccttcg gcgcgatttt gccggttact gcgctgtacc aaatgcggga caacgtaagc	480
actacatttc gctcatcgcc agcccagtcg ggcggcgagt tccatagcgt taaggtttca	540
tttagcgcct caaatagatc ctgttcagga accgatcaa agagttcctc cgccgctgga	600
cctaccaagg caacgctatg ttctcttgct ttgtcagca agatagccag atcaatgtcg	660
atcgtggctg gctcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc	720
agttcgcgct tagctggata acgccacgga atgatgtcgt cgtgcacaac aatggtgact	780

tctacagcgc ggagaatctc gctctctcca ggggaagccg aagtttccaa aaggctcgttg	840
atcaaagctc gccgcgttgt ttcataaagc cttacggtca ccgtaaccag caaatcaata	900
tcactgtgtg gcttcaggcc gccatccact gcggagccgt acaaatgtac ggccagcaac	960
gtcggttcga gatggcgctc gatgacgcc aactacctctg atagttgagt cgatacttcg	1020
gcgatcacccg cttccctcat gatgtttaac tttgttttag ggcgactgcc ctgctgcgta	1080
acatcgttgc tgctccataa catcaaacat cgacccacgg cgtaacgcgc ttgctgcttg	1140
gatgcccagag gcatagactg taccctaaaa aaacagtcac aacaagccat gaaaaccgcc	1200
actgcgcctg taccaccgct gcgttcggtc aaggttctgg accagttgcg tgagcgcata	1260
cgctacttgc attacagctt acgaaccgaa caggcttatg tccactgggt tcgtgccttc	1320
atccgtttcc acggtgtgcg tcacccggca accttgggca gcagcgaagt cgaggcattt	1380
ctgtcctggc tggcgaacga gcgcaagggt tcggtctcca cgcacgtca ggcattggcg	1440
gccttgctgt tcttctacgg caagggtgctg tgcacggatc tgccctggct tcaggagatc	1500
ggaagacctc ggcgcgcgcg gcgcttgccg gtggtgctga ccccgatga agtggttcgc	1560
atcctcggtt ttctggaagg cgagcatcgt ttgttcgccc agcttctgta tggaacgggc	1620
atgcggatca gtgagggttt gcaactgcgg gtcaaggatc tggatttcga tcacggcacg	1680
atcatcgtgc gggagggcaa gggctccaag gatcgggcct tgatgttacc cgagagcttg	1740
gcacccagcc tgcgcgagca ggggaattaa ttccacggg ttttgctgcc cgcaaacggg	1800
ctgttctggt gttgctagtt tggtatcaga atcgcagatc cggcttcagc cggtttgccg	1860
gctgaaagcg ctatttcttc cagaattgcc atgatttttt cccacggga ggcgtcactg	1920
gctcccggtg tgcggcagc tttgattcga taagcagcat cgcctgtttc aggctgtcta	1980
tgtgtgactg ttgagctgta acaagttgtc tcagggtgtc aatttcattg tctagttgct	2040
ttgttttact ggtttcacct gttctattag gtgttacatg ctgttcattc gttacattgt	2100
cgatctgttc atggtgaaca gctttgaatg caccaaaaac tcgtaaaagc tctgatgtat	2160
ctatcttttt tacaccgttt tcatctgtgc atatggacag tttcccttt gatatgtaac	2220
ggtgaacagt tgttctactt ttgtttgtta gtcttgatgc ttcactgata gatacaagag	2280
ccataagaac ctcatgcct tccgtattta gccagtatgt tctctagtgt ggttcgttgt	2340
ttttgcgtga gccatgagaa cgaaccattg agatcatact tactttgcat gtcactcaaa	2400
aattttgcct caaaactggt gagctgaatt tttgcagtta aagcatcgtg tagtgttttt	2460
cttagtccgt tatgtaggta ggaatctgat gtaatgggtt ttggtatttt gtcaccattc	2520
atttttatct ggttggtctc aagttcggtt acgagatcca tttgtctatc tagttcaact	2580
tggaaaatca acgtatcagt cgggcggcct cgcttatcaa ccaccaattt catattgctg	2640
taagtgttta aatctttact tattggtttc aaaaccattt ggttaagcct tttaaactca	2700

tggtagttat tttcaagcat taacatgaac ttaaattcat caaggctaat ctctatatatt	2760
gccttgtgag ttttcttttg tgtagttct ttttaataacc actcataaat cctcatagag	2820
tatttgtttt caaaagactt aacatgttcc agattatatt ttatgaattt ttttaactgg	2880
aaaagataag gcaatatctc ttcactaaaa actaattcta atttttcgct tgagaacttg	2940
gcatagtttg tccactggaa aatctcaaag cctttaacca aaggattcct gatttccaca	3000
gttctcgtca tcagctctct ggtagcttta gctaatacac cataagcatt ttccctactg	3060
atgttcatca tctgagcgta ttggttataa gtgaacgata ccgccggtc tttccttgta	3120
gggttttcaa tcgtggggtt gagtagtgc acacagcata aaattagctt ggtttcatgc	3180
tccgttaagt catagcgact aatcgctagt tcatttgctt tgaaaacaac taattcagac	3240
atacatctca attggtctag gtgattttta tcaactatacc aattgagatg ggctagtcaa	3300
tgataattac tagtcctttt cctttgagtt gtgggtatct gttaaattctg ctagacctt	3360
gctggaaaac ttgtaaattc tgctagacct tctgtaaatt ccgctagacc tttgtgtgtt	3420
ttttttgttt atattcaagt gggtataatt tatagaataa agaaagaata aaaaaagata	3480
aaaagaatag atcccagccc tgtgtataac tcaactactt agtcagttcc gcagtattac	3540
aaaaggatgt cgcaaagcgt gtttgctcct ctacaaaaca gaccttaaaa ccctaaaggc	3600
ttaagtagca ccctcgcaag ctccggcaaa tcgctgaata ttccttttgt ctccgacct	3660
caggcacctg agtcgctgtc ttttctgtga cattcagttc gctgcgctca cggctctggc	3720
agtgaatggg ggtaaattggc actacaggcg ccttttatgg attcatgcaa ggaaactacc	3780
cataatacaa gaaaagcccg tcacgggctt ctacgggctt tttatggcggt gtctgctatg	3840
tggtgctatc tgactttttg ctgttcagca gttcctgccc tctgattttc cagtctgacc	3900
acttcggatt atcccgtgac aggtcattca gactggctaa tgcacccagt aaggcagcgg	3960
tatcatcaac aggcttacct gtcttactgt cgggaattca tttaaatagt caaaagcctc	4020
cgaccggagg cttttgactg ctaggcgatc tgtgctgttt gccacggtat gcagcaccag	4080
cgcgagatta tgggctcgca cgctcgactg tcggacgggg gcaactggaac gagaagtacg	4140
gcgagccgtc acgcccttga ctatgccaca tcctgagcaa ataattcaac cactaaacaa	4200
atcaaccgcg tttcccggag gtaaccaagc ttgcgggaga gaatgatgaa caagagccaa	4260
caagttcaga caatcaccct ggccgccgcc cagcaaatgg cggcggcggt ggaaaaaaa	4320
gccactgaga tcaactggc ggtggtgttt tccgtagttg accgcggagg caacacgctg	4380
cttatccagc ggatggacga ggccttcgtc tccagctgcg atatttccct gaataaagcc	4440
tggagcgcct gcagcctgaa gcaaggtacc catgaaatta cgtcagcgggt ccagccagga	4500
caatctctgt acggtctgca gctaaccaac caacagcgaa ttattatttt tggcggcggc	4560

ctgccagtta tttttaatga gcaggtaatt ggcgcgctcg gcgttagcgg cggtagcgtc	4620
gagcaggatc aattattagc ccagtgcgcc ctggattggt tttccgcatt ataacctgaa	4680
gcgagaaggt atattatgag ctatcgtatg ttccgccagg cattctgagt gttaacgagg	4740
ggaccgtcat gtcgctttca ccgccaggcg tacgcctggt ttacgatccg cgcgggcacc	4800
atgccggcgc catcaatgag ctgtgctggg ggctggagga gcagggggtc ccctgccaga	4860
ccataacctg tgacggaggc ggtgacgccg ctgcgctggg cgccctggcg gccagaagct	4920
cgcacctgcg ggtgggtatc gggctcagcg cgccggcgga gatagccctc actcatgccc	4980
agctgccggc ggacgcgccg ctggctaccg gacacgtcac cgatagcgac gatcaactgc	5040
gtacgctcgg cgccaacgcc gggcagctgg ttaaagtcct gccgttaagt gagagaaact	5100
gaatgtatcg tatctatacc cgcaccgggg ataaaggcac caccgccctg tacggcggca	5160
gccgcatcga gaaagaccat attcgcgtcg aggcctacgg caccgtcgat gaactgatat	5220
cccagctggg cgtctgctac gccacgacct gcgacgccgg gctgcgggaa agcctgcacc	5280
atattcagca gacgctgttc gtgctggggg ctgaactggc cagcgatgcg cggggcctga	5340
cccgctgag ccagacgatc ggccaagagg agatcaccgc cctggagcgg cttatcgacc	5400
gcaatatggc cgagagcggc ccgttaaaac agttcgtgat cccggggagg aatctcgct	5460
ctgcccagct gcacgtggcg cgcaccagct cccgtcggct cgaacgcctg ctgacggcca	5520
tggaccgcgc gcatccgctg cgcgacgcgc tcaaacgcta cagcaatcgc ctgtcggatg	5580
ccctgttctc catggcgcgga atcgaagaga ctaggcctga tgcttgcgct tgaactggcc	5640
tagcaaacac agaaaaaagc ccgcacctga cagtgcgggc ttttttttct ctaggcgatc	5700
tgtgtgtttt gccacggtat gcagcaccag cgcgagatta tgggctcgca cgctcgactg	5760
tcggacgggg gactggaac gagaagtcag gcgagccgtc acgcccttga ctatgccaca	5820
tcctgagcaa ataattcaac cactaaacaa atcaaccgcg tttccggag gtaaccaagc	5880
ttcacctttt gagccgatga acaatgaaaa gatcaaacg atttgagta ctggcccagc	5940
gccccgtcaa tcaggacggg ctgattggcg agtggcctga agaggggctg atcgccatgg	6000
acagcccctt tgacctggtc tottcagtaa aagtggacaa cggctctgatc gtcgaactgg	6060
acggcaaacg ccgggaccag tttgacatga tcgaccgatt tatcgccgat tacgcgatca	6120
acgttgagcg cacagagcag gcaatgcgcc tggaggcggt ggaaatagcc cgtatgctgg	6180
tggatattca cgtcagccgg gaggagatca ttgccatcac taccgccatc acgccggcca	6240
aagcggtcga ggtgatggcg cagatgaacg tgggtggagat gatgatggcg ctgcagaaga	6300
tgcttgcccg ccggaccccc tccaaccagt gccacgtcac caatctcaaa gataatccgg	6360
tgagattgac cgctgacgcc gccgaggccg ggatccgcgg cttctcagaa caggagacca	6420
cggtcggtat cgcgcgctac gcgcggttta acgccctggc gctgttggtc ggttcgcagt	6480

gcggccgccc cggcgtgttg acgcagtgct cgggtggaaga ggccaccgag ctggagctgg	6540
gcatgcgtgg cttaaccagc tacgccgaga cgggtgcggt ctacggcacc gaagcggat	6600
ttaccgacgg cgatgatacg ccgtgggtcaa aggcgttcct cgcctcggcc tacgcctccc	6660
gcgggttgaa aatgcgctac acctccggca ccggatccga agcgtgatg ggctattcgg	6720
agagcaagtc gatgctctac ctogaatcgc gctgcatctt cattactaaa ggcgccgggg	6780
ttcagggact gcaaaacggc gcgggtgagct gtatcggcat gaccggcgct gtgccgtcgg	6840
gcattcgggc ggtgctggcg gaaaacctga tcgcctctat gctcgacctc gaagtggcgt	6900
ccgccaacga ccagactttc tcccactcgg atattcgccg caccgcgcgc accctgatgc	6960
agatgctgcc gggcaccgac tttattttct ccggctacag cgcggtgccg aactacgaca	7020
acatgttcgc cggctcgaac ttcgatgcgg aagattttga tgattacaac atcctgcagc	7080
gtgacctgat ggttgacggc ggctgcgctc cggtgaccga ggcggaacc attgccattc	7140
gccagaaagc ggcgcgggcg atccaggcgg tttccgcga gctggggctg ccgccaatcg	7200
ccgacgagga ggtggaggcc gccacctacg cgcacggcag caacgagatg ccgccgcgta	7260
acgtggtgga ggatctgagt gcgggtggaag agatgatgaa gcgcaacatc accggcctcg	7320
atattgtcgg cgcgctgagc cgcagcggct ttgaggatat cgccagcaat attctcaata	7380
tgctgcgcca gcgggtcacc ggcgattacc tgcgacctc ggccattctc gatcggcagt	7440
tcgaggtggt gagtgcggtc aacgacatca atgactatca ggggccgggc accggctatc	7500
gcatctctgc cgaacgctgg gcggagatca aaaatattcc gggcgtggtt cagcccgaca	7560
ccattgaata aggcggtatt cctgtgcaac agacaaccca aattcagccc tcttttacct	7620
tgaaaacccg cgagggcggg gtagcttctg ccgatgaacg cgcgatgaa gtggtgatcg	7680
gcgtcggccc tgccttcgat aaacaccagc atcacactct gatcgatatg ccccatggcg	7740
cgatcctcaa agagctgatt gccggggtgg aagaagaggg gcttcacgcc cgggtggtgc	7800
gcattctgcg cacgtccgac gtctccttta tggcctggga tgcggccaac ctgagcggct	7860
cggggatcgg catcggtatc cagtcgaagg ggaccacggt catccatcag cgcgatctgc	7920
tgccgctcag caacctggag ctgttctccc aggcgccgct gctgacgctg gagacctacc	7980
ggcagattgg caaaaacgct gcgcgctatg cgcgcaaaga gtcaccttcg ccggtgccgg	8040
tggtgaacga tcagatggtg cggccgaaat ttatggccaa agccgcgcta tttcatatca	8100
aagagaccaa acatgtggtg caggacgccg agcccgctac cctgcacatc gacttagtaa	8160
gggagtgacc atgagcgaga aaaccatgcg cgtgcaggat tatccgttag ccaccgctg	8220
cccgagcat atcctgacgc ctaccggcaa accattgacc gatattacct tcgagaaggt	8280
gctctctggc gaggtgggcc cgcaggatgt gcggatctcc cgccagacct ttgagtacca	8340

