

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 426 070**

51 Int. Cl.:

C12P 7/16 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **16.04.2008 E 08745966 (5)**

97 Fecha y número de publicación de la concesión europea: **10.07.2013 EP 2134849**

54 Título: **Producción fermentativa de isobutanol utilizando enzimas cetol-ácido reductoisomerasas muy activas**

30 Prioridad:

18.04.2007 US 912491 P

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

21.10.2013

73 Titular/es:

**BUTAMAX (TM) ADVANCED BIOFUELS LLC
(100.0%)
200 POWDER MILL ROAD
WILMINGTON DE 19803, US**

72 Inventor/es:

**LIAO, DER-ING;
NELSON, MARK J. y
BRAMUCCI, MICHAEL G.**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 426 070 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

Producción fermentativa de isobutanol utilizando enzimas cetol-ácido reductoisomerasas muy activas

Campo de la invención

5 La invención se refiere al campo de la microbiología industrial y a la producción de alcoholes. De modo más específico, se produce isobutanol a través de la fermentación industrial de un microorganismo recombinante utilizando enzimas cetol-ácido reductoisomerasas (KARI) especiales con un número de recambio alto. Esta invención también se refiere a métodos para descubrir enzimas KARI muy activas a partir de microorganismos naturales.

Antecedentes de la invención

10 El butanol es un importante producto químico, útil como aditivo de combustibles, como materia prima química en la industria del plástico, y como extractante de calidad alimentaria en la industria alimentaria y de aromatizantes. Cada año, de diez a doce mil millones de libras de butanol son producidas por medios petroquímicos, y es probable que aumente la necesidad de esta materia prima química.

15 Los métodos para la síntesis química del isobutanol son conocidos, tales como la síntesis oxo, la hidrogenación catalítica del monóxido de carbono (Ullmann's Encyclopedia of Industrial Chemistry, 6ª edición, 2003, Wiley-VCH Verlag GmbH and Co., Weinheim, Alemania, vol. 5, pp. 716-719), y la condensación de Guerbet del metanol con n-propanol (Carlini *et ál.*, J. Molec. Catal. A:Chem., 220, 215-220, 2004). Estos procesos emplean materiales de partida derivados de petroquímicos y, en general, son caros y perjudiciales para el medioambiente. La producción de isobutanol a partir de materias primas derivadas de plantas podría minimizar las emisiones de gases causantes del efecto invernadero y representarán un avance en la técnica.

20 El isobutanol se produce de modo biológico como subproducto de la fermentación de levaduras. Es un componente del "aceite de fusel" que se forma como resultado del metabolismo incompleto de los aminoácidos por este grupo de hongos. El isobutanol se produce específicamente a partir del catabolismo de la L-valina. Después de que el grupo amina de la L-valina se haya recogido como fuente de nitrógeno, el α -cetoácido resultante es descarboxilado y reducido a isobutanol por enzimas de la denominada vía de Ehrlich (Dickinson *et ál.*, J. Biol. Chem., 273, 25752-25756, 1998). Los rendimientos del aceite de fusel y/o sus componentes logrados durante la fermentación de bebidas generalmente son bajos. Por ejemplo, se ha indicado que la concentración de isobutanol producido en la fermentación de la cerveza es menor que 16 partes por millón (García *et ál.*, Process Biochemistry, 29, 303-309, 1994). La adición de L-valina exógena a la fermentación aumenta el rendimiento de isobutanol, según describen Dickinson *et ál.*, *supra*, en donde se indica que se obtiene un rendimiento de isobutanol de 3 g/l proporcionando L-valina a una concentración de 20 g/l en la fermentación. Además, se ha demostrado la producción de n-propanol, isobutanol y alcohol isoamílico por células inmovilizadas por alginato de calcio de *Zymomonas mobilis*. Se empleó un medio que contenía glucosa al 10% suplementado con L-Leu, L-Ile, L-Val, ácido alfa-cetoisocaproico (alfa-KCA), ácido alfa-cetobutírico (alfa-KBA) o ácido alfa-cetoisovalérico (alfa-KVA) (Oaxaca, *et ál.*, Acta Biotechnol., 11, 523-532, 1991). El alfa-KCA aumenta los niveles de isobutanol. Los aminoácidos también producen los correspondientes alcoholes, pero en un grado menor que los cetoácidos. Se demostró un aumento en el rendimiento de alcoholes C3-C5 a partir de carbohidratos cuando se añaden los aminoácidos leucina, isoleucina y/o valina al medio de crecimiento como fuente de nitrógeno (documento WO 20055040392).

35 Aunque los métodos descritos anteriormente indican el potencial de producción de isobutanol a través de medios biológicos, estos métodos tienen un coste prohibitivo a escala industrial para la producción de isobutanol. La biosíntesis de isobutanol directamente desde azúcares sería económicamente viable y representaría un avance en la técnica. Sin embargo, esta producción se ve gravemente obstaculizada por la lenta enzima cetol-ácido reductoisomerasa (KARI), que cataliza la segunda etapa en la vía del isobutanol que convierte el acetolactato en dihidroxivalerato. Debido a que esta enzima ya se expresa a niveles altos (S. Epelbaum *et ál.*, J. Bacteriol., 180, 4056-4067, 1998), resulta necesario aumentar la actividad de KARI sin aumentar la cantidad de proteína, es decir, aumentar la actividad específica de la enzima.

Los solicitantes han resuelto el problema mencionado a través del descubrimiento de una enzima KARI que tiene una alta actividad específica que puede utilizarse para potenciar la producción biológica del isobutanol.

Resumen de la invención

50 La invención se refiere a organismos recombinantes que expresan enzimas KARI muy activas. El microorganismo modificado tendrá niveles altos de la forma corta de la enzima KARI que posee una actividad específica significativamente mayor (6-8 veces mayor que la enzima KARI en *E. coli*) y puede utilizarse para la producción comercial de isobutanol. Por consiguiente, en una realización, la invención proporciona un método para la conversión de acetolactato a dihidroxisovalerato, que comprende:

55 a) proporcionar una célula hospedante microbiana que comprende una construcción genética que codifica un polipéptido que tiene una actividad específica de cetol-ácido reductoisomerasa mayor que la actividad específica de

una cetol-ácido reductoisomerasa de *E. coli*; y

b) poner en contacto la célula hospedante de a) con acetolactato, por lo que se produce 2,3-dihidroxiisovalerato.