ggcgcagatt gccgagcaga tgcagcgcca tgcggtggcg cgcaattttcc gccgcgcggc	8400
ggagcttata gccattcctg acgagcgcat tctggctata tataacgcgc tgcgcccgtt	8460
ccgctcctcg caggcggagc tgctggcgat cgccgacgag ctggagcaca cctggcatgc	8520
gacagtgaat gccgcctttg tccgggagtc ggcggaagtg tatcagcagc ggcataagct	8580
gcgtaaagga agctaagcgg aggtcagcat gccgttaata gccgggattg atatcggcaa	8640
cgccaccacc gaggtggcgc tggcgtccga ctaccgcgag gcgagggcgt ttgttgccag	8700
cgggatcgtc gcgacgacgg gcatgaaagg gacgcgggac aatatcgccg ggaccctcgc	8760
cgcgctggag caggccctgg cgaaaacacc gtggctcgatg agcgatgtct ctcgcatcta	8820
tcttaacgaa gccgcgccgg tgattggcga tgtggcgatg gagaccatca ccgagaccat	8880
tatcaccgaa tcgaccatga tcggtcataa cccgcagacg ccgggcgggg tgggcgttgg	8940
cgtggggacg actatcgccc tcgggcggct ggcgacgctg ccggcggcgc agtatgccga	9000
ggggtggatc gtactgattg acgacgccgt cgatttcctt gacgccgtgt ggtggctcaa	9060
tgaggcgctc gaccggggga tcaacgtggt ggcggcgatc ctcaaaaagg acgacggcgt	9120
gctggtgaac aaccgcctgc gtaaaacctt gccggtggtg gatgaagtga cgctgctgga	9180
gcaggtcccc gagggggtaa tggcggcggt ggaagtggcc gcgccgggcc aggtggtgcg	9240
gatcctgtcg aatccctacg ggatcgccac cttcttcggg ctaagcccg aagagacca	9300
ggccatcgtc cccatcgccc gcgccctgat tggcaaccgt tccgcggtgg tgctcaagac	9360
cccgcagggg gatgtgcagt cgcgggtgat cccggcgggc aacctctaca ttagcggcga	9420
aaagcgccgc ggagaggccg atgtcgccga ggcgcgggaa gccatcatgc aggcgatgag	9480
cgcctgcgct ccggtacgcg acatccgcgg cgaaccgggc acccacgccg gcggcatgct	9540
tgagcgggtg cgcaaggtaa tggcgtccct gaccggccat gagatgagcg cgatatacat	9600
ccaggatctg ctggcggttg atacgtttat tccgcgcaag gtgcagggcg ggatggccgg	9660
cgagtgcgcc atggagaatg ccgtcgggat ggcggcgatg gtgaaagcgg atcgtctgca	9720
aatgcaggtt atcgcccgcg aactgagcgc ccgactgcag accgaggtgg tgggtggcgg	9780
cgtggaggcc aacatggcca tcgccggggc gttaaccact cccggctgtg cggcgccgct	9840
ggcgatcctc gacctcggcg ccggctcgac ggatgcggcg atcgtcaacg cggaggggca	9900
gataacggcg gtccatctcg ccggggcggg gaatatggtc agcctgttga ttaaaaccga	9960
gctgggcctc gaggatcttt cgctggcgga agcgataaaa aaataccgcg tggccaaagt	10020
ggaaagcctg ttcagtattc gtcacgagaa tggcgcgggtg gaggttctttc gggaagccct	10080
cagcccgcg gtgttcgccca aagtgggtga catcaaggag ggcgaactgg tgccgatcga	10140
taacgccagc ccgctggaag aaattcgtct cgtgcgccgg caggcgaaag agaaagtgtt	10200
tgtcaccaac tgcctgcgcg cgctgcgcca ggtctcacc gccggttcca ttcgcgatat	10260

cgcctttgtg	gtgctggtgg	gcggctcatc	gctggacttt	gagatcccgc	agcttatcac	10320
ggaagccttg	tcgcactatg	gcgtggctcg	cgggcagggc	aatattcggg	gaacagaagg	10380
gccgcgcaat	gcggtcgcca	ccgggctgct	actggccggg	caggcgaatt	aaacggggcg	10440
tcgcgccagc	ctctaggtac	aaataaaaaa	ggcacgtcag	atgacgtgcc	ttttttcttg	10500
tctagcgtgc	accaatgctt	ctggcgtcag	gcagccatcg	gaagctgtgg	tatggctgtg	10560
caggtcgtaa	atcactgcat	aattcgtgtc	gctcaaggcg	cactcccgtt	ctggataatg	10620
ttttttgcgc	cgacatcata	acggttctgg	caaataattct	gaaatgagct	gttgacaatt	10680
aatcatccgg	ctcgtataat	gtgtggaatt	gtgagcggat	aacaatttca	cacaggaaac	10740
agaccatgac	tagtaaggag	gacaattcca	tggctgctgc	tgctgataga	ttaaacttaa	10800
cttcgggcca	cttgaatgct	ggtagaaaga	gaagttcctc	ttctgtttct	ttgaaggctg	10860
ccgaaaagcc	tttcaagggt	actgtgattg	gatctggtaa	ctggggact	actattgcc	10920
aggtggttgc	cgaaaattgt	aagggatacc	cagaagtttt	cgctccaata	gtacaaatgt	10980
gggtgttcga	agaagagatc	aatggtgaaa	aattgactga	aatcataaat	actagacatc	11040
aaaacgtgaa	atacttgctt	ggcatcactc	taccgcagaa	tttggttgct	aatccagact	11100
tgattgattc	agtcaaggat	gtcgacatca	tcgttttcaa	cattccacat	caatttttgc	11160
cccgtatctg	tagccaattg	aaaggctcatg	ttgattcaca	cgtcagagct	atctcctgtc	11220
taaagggttt	tgaagttggt	gctaaagggt	tccaattgct	atcctcttac	atcactgagg	11280
aactaggtat	tcaatgtggt	gctctatctg	gtgctaacat	tgccaccgaa	gtcgtcaag	11340
aacactggtc	tgaaacaaca	gttgcttacc	acattocaaa	ggatttcaga	ggcgagggca	11400
aggacgtcga	ccataagggt	ctaaaggcct	tggtccacag	accttacttc	cacgttagtg	11460
tcacgaaga	tggtgctggt	atctccatct	gtgggtgctt	gaagaacgtt	gttgcccttag	11520
gttggtggtt	cgtcgaagggt	ctaggctggg	gtaacaacgc	ttctgctgcc	atccaaagag	11580
tcgggttggg	tgagatcatc	agattcggtc	aaatgttttt	cccagaatct	agagaagaaa	11640
catactacca	agagtctgct	ggtgttgctg	atttgatcac	cacctgcgct	ggtggtagaa	11700
acgtcaagggt	tgctaggcta	atggctactt	ctggtaagga	cgcctgggaa	tgtgaaaagg	11760
agttgttgaa	tgccaatcc	gctcaagggt	taattacctg	caaagaagtt	cacgaatggt	11820
tggaacatg	tggtctgtgc	gaagacttcc	cattatttga	agccgtatac	caaatcgttt	11880
acaacaacta	cccaatgaag	aacctgccgg	acatgattga	agaattagat	ctacatgaag	11940
attagattta	ttggatccag	gaaacagact	agaattatgg	gattgactac	taaacctcta	12000
tctttgaaag	ttaacgcgcg	tttggttcgac	gtcgacggta	ccattatcat	ctctcaacca	12060
gccattgctg	cattctggag	ggatttcggg	aaggacaaac	cttatttcga	tgctgaacac	12120

gttatccaag tctcgcatgg ttggagaacg tttgatgcca ttgctaagtt cgctccagac 12180
 ttgccaatg aagagtatgt taacaaatta gaagctgaaa ttccggtcaa gtacggtgaa 12240
 aaatccattg aagtcccagg tgcagttaag ctgtgcaacg ctttgaacgc tctacaaaaa 12300
 gagaaatggg ctgtggcaac ttccggtacc cgtgatatgg cacaaaaatg gttcgagcat 12360
 ctgggaatca ggagacaaaa gtacttcatt accgctaatt atgtcaaaca gggtaagcct 12420
 catccagaac catatctgaa gggcaggaat ggcttaggat atccgatcaa tgagcaagac 12480
 ccttccaaat ctaaggtagt agtatttgaa gacgctccag caggtattgc cgccggaaaa 12540
 gccgccggtt gtaagatcat tggatttgcc actactttcg acttggactt cctaaaggaa 12600
 aaaggctgtg acatcattgt caaaaaccac gaatccatca gagttggcgg ctacaatgcc 12660
 gaaacagacg aagttgaatt catttttgac gactacttat atgctaagga cgatctgttg 12720
 aaatggtaac ccgggctgca ggcatgcaag cttggctgtt ttggcggatg agagaagatt 12780
 ttcagcctga tacagattaa atcagaacgc agaagcggtc tgataaaaca gaatttgcct 12840
 gggggcagta gcgcggtggt cccacctgac cccatgccga actcagaagt gaaacgccgt 12900
 agcgccgatg gtagtgtggg gtctcccat gcgagagtag ggaactgcca ggcataaat 12960
 aaaacgaaag gctcagtcga aagactgggc ctttcgtttt atctgttggt tgcggtgaa 13020
 cgctctcctg agtaggacaa atccgccggg agcggatttg aacgttgcca agcaacggcc 13080
 cggaggggtg cgggcaggac gcccgccata aactgccagg catcaaatta agcagaaggc 13140
 catcctgacg gatggccttt ttgcgtttct acaaactcca gctggatcgg gcgctagagt 13200
 atacatttaa atggtaccct ctagtcaagg ccttaagtga gtcgtattac ggactggccg 13260
 tcgttttaca acgtcgtgac tgggaaaacc ctggcggttac ccaacttaat cgccttgacg 13320
 cacatcccc ttctgccagc tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc 13380
 aacagttgcg cagcctgaat ggcgaatggc gcctgatgcg gtattttctc cttacgcac 13440
 tgtgcggtat ttcacaccgc atatggtgca ctctcagtac aatctgctct gatgccgcac 13500
 agttaagcca gccccgacac ccgccaacac ccgctgacga gct 13543

<210> 53
 <211> 13543
 <212> DNA
 <213> Secuencia Artificial

5

<220>
 <223> Plásmido

<400> 53
 tagtaaagcc ctgcctagat tttaatgcgg atgttgcgat tacttcgcca actattgcga 60
 taacaagaaa aagccagcct ttcatgatat atctcccaat ttgtgtaggg cttattatgc 120
 acgcttaaaa ataataaaag cagacttgac ctgatagttt ggctgtgagc aattatgtgc 180

ttagtgcac	taacgcttga	gttaagccgc	gccgcgaagc	ggcgtcggct	tgaacgaatt	240
gtagacatt	atttgccgac	taccttggtg	atctcgccct	tcacgtagt	gacaaattct	300
tccaactgat	ctgcgcgcga	ggccaagcga	tcttcttctt	gtccaagata	agcctgtcta	360
gcttcaagta	tgacgggctg	atactgggcc	ggcaggcgct	ccattgcca	gtcggcagcg	420
acatccttcg	gcgcgatttt	gccggttact	gcgctgtacc	aaatgcggga	caacgtaagc	480
actacatttc	gctcatcgcc	agcccagtcg	ggcggcgagt	tccatagcgt	taaggtttca	540
tttagcgcct	caaatagatc	ctgttcagga	accggatcaa	agagttcctc	cgccgctgga	600
cctaccaagg	caacgctatg	ttctcttgct	tttgtcagca	agatagccag	atcaatgtcg	660
atcgtggctg	gctcgaagat	acctgcaaga	atgtcattgc	gctgccattc	tccaaattgc	720
agttcgcgct	tagctggata	acgccacgga	atgatgtcgt	cgtgcacaac	aatggtgact	780
tctacagcgc	ggagaatctc	gctctctcca	ggggaagccg	aagtttccaa	aaggctcgtt	840
atcaaagctc	gccgcgttgt	ttcatcaagc	cttacggtca	ccgtaaccag	caaatacata	900
tcaactgtgtg	gcttcaggcc	gccatccaact	gcggagccgt	acaaatgtac	ggccagcaac	960
gtcggttcga	gatggcgctc	gatgacgcca	actacctctg	atagttgagt	cgatacttcg	1020
gcgatcaccg	cttccctcat	gatgtttaac	tttgttttag	ggcgactgcc	ctgctgcgta	1080
acatcggttc	tgctccataa	catcaaactc	cgaccacgga	cgtaacgcgc	ttgctgcttg	1140
gatgcccgag	gcatagactg	tacccccaaa	aaacagtcac	aacaagccat	gaaaaccgcc	1200
actgcgccgt	taccaccgct	gcgttcggtc	aaggttctgg	accagttgcg	tgagcgcata	1260
cgtacttgc	attacagctt	acgaaccgaa	caggcttatg	tccactgggt	tcgtgccttc	1320
atccgtttcc	acggtgtgcg	tcaccgcgga	accttgggca	gcagcgaagt	cgaggcattt	1380
ctgtcctggc	tggcgaacga	gcgcaagggt	tcggtctcca	cgcacgtca	ggcattggcg	1440
gccttgctgt	tcttctacgg	caagggtgctg	tgacgggata	tgccctggct	tcaggagatc	1500
ggaagacctc	ggccgtcgcg	gcgcttgccg	gtggtgctga	ccccggatga	agtggttcgc	1560
atcctcggtt	ttctggaagg	cgagcatcgt	ttgttcgccc	agcttctgta	tggaacgggc	1620
atgcggatca	gtgaggggtt	gcaactgcgg	gtcaaggata	tggaattcga	tcacggcacg	1680
atcatcgtgc	gggagggcaa	gggtccaag	gatcgggcct	tgatgttacc	cgagagcttg	1740
gcaccagcc	tgcgcgagca	gggaattaa	ttcccacggg	ttttgctgcc	cgcaaacggg	1800
ctgttctggt	gttgctagtt	tgttatcaga	atcgagata	cggcttcagc	cggtttgccg	1860
gctgaaagcg	ctatttcttc	cagaattgcc	atgatttttt	ccccacggga	ggcgtcactg	1920
gctcccgtgt	tgtcggcagc	tttgattcga	taagcagcat	cgcctgttcc	aggctgtcta	1980
tgtgtgactg	ttgagctgta	acaagttgtc	tcagggtgtc	aatttcatgt	tctagttgct	2040

ttgttttact ggtttcacct gttctattag gtgttacatg ctgttcatct gttacattgt	2100
cgatctgttc atggtgaaca gctttgaatg caccaaaaac tcgtaaaagc tctgatgtat	2160
ctatcttttt tacaccggtt tcatctgtgc atatggacag ttttcccttt gatatgtaac	2220
ggtgaacagt tgttctactt ttgtttgtta gtcttgatgc ttcactgata gatacaagag	2280
ccataagaac ctcatatcct tccgtattta gccagtatgt tctctagtgt ggttcggtgt	2340
ttttgcgtga gccatgagaa cgaaccattg agatcatact tactttgcat gtcactcaaa	2400
aattttgcct caaaactggt gagctgaatt tttgcagtta aagcatcgtg tagtgttttt	2460
cttagtccgt tatgtaggta ggaatctgat gtaatggttg ttggtatttt gtcaccattc	2520
atttttatct ggttgttctc aagttcggtt acgagatcca tttgtctatc tagttcaact	2580
tggaaaatca acgtatcagt cgggcggcct cgcttatcaa ccaccaattt catattgctg	2640
taagtgttta aatctttact tattggtttc aaaaccatt ggtaagcct tttaaactca	2700
tggtagtatt tttaagcat taacatgaac ttaaattcat caaggctaatt ctctatatatt	2760
gccttgtagag ttttcttttg tgtagttct ttttaataacc actcataaat cctcatagag	2820
tatttgtttt caaaagactt aacatgttcc agattatatt ttatgaattt ttttaactgg	2880
aaaagataag gcaatatctc ttcactaaaa actaattcta atttttcgct tgagaacttg	2940
gcatagtttg tccactggaa aatctcaaag cttttaacca aaggattcct gatttccaca	3000
gttctcgtca tcagctctct ggttgcttta gctaatacac cataagcatt ttccctactg	3060
atgttcatca tctgagcgta ttggtataaa gtgaacgata ccgtccgttc tttccttgta	3120
gggttttcaa tcgtggggtt gagtagtgc acacagcata aaattagctt ggtttcatgc	3180
tccgttaagt catagcgact aatcgctagt tcatttgctt tgaaaacaac taattcagac	3240
atacatctca attggtctag gtgattttaa tcactatacc aattgagatg ggctagtcaa	3300
tgataattac tagtcccttt cctttgagtt gtgggtatct gtaaattctg ctagaccttt	3360
gctggaaaac ttgtaaattc tgctagacct tctgtaaatt ccgctagacc tttgtgtgtt	3420
ttttttgttt atattcaagt ggttataatt tatagaataa agaaagaata aaaaaagata	3480
aaaagaatag atcccagccc tgtgtataac tcactacttt agtcagttcc gcagtattac	3540
aaaaggatgt cgcaaacgct gtttgctcct ctacaaaaca gaccttaaaa ccctaaaggc	3600
ttaagtagca cctcgcgaag ctcgggcaaa tcgctgaata ttccttttgt ctccgaccat	3660
caggcacctg agtcgctgtc tttttcgtga cattcagttc gctgcgctca cggtctctggc	3720
agtgaatggg ggtaaatggc actacaggcg ccttttatgg attcatgcaa ggaaactacc	3780
cataatacaa gaaaagcccg tcacgggctt ctacgggcgt tttatggcgg gtctgctatg	3840
tggtgctatc tgactttttg ctgttcagca gttcctgccc tctgattttc cagtctgacc	3900
acttcggatt atcccgtgac aggtcattca gactggctaa tgcaccagc aaggcagcgg	3960

tatcatcaac	aggcttacct	gtcttactgt	cggaattca	tttaaatagt	caaaagcctc	4020
cgaccggagg	cttttgactg	ctaggcgatc	tgtgctgttt	gccacggtat	gcagcaccag	4080
cgcgagatta	tgggctcgca	cgctcgactg	tcggacgggg	gcactggaac	gagaagtcag	4140
gcgagccgtc	acgcccttga	caatgccaca	tcctgagcaa	ataattcaac	cactaaacaa	4200
atcaaccgcg	tttcccgag	gtaaccaagc	ttgcgggaga	gaatgatgaa	caagagccaa	4260
caagttcaga	caatcacct	ggccgcgccc	cagcaaattg	cggcggcggt	ggaaaaaaaa	4320
gccactgaga	tcaacgtggc	ggtggtgttt	tccgtagtgt	accgcggagg	caacacgctg	4380
cttatccagc	ggatggacga	ggccttcgtc	tccagctgcg	atatttccct	gaataaagcc	4440
tggagcgcct	gcagcctgaa	gcaaggtagc	catgaaatta	cgtcagcggt	ccagccagga	4500
caatctctgt	acggtctgca	gctaaccaac	caacagcgaa	ttattatatt	tggcggcggc	4560
ctgccagtta	tttttaatga	gcaggtaatt	ggcgccgtcg	gcgttagcgg	cggtacggtc	4620
gagcaggatc	aattattagc	ccagtgcgcc	ctggattgtt	ttccgcatt	ataacctgaa	4680
gcgagaaggt	atattatgag	ctatcgtatg	ttccgccagg	cattctgagt	gttaacgagg	4740
ggaccgtcat	gtcgctttca	ccgccaggcg	tacgcctgtt	ttacgatccg	cgcgggcacc	4800
atgccggcgc	catcaatgag	ctgtgctggg	ggctggagga	gcagggggtc	ccctgccaga	4860
ccataacct	tgacggaggc	ggtgacgccg	ctgcgctggg	cgccctggcg	gccagaagct	4920
cgccctcgcg	ggtgggtatc	gggtcagcg	cgtccggcga	gatagccctc	actcatgccc	4980
agctgccggc	ggacgcgccg	ctggctaccg	gacacgtcac	cgatagcgac	gatcaactgc	5040
gtacgctcgg	cgccaacgcc	gggcagctgg	ttaaagtcc	gccgttaagt	gagagaaact	5100
gaatgtatcg	tatctatacc	cgcaccgggg	ataaaggcac	caccgccctg	tacggcggca	5160
gccgcacga	gaaagaccat	attcgcgtcg	aggcctacgg	caccgtcgat	gaactgatat	5220
cccagctggg	cgtctgctac	gccacgaccc	gcgacgccgg	gctgcgggaa	agcctgcacc	5280
atattcagca	gacgctgttc	gtgctggggg	ctgaactggc	cagcgatgcg	cggggcctga	5340
cccgctgag	ccagacgatc	ggcgaagagg	agatcacccg	cctggagcgg	cttatcgacc	5400
gcaatatggc	cgagagcggc	ccgttaaaac	agttcgtgat	cccggggagg	aatctcgct	5460
ctgccagct	gcacgtggcg	cgcacccagt	cccgtcggct	cgaacgcctg	ctgacggcca	5520
tggaccgcgc	gcacccgtg	cgcgacgcgc	tcaaacgcta	cagcaatcgc	ctgtcggtg	5580
ccctgttctc	catggcgcg	atcgaagaga	ctaggcctga	tgcttgcgct	tgaactggcc	5640
tagcaaacac	agaaaaagc	ccgcacctga	cagtgcgggc	tttttttttc	ctaggcgatc	5700
tgtgctgttt	gccacggtat	gcagcaccag	cgcgagatta	tgggctcgca	cgctcgactg	5760
tcggacgggg	gcactggaac	gagaagtcag	gcgagccgtc	acgcccttga	caatgccaca	5820

tcctgagcaa ataattcaac cactaaacaa atcaaccgcg tttcccgag gtaaccaagc	5880
ttcacctttt gagccgatga acaatgaaaa gatcaaaacg atttgagta ctggcccagc	5940
gccccgtcaa tcaggacggg ctgattggcg agtggcctga agaggggctg atcgccatgg	6000
acagcccctt tgacccggtc tcttcagtaa aagtggacaa cggctctgatc gtcgaactgg	6060
acggcaaacg ccgggaccag tttgacatga tcgaccgatt tatcgccgat tacgggatca	6120
acgttgagcg cacagagcag gcaatgcgcc tggaggcggt ggaaatagcc cgtatgctgg	6180
tggatattca cgtcagccgg gaggagatca ttgccatcac taccgccatc acgcccggca	6240
aagcggtcga ggtgatggcg cagatgaacg tggaggagat gatgatggcg ctgcagaaga	6300
tgcgtgcccg ccggaccccc tccaaccagt gccacgtcac caatctcaaa gataatccgg	6360
tgcagattgc cgctgacgcc gccgaggccg ggatccgcgg cttctcagaa caggagacca	6420
cggtcgggat cgcgcgctac gcgccgttta acgccctggc gctggtggtc ggttcgcagt	6480
gcggccgccc cggcgtggtg acgcagtgtc cgggtggaaga ggccaccgag ctggagctgg	6540
gcatgcgtgg cttaaccagc tacgccgaga cgggtgtcgg ctacggcacc gaagcgggat	6600
ttaccgacgg cgatgatacg ccgtggtcaa aggcgttcct cgcctcggcc tacgcctccc	6660
gcgggttgaa aatgcgctac acctccggca ccggatccga agcgtgatg ggctattcgg	6720
agagcaagtc gatgctctac ctgcaatcgc gctgcattct cattactaaa ggcgccgggg	6780
ttcagggact gcaaaacggc gcggtgagct gtatcgcat gaccggcgct gtgccgtcgg	6840
gcattcgggc ggtgctggcg gaaaacctga tcgcctctat gctcgacctc gaagtggcgt	6900
ccgccaacga ccagaacttc tcccactcgg atattcgccg caccgcgcgc accctgatgc	6960
agatgctgcc ggccaccgac tttattttct ccggctacag cgcggtgccg aactacgaca	7020
acatgttcgc cggtcgaac ttcgatgcgg aagattttga tgattacaac atcctgcagc	7080
gtgacctgat ggttgacggc ggctgcgtc cggtgaccga ggcggaacc attgccattc	7140
gccagaaagc ggcgcggggc atccaggcgg tttccgcga gctggggctg ccgccaatcg	7200
ccgacgagga ggtggaggcc gccacctacg cgcacggcag caacgagatg ccgccgcgt	7260
acgtggtgga ggatctgagt gcggtggaag agatgatgaa gcgcaacatc accggcctcg	7320
atattgtcgg cgcgctgagc cgcagcggct ttgaggatat cgccagcaat attctcaata	7380
tgctgcgcca gcgggtcacc ggcgattacc tgcagacctc ggccattctc gatcggcagt	7440
tcgaggtggt gagtgcggtc aacgacatca atgactatca ggggccgggc accggctatc	7500
gcatctctgc cgaacgctgg gcggagatca aaaatattcc gggcgtggtt cagcccagca	7560
ccattgaata aggcgggtatt cctgtgcaac agacaaccca aattcagccc tcttttacc	7620
tgaaaacccg cgagggcggg gtagcttctg ccgatgaacg cgccgatgaa gtggtgatcg	7680
gcgtcggccc tgccttcgat aaacaccagc atcacactct gatcgatatg ccccatggcg	7740