5 En una realización preferida, la construcción genética codifica un polipéptido que tiene una actividad específica de cetol-ácido reductoisomerasa mayor que 1,1 $\mu\text{mol}/\text{min}/\text{mg}$ basada en la proteína purificada, medida mediante el ensayo del consumo de NADPH, ejecutado bajo las siguientes condiciones:

a) pH de aproximadamente 7,5;

b) una temperatura de aproximadamente 22,5 °C; y

c) potasio mayor que aproximadamente 10 mM.

En otra realización, la invención proporciona un método para la producción de isobutanol, que comprende:

10 a) proporcionar una célula hospedante microbiana recombinante que comprende las siguientes construcciones genéticas:

1) al menos una construcción genética que codifica una enzima acetolactato sintasa para la conversión del piruvato a acetolactato (etapa a de la vía);

15 2) al menos una construcción genética que codifica una enzima cetol-ácido reductoisomerasa con una actividad específica mayor que 1,1 $\mu\text{mol}/\text{min}/\text{mg}$ basada en la proteína purificada, medida mediante el ensayo del consumo de NADPH, ejecutado bajo las siguientes condiciones:

i) pH de aproximadamente 7,5;

ii) una temperatura de aproximadamente 22,5 °C; y

20 iii) potasio mayor que aproximadamente 10 mM, para la conversión de (S)-acetolactato a 2,3-dihidroxiisovalerato (etapa b de la vía);

3) al menos una construcción genética que codifica una acetohidroxiácido deshidratasa para la conversión de 2,3-dihidroxiisovalerato a α -cetoisovalerato (etapa c de la vía);

4) al menos una construcción genética que codifica una cetoácido de cadena ramificada descarboxilasa para la conversión de α -cetoisovalerato a isobutiraldehído (etapa d de la vía);

25 5) al menos una construcción genética que codifica una alcohol de cadena ramificada deshidrogenasa para la conversión de isobutiraldehído a isobutanol (etapa e de la vía); y

b) cultivar la célula hospedante de a) bajo condiciones en las que se produce isobutanol.

30 En otra realización, la invención proporciona una célula hospedante recombinante, que comprende una enzima cetol-ácido reductoisomerasa que tiene una actividad específica mayor que la actividad específica de una cetol-ácido reductoisomerasa de *E. coli*.

En otra realización, la invención proporciona un método para la identificación y el aislamiento de una construcción genética, que codifica una enzima cetol-ácido reductoisomerasa que tiene una actividad específica mayor que 1,1 $\mu\text{mol}/\text{min}/\text{mg}$ basada en la proteína purificada, medida mediante el ensayo del consumo de NADPH, ejecutado bajo las siguientes condiciones:

35 i) pH de aproximadamente 7,5;

ii) una temperatura de aproximadamente 22,5 °C; y

iii) potasio mayor que aproximadamente 10 mM;

que comprende las etapas de:

40 a) identificar especies bacterianas que tengan un tiempo de duplicación más corto que *E. coli* cuando se cultivan en medio mínimo M9;

b) seleccionar las especies bacterianas de (a) para la actividad cetol-ácido reductoisomerasa para identificar las especies bacterianas activas;

45 c) sondear el ADN genómico de las especies bacterianas activa des b) con secuencias de ácidos nucleicos que tengan homología con construcciones genéticas conocidas por codificar una cetol-ácido reductoisomerasa, para identificar y aislar las construcciones genéticas que codifican una cetol-ácido reductoisomerasa a partir de dichas

especies bacterianas activas; y

d) amplificar y expresar las construcciones genéticas que codifican una cetol-ácido reductoisomerasa a partir de dichas especies bacterianas activas; y

5 e) seleccionar las construcciones genéticas expresadas de la etapa d) para detectar las que tengan una actividad específica mayor que 1,1 $\mu\text{mol}/\text{min}/\text{mg}$ basada en la proteína purificada, medida mediante el ensayo del consumo de NADPH, ejecutado bajo las siguientes condiciones:

i) pH de aproximadamente 7,5;

ii) una temperatura de aproximadamente 22,5 °C; y

iii) potasio mayor que aproximadamente 10 mM.

10 **Breve descripción de las figuras y las secuencias de la invención**

La invención puede entenderse más a fondo a partir de la siguiente descripción detallada, la figura y las descripciones de las secuencias adjuntas, que forman parte de esta solicitud.

La figura 1 muestra cuatro vías biosintéticas del isobutanol diferentes. Las etapas denominadas "a", "b", "c", "d", "e", "f", "g", "h", "i", "j" y "k" representan el sustrato para producir las conversiones descritas a continuación.

15 Las siguientes secuencias se ajustan a 37 C.F.R. 1.821-1.825 ("Requirements for Patent Applications Containing Nucleotide Sequences and/or Amino Acid Sequence Disclosures - the Sequence Rules") y son coherentes con la norma ST.25 (1998) de la World Intellectual Property Organization (WIPO) y los requisitos para el listado de secuencias de la EPO y PCT (regla 5.2 y 49.5(a-bis), y la sección 208 y el anexo C de las instrucciones administrativas). Los símbolos y el formato utilizado para los datos de las secuencias de nucleótidos y aminoácidos se ajustan a las reglas indicadas en 37 C.F.R. §1.822.

Tabla 1

Resumen de SEQ ID NO: de los genes y proteínas de la vía del isobutanol preferida		
Descripción	SEQ ID NO: ácido nucleico	SEQ ID NO: péptido
<i>Klebsiella pneumoniae budB</i> (acetolactato sintasa)	1	2
<i>E. coli ilvC</i> (acetohidroxiácido reductoisomerasa)	3	4
<i>E. coli ilvD</i> (acetohidroxiácido deshidratasa)	5	6
<i>Lactococcus lactis kivD</i> (α -cetoácido de cadena ramificada descarboxilasa), con codones optimizados	7	8
<i>E. coli yqhD</i> (alcohol de cadena ramificada deshidrogenasa)	9	10

Las SEQ ID NO:11-22 son las secuencias de nucleótidos de los cebadores oligonucleotídicos utilizados para generar las construcciones en el ejemplo 1.

25 Las SEQ ID NO:23-30 son las secuencias de nucleótidos de los cebadores oligonucleotídicos utilizados para generar las construcciones en el ejemplo 2.

Las SEQ ID NO:11 y 12 son las secuencias de ADN de los cebadores utilizados en el ejemplo 1 para la amplificación mediante PCR del gen *ilvC*.

La SEQ ID NO:13 es la secuencia de ADN directa para el cebador utilizado para clonar el gen KARI en el vector *pBAD* en *E. coli*.

30 La SEQ ID NO:14 es la secuencia de ADN inversa para el cebador utilizado para clonar el gen KARI en el vector *pBAD* en *E. coli*.