cgatcctcaa	agagctgatt	gccgggggtgg	aagaagaggg	gcttcacgcc	cggttggtgc	7800
gcattctgcg	cacgtccgac	gtctccttta	tggcctggga	tgcggccaac	ctgagcggct	7860
cggggatcgg	catcgggtatc	cagtcgaagg	ggaccacggt	catccatcag	cgcgatctgc	7920
tgccgctcag	caacctggag	ctgttctccc	aggcgccgct	gctgacgctg	gagacctacc	7980
ggcagattgg	caaaaacgct	gcgcgctatg	cgcgcaaaga	gtcaccttcg	ccggtgccgg	8040
tggtgaacga	tcagatgggt	cgcccgaaat	ttatggccaa	agccgcgcta	tttcatatca	8100
aagagaccaa	acatgtgggt	caggacgccg	agcccgtcac	cctgcacatc	gacttagtaa	8160
gggagtgacc	atgagcgaga	aaaccatgcg	cgtgcaggat	tatccgttag	ccaccgctg	8220
cccgagcat	atcctgacgc	ctaccggcaa	accattgacc	gatattaccc	tcgagaaggt	8280
gctctctggc	gaggtgggcc	cgcaggatgt	gcggatctcc	cgccagaccc	ttgagtacca	8340
ggcgagatt	gccgagcaga	tgcagcgcca	tgcggtggcg	cgcaatttcc	gccgcggggc	8400
ggagcttata	gccattcctg	acgagcgcat	tctggctatc	tataacgcgc	tgcgcccgtt	8460
ccgctcctcg	caggcggagc	tgctggcgat	cgcgcagcag	ctggagcaca	cctggcatgc	8520
gacagtgaat	gccgcctttg	tcogggagtc	ggcggaagtg	tatcagcagc	ggcataagct	8580
gcgtaaagga	agctaagcgg	aggtcagcat	gccgttaata	gccgggattg	atatcggcaa	8640
cgccaccacc	gaggtggcgc	tggcgtccga	ctaccgcgag	gcgagggcgt	ttgttgccag	8700
cgggatcgtc	gcgacgacgg	gcatgaaagg	gacgcgggac	aatatcgccg	ggaccctcgc	8760
cgcgctggag	caggccctgg	cgaaaacacc	gtggtcgatg	agcgatgtct	ctcgcatcta	8820
tcttaacgaa	gccgcgccgg	tgattggcga	tgtggcgatg	gagaccatca	ccgagaccat	8880
tatcaccgaa	tcgaccatga	tcggtcataa	cccgcagacg	ccgggcgggg	tgggcgttgg	8940
cgtggggacg	actatcgccc	tcgggcggct	ggcgacgctg	ccggcggcgc	agtatgccga	9000
ggggtggatc	gtactgattg	acgacgccgt	cgatttcctt	gacgccgtgt	ggtggctcaa	9060
tgaggcgctc	gaccggggga	tcaacgtggt	ggcggcgac	ctcaaaaagg	acgacggcgt	9120
gctggtgaac	aaccgcctgc	gtaaaacct	gccggtgggt	gatgaagtga	cgtgctgga	9180
gcaggtcccc	gagggggtaa	tggcggcggt	ggaagtggcc	gcgcggggcc	aggtggtgcg	9240
gatcctgtcg	aatccctacg	ggatcgccac	cttcttcggg	ctaagcccgg	aagagaccca	9300
ggccatcgtc	cccatcgccc	gcgcctgat	tggcaaccgt	tcgcggtgg	tgtcaagac	9360
ccgcagggg	gatgtgcagt	cgcgggtgat	ccggcgggc	aacctctaca	ttagcggcga	9420
aaagcgccgc	ggagaggccg	atgtcgccga	ggcgcgga	gccatcatgc	aggcgatgag	9480
cgcctgcgct	ccggtacgcg	acatccgcgg	cgaaccgggc	acccacgccg	gcggcatgct	9540
tgagcgggtg	cgcaaggtaa	tggcgtccct	gaccggccat	gagatgagcg	cgatatacat	9600

ccaggatctg ctggcgggtg atacgtttat tccgcgcaag gtgcagggcg ggatggccgg	9660
cgagtgcgcc atggagaatg ccgtcgggat ggcggcgatg gtgaaagcgg atcgtctgca	9720
aatgcaggtt atcgcccgcg aactgagcgc ccgactgcag accgaggtgg tggtagggcg	9780
cgtggaggcc aacatggcca tcgccccggc gttaccact cccggctgtg cggcgccgct	9840
ggcgatcctc gacctggcg ccggctcgac ggatgccccg atcgtcaacg cggaggggca	9900
gataacggcg gtccatctcg cccggggcgg gaatatggtc agcctgttga ttaaaaccga	9960
gctgggcctc gaggatcttt cgctggcgga agcgataaaa aaatacccg cggccaaagt	10020
ggaaagcctg ttcagtattc gtcacgagaa tggcgcggtg gaggcttttc gggaagccct	10080
cagccccggc gtgttcgcca aagtgggtg catcaaggag ggcgaactgg tgccgatcga	10140
taacgccagc ccgctggaaa aaattcgtct cgtgcgccgg caggcgaaa agaaagtgtt	10200
tgtcaccaac tgctgcgcg cgctgcgcca ggtctcacc cgcgggtcca ttcgcgatat	10260
cgcctttgtg gtgctgggtg gcggctcatc gctggacttt gagateccgc agcttatcac	10320
ggaagccttg tcgcactatg gcgtgggtgc cgggcagggc aatattcggg gaacagaagg	10380
gccgcgcaat gcggctcgcca ccgggctgct actggccggg caggcgcaatt aaacgggcgc	10440
tcgcgccagc ctctaggtac aaataaaaaa ggcacgtcag atgacgtgcc ttttttcttg	10500
tctagcgtgc accaatgctt ctggcgctcag gcagccatcg gaagctgtgg tatggctgtg	10560
caggtcgtaa atcactgcat aattcgtgtc gctcaaggcg cactcccgtt ctggataatg	10620
ttttttgcgc cgacatcata acggttctgg caaatattct gaaatgagct gttgacaatt	10680
aatcatccgg ctctgtataat gtgtggaatt gtgagcggat aacaatttca cacaggaaac	10740
agaccatgac tagtaaggag gacaattcca tggctgctgc tgctgataga ttaaaactaa	10800
cttcggcca cttgaatgct ggtagaaaga gaagttcctc ttctgtttct ttgaaggctg	10860
ccgaaaagcc tttcaagggt actgtgattg gatctggtaa ctggggtact actattgcca	10920
aggtggttgc cgaaaattgt aaggatacc cagaagtttt cgctccaata gtacaaatgt	10980
gggtgttcga agaagagatc aatggtgaaa aattgactga aatcataaat actagacatc	11040
aaaacgtgaa atacttgctt ggcatcactc tacccgacaa tttggttgct aatccagact	11100
tgattgattc agtcaaggat gtcgacatca tcgttttcaa cattccacat caatttttgc	11160
ccgatatctg tagccaattg aaaggctcatg ttgattcaca cgtcagagct atctcctgtc	11220
taaagggttt tgaagttggg gctaaagggt tccaattgct atcctcttac atcactgagg	11280
aactaggtat tcaatgtggg gctctatctg gtgctaacat tgccaccgaa gtcgctcaag	11340
aacactggtc tgaaacaaca gttgcttacc acattccaaa ggatttcaga ggcgagggca	11400
aggacgtcga ccataagggt ctaaaggcct tgttcacag accttacttc cacgttagtg	11460
tcacgaaga tgttgctggg atctccatct gtggtgcttt gaagaacgtt gttgccttag	11520

gttggtggtt cgtcgaaggt ctaggctggg gtaacaacgc ttctgctgcc atccaaagag	11580
tcggtttggg tgagatcatc agattcggtc aaatgttttt cccagaatct agagaagaaa	11640
catactacca agagtctgct ggtgttgctg atttgatcac cacctgcgct ggtggtagaa	11700
acgtcaaggt tgctaggcta atggctactt ctggttaagga cgcctgggaa tgtgaaaagg	11760
agttgttgaa tggccaatcc gctcaaggtt taattacctg caaagaagtt cacgaatggt	11820
tggaacatg tggctctgtc gaagacttcc cattatttga agccgtatac caaatcgttt	11880
acaacaacta cccaatgaag aacctgccgg acatgattga agaattagat ctacatgaag	11940
attagattta ttggatccag gaaacagact agaattatgg gattgactac taaacctcta	12000
tctttgaaag ttaacgccgc tttgttcgac gtcgacggtta ccattatcat ctctcaacca	12060
gccattgctg cattctggag ggatttcggt aaggacaaac cttatttcga tgctgaacac	12120
gttatccaag tctcgcattg ttggagaacg tttgatgcca ttgctaagtt cgctccagac	12180
tttgccaatg aagagtatgt taacaaatta gaagctgaaa ttccggtcaa gtacggtgaa	12240
aaatccattg aagtcaccag tgcatgtaag ctgtgcaacg ctttgaaacgc tctaccaaaa	12300
gagaaatggg ctgtggcaac ttccggtacc cgtgatatgg cacaaaaatg gttcgagcat	12360
ctgggaatca ggagaccaa gtacttcatt accgctaata atgtcaaaca gggtaagcct	12420
catccagaac catatctgaa gggcaggaat ggcttaggat atccgatcaa tgagcaagac	12480
ccttccaaat ctaaggtagt agtatttgaa gacgctccag caggtattgc cgccggaaaa	12540
gccgccggtt gtaagatcat tggattgcc actactttcg acttggactt cctaaaggaa	12600
aaaggctgtg acatcattgt caaaaaccac gaatccatca gagttggcgg ctacaatgcc	12660
gaaacagacg aagttgaatt catttttgac gactacttat atgctaagga cgatctgttg	12720
aaatggtaac ccgggctgca ggcatgcaag cttggctggt ttggcggatg agagaagatt	12780
ttcagcctga tacagattaa atcagaacgc agaagcggtc tgataaaaca gaatttgccct	12840
ggcggcagta gcgcggtggt cccacctgac cccatgccga actcagaagt gaaacgccgt	12900
agcgccgatg gtagtggtgg gtctcccat gcgagagtag ggaactgcca ggcacaaat	12960
aaaacgaaag gctcagtcga aagactgggc ctttcgtttt atctgttggt tgctcgtgaa	13020
cgctctcctg agtaggacaa atccgccggg agcggatttg aacgttgcca agcaacggcc	13080
cggagggtgg cgggcaggac gcccgccata aactgccagg catcaaatta agcagaaggc	13140
catcctgacg gatggccttt ttgcgtttct acaaactcca gctggatcgg gcgctagagt	13200
atacatttaa atggtaccct ctagtcaagg ccttaagtga gtcgtattac ggactggccg	13260
tcgttttaca acgtcgtgac tgggaaaacc ctggcgttac ccaacttaat cgccttgacg	13320
cacatcccc tttcgccagc tggcgtaata gcgaagaggc ccgcaccgat cgcccttccc	13380

ES 2 552 317 T3

aacagttgcg cagcctgaat ggcgaatggc gcctgatgcg gtattttctc cttacgcac 13440
 tgtgcggtat ttcacaccgc atatggtgca ctctcagtac aatctgctct gatgccgcat 13500
 agttaagcca gcccgcacac ccgccaacac ccgctgacga gct 13543

<210> 54
 <211> 13402
 <212> DNA
 <213> Secuencia Artificial

5

<220>
 <223> Plásmido

<400> 54
 tagtaaagcc ctgctagat tttaatgcmg atggtgcmg tacttcgcca actattgcmg 60
 taacaagaaa aagccagcct ttcattgatat atctcccaat ttgtgtaggg cttattatgc 120
 acgcttaaaa ataataaaaag cagacttgac ctgatagttt ggctgtgagc aattatgtgc 180
 ttagtgcatc taacgcttga gttaagccgc gccgcgaagc ggcgtcggt tgaacgaatt 240
 gttagacatt atttgccgac taccttggtg atctcgccct tcacgtagtg gacaaattct 300
 tccaactgat ctgcmgcmg gcccaagcmg tcttcttctt gtccaagata agcctgtcta 360
 gcttcaagta tgacgggctg atactgggcc ggcaggcgct ccattgcccc gtcggcagcg 420
 acatccttcg gcmgcmgttt gccggttact gcmgtgtacc aaatgcmgga caacgtaagc 480
 actacatttc gctcatcgcc agcccagtcg ggcggcmgagt tccatagcmg taaggtttca 540
 tttagcmgct caaatagatc ctgttcagga accggatcaa agagtctctc cgcgcmgtga 600
 cctaccaagg caacgctatg ttctcttgct ttgttcagca agatagccag atcaatgtcg 660
 atcgtggctg gctcgaagat acctgcaaga atgtcattgc gctgccattc tccaaattgc 720
 agttcgcgct tagctggata acgcccagga atgatgtcgt cgtgcacaac aatggtgact 780
 tctacagcmg ggagaatctc gctctctcca ggggaagcmg aagtttccaa aaggtcgttg 840
 atcaaagctc gccgcmgtgt ttcattcaagc cttacggtca ccgtaaccag caaatcaata 900
 tcaactgttg gcttcaggcc gccatccact gcggagcmgt acaaatgtac ggccagcaac 960
 gtcggttcga gatggcmgtc gatgacgcca actacctctg atagttgagt cgatacttcg 1020
 gcgatcaccg cttccctcat gatgtttaac ttgttttag ggcgactgcc ctgctgcgta 1080
 acatcgttgc tgctccataa catcaaacat cgaccaccmg cgtaacgcmg ttgctgcttg 1140
 gatgcccagag gcatagactg taccctaaaa aaacagtcatt aacaagccat gaaaaccgcm 1200
 actgcmgcmgt taccaccgct gcgttcggtc aaggttctgg accagttgcm tgagcmgata 1260
 cgctacttgc attacagctt acgaaccgaa caggcttatg tccactgggt tcgtgccttc 1320
 atccgtttcc acggtgtgcm taccctggga accttgggca gcagcmgaagt cgaggcattt 1380
 ctgtcctggc tggcgaacga gcmgaaggtt tcggtctcca cgcctcgtca ggcattggcm 1440