La SEQ ID NO:15 es la secuencia de ADN directa para *ilvC-trc-SacI-F* utilizada para amplificar el gen KARI.

La SEQ ID NO:16 es la secuencia de ADN inversa para *ilvC-trc-HindIII-R* utilizada para amplificar el gen KARI.

35 Las SEQ ID NO:17 a 22 son las secuencias de nucleótidos de los cebadores utilizados para confirmar la presencia de la inserción de *ilvC* de *E. coli* mediante digestión con *SacI* y secuenciación del ADN.

SEQ ID NO:17 - *ilvC-trc-F3*

SEQ ID NO:18 - *ilvC-trc-F5*

SEQ ID NO:19 - *ilvC-trc-R2*

SEQ ID NO:20 - *ilvC-trc-R4*

SEQ ID NO:21 - *pBAD-eF1*

5 SEQ ID NO:22 - *PALPK-R1*

Las SEQ ID NOS:23 a 26 son las secuencias de nucleótidos de los cebadores oligonucleotídicos utilizados en la dirección directa e inversa para amplificar los genes *ilvC* a partir del ADN genómico de *Pseudomonas aeruginosa* (PAO1) y *Pseudomonas fluorescens* (PF5) mediante PCR.

SEQ ID NO:23 - PAO1-C-F1

10 SEQ ID NO:24 - PAO1-C-R1

SEQ ID NO:25 - PF5-C-F1

SEQ ID NO:26 - PF5-C-R1

Las SEQ ID NO:21, 22, 27 y 28 son las secuencias de nucleótidos de los cebadores oligonucleotídicos utilizados para validar las secuencias de ADN de los clones positivos de los genes *ilvC* de *Pseudomonas*.

15 SEQ ID NO:27 - PF5-S-F2

SEQ ID NO:28 - PF5-S-R2

Las siguientes SEQ ID NO: se corresponden con las secuencias de ADN de los genes KARI utilizados en esta invención:

SEQ ID NO:29 - *E. coli* K12 - *ilvC*

20 SEQ ID NO:30 - KARI con codones optimizados de *Vibrio* para la expresión en *E. coli*

SEQ ID NO:31 - *Pseudomonas aeruginosa* - PAO1 - *ilvC*

SEQ ID NO:32 - *Pseudomonas fluorescens* - PF5 - *ilvC*

Las siguientes SEQ ID NO: son las secuencias de aminoácidos que corresponden con SEQ ID NO:29-32, respectivamente:

25 SEQ ID NO:33 - *E. coli* K12 - *ilvC* - [KARI de *E. coli* K12]

SEQ ID NO:34 - KARI de *Vibrio cholerae*

SEQ ID NO:35 - PAO1-*ilvC* (1aa-338aa)

SEQ ID NO:36 - PF5-*ilvC* (1aa-338aa)

La SEQ ID NO:37 es el cebador directo PAL-F1.

30 La SEQ ID NO:38 es el cebador inverso (PAL-R1).

La SEQ ID NO:39 es el cebador directo (PAL-EcoR1-F1).

La SEQ ID NO:40 es el cebador inverso (PAL-EcoR1-R1).

Descripción detallada de la invención

35 La presente invención se refiere a un método para la conversión de acetolactato a 2,3-dihidroisovalerato utilizando células hospedantes microbianas que contienen una enzima KARI muy activa. El 2,3-dihidroisovalerato formado de esta manera después se convierte, a través de las etapas mostradas en la figura 1, en isobutanol. La invención también describe métodos para descubrir enzimas KARI más rápidas en sus microorganismos hospedantes naturales y la evolución molecular de estas enzimas para mejorar aún más su actividad catalítica.

40 La presente invención satisface una serie de necesidades comerciales e industriales. El butanol es una importante materia prima química industrial con una diversidad de aplicaciones, siendo su potencial como combustible o aditivo de combustibles particularmente significativo. Aunque es un alcohol de solo cuatro átomos de carbono, el butanol tiene un contenido energético similar al de la gasolina y puede mezclarse con cualquier combustible fósil. El butanol

se prefiere como combustible o aditivo de combustibles porque sólo produce CO₂ y pocos o ningún SO_x o NO_x cuando se quema en un motor de combustión interna convencional. Además, el butanol es menos corrosivo que el etanol, el aditivo para combustibles más preferido hasta la fecha.

5 Además de su utilidad como biocombustible o aditivo de combustibles, el butanol tiene el potencial de afectar a los problemas de distribución del hidrógeno en la emergente industria de las células de combustible. A día de hoy, las células de combustible están plagadas de problemas de seguridad asociados con el transporte y la distribución del hidrógeno. El butanol puede reformarse con facilidad con respecto a su contenido en hidrógeno, y puede distribuirse a través de las gasolineras existentes con la pureza requerida para las células de combustible o los vehículos.

10 Las siguientes definiciones y abreviaturas deben emplearse para la interpretación de las reivindicaciones y la memoria descriptiva.

El término "invención" o la expresión "presente invención", tal como se emplean en la presente, se aplican en general a todas las realizaciones de la invención según se define en las reivindicaciones según están presentadas o tal como se modificarán y suplementarán posteriormente, o en la memoria descriptiva.

15 La expresión "vía biosintética del isobutanol" se refiere a la vía enzimática para producir isobutanol. Las vías biosintéticas del isobutanol preferidas se ilustran en la figura 1 y se describen en la presente.

La expresión "ensayo de consumo de NADPH" se refiere a un ensayo enzimático para la determinación de la actividad específica de la enzima KARI, que implica medir la desaparición del cofactor de KARI, NADPH, de la reacción enzimática, según se describe en Aulabaugh *et ál.*, Biochemistry, 29, 2824-2830, 1990).

"KARI" es la abreviatura de la enzima cetol-ácido reductoisomerasa.

20 Las expresiones "acetohidroxiácido isomeroreductasa" y "cetol-ácido reductoisomerasa" se utilizarán de modo intercambiable y se refieren a la enzima que tiene el número EC, EC 1.1.1.86 (Enzyme Nomenclature 1992, Academic Press, San Diego. La cetol-ácido reductoisomerasa cataliza la reacción de (S)-acetolactato a 2,3-dihidroxiisovalerato, tal como se describe con más detalle a continuación. Estas enzimas están disponibles en una serie de fuentes que incluyen, pero no se limitan a *E. coli* n.º de registro de GenBank NC_000913 REGIÓN: 3955993.3957468, *Vibrio cholerae* n.º de registro de GenBank NC_002505 REGIÓN: 157441..158925, *Pseudomonas aeruginosa* n.º de registro de GenBank NC_002516 REGIÓN: 5272455..5273471, y *Pseudomonas fluorescens* n.º de registro de GenBank C_004129 REGIÓN: 6017379..6018395.