gccttgctgt tcttctacgg caaggtgctg tgcacggatc tgccctggct tcaggagatc	1500
ggaagacctc ggccgtcgcg gcgcttgccg gtggtgctga ccccgatga agtggttcgc	1560
atcctcggtt ttctggaagg cgagcatcgt ttgttcgccc agcttctgta tggaacgggc	1620
atgcggatca gtgaggggtt gcaactgcgg gtcaaggatc tggatttcga tcacggcacg	1680
atcatcgtgc gggagggcaa gggctccaag gatcgggcct tgatgttacc cgagagcttg	1740
gcaccagcc tgcgcgagca ggggaattaa ttcccacggg ttttgctgcc cgcaaacggg	1800
ctgttctggt gttgctagtt tgttatcaga atcgagatc cggcttcagc cggtttgccg	1860
gctgaaagcg ctatttcttc cagaattgcc atgatttttt cccacggga ggcgtcactg	1920
gctcccggtg tgcggcagc tttgattcga taagcagcat cgctgtttc aggtgtcta	1980
tgtgtgactg ttgagctgta acaagttgtc tcagggtgtc aatttcattg tctagttgct	2040
ttgttttact ggtttcacct gttctattag gtgttacatg ctgttcattc gttacattgt	2100
cgatctgttc atggtgaaca gctttgaatg caccaaaaac togtaaaagc tctgatgtat	2160
ctatcttttt tacaccgttt tcatctgtgc atatggacag ttttcccttt gatatgtaac	2220
ggtgaacagt tgttctactt ttgtttgtta gtcttgatgc ttcactgata gatacaagag	2280
ccataagaac ctcagatcct tccgtattta gccagtatgt tctctagtgt ggttcggtgt	2340
ttttgcgtga gccatgagaa cgaaccattg agatcatact tactttgcat gtcactcaaa	2400
aattttgcct caaaactggt gagctgaatt tttgcagtta aagcatcgtg tagtgttttt	2460
cttagtccgt tatgtaggta ggaatctgat gtaatggttg ttggtatttt gtcaccattc	2520
atttttatct ggttggttctc aagttcgggt acgagatcca tttgtctatc tagttcaact	2580
tggaaaatca acgtatcagt cgggcggcct cgcttatcaa ccaccaattt catattgctg	2640
taagtgttta aatctttact tattggtttc aaaaccatt ggtaagcct tttaaactca	2700
tggtagttaa tttcaagcat taacatgaac tttaattcat caaggctaatt ctctatattt	2760
gccttgtagg ttttcttttg tgttagttct ttaataaacc actcataaat cctcatagag	2820
tatttgtttt caaaagactt aacatgttcc agattatatt ttatgaattt ttttaactgg	2880
aaaagataag gcaatatctc ttcactaaaa actaattcta atttttcgct tgagaacttg	2940
gcatagtttg tccactggaa aatctcaaag cctttaacca aaggattcct gatttcocac	3000
gttctcgtca tcagctctct ggttgcttta gctaatacac cataagcatt ttccctactg	3060
atgttcatca tctgagcgta ttggttataa gtgaacgata ccgtccgttc tttccttgta	3120
gggttttcaa tcgtgggggt gagtagtgcc acacagcata aaattagctt ggtttcatgc	3180
tccgttaagt catagcgact aatcgctagt tcatttgctt tgaaaacaac taattcagac	3240
atacatctca attggtctag gtgattttta tcactataacc aattgagatg ggctagtcaa	3300

tgataattac tagtcctttt cctttgagtt gtgggtatct gttaaattctg ctagaccttt	3360
gctggaaaac ttgtaaattc tgctagaccc tctgtaaatt ccgctagacc tttgtgtgtt	3420
ttttttgttt atattcaagt ggttataatt tatagaataa agaaagaata aaaaaagata	3480
aaaagaatag atcccagccc tgtgtataac tcactacttt agtcagttcc gcagtattac	3540
aaaaggatgt cgcaaacgct gtttgctcct ctacaaaaca gaccttaaaa ccctaaaggc	3600
ttaagtagca cctcgcgaag ctccgggcaaa tcgctgaata ttccttttgt ctccgaccat	3660
caggcacctg agtcgctgtc tttttcgtga cattcagttc gctgogctca cggctctggc	3720
agtgaatggg ggtaaatggc actacaggcg ccttttatgg attcatgcaa ggaaactacc	3780
cataatacaa gaaaagcccc tcacgggctt ctacgggctt tttatggcgg gtctgctatg	3840
tggtgctatc tgactttttg ctgttcagca gttcctgccc tctgattttc cagtctgacc	3900
acttcggatt atcccgtgac aggtcattca gactggctaa tgcacccagt aaggcagcgg	3960
tatcatcaac aggcttacc gtcttactgt cgggaattca tttaaatagt caaaagcctc	4020
cgaccggagg cttttgactg ctaggcgatc tgtgctgttt gccacggtat gcagcaccag	4080
cgcgagatta tgggctcgca cgctcgactg tcggacgggg gcaactggaac gagaagtcag	4140
gcgagccgtc acgcccttga caatgccaca tcctgagcaa ataattcaac cactaaacaa	4200
atcaaccgcg tttcccgag gtaaccaagc ttgcgggaga gaatgatgaa caagagccaa	4260
caagttcaga caatcacctt gcccgccgcc cagcaaattg cggcggcggt ggaaaaaaaa	4320
gccactgaga tcaacgtggc ggtggtgttt tccgtagttg acccgggagg caacacgtg	4380
cttatccagc ggatggacga ggccttcgtc tccagctgcg atatttcctt gaataaagcc	4440
tggagcgctt gcagcctgaa gcaaggatcc catgaaatta cgtcagcggg ccagccagga	4500
caatctctgt acggtctgca gctaaccaac caacagcgaa ttattatttt tggcggcggc	4560
ctgccagtta tttttaatga gcaggtaatt ggcgccgtcg gcgttagcgg cggtagcgtc	4620
gagcaggatc aattattagc ccagtgcgcc ctggattgtt tttccgcatt ataacctgaa	4680
gcgagaaggt atattatgag ctatcgtatg ttccgccagg cattctgagt gttaacgagg	4740
ggaccgtcat gtgcgtttca ccgccaggcg tacgcctgtt ttacgatccg cgcgggcacc	4800
atgccggcgc catcaatgag ctgtgctggg ggctggagga gcagggggc ccctgccaga	4860
ccataacctt tgacggaggc ggtgacgccg ctgcgctggg cgccctggcg gccagaagct	4920
cgcctctgcg ggtgggtatc gggctcagcg cgtccggcga gatagccctc actcatgccc	4980
agctgccggc ggacgcgccg ctggctaccg gacacgtcac cgatagcgac gatcaactgc	5040
gtacgctcgg cgccaacgcc gggcagctgg ttaaagtcct gccgttaagt gagagaaact	5100
gaatgtatcg tatctatacc cgcacccggg ataaaggcac caccgcctg tacggcggca	5160
gccgcatcga gaaagaccat attcgcgtcg aggcctacgg caccgtcgat gaactgatat	5220

cccagctggg cgtctgctac gccacgaccc gcgacgccgg gctgcgggaa agcctgcacc 5280
 atattcagca gacgctgttc gtgctggggg ctgaactggc cagcgatgcg cggggcctga 5340
 cccgcctgag ccagacgata ggcgaagagg agatcacccg cctggagcgg cttatcgacc 5400
 gcaatatggc cgagagcggc ccgttaaaac agttcgtgat cccggggagg aatctcgcct 5460
 ctgcccagct gcaccctgat gcttgcgctt gaactggcct agcaaacaca gaaaaagcc 5520
 cgcacctgac agtgccgggt ttttttttcc taggcgatct gtgctgtttg ccacggtatg 5580
 cagcaccagc gcgagattat gggctcgcac gctcgactgt cggacggggg cactggaacg 5640
 agaagtcagg cgagccgtca cgcccttgac aatgccacat cctgagcaaa taattcaacc 5700
 actaaacaaa tcaaccgcgt ttcccggagg taaccaagct tcaccttttg agccgatgaa 5760
 caatgaaaag atcaaaacga tttgcagtac tggcccagcg ccccgctcaat caggacgggc 5820
 tgattggcga gtggcctgaa gaggggctga tcgccatgga cagccccttt gaccgggtct 5880
 cttcagtaaa agtggacaac ggtctgatcg tcgaactgga cggcaaacgc cgggaccagt 5940
 ttgacatgat cgaccgattt atcgccgatt acgcgatcaa cgttgagcgc acagagcagg 6000
 caatgcgcct ggagggcgtg gaaatagccc gtatgctggt ggatattcac gtcagccggg 6060
 aggagatcat tgccatcact accgccatca cgccggccaa agcggtcgag gtgatggcgc 6120
 agatgaacgt ggtggagatg atgatggcgc tgcagaagat gcgtgcccgc cggaccccct 6180
 ccaaccagtg ccacgtcacc aatctcaaag ataatccggt gcagattgcc gctgacgccg 6240
 ccgaggccgg gatccgcggc ttctcagaac aggagaccac ggtcggtatc gcgcgctacg 6300
 cgccgtttaa cgccctggcg ctgttggtcg gttcgcagtg cggccgcccc ggcggttga 6360
 cgcagtgtc ggtggaagag gccaccgagc tggagctggg catgcgtggc ttaaccagct 6420
 acgccgagac ggtgtcggtc tacggcaccg aagcgggtatt taccgacggc gatgatacgc 6480
 cgtggtcaaa ggcgttcctc gcctcggcct acgcctcccg cgggttgaaa atgcgctaca 6540
 cctccggcac cgatccgaa gcgctgatgg gctattcgga gagcaagtcg atgctctacc 6600
 tcgaatcgcg ctgcatcttc attactaaag gcgccgggggt tcagggactg caaaacggcg 6660
 cggtagactg tatcgcatg accggcgctg tgccgtcggg cattcgggcg gtgctggcgg 6720
 aaaacctgat cgctctatg ctcgacctcg aagtggcgct cgccaacgac cagactttct 6780
 cccactcgga tattcgccgc accgcgcgca ccctgatgca gatgctgccg ggcaccgact 6840
 ttattttctc cggctacagc gcggtgccga actacgacaa catgttcgcc ggctcgaact 6900
 tcgatgcgga agattttgat gattacaaca tcctgcagcg tgacctgatg gttgacggcg 6960
 gcctgcgtcc ggtgaccgag gcggaaacca ttgccattcg ccagaaagcg gcgcgggcga 7020
 tccaggcggt tttccgcgag ctggggctgc cgccaatcgc cgacgaggag gtggaggccg 7080

ccacctacgc gcacggcagc aacgagatgc cgccgcgtaa cgtggtggag gatctgagtg	7140
cggtggaaga gatgatgaag cgcaacatca ccggcctcga tattgtcggc gcgctgagcc	7200
gcagcggcctt tgaggatata gccagcaata ttctcaatat gctgcgccag cgggtcaccg	7260
gcgattacct gcagacctcg gccattctcg atcggcagtt cgagggtggtg agtgcgggtca	7320
acgacatcaa tgactatcag gggccgggca ccggctatcg catctctgcc gaacgctggg	7380
cggagatcaa aaatattccg ggcgtggttc agcccgacac cattgaataa ggcgggtattc	7440
ctgtgcaaca gacaacccaa attcagccct cttttaccct gaaaaccgc gagggcgggg	7500
tagcttctgc cgatgaacgc gccgatgaag tggatgatcg cgtcggccct gccttcgata	7560
aacaccagca tcacactctg atcgatatgc cccatggcgc gatcctcaaa gagctgattg	7620
ccggggtgga agaagagggg cttcacgccc ggggtggtcg cattctgcgc acgtccgacg	7680
tctcctttat ggcctgggat gcggccaacc tgagcggctc ggggatcggc atcggtatcc	7740
agtcgaaggg gaccacggtc atccatcagc gcgatctgct gccgctcagc aacctggagc	7800
tgttctccca ggcgcgctg ctgacgctgg agacctaccg gcagattggc aaaaacgctg	7860
cgcgctatgc gcgcaaagag tcaccttcgc cggtgccggt ggtgaacgat cagatggtgc	7920
ggccgaaatt tatggccaaa gccgcgctat ttcatatcaa agagaccaa catgtggtgc	7980
aggacgccga gcccgtcacc ctgcacatcg acttagtaag ggagtgacca tgagcgagaa	8040
aaccatgcgc gtgcaggatt atccgttagc caccgctgc ccggagcata tcctgacgcc	8100
taccggcaaa ccattgaccg atattaccct cgagaagggtg ctctctggcg aggtgggcc	8160
gcaggatgtg cggatctccc gccagaccct tgagtaccag gcgcagattg ccgagcagat	8220
gcagcgccat gcggtggcgc gcaatttcgc ccgcgcggcg gagcttatcg ccattcctga	8280
cgagcgcatc ctggctatct ataacgcgct gcgccggttc cgctcctcgc aggcggagct	8340
gctggcgatc gccgacgagc tggagcacac ctggcatgcg acagtgaatg ccgcctttgt	8400
ccgggagtcg gcggaagtgt atcagcagcg gcataagctg cgtaaaggaa gctaagcgga	8460
ggtcagcatg ccgttaatag ccgggattga tatcggaac gccaccaccg aggtggcgct	8520
ggcgtccgac taccgcgagc cgagggcggtt tgttgccagc gggatcgtcg cgacgacggg	8580
catgaaaggg acgcgggaca atatcgccgg gaccctcgcc gcgctggagc aggccctggc	8640
gaaaacaccg tggtcgatga gcgatgtctc tcgcatctat cttaacgaag ccgcgcgggt	8700
gattggcgat gtggcgatgg agaccatcac cgagaccatt atcaccgaat cgaccatgat	8760
cggtcataac ccgcagacgc cgggcgggggt gggcggttggc gtggggacga ctatcgccct	8820
cgggcggctg gcgacgctgc cggcggcgca gtatgccgag ggggtggatcg tactgattga	8880
cgacgccgtc gatttccttg acgccgtgtg gtggctcaat gaggcgctcg accgggggat	8940
caacgtggtg gcggcgatcc tcaaaaagga cgacggcggtg ctggtgaaca accgcctgcg	9000

taaaaccctg	ccggtggtgg	atgaagtgac	gctgctggag	caggtccccg	agggggtaat	9060
ggcggcggtg	gaagtggccg	cgccgggcca	ggtggtgcgg	atcctgtcga	atccctacgg	9120
gatcgccacc	ttcttcgggc	taagcccga	agagaccag	gccatcgtec	ccatcgcccc	9180
cggcctgatt	ggcaaccgtt	ccgcggtggt	gctcaagacc	ccgcaggggg	atgtgcagtc	9240
gcgggtgatc	ccggcgggca	acctctacat	tagcggcgaa	aagcgccgcg	gagaggccga	9300
tgtcgccgag	ggcgcggaag	ccatcatgca	ggcgatgagc	gcctgcgctc	cggtacgcga	9360
catccgcggc	gaaccgggca	cccacgccgg	cggcatgctt	gagcgggtgc	gcaaggtaat	9420
ggcgtccctg	accggccatg	agatgagcgc	gatatacatc	caggatctgc	tggcgggtga	9480
tacgtttatt	ccgcgcaagg	tgcagggcgg	gatggccggc	gagtgcgcca	tggagaatgc	9540
cgtcgggatg	gcggcgatgg	tgaaagcggg	tcgtctgcaa	atgcaggtta	tcgcccgcga	9600
actgagcgcc	cgactgcaga	ccgaggtggt	ggtgggcggc	gtggaggcca	acatggccat	9660
cgcggggcgg	ttaaccactc	ccggctgtgc	ggcgccgctg	gcgatccctg	acctcggcgc	9720
cggctcgacg	gatcgggcga	tcgtcaacgc	ggaggggcag	ataacggcgg	tccatctcgc	9780
cggggcgggg	aatatggtca	gcctgttgat	taaaaccgag	ctgggcctcg	aggatctttc	9840
gctggcgga	gcgataaaaa	aatacccgt	ggccaaagt	gaaagcctgt	tcagtattcg	9900
tcacgagaat	ggcgcggtgg	agttctttcg	ggaagccctc	agcccggcgg	tgttcgccaa	9960
agtggtgtac	atcaaggagg	gcgaactggt	gccgatcgat	aacgccagcc	cgctggaaaa	10020
aattcgtctc	gtgcgccggc	aggcgaaga	gaaagtgttt	gtcaccaact	gcctgcgcgc	10080
gctgcgccag	gtctcaccgc	gcggttccat	tcgcgatatc	gcctttgtgg	tgctggtggg	10140
cggctcatcg	ctggactttg	agatcccga	gcttatcacg	gaagccttgt	cgcactatgg	10200
cgtggtcgcc	gggcagggca	atattcgggg	aacagaaggg	ccgcgcaatg	cggtcgccac	10260
cgggctgcta	ctggccggtc	aggcgaatta	aacgggcgct	cgcgccagcc	tctaggtaca	10320
aataaaaaag	gcacgtcaga	tgacgtgcct	tttttcttgt	ctagcgtgca	ccaatgcttc	10380
tggcgctcag	cagccatcgg	aagctgtggt	atggctgtgc	aggctgtaaa	tcactgcata	10440
attcgtgtcg	ctcaaggcgc	actcccgttc	tggataatgt	tttttgcgcc	gacatcataa	10500
cggttctggc	aaatattctg	aatgagctg	ttgacaatta	atcatccggc	tcgtataatg	10560
tgtggaattg	tgagcggata	acaatttcac	acaggaaaca	gaccatgact	agtaaggagg	10620
acaattccat	ggctgctgct	gctgatagat	taaacttaac	ttccggccac	ttgaatgctg	10680
gtagaaagag	aagttcctct	tctgtttctt	tgaaggctgc	cgaaaagcct	ttcaaggtta	10740
ctgtgattgg	atctggtaac	tggggtacta	ctattgccaa	ggtggttgcc	gaaaattgta	10800
agggataccc	agaagttttc	gctccaatag	tacaaatgtg	ggtgttcgaa	gaagagatca	10860

atggtgaaaa attgactgaa atcataaata ctagacatca aaacgtgaaa tacttgacctg	10920
gcatcactct acccgacaat ttggttgcta atccagactt gattgattca gtcaaggatg	10980
tcgacatcat cgttttcaac attccacatc aatttttgcc ccgtatctgt agccaattga	11040
aaggtcatgt tgattcacac gtcagagcta tctcctgtct aaagggtttt gaagttggtg	11100
ctaaagggtgt ccaattgcta tcctcttaca tcaactgagga actaggtatt caatgtggtg	11160
ctctatctgg tgctaacatt gccaccgaag tcgctcaaga acaactggtct gaaacaacag	11220
ttgcttacca cattccaaag gatttcagag gcgagggcaa ggacgtcgac cataagggtc	11280
taaaggcctt gttccacaga ccttacttcc acgttagtgt catcgaagat gttgctggta	11340
tctccatctg tgggtgcttg aagaacgttg ttgccttagg ttgtggtttc gtcgaaggtc	11400
taggtgggg taacaacgct tctgctgcc tccaaagagt cggtttgggt gagatcatca	11460
gattcgggtca aatgtttttc ccagaatcta gagaagaaac atactaccaa gagtctgctg	11520
gtgttgctga tttgatcacc acctgcgctg gtggtagaaa cgtcaagggt gctaggctaa	11580
tggctacttc tggtaaggac gcctgggaat gtgaaaagga gttgttgaat ggccaatccg	11640
ctcaagggtt aattacctgc aaagaagttc acgaatggtt ggaaacatgt ggctctgtcg	11700
aagacttccc attatttgaa gccgtatacc aaatcgttta caacaactac ccaatgaaga	11760
acctgccgga catgattgaa gaattagatc tacatgaaga ttagatttat tggatccagg	11820
aaacagacta gaattatggg attgactact aaacctctat ctttgaaagt taacgccgct	11880
ttgttcgacg tcgacggtac cattatcatc tctcaaccag ccattgctgc attctggagg	11940
gatttcggtg aggacaaacc ttatttcgat gctgaacacg ttatccaagt ctgcgatggt	12000
tggagaacgt ttgatgccat tgctaagttc gctccagact ttgccaatga agagtatgtt	12060
aacaaattag aagctgaaat tccggtcaag tacggtgaaa aatccattga agtcccagg	12120
gcagttaagc tgtgcaacgc tttgaacgct ctacccaaaag agaaatgggc tgtggcaact	12180
tccggtaccc gtgatatggc acaaaaatgg ttcgagcatc tgggaatcag gagaccaaag	12240
tacttcatta ccgctaata tgtcaaacag ggtaagcctc atccagaacc atatctgaag	12300
ggcaggaatg gcttaggata tccgatcaat gagcaagacc cttccaaatc taaggtagta	12360
gtatttgaag acgctccagc aggtattgcc gccggaaaag ccgccggttg taagatcatt	12420
ggtattgcca ctactttcga cttggacttc cttaaaggaaa aaggctgtga catcattgtc	12480
aaaaaccacg aatccatcag agttggcggc tacaatgccg aaacagacga agttgaattc	12540
atttttgacg actacttata tgctaaggac gatctgttga aatggtaacc cgggctgcag	12600
gcatgcaagc ttggctgttt tggcggtatg gagaagattt tcagcctgat acagattaaa	12660
tcagaacgca gaagcgtct gataaaacag aatttgctg gcggcagtag cgcgggtggtc	12720
ccacctgacc ccatgccgaa ctcagaagtg aaacgccgta gcgccgatgg tagtgtggg	12780

tctcccatg cgagagtagg gaactgccag gcatcaaata aaacgaaagg ctcagtcgaa	12840
agactggggc tttcgtttta tctgttggtt gtcggtgaac gctctcctga gtaggacaaa	12900
tccgccggga gcggatttga acgttgcgaa gcaacggccc ggaggggtggc gggcaggacg	12960
cccgccataa actgccaggc atcaaattaa gcagaaggcc atcctgacgg atggcctttt	13020
tgcgtttcta caaactccag ctggatcggg cgctagagta tacatttaaa tggtagcctc	13080
tagtcaaggc cttaagttag tcgtattacg gactggccgt cgttttacaa cgtcgtgact	13140
gggaaaaccc tggcgttacc caacttaatc gccttgtagc acatccccct ttcgccagct	13200
ggcgtaatag cgaagaggcc cgcaccgacg gcccttccca acagttgcgc agcctgaatg	13260
gcgaatggcg cctgatgcgg tattttctcc ttacgcacgt gtgcggtatt tcacaccgca	13320
tatggtgcac tctcagtaca atctgctctg atgccgcata gttaagccag ccccgacacc	13380
cgccaacacc cgctgacgag ct	13402

<210> 55
 <211> 14443
 <212> DNA
 <213> Secuencia Artificial

5

<220>
 <223> Plásmido

<400> 55	
ttctgataac aaactagcaa caccagaaca gcccgtttgc gggcagcaaa acccgtggga	60
attaattccc ctgctcgcgc aggctgggtg ccaagctctc gggtaacatc aaggcccgat	120
ccttgagacc cttcttacag agatgaaaaa caaacgcga cgccaggcgg catcgcggtc	180
tcagagatat gtttacgtag atcgaagagc accggtgttt aaacgccctt gacgatgcca	240
catcctgagc aaataattca accactaaac aaatcaaccg cgtttcccgg aggtaacgga	300
gctcatgacg ctgtgttggt gtgaagccct gatcgacatg ctgccccggc agacgacgct	360
gggtgagggc ggctttgccc cttacgcagg cggagcggtc ttcaacacgg caattgcgct	420
ggggcgcttt ggcgtccctt cagccttttt taccggtctt tccgacgaca tgatgggcca	480
tatcctgcgg gagaccctgc gggccagcaa ggtggatttc agctattgcg ccaccctgtc	540
gcgccccacc accattgcgt tcgttaagct ggttgatggc catgcgacct acgcttttta	600
cgacgagaac accgccggcc ggatgatcac cgaggccgaa cttccggcct tgggagcgga	660
ttgcgaagcg ctgcatttcg gcgccatcag ccttattccc gaaccctgcg gcagcaccta	720
tgaggcgctg atgacgcgcg agcatgagac ccgcgtcacc tcgctcgatc cgaacattcg	780
tcccggtctt atccagaaca agcagtcgca catggccggc atccgccgca tggcgccgat	840
gtctgacatc gtcaagttct cggatgagga cctggcggtg ttcggtctgg aaggcgacga	900

ggacacgctt gcccgccact ggctgcacca cggtgcaaaa ctcgctggtg tcacccgtgg	960
cgccaagggt gccgtgggtt acagcgccaa tctcaagggt gaagtggcct ccgagcgct	1020
cgaagtggtc gatacggtcg gcgcggcgca tacgttcgat gccggcattc ttgcttcgct	1080
gaaaatgcag ggcctgctga ccaaagcgca ggtggcttcg ctgagcgaag agcagatcag	1140
aaaagctttg gcgcttgcg cgaaagccgc tgcggtcact gtctcgggg ctggcgcaaa	1200
tccgcctttc gcgcatgaaa tccggttggtg attaatataa gcacgcagtc aaacaaaaa	1260
cccgcgccat tgcgcgggtt tttttatgcc cgaaggcgcg ccagcacgca gtcaaacaaa	1320
aaaccgcgc cattgcgcgg gtttttttat gccgaacgg ccgaggtctt ccgatctcct	1380
gaagccaggg cagatccgtg cacagcacct tgccgtagaa gaacagcaag gccgccaatg	1440
cctgacgatg cgtggagacc gaaaccttgc gctcgttcgc cagccaggac agaaatgcct	1500
cgacttcgct gctgccaag gttgccgggt gacgcacacc gtggaaacgg atgaaggcac	1560
gaacccagtg gacataagcc tggtcggttc gtaagctgta atgcaagtag cgtatgcgct	1620
cacgcaactg gtccagaacc ttgaccgaac gcagcggtag taacggcgca gtggcggttt	1680
tcatggcttg ttatgactgt tttttgggg tacagtctat gcctcgggca tccaagcagc	1740
aagcgcgtta cgcctgggt cgatgtttga tggtatggag cagcaacgat gttacgcagc	1800
agggcagtcg ccctaaaaca aagttaaaca tcatgaggga agcggtgatc gccgaagtat	1860
cgactcaact atcagaggta gttggcgta tcgagcgcca tctcgaaccg acgttgctgg	1920
ccgtacattt gtacggctcc gcagtggatg gcggcctgaa gccacacagt gatattgatt	1980
tgctggttac ggtgaccgta aggcttgatg aaacaacgcg gcgagctttg atcaacgacc	2040
ttttggaaac ttcggcttcc cctggagaga gcgagattct ccgcgctgta gaagtcacca	2100
ttgttgta caacgacatc attccgtggc gttatccagc taagcgcgaa ctgcaatttg	2160
gagaatggca gcgcaatgac attcttgacg gtatcttcga gccagccacg atcgacattg	2220
atctggctat cttgctgaca aaagcaagag aacatagcgt tgccttggtg ggtccagcgg	2280
cggaggaact ctttgatccg gttcctgaac aggatctatt tgaggcgcta aatgaaacct	2340
taacgctatg gaactcgccg cccgactggg ctggcgatga gcgaaatgta gtgcttacgt	2400
tgtccgcat ttggtacagc gcagtaaccg gcaaaatcgc gccgaaggat gtcgctgccg	2460
actgggcaat ggagcgctg ccggcccagt atcagcccg cactctgaa gctagacagg	2520
cttatcttg acaagaagaa gatcgcttg cctcgcgcg agatcagttg gaagaatttg	2580
tccactacgt gaaaggcgag atcaccaagg tagtcggcaa ataattgtcta acaattcggt	2640
caagccgacg ccgcttcgcg gcgcggctta actcaagcgt tagatgcact aagcacataa	2700
ttgctcacag ccaaactatc aggtcaagtc tgcttttatt atttttaagc gtgcataata	2760
agccctacac aaattgggag atatatcatg aaaggctggc tttttcttgt tatcgcaata	2820