25

La expresión "acetolactato sintasa" se refiere a una enzima que cataliza la conversión de piruvato a acetolactato y CO₂. El acetolactato tiene dos estereoisómeros, (R) y (S); la enzima prefiere el (S)-isómero, que es el producido por los sistemas biológicos. Las acetolactato sintasas preferidas se conocen con el número EC 2.2.1.6 9 (Enzyme Nomenclature 1992, Academic Press, San Diego). Estas enzimas están disponibles en una serie de fuentes que incluyen, pero no se limitan a *Bacillus subtilis* (GenBank n.ºs: CAB15618, Z99122, NCBI (National Center for Biotechnology Information), secuencia de aminoácido, secuencias de nucleótidos NCBI, respectivamente), *Klebsiella pneumoniae* (GenBank n.ºs: AAA25079 (SEQ ID NO:2), M73842 (SEQ ID NO:1)), y *Lactococcus lactis* (GenBank n.ºs: AAA25161, L16975).

30

35

La expresión "acetohidroxiácido deshidratasa" se refiere a una enzima que cataliza la conversión de 2,3-dihidroxiisovalerato a α -cetoisovalerato. Las acetohidroxiácido deshidratasas preferidas se conocen con el número EC 4.2.1.9. Estas enzimas están disponibles en una amplia gama de microorganismos que incluyen, pero no se limitan a *E. coli* (GenBank n.ºs: YP_026248 (SEQ ID NO:6), NC_000913 (SEQ ID NO:5)), *S. cerevisiae* (GenBank n.ºs: NP_012550, NC_001142), *M. maripaludis* (GenBank n.ºs: CAF29874, BX957219), y *B. subtilis* (GenBank n.ºs: CAB14105, Z99115).

40

La expresión " α -cetoácido de cadena ramificada descarboxilasa" se refiere a una enzima que cataliza la conversión de α -cetoisovalerato a isobutiraldehído y CO₂. Las α -cetoácido de cadena ramificada descarboxilasas preferidas se conocen con el número EC 4.1.1.72, y están disponibles en una serie de fuentes que incluyen, pero no se limitan a *Lactococcus lactis* (GenBank n.ºs: AAS49166, AY548760, CAG34226 (SEQ ID NO:8), AJ746364), *Salmonella typhimurium* (GenBank n.ºs: NP_461346, C_003197), y *Clostridium acetobutylicum* (GenBank n.ºs: NP_149189, NC_001988).

45

La expresión "alcohol de cadena ramificada deshidrogenasa" se refiere a una enzima que cataliza la conversión de isobutiraldehído a isobutanol. Las alcohol de cadena ramificada deshidrogenasas preferidas se conocen con el número EC 1.1.1.265, pero también pueden clasificarse bajo otras alcohol deshidrogenasas (de modo específico, EC 1.1.1.1 y EC 1.1.1.2). Estas enzimas utilizand NADH (nicotinamida adenina dinucleótido reducido) y/o NADPH como donante de electrones, y están disponibles en una serie de fuentes que incluyen, pero no se limitan a *S. cerevisiae* (GenBank n.ºs: NP_010656, NC_001136; NP_014051, NC_001145), *E. coli* (GenBank n.ºs: NP_417484 (SEQ ID NO:10), NC_000913 (SEQ ID NO:9)), y *C. acetobutylicum* (GenBank n.ºs: NP_349892, NC_003030).

50

55 La expresión "cetoácido de cadena ramificada deshidrogenasa" se refiere a una enzima que cataliza la conversión de α -cetoisovalerato a isobutiril-CoA (isobutiril-coenzima A), empleando NAD⁺ (nicotinamida adenina dinucleótido)

como aceptor de electrones. Las cetoácido de cadena ramificada deshidrogenasas preferidas se conocen con el número EC 1.2.4.4. Estas cetoácido de cadena ramificada deshidrogenasas están formadas por cuatro subunidades, y las secuencias de todas las subunidades están disponibles en una amplia gama de microorganismos que incluyen, pero no se limitan a *B. subtilis* (GenBank n.^{os}: CAB14336, Z99116; CAB14335, Z99116; CAB14334, Z99116; y CAB14337, Z99116), y *Pseudomonas putida* (GenBank n.^{os}: AAA65614, M57613; AAA65615, M57613; AAA65617, M57613; y AAA65618, M57613).

La expresión "sustrato de carbono" o "sustrato de carbono fermentable" se refiere a una fuente de carbono capaz de ser metabolizada por los organismos hospedantes de la presente invención y, en particular, a fuentes de carbono seleccionadas del grupo que consiste en monosacáridos, oligosacáridos, polisacáridos y sustratos de un solo carbono o sus mezclas.

Los términos " k_{cat} " y " K_m " son conocidos por los expertos en la técnica, y se describen en *Enzyme Structure and Mechanism*, 2^a ed. (Ferst, W. H. Freeman, NY, 1985, pp. 98-120). El término " k_{cat} ", a menudo denominado "número de recambio", se define como el número máximo de moléculas de sustrato convertidas en producto por sitio activo por unidad de tiempo, o el número de veces que la enzima sufre un recambio por unidad de tiempo. $k_{cat} = V_{max}/[E]$, siendo $[E]$ la concentración de enzima (Ferst, *supra*). Las expresiones "recambio total" y "número de recambio total" se emplean en la presente para indicar la cantidad de producto formado mediante la reacción de una enzima KARI con el sustrato.

La expresión "eficacia catalítica" se define como la k_{cat}/K_M de una enzima. La "eficacia catalítica" se emplea para cuantificar la especificidad de una enzima por un sustrato.

La expresión "actividad específica" significa las unidades de enzima/mg de proteína, definiéndose la unidad de enzima como los moles de producto formado/minuto bajo condiciones especificadas de temperatura, pH, $[S]$, etc.

Los términos "lento" o "rápido", cuando se emplean con respecto a una actividad enzimática, se refieren al número de recambio de la enzima comparado con un patrón.

Las expresiones "molécula de ácido nucleico aislada", "fragmento de ácido nucleico aislado" y "construcción genética" se emplean de modo intercambiable y significan un polímero de ARN o ADN que es monocatenario o bicatenario, que contiene opcionalmente bases nucleotídicas sintéticas, no naturales o alteradas. Un fragmento de ácido nucleico aislado en forma de un polímero de ADN puede estar formado por uno o más segmentos de ADNc, ADN genómico o ADN sintético.