gttggcgaag taatcgcaac atccgcatta aaatctagcg agggctttac taagctcgtc	2880
agcgggtggt ggcggtgtgc ggggctggct taactatgcg gcatcagagc agattgtact	2940
gagagtgcac catatgcggt gtgaaatacc gcacagatgc gtaaggagaa aataccgcat	3000
caggcgccat tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc	3060
ttcgctatta cgccagctgg cgaaaggggg atgtgctgca aggcgattaa gttgggtaac	3120
gccagggttt tcccagtcac gacgttgtaa aacgacggcc agtccgtaat acgactcact	3180
taaggccttg actagagggt accatttaaa tgtatactct agcgcccgat ccagctggag	3240
ttttagaaaa cgcaaaaagg ccatccgtca ggatggcctt ctgcttaatt tgatgcctgg	3300
cagtttatgg cgggcgtcct gcccgccacc ctccgggccc ttgcttcgca acgttcaaat	3360
ccgctcccgg cggatttgtc ctactcagga gagcgttcac cgacaaacaa cagataaaac	3420
gaaaggccca gtctttcgac tgagcctttc gttttatttg atgcctggca gttccctact	3480
ctcgcatggg gagacccac actaccatcg gcgctacggc gtttcacttc tgagtccggc	3540
atggggtcag gtgggaccac cgcgctactg ccgcaggca aattctgttt tatcagaccg	3600
cttctgcgtt ctgatttaat ctgtatcagg ctgaaaatct tctctcatcc gccaaaacag	3660
ccaagcttgc atgcctgcag cccgggttac catttcaaca gatcgtcctt agcatataag	3720
tagtcgtcaa aaatgaattc aacttcgtct gtttcggcat tgtagccgcc aactctgatg	3780
gattcgtggt ttttgacaat gatgtcacag cctttttcct ttaggaagtc caagtcgaaa	3840
gtagtggaac taccaatgat cttacaaccg gcggcttttc cggcggcaat acctgctgga	3900
gcgtcttcaa atactactac cttagatttg gaagggtcct gctcattgat cggatatcct	3960
aagccattcc tgccttcag atatggttct ggatgaggct taccctgttt gacatcatta	4020
gcggtaatga agtactttgg tctcctgatt ccagatgct cgaaccattt ttgtgccata	4080
tcacgggtac cggaagttgc cacagcccat ttctcttttg gtagagcgtt caaagcgttg	4140
cacagcttaa ctgcacctgg gacttcaatg gatttttcac cgtacttgac cggaatttca	4200
gcttctaatt tgtaacata ctcttcattg gcaaagtctg gagcgaactt agcaatggca	4260
tcaaacgttc tccaaccatg cgagacttgg ataacgtgtt cagcatcgaa ataaggtttg	4320
tccttaccga aatccctcca gaatgcagca atggctgggt gagagatgat aatggtaccg	4380
tgcagtcga acaaagcggc gttaactttc aaagatagag gtttagtagt caatcccata	4440
attctagtct gtttcctgga tccaataaat ctaatcttca tgtagatcta attcttcaat	4500
catgtccggc aggttcttca ttgggtagtt gttgtaaacy atttggtata cggcttcaaa	4560
taatgggaag tcttcgacag agccacatgt ttccaaccat tcgtgaactt ctttgcagg	4620
aattaaacct tgagcggatt ggccattcaa caactccttt tcacattccc aggcgtcctt	4680

accagaagta gccattagcc tagcaacott gacgtttcta ccaccagcgc aggtggtgat	4740
caaatcagca acaccagcag actcttggtta gtatgtttct tctctagatt ctgggaaaaa	4800
catttgaccg aatctgatga tctcacccaa accgactctt tggatggcag cagaagcggt	4860
gttaccaccag cctagacctt cgacgaaacc acaacctaag gcaacaacgt tcttcaaagc	4920
accacagatg gagataccag caacatcttc gatgacacta acgtggaagt aaggtctgtg	4980
gaacaaggcc tttagaacct tatggtcgac gtcttggccc tcgcctctga aatcctttgg	5040
aatgtggtta gcaactgttg tttcagacca gtgttcttga gcgacttcgg tggcaatgtt	5100
agcaccagat agagcaccac attgaatacc tagttcctca gtgatgtaag aggatagcaa	5160
ttggacacct ttagcaccaa cttcaaaacc ctttagacag gagatagctc tgacgtgtga	5220
atcaacatga cctttcaatt ggctacagat acggggcaaa aattgatgtg gaatgttgaa	5280
aacgatgatg tcgacatcct tgactgaatc aatcaagtct ggattagcaa ccaaattgtc	5340
gggtagagtg atgccaggca agtatttcac gttttgatgt ctagtattta tgatttcagt	5400
caatttttca ccattgatct cttcttcgaa caccacatt tgtactattg gagcgaaaac	5460
ttctgggtat cccttacaat tttcggcaac caccttggca atagtagtac cccagttacc	5520
agatccaatc acagtaacct tgaaaggctt ttcggcagcc ttcaaagaaa cagaagagga	5580
acttctcttt ctaccagcat tcaagtggcc ggaagttaag tttaatctat cagcagcagc	5640
agccatggaa ttgtcctcct tactagtcac ggtctgtttc ctgtgtgaaa ttgttatccg	5700
ctcacaattc cacacattat acgagccgga tgattaattg tcaacagctc atttcagaat	5760
atttgccaga accgttatga tgtcggcgca aaaaacatta tccagaacgg gagtgcgcct	5820
tgagcgacac gaattatgca gtgatttacg acctgcacag ccataccaca gcttccgatg	5880
gctgcctgac gccagaagca ttggtgcacg ctagacaaga aaaaaggcac gtcactctgac	5940
gtgccttttt tatttgtacc tagaggctgg cgcgagcgcc cgtttaattc gcctgaccgg	6000
ccagtagcag cccggtggcg accgcattgc gcggcccttc tgttccccga atattgccct	6060
gcccggcgac cagcccatag tgcgacaagg cttccgtgat aagctgcggg atctcaaagt	6120
ccagcgatga gccgcccacc agcaccacaa aggcgatatc gcgaatggaa ccgccgggtg	6180
agacctggcg cagcgcgcg caggcagttgg tgacaaacac tttctctttc gcctgccggc	6240
gcacgagacg aattttttcc agcgggctgg cgttatcgat cggcaccagt tcgccctcct	6300
tgatgtacac cactttggcg aacaccgccg ggctgagggc ttcccgaag aactccaccg	6360
cgccattctc gtgacgaata ctgaacaggc tttccacttt ggccagcggg tattttttta	6420
tcgcttccgc cagcgaaaga tcctcgaggc ccagctcggg tttaataaac aggctgacca	6480
tattccccgc cccggcgaga tggaccgccg ttatctgccc ctccgcgttg acgatcgccg	6540
catccgtcga gccggcgccg aggtcgagga tcgccagcgg cgccgcacag ccgggagtg	6600

ttaacgcccc ggcgatggcc atgttggcct ccaogccgcc caccaccacc tcggtctgca	6660
gtcggggcgt cagttcgcgg gcgataacct gcatttgcag acgatccgct ttcaccatcg	6720
ccgccatccc gacggcattc tccatggcgc actcgccggc catcccggcc tgcaccttgc	6780
gcggaataaa cgtatccacc gccagcagat cctggatgta tatcgcgctc atctcatggc	6840
cggtcaggga cgccattacc ttgcgcaccc gctcaagcat gccgccggcg tgggtgcccg	6900
gttcgccggc gatgtcgcgt accggagcgc aggcgctcat cgcctgcatg atggcttccg	6960
cggcctcggc gacatcggcc tctccgcggc gcttttcgcc gctaagttag aggttgcccg	7020
ccgggatcac ccgcgactgc acatccccct gcgggggtctt gagcaccacc gcggaacggt	7080
tgccaatcag ggcgcggggc atggggacga tggcctgggt ctcttcgggg cttagcccga	7140
agaaggtggc gatcccgtag ggattcgaca ggatccgcac cacctggccc ggcgcggcca	7200
cttccaccgc cgccattacc ccctcgggga cctgctccag cagcgtcact tcatccacca	7260
ccggcagggt tttacgcagg cggttgttca ccagcacgcc gtcgtccttt ttgaggatcg	7320
ccgccaccac gttgatcccc cggtcgagcg cctcattgag ccaccacacg gcgtcaagga	7380
aatcgacggc gtcgtcaatc agtacgatcc acccctcggc atactgcgcc gccggcagcg	7440
tcgccagccg cccgagggcg atagtcgtcc ccacgccaac gccacccccg cccggcgtct	7500
gcgggttatg accgatcatg gtcgattcgg tgataatggt ctcggtgatg gtctccatcg	7560
ccacatcgcc aatcaccggc gcggcttcgt taagatagat gcgagagaca tcgctcatcg	7620
accacggtgt tttcgccagg gcctgctcca gcgcggcgag ggtcccggcg atattgtccc	7680
gcgtcccttt catgcccgtc gtcgcgacga tcccgctggc aacaaacgcc ctgcctcgcg	7740
ggtagtcgga cgccagcgcc acctcggtag tggcgttgcc gatatcaatc ccggctatta	7800
acggcatgct gacctccgct tagcttcctt tacgcagctt atgccgctgc tgatacactt	7860
ccgccgactc ccggacaaag gcggcattca ctgtcgcatg ccagggtgtgc tccagctcgt	7920
cggcgatcgc cagcagctcc gcctgcgagg agcggaacgg gcgcagcgcg ttatagatag	7980
ccagaatgcg ctcgtcagga atggcgataa gtcgcccgcc gcggcggaag ttgcgcgcca	8040
ccgcatggcg ctgcatctgc tcggcaatct gcgcctggta ctcaagggtc tggcgggaga	8100
tccgcacatc ctgcggggcc acctcgccag agagcacctt ctcgagggtg atatcggtca	8160
atggtttgcc ggtaggcgtc aggatatgct ccgggcagcg ggtggctaac ggataatcct	8220
gcacgcgcat ggttttctcg ctcatggtca ctcccttact aagtcgatgt gcagggtgac	8280
gggctcggcg tcctgcacca catgtttggt ctctttgata tgaaatagcg cggctttggc	8340
cataaatttc ggccgcacca tctgatcggt caccaccggc accggcggaag gtgactcttt	8400
gcgcgcatag cgcgcagcgt ttttgccaat ctgccggtag gtctccagcg tcagcagcgg	8460

cgccctgggag aacagctcca ggttgctgag cggcagcaga tcgcgctgat ggatgaccgt	8520
ggtccccttc gactggatac cgatgccgat ccccgagccg ctgaggttgg ccgcatccca	8580
ggccataaag gagacgtcgg acgtgcgcag aatgcgcacc acccgggcgt gaagcccctc	8640
ttcttccacc ccggcaatca gctctttgag gatcgcgcga tggggcatat cgatcagagt	8700
gtgatgctgg tgtttatcga aggcagggcc gacgcgcgatc accacttcat cggcgcgttc	8760
atcggcagaa gctaccccg cctcgcgggt tttcagggta aaagagggct gaatttgggt	8820
tgtctgttgc acaggaatac cgccttgctc aatggtgtcg ggctgaacca cggccggaat	8880
atctttgatc tccgccagc gttcggcaga gatgcgatag ccggtgcccg gccctgata	8940
gtcattgatg tcgttgaccg cactcaccac ctggaactgc cgatcgaaaa tggccgaggt	9000
ctgcaggtaa tcgcgggtga cccgctggcg cagcatattg agaatttgc tggcgatatc	9060
ctcaaagccg ctgcggctca gcgcgccgac aatatcgagg ccggtgatgt tgcgcttcat	9120
catctcttcc accgcactca gatcctccac cacgttacgc ggcgccatct cgttgcctgcc	9180
gtgcgcgtag gtggcggcct ccacctctc gtcggcgatt ggcggcagcc ccagctcgcg	9240
gaaaaccgcc tggatcgccc gcgcgcctt ctggcgaatg gcaatggttt ccgcctcgg	9300
caccggacgc aggcgcgcgt caaccatcag gtcacgctgc aggatgttgt aatcatcaaa	9360
atcttccgca tcgaagtctg agccggcgaa catgttgtcg tagttcggca ccgcgctgta	9420
gccggagaaa ataaagtcgg tgcccggcag catctgcac aggtgcgcg cggtcggcg	9480
aatatccgag tgggagaaa tctggtcgtt ggcggcagcc acttcgaggt cgagcataga	9540
ggcgatcagg ttttccgcca gcaccgccc aatgcccagc ggcacagcgc cggatcatgcc	9600
gatacagctc accgcgcctt tttgcagtcc ctgaacccc gcgcctttag taatgaagat	9660
gcagcgcgat tcgaggtaga gcatcgactt gctctccgaa tagcccatca gcgcttcgga	9720
tccggtgccg gaggtgtagc gcattttcaa cccgcgggag gcgtaggccg aggcgaggaa	9780
cgcctttgac cacggcgat catcgccgct ggtaaatacc gcttcgggtgc cgtagaccga	9840
caccgtctcg gcgtagctgg ttaagccacg catgccagc tccagctcgg tggcctcttc	9900
caccgagcac tcggtcaaca cgcggggcg gccgcactgc gaaccgacca acagcgccag	9960
ggcgttaaac ggcgcgtagc gcgcgatacc gaccgtggtc tcctgttctg agaagccgcg	10020
gatcccgcc tcggcgcgct cagcggcaat ctgcacggga ttatctttga gattggtgac	10080
gtggcactgg ttggaggggg tccggcgggc acgcatcttc tgcagcgcca tcatcatctc	10140
caccacgttc atctgcgcca tcacctgac cgctttggcc ggcgatgatg cggtagtgat	10200
ggcaatgac tcctcccgcc tgacgtgaat atccaccagc atacgggcta ttccaccgc	10260
ctccagggcg attgcctgct ctgtgcgctc aacgttgatc gcgtaatcgg cgataaatcg	10320
gtcgatcatg tcaaactgg cccggcggtt gccgtccagt tcgacgatca gaccgttgc	10380