El término "gen" se refiere a un fragmento de ácido nucleico que es capaz de ser expresado como una proteína específica, que incluye opcionalmente secuencias reguladoras anteriores (secuencias no codificadoras 5') y posteriores (secuencias no codificadoras 3') a la secuencia codificadora. Un "gen nativo" se refiere a un gen según se encuentra en la naturaleza con sus propias secuencias reguladoras. Un "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. Por consiguiente, un gen quimérico puede comprender secuencias reguladoras y secuencias codificadoras que se derivan de fuentes diferentes, o secuencias reguladoras y secuencias codificadoras derivadas de la misma fuente pero dispuestas de una manera diferente a la que se encuentra en la naturaleza. Un "gen endógeno" se refiere a un gen nativo en su localización natural en el genoma de un organismo. Un gen "extraño" se refiere a un gen que no se encuentra normalmente en el organismo hospedante, pero que se introduce en el organismo hospedante mediante transferencia de genes. Los genes extraños pueden comprender genes nativos insertados en un organismo no nativo, o genes quiméricos. Un "transgén" es un gen que se ha introducido en el genoma mediante un procedimiento de transformación.

Tal como se emplea en la presente, la expresión "secuencia codificadora" se refiere a una secuencia de ADN que codifica una secuencia de aminoácidos específica. Las "secuencias reguladoras adecuadas" se refieren a secuencias de nucleótidos localizadas cadena arriba (secuencias no codificadoras 5'), dentro, o cadena abajo (secuencias no codificadoras 3') de una secuencia codificadora, y que influyen en la transcripción, el procesamiento o la estabilidad del ARN, o la traducción de la secuencia codificadora asociada. Las secuencias reguladoras pueden incluir promotores, secuencias conductoras de la traducción, intrones, secuencias de reconocimiento de la poliadenilación, sitios de procesamiento del ARN, sitios de unión de efectores y estructuras de tallo-bucle.

El término "promotor" se refiere a una secuencia de ADN capaz de controlar la expresión de una secuencia codificadora o ARN funcional. En general, una secuencia codificadora se localiza 3' a la secuencia del promotor. Los promotores pueden derivarse en su totalidad de un gen nativo, o estar compuestos por diferentes elementos derivados de diferentes promotores que se encuentran en la naturaleza, o incluso comprender segmentos de ADN sintético. Los expertos en la técnica entienden que diferentes promotores pueden dirigir la expresión de un gen en diferentes tejidos o tipos celulares, o en diferentes etapas del desarrollo, o en respuesta a diferentes condiciones ambientales o fisiológicas. Los promotores que provocan que un gen se exprese en la mayoría de los tipos celulares la mayor parte del tiempo se denominan habitualmente "promotores constitutivos". También se reconocen que, puesto que en la mayoría de los casos, no se han definido por completo los límites exactos de las secuencias reguladoras, unos fragmentos de ADN de diferente longitud pueden tener idéntica actividad promotora.

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

El término "expresión", tal como se emplea en la presente, se refiere a la transcripción y la acumulación estable de ARN sentido (ARNm) o antisentido derivado del fragmento de ácido nucleico de la invención. La expresión también puede referirse a la traducción de ARNm en un polipéptido.

Tal como se emplea en la presente, el término "transformación" se refiere a la transferencia de un fragmento de ácido nucleico hacia el genoma de un organismo hospedante, dando como resultado una herencia genéticamente estable. Los organismos hospedantes que contienen los fragmentos de ácidos nucleicos transformados se denominan organismos "transgénicos", "recombinantes" o "transformados".

Los términos "plásmido", "vector" y "módulo" se refieren a un elemento extracromosómico que a menudo porta genes que no son parte del metabolismo central de la célula, y normalmente está en forma de un fragmento de ADN bicatenario circular. Estos elementos pueden ser secuencias de replicación autónoma, secuencias de integración en el genoma, fagos o secuencias de nucleótidos, lineales o circulares, de un ARN o ADN monocatenario o bicatenario, derivadas de cualquier fuente, en las que una serie de secuencias de nucleótidos se han unido o recombinado en una construcción exclusiva que es capaz de introducir un fragmento de promotor y una secuencia de ADN para un producto génico seleccionado, junto con secuencias no traducidas 3' apropiadas, en una célula. Un "módulo de transformación" se refiere a un vector específico que contiene un gen extraño y que tiene elementos, además del gen extraño, que facilitan la transformación de una célula hospedante concreta. Un "módulo de expresión" se refiere a un vector específico que contiene un gen extraño y que tiene elementos, además del gen extraño, que permiten una expresión potenciada de este gen en un hospedante extraño.

Tal como se emplea en la presente, la expresión "degeneración de codones" se refiere a la naturaleza del código genético que permite variaciones en la secuencia de nucleótidos que no afectan a la secuencia de aminoácidos del polipéptido codificado. Los expertos en la técnica conocerán el "sesgo de codones" que muestra una célula hospedante específica en cuanto a la utilización de codones de nucleótidos para especificar un aminoácido concreto. Por tanto, cuando se sintetiza un gen para una mayor expresión en una célula hospedante, resulta deseable diseñar el gen de modo que su frecuencia de utilización de codones se parezca a la frecuencia de utilización de codones preferida de la célula hospedante.

La expresión "con codones optimizados", con respecto a los genes o las regiones codificadoras de moléculas de ácidos nucleicos para la transformación de diversos hospedantes, se refiere a la alteración de codones en el gen o las regiones codificadoras de las moléculas de ácidos nucleicos para que reflejen la utilización de codones típica del organismo hospedante sin alterar el polipéptido codificado por el ADN.

Las técnicas de ADN recombinante y de clonación molecular convencionales utilizadas en la presente son muy conocidas en la técnica y se describen en Sambrook *et ál.* (Sambrook, Fritsch, y Maniatis, Molecular Cloning: A Laboratory Manual, 2ª edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989) (en lo sucesivo "Maniatis"); y en Silhavy *et ál.* (Silhavy, *et ál.*, Experiments with Gene Fusions, Cold Spring Harbor Laboratory Press Cold Spring Harbor, NY, 1984); y en Ausubel, F. M. *et ál.*, (Ausubel, *et ál.*, Current Protocols in Molecular Biology, publicado por Greene Publishing Assoc. and Wiley-Interscience, 1987).

La presente invención produce isobutanol a partir de fuentes de carbono derivadas de plantas, evitando el impacto ambiental negativo asociado con los procesos petroquímicos convencionales para la producción de butanol. En una realización, la presente invención proporciona un método para la identificación de enzimas KARI con alta eficacia catalítica que podrían eliminarlo como etapa limitante de la velocidad en la conversión de carbohidratos en isobutanol. El método comprende la identificación y las enzimas KARI adecuadas y su mutagénesis utilizando métodos muy conocidos en la técnica, para aumentar la actividad específica de la enzima, tal como se describe en detalle a continuación.

Enzimas cetoácido reductoisomerasas (KARI)

La acetohidroxiácido isomeroreductasa o cetol-ácido reductoisomerasa (KARI, EC 1.1.1.86) cataliza dos etapas de la biosíntesis de aminoácidos de cadena ramificada, y es una enzima clave en su biosíntesis. KARI se encuentra en una diversidad de organismos, y la comparación de secuencias de aminoácidos a través de especies ha revelado que existen dos tipos de esta enzima: una forma corta (clase I) que se encuentra en hongos y la mayoría de las bacterias, y una forma larga (clase II) típica de plantas.

La forma corta de KARI generalmente tiene entre 330-340 restos aminoácidos. La forma larga de KARI tiene aproximadamente 490 restos aminoácidos. Sin embargo, algunas bacterias, tales como *Escherichia coli*, poseen una forma larga, en la que la secuencia de aminoácidos se diferencia de forma apreciable de la que se encuentra en plantas. KARI es codificada por el gen *ilvC* y es una enzima fundamental para el crecimiento de *E. coli* y otras

bacterias en un medio mínimo. KARI utiliza NADPH como cofactor y requiere un catión divalente, tal como Mg^{++} , para su actividad. Además de la utilización de acetolactato en la vía de la valina, KARI también convierte acetohidroxibutanoato en dihidroximetilpentanoato en la vía de producción de isoleucina.

5 Se ha resuelto la estructura cristalina de la enzima KARI de *E. coli* a una resolución de 2,6 Å (Tyagi, *et ál.*, Protein Science, 14, 3089-3100, 2005). Esta enzima está formada por dos dominios, uno con una estructura α/β mixta, que es similar al que se encuentra en otras deshidrogenasas dependientes de nucleótidos de piridina. El segundo dominio es fundamentalmente una α -hélice y muestra bastantes pruebas de duplicación interna. La comparación de los sitios activos de KARI de *E. coli*, *Pseudomonas aeruginosa* y la espinaca demuestra que la mayoría de los restos en el sitio activo de la enzima ocupan posiciones conservadas. Mientras que KARI de *E. coli* cristaliza como un tetrámero, que posiblemente sea la probable unidad biológicamente activa, la KARI de *P. aeruginosa* (Ahn, *et ál.*, J. Mol. Biol., 328, 505-515, 2003) forma un dodecámero, y la enzima de la espinaca forma un dímero. Las KARI conocidas son enzimas lentas con el número de recambio (k_{cat}) indicado de 2 s^{-1} (Aulabaugh *et ál.*, Biochemistry, 29, 2824-2830, 1990) o $0,12\text{ s}^{-1}$ (Rane *et ál.*, Arch. Biochem. Biophys., 338, 83-89, 1997) para el acetolactato. Los estudios han demostrado que el control genético de la biosíntesis de isoleucina-valina en *E. coli* es diferente del que se encuentra en *Ps. aeruginosa* (Marinus, *et ál.*, Genetics, 63, 547-556, 1969).

Identificación e aislamiento de enzimas KARI de alta actividad

Se realizó un examen de organismos con unas velocidades de duplicación mayores que *E. coli*. Se identificaron tres microorganismos, *Pseudomonas aeruginosa* (PAO1), *Pseudomonas fluorescens* (PF5), y *Vibrio cholerae* (N16961), con unos tiempos de duplicación más rápidos que *E. coli* cuando se cultivan en medio mínimo M9. Los genes que codifican una enzima KARI se aislaron de cada una de estas variedades, y las proteínas codificadas se expresaron y se purificaron parcialmente. La actividad específica de las enzimas aisladas de los organismos con alta velocidad de duplicación se compararon frente a la de KARI de *E. coli* empleando el método de ensayo del consumo de NADPH que mide la desaparición del cofactor, NADPH, durante la conversión enzimática de acetolactato a α,β -dihidroxiisovalerato a 340 nm. La actividad se calcula utilizando el coeficiente de extinción molar de $6220\text{ M}^{-1}\text{ cm}^{-1}$ para NADPH y se indica como μmol de NADPH consumido por min por mg de proteínas totales en extractos celulares (véase Aulabaugh y Schloss, Biochemistry, 29, 2824-2830, 1990).

Un objeto de la presente invención es proporcionar una enzima KARI que tenga una actividad específica mayor que $1,1\text{ }\mu\text{mol}/\text{min}/\text{mg}$ de KARI, medida utilizando la proteína purificada según el ensayo de consumo de NADPH descrito en la presente. La KARI de *E. coli* es una enzima lenta, y resulta fundamental para la vía sintética de aminoácidos de cadena ramificada. El gen que codifica KARI (*ilvC*) está desactivado cuando las células crecen en un medio rico, pero se expresa a niveles elevados (aproximadamente 10% de las proteínas solubles) cuando se cultivan en un medio mínimo (S. Epelbaum *et ál.*, *supra*).

El proceso de selección de una enzima KARI adecuada implica dos estrategias. La primera consiste en buscar una nueva KARI entre la diversidad natural. Esta búsqueda implica aislar homólogos de las enzimas disponibles en general de otros organismos, utilizando técnicas muy conocidas en la técnica. Esta búsqueda se fundamenta en una hipótesis acerca de los organismos que más probabilidad tienen las KARI adecuadas, basada en el tiempo de duplicación del organismo. Una segunda estrategia implica crear y buscar una diversidad artificial mediante la construcción de un vector de expresión fuerte, una mutagénesis y evolución de la secuencia codificadora de KARI y, por último, una selección de los variantes con mejor actividad KARI.

40 Utilizando los anteriores métodos, se aislaron enzimas KARI de *Pseudomonas fluorescens* (SEQ ID NO:35 [*PAO1-ilvC*], SEQ ID NO:36 [*PF5-ilvC*]) y *Vibrio cholerae* (SEQ ID NO:34) que tienen unas actividades específicas que son mayores que las de la enzima KARI aislada a partir de *E. coli* (SEQ ID NO:33 [*E. coli* K12 - *ilvC*]). En la presente invención se prefieren las enzimas KARI con unas actividades específicas mayores que aproximadamente $1,1\text{ }\mu\text{mol}/\text{min}/\text{mg}$, siendo unas actividades específicas de aproximadamente $5\text{-}40\text{ }\mu\text{mol}/\text{min}/\text{mg}$ particularmente adecuadas. Resulta preferible que la actividad específica de la KARI se mida utilizando proteínas purificadas e incorporando el ensayo de consumo de NADPH (Aulabaugh, *supra*), ejecutado entre $20\text{ }^{\circ}\text{C}$ y $25\text{ }^{\circ}\text{C}$, siendo preferido aproximadamente $22,5\text{ }^{\circ}\text{C}$, a un pH entre 7,0 y 8,0, siendo preferido un pH de aproximadamente 7,5, y en un tampón que tenga potasio al menos 10 mM, siendo adecuado al menos aproximadamente 10 mM a aproximadamente 50 mM. Algunas de las enzimas específicas útiles en la invención se listan a continuación en la tabla 2.