cacttttact gaagagaccg ggtcaaaggg gctgtccatg gcgatcagcc cctcttcagg	10440
ccactcgcca atcagcccgt cctgattgac ggggcgctgg gccagtactg caaatcgttt	10500
tgatcttttc attgttcata ggctcaaaag gtgaagcttg gttacctcog ggaaacgcgg	10560
ttgatttggt tagtggttga attatttgct caggatgtgg cattgtcaag ggcgtgacgg	10620
ctcgctgac ttctcgttcc agtgcccccg tccgacagtc gagcgtgcca gcccataatc	10680
tcgcgctggg gctgcatacc gtggcaaaaca gcacagatcg cctaggaaaa aaaaagcccg	10740
cactgtcagg tgcgggcttt tttctgtgtt tgctaggcca gttcaagcgc aagcatcagg	10800
gtgcagctgg gcagaggcga gattcctccc cgggatcacg aactgtttta acgggccgct	10860
ctcgccata ttgcggtcga taagccgctc cagggcggtg atctcctctt cgccgatcgt	10920
ctggctcagg cgggtcaggc cccgcgcata gctggccagt tcagccccca gcacgaacag	10980
cgtctgctga atatggtgca ggctttcccg cagcccgcg tcgcgggtcg tggcgtagca	11040
gacgcccagc tgggatataca gtcatcgac ggtgccgtag gcctcgacgc gaatatggtc	11100
tttctcgatg cggtgcgcgc cgtacagggc ggtggtgcct ttatccccgg tgcgggtata	11160
gatacgatac attcagtttc tctcacttaa cggcaggact ttaaccagct gcccgcgctt	11220
ggcgccgagc gtacgcagtt gatcgctcgt atcgggtgacg tgtccggtag ccagcggcgc	11280
gtccgccggc agctgggcat gagtgagggc tatctcgccg gacgcgctga gcccgatacc	11340
caaccgcagg ggcgagcttc tggccgccag ggcgccagc gcagcggcgt caccgcctcc	11400
gtcatagggt atggtctggc aggggacccc ctgctcctcc agccccagc acagctcatt	11460
gatggcgccg gcatggtgcc cgcgcggatc gtaaaacagg cgtacgcctg gcggtgaaag	11520
cgacatgacg gtccccctcg taacaactcag aatgcctggc ggaacatacg atagctcata	11580
atataccttc tcgcttcagg ttataatgcg gaaaaacaat ccagggcgca ctgggctaata	11640
aattgatcct gctcgaccgt accgccgcta acgcgcagcg cgccaattac ctgctcatta	11700
aaaataactg gcaggccgcc gccaaaaata ataattcgct gttggttggg tagctgcaga	11760
ccgtacagag attgtcctgg ctggaccgct gacgtaattt catgggtacc ttgcttcagg	11820
ctgcaggcgc tccaggcttt attcaggga atatcgagc tggagacgaa ggcctcgtcc	11880
atccgctgga taagcagcgt gttgcctccg cggtaacta cggaaaacac caccgccagc	11940
ttgatctcag tggttttttt ttccaccgcc gccgccattt gctgggcggc ggccagggtg	12000
attgtctgaa cttgttggtt cttgttcata attctctccc gcaagcttgg ttacctccgg	12060
gaaacgcggt tgatttggtt agtggttgaa ttatttgctc aggatgtggc attgtcaagg	12120
gcgtgacggc tcgcctgact tctcgttcca gtgccccgt cgcacagtcg agcgtgcgag	12180
cccataatct cgcgctggtg ctgcataacc tggcaaacag cacagatcgc ctagcagtc	12240

aaagcctccg gtcggaggct tttgactatt taaatgaatt cccgacagta agacgggtaa	12300
gcctgttgat gataccgctg ccttactggg tgcattagcc agtctgaatg acctgtcacg	12360
ggataatccg aagtggtcag actggaaaat cagagggcag gaactgctga acagcaaaaa	12420
gtcagatagc accacatagc agacccgcca taaaacgccc tgagaagccc gtgacgggct	12480
tttcttgat tatgggtagt ttccttgcat gaatccataa aaggcgctg tagtgccatt	12540
tacccccatt cactgccaga gccgtgagcg cagcgaactg aatgtcacga aaaagacagc	12600
gactcaggtg cctgatggtc ggagacaaaa ggaatattca gcgatttgcc cgagcttgcg	12660
aggggtgctac ttaagccttt aggggtttta ggtctgtttt gtagaggagc aaacagcgtt	12720
tgcgacatcc ttttgtaata ctgcggaact gactaaagta gtgagttata cacagggctg	12780
ggatctattc tttttatctt tttttattct ttctttattc tataaattat aaccacttga	12840
atataaacia aaaaaacaca caaaggctca gcggaattta cagagggctc agcagaattt	12900
acaagttttc cagcaaaggc ctagcagaat ttacagatac ccacaactca aaggaaaagg	12960
actagtaatt atcattgact agcccatctc aattgggtata gtgattaaaa tcacctagac	13020
caattgagat gtatgtctga attagttggt ttcaaagcaa atgaactagc gattagtgcg	13080
tatgacttaa cggagcatga aaccaagcta attttatgct gtgtggcact actcaacccc	13140
acgattgaaa accctacaag gaaagaacgg acggtatcgt tcacttataa ccaatacgtc	13200
cagatgatga acatcagtag ggaaaatgct tatggtgtat tagctaaagc aaccagagag	13260
ctgatgacga gaactgtgga aatcaggaat cctttgggta aaggctttga gattttccag	13320
tggacaaact atgccaagtt ctcaagcgaa aaattagaat tagtttttag tgaagagata	13380
ttgccttatc ttttccagtt aaaaaattc ataaaatata atctggaaca tgtaagtct	13440
tttgaaaaca aatactctat gaggatttat gagtgttat taaaagaact aacacaaaag	13500
aaaactcaca aggcaaatat agagattagc cttgatgaat ttaagttcat gttaatgctt	13560
gaaaataact accatgagtt taaaaggctt aaccaatggg ttttgaaacc aataagtaaa	13620
gatttaacaa cttacagcaa tatgaaattg gtggttgata agcgaggccg cccgactgat	13680
acgttgattt tccaagttga actagataga caaatggatc tcgtaaccga acttgagaac	13740
aaccagataa aatgaatgg tgacaaaata ccaacaacca ttacatcaga ttcctaccta	13800
cataacggac taagaaaaac actacacgat gctttaactg caaaaattca gctcaccagt	13860
tttgaggcaa aatttttgag tgacatgcaa agtaagtatg atctcaatgg ttcgttctca	13920
tggctcacgc aaaaacaacg aaccacacta gagaacatac tggctaaata cggaaggatc	13980
tgaggttctt atggctcttg tatctatcag tgaagcatca agactaacia acaaaagtag	14040
aacaactgtt caccgttaca tatcaaaggg aaaactgtcc atatgcacag atgaaaacgg	14100
tgtaaaaaag atagatacat cagagctttt acgagttttt ggtgcattca aagctgttca	14160

	ccatgaacag atcgacaatg taacagatga acagcatgta acacctaata gaacaggtga	14220
	aaccagtaaa acaaagcaac tagaacatga aattgaacac ctgagacaac ttgttacagc	14280
	tcaacagtca cacatagaca gcctgaaaca ggcgatgctg cttatcgaat caaagctgcc	14340
	gacaacacgg gagccagtga cgcctcccgt ggggaaaaaa tcatggcaat tctggaagaa	14400
	atagcgcttt cagccggcaa accggctgaa gccgatctg cga	14443
	<210> 56	
	<211> 28	
	<212> DNA	
5	<213> Secuencia Artificial	
	<220>	
	<223> Cebador	
	<400> 56	
	attaatcatg cttggcaaaa gcatgaca	28
10	<210> 57	
	<211> 26	
	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
15	<223> Cebador	
	<400> 57	
	ctcgagagat aaccctcttg tttttg	26
	<210> 58	
	<211> 100	
20	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
	<223> Cebador	
	<400> 58	
	cgtctaccct tgttatacct cacaccgcaa ggagacgatc atgaccaata atcccccttc	60
25	agcacagatt aagcccggcg gtgtaggctg gagctgcttc	100
	<210> 59	
	<211> 100	
	<212> DNA	
	<213> Secuencia Artificial	
30	<220>	
	<223> Primer	
	<400> 59	
	gcatcaggca atgaataccc aatgcgacca gcttcttata tcagaacagc cccaacggtt	60
	tatccgagta gctcaccagc catatgaata tcctccttag	100
	<210> 60	
35	<211> 22	
	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
	<223> Cebador	

ES 2 552 317 T3

	<400> 60	
	atgaccaata atcccccttc ag	22
	<210> 61	
	<211> 21	
5	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
	<223> Cebador	
	<400> 61	
10	gcttcttata tcagaacagc c	21

REIVINDICACIONES

1. Una bacteria recombinante que comprende en su genoma o en al menos una construcción recombinante:
 - (a) una secuencia de nucleótidos heteróloga que codifica un polipéptido que tiene actividad transportadora de sacarosa, teniendo dicho polipéptido al menos un 95% de identidad de secuencia, basado en el método de alineamiento Clustal V, cuando se compara con una secuencia de aminoácidos como la representada en SEQ ID NO: 24; y
 - (b) una secuencia de nucleótidos heteróloga que codifica un polipéptido que tiene actividad hidrolasa de sacarosa, teniendo dicho polipéptido al menos un 95% de identidad de secuencia, basado en el método de alineamiento Clustal V, cuando se compara con una secuencia de aminoácidos como la representada en SEQ ID NO: 26;
- en donde (a) y (b) están cada uno ligados operativamente al mismo o a diferente promotor, en el que además dicha bacteria recombinante es capaz de metabolizar sacarosa.
2. La bacteria recombinante de la reivindicación 1 en donde la secuencia de nucleótidos que codifica el polipéptido que tiene actividad transportadora de sacarosa tiene al menos un 95% de identidad de secuencia, basado en el método de alineamiento BLASTN, cuando se compara con una secuencia de nucleótidos como la representada en SEQ ID NO: 23.
3. La bacteria recombinante de la reivindicación 1 o de la reivindicación 2 en donde la secuencia de nucleótidos que codifica el polipéptido que tiene actividad hidrolasa de sacarosa tiene al menos un 95% de identidad de secuencia, basado en el método de alineamiento BLASTN, cuando se compara con una secuencia de nucleótidos como la representada en SEQ ID NO: 25.
4. La bacteria recombinante de cualquiera de las reivindicaciones de 1 a 3 que comprende además en su genoma o en al menos una construcción recombinante, una secuencia de nucleótidos que codifica el polipéptido que tiene actividad fructoquinasa.
5. La bacteria recombinante de la reivindicación 4, en donde el polipéptido que tiene actividad fructoquinasa tiene al menos un 95% de identidad de secuencia, basado en el método de alineamiento Clustal V, con una secuencia de aminoácidos como la representada en SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 36, SEQ ID NO: 38, SEQ ID NO: 40, SEQ ID NO: 42, SEQ ID NO: 44, o SEQ ID NO: 46.
6. La bacteria recombinante de la reivindicación 4, en donde el polipéptido que tiene actividad fructoquinasa tiene una secuencia de aminoácidos como la representada en SEQ ID NO: 40.
7. La bacteria recombinante de cualquiera de las reivindicaciones de 1 a 6 en donde dicha bacteria se selecciona de un grupo que consiste en los géneros: *Escherichia*, *Klebsiella*, *Citrobacter*, y *Aerobacter*.
8. La bacteria recombinante de la reivindicación 7, en donde dicha bacteria es *Escherichia coli*.
9. La bacteria recombinante de cualquiera de las reivindicaciones de 1 a 8 en donde la bacteria recombinante produce 1,3-propanodiol, glicerol y/o ácido 3-hidroxipropiónico.
10. Un proceso para fabricar glicerol, 1,3-propanodiol y/o ácido 3-hidroxipropiónico a partir de sacarosa que comprende:
 - (a) cultivar la bacteria recombinante de la reivindicación 9 en presencia de sacarosa; y
 - (b) opcionalmente, recuperar el glicerol, 1,3-propanodiol y/o ácido 3-hidroxipropiónico producidos.