50

Tabla 2

Enzimas KARI de la presente invención	
Gen	Índice de GenBank
<i>E. coli</i> K12 <i>ilvC</i>	n.º de registro de GenBank NC_000913 REGION: 3955993..3957468
Con codones optimizados para la expresión en <i>E. coli</i> de KARI de <i>Vibrio cholerae</i>	n.º de registro de GenBank NC_002505 REGIÓN: 157441..158925

Enzimas KARI de la presente invención	
Gen	Índice de GenBank
<i>Pseudomonas aeruginosa</i> PA01 <i>ilvC</i>	n.º de registro de GenBank NC_002516 REGIÓN: 5272455..5273471
<i>Pseudomonas fluorescens</i> PF5 <i>ilvC</i>	n.º de registro de GenBank NC_004129 REGIÓN: 6017379..6018395

La presente invención no se limita a las enzimas de *Pseudomonas* y *Vibrio* específicas descritas en la presente. Por ejemplo, estos polipéptidos pueden utilizarse como base para descubrir homólogos que tengan una actividad similar, o como moldes para la mutagénesis y evolución de proteínas.

Aislamiento de homólogos de KARI

5 El fragmento de ácido nucleico de la presente invención puede utilizarse para aislar genes que codifican proteínas homólogas a partir de la misma especie microbiana o de otra distinta. El aislamiento de genes KARI homólogos empleando protocolos dependientes de la secuencia es muy conocido en la técnica. Los ejemplos de protocolos dependientes de la secuencia incluyen, pero no se limitan a métodos de hibridación de ácidos nucleicos, y métodos de amplificación de ADN y ARN, según se ejemplifica en diversos usos de la tecnología de la amplificación de ácidos nucleicos, por ejemplo, la reacción en cadena de la polimerasa (PCR) (Mullis *et ál.*, patente de EEUU n.º 4.683.202), la reacción en cadena de la ligasa (LCR) (Tabor, *et ál.*, Proc. Acad. Sci. USA, 82, 1074, 1985) o la amplificación de desplazamiento de hebra (SDA) (Walker, *et ál.*, Proc. Natl. Acad. Sci. USA., 89, 392, 1992).

15 Por ejemplo, los genes que codifican proteínas o polipéptidos similares a los de la presente invención pueden aislarse directamente utilizando todo el fragmento, o una porción, del presente ácido nucleico como sonda de hibridación de ADN para seleccionar bancos de cualquier bacteria deseada utilizando metodologías muy conocidas por los expertos en la técnica. Pueden diseñarse sondas oligonucleotídicas específicas basadas en la presente secuencia de ácido nucleico y sintetizarse mediante métodos conocidos en la técnica (Maniatis, *supra*). Además, la secuencia completa puede utilizarse directamente para sintetizar sondas de ADN mediante métodos conocidos por los expertos en la técnica, tales como marcaje de ADN de cebadores aleatorio, traducción de desplazamiento de mella, o técnicas de marcaje terminal, o sondas de ARN que empleen sistemas de transcripción *in vitro* disponibles. Además, pueden diseñarse cebadores específicos y emplearse para amplificar una parte o la longitud completa de la presente secuencia. Los productos de la amplificación resultantes pueden marcarse directamente durante las reacciones de amplificación o marcarse después de las reacciones de amplificación, y emplearse como sondas para aislar fragmentos de ADN de longitud completa bajo condiciones de rigurosidad apropiadas.

25 Generalmente, en las técnicas de amplificación de tipo PCR, los cebadores tienen diferentes secuencias y no son complementarios entre sí. Dependiendo de las condiciones de ensayo deseadas, las secuencias de los cebadores deben diseñarse para proporcionar una replicación eficaz y fiel del ácido nucleico diana. Los métodos de diseño de cebadores de PCR son habituales y muy conocidos en la técnica, por ejemplo, Thein *et ál.* (Thein, *et ál.*, "The use of oligonucleotide as specific hybridization probes in the Diagnosis of Genetic Disorders", en Human Genetic Diseases: A Practical Approach, K. E. Davis ed., 1986, pp. 33-50, IRL Press, Hemdon, Virginia); y Rychlik (Rychlik, 1993, en White, B. A. (ed.), Methods in Molecular Biology, vol. 15, pp. 31-39, PCR Protocols: Current Methods and Applications, Humana Press, Inc., Totowa, NJ).

35 En general, pueden utilizarse dos segmentos cortos de la presente secuencia en los protocolos de la reacción en cadena de la polimerasa para amplificar fragmentos de ácidos nucleicos más largos que codifican genes homólogos a partir del ADN o ARN. La reacción en cadena de la polimerasa también puede realizarse con un banco de fragmentos de ácidos nucleicos clonados, en los que la secuencia de un cebador se deriva del presente fragmento de ácido nucleico, y la secuencia del otro cebador aprovecha la presencia de los tramos de poli(ácido adenílico) hacia el extremo 3' del ARNm precursor que codifica genes microbianos.

40 Como alternativa, la segunda secuencia del cebador puede basarse en secuencias derivadas del vector de clonación. Por ejemplo, los expertos en la técnica pueden seguir el protocolo RACE (Frohman *et ál.*, Proc. Natl. Acad. Sci. USA, 85, 8998, 1988) para generar ADNc utilizando la PCR para amplificar copias de la región entre un punto individual de la transcripción y el extremo 3' o 5'. Pueden diseñarse cebadores orientados en las direcciones 3' y 5' a partir de la presente secuencia. Utilizando sistemas RACE 3' o RACE 5' disponibles en el mercado (Life Technologies, Rockville, MD), pueden aislarse fragmentos de ADNc 3' o 5' específicos (Ohara *et ál.*, Proc. Natl. Acad. Sci. USA, 86, 5673, 1989; y Loh *et ál.*, Science, 243, 217-220, 1989).

45 Como alternativa, la presente secuencia puede emplearse como reactivo de hibridación para la identificación de homólogos. Los componentes básicos de un ensayo de hibridación de ácidos nucleicos incluyen una sonda, una muestra de la cual se sospecha que contiene el gen o fragmento del gen de interés, y un método de hibridación específico. Las sondas de la presente invención generalmente son secuencias de ácidos nucleicos monocatenarios que son complementarias con las secuencias de los ácidos nucleicos que se van a detectar. Las sondas son

"hibridables" con la secuencia del ácido nucleico que se va a detectar. La longitud de la sonda puede variar desde 5 bases a decenas de miles de bases, y dependerá del ensayo específico que se va a realizar. Generalmente, resulta adecuada una longitud de la sonda de aproximadamente 15 bases a aproximadamente 30 bases. Solo es necesario que una parte de la molécula de la sonda sea complementaria con la secuencia del ácido nucleico que se va a detectar. Además, no es necesario que la complementariedad entre la sonda y la secuencia diana sea perfecta. La hibridación se produce entre moléculas imperfectamente complementarias, con el resultado de que una determinada fracción de las bases en la región hibridada no están apareadas con la base complementaria adecuada.

Los métodos de hibridación están bien definidos. Generalmente, la sonda y la muestra deben mezclarse bajo condiciones que permitan la hibridación de los ácidos nucleicos. Esto implica poner en contacto la sonda y la muestra en presencia de una sal inorgánica u orgánica bajo las condiciones de concentración y temperatura adecuadas. Los ácidos nucleicos de la muestra y la sonda deben ponerse en contacto durante un tiempo suficiente para que se produzca cualquier hibridación posible entre el ácido nucleico de la muestra y la sonda. La concentración de la sonda o diana en la mezcla determinará el tiempo necesario para que se produzca la hibridación. Cuanta mayor sea la concentración de la sonda o diana, más corto será el tiempo necesario de incubación de la hibridación. Opcionalmente, puede añadirse un agente caotrópico. El agente caotrópico estabiliza los ácidos nucleicos inhibiendo la actividad nucleasa. Además, el agente caotrópico permite una hibridación sensible y rigurosa de sondas oligonucleotídicas cortas a temperatura ambiente (Van Ness *et al.*, Nucl. Acids Res., 19, 5143-5151, 1991). Los agentes caotrópicos adecuados incluyen cloruro de guanidinio, tiocianato de guanidinio, tiocianato de sodio, tetracloroacetato de litio, perclorato de sodio, tetracloroacetato de rubidio, yoduro de potasio, y trifluoroacetato de cesio, entre otros. Generalmente, el agente caotrópico estará presente a una concentración final de aproximadamente 3 M. Si se desea, se puede añadir formamida a la mezcla de hibridación, generalmente al 30-50% (en v/v).

Pueden emplearse diversas disoluciones de hibridación. Generalmente, estas comprenden de aproximadamente 20% a 60% en volumen, preferiblemente 30% de un disolvente orgánico polar. Una disolución de hibridación habitual emplea aproximadamente 30-50% en v/v de formamida, cloruro de sodio de aproximadamente 0,15 a 1,0 M, tampones de aproximadamente 0,05 a 0,1 M, tales como citrato de sodio, Tris-HCl, PIPES o HEPES (pH en el intervalo de aproximadamente 6-9), detergente de aproximadamente 0,05% al 0,2%, tal como dodecilsulfato de sodio, o EDTA entre 0,5-20 mM, FICOLL (Pharmacia Inc.) (aproximadamente 300-500 kiloDaltons), polivinilpirrolidona (aproximadamente 250-500 kiloDaltons), y albúmina de suero. En la disolución de hibridación típica también se incluyen ácidos nucleicos vehículo sin marcar, de aproximadamente 0,1 a 5 mg/ml, ADN nucleico fragmentado, por ejemplo, ADN de timo de ternera o de esperma de salmón, o ARN de levadura, y opcionalmente de aproximadamente 0,5% al 2% en p/v de glicina. También pueden incluirse otros aditivos, tales como agentes de exclusión de volumen, que incluyen una diversidad de agentes que se hinchan en agua o hidrosolubles polares, tales como polietilenglicol, polímeros aniónicos, tales como poliácido acrílico o polimetacrilato, y polímeros sacarídicos aniónicos, tales como sulfato de dextrano.

La hibridación de ácidos nucleicos puede adaptarse a una diversidad de formatos de ensayo. Uno de los más adecuados es el formato de ensayo de "sandwich". El ensayo de "sandwich" puede adaptarse particularmente a la hibridación bajo condiciones no desnaturizantes. Un componente principal de un ensayo de tipo "sandwich" es un soporte sólido. El soporte sólido tiene adsorbido sobre él o unido covalentemente a él una sonda de ácido nucleico inmovilizada que no está marcada y es complementaria con una porción de la secuencia.

Vías biosintéticas del isobutanol

Uno de los principales usos de las presentes enzimas KARI de alta actividad será como elemento en las vías metabólicas útiles para la producción de isobutanol. Se ha aclarado y caracterizado una serie de estas vías.

Los microorganismos que utilizan carbohidratos emplean la vía de Embden-Meyerhof-Parnas (EMP), la vía de Entner y Doudoroff, y el ciclo de pentosa fosfato como las principales vías metabólicas para proporcionar energía y precursores celulares para el crecimiento y el mantenimiento. Estas vías tienen en común el intermedio gliceraldehído-3-fosfato y, en último término, se forma piruvato directamente o en combinación con la vía EMP. Después, el piruvato se transforma en acetil-coenzima A (acetil-CoA) a través de una diversidad de medios. El acetil-coA actúa como intermedio clave, por ejemplo, para generar ácidos grasos, aminoácidos y metabolitos secundarios. Las reacciones combinadas de la conversión de azúcares en piruvato producen energía (por ejemplo, adenosina-5'-trifosfato, ATP) y equivalentes reductores (por ejemplo, nicotinamida adenina dinucleótido reducido, NADH, y nicotinamida adenina dinucleótido fosfato reducido, NADPH). El NADH y el NADPH deben reciclarse a sus formas oxidadas (NAD⁺ y NADP⁺, respectivamente). En presencia de aceptores de electrones inorgánicos (por ejemplo, O₂, NO₃ y SO₄²⁻), los equivalentes reductores pueden utilizarse para aumentar la agrupación de energía; como alternativa, puede formarse un subproducto de carbono reducido.

Existen cuatro vías potenciales para la producción de isobutanol a partir de fuentes de carbohidratos con microorganismos recombinantes, tal como se muestra en la figura 1. Todas las vías potenciales para la conversión de carbohidratos en isobutanol se han descrito en la solicitud de patente de EEUU de propiedad de los solicitantes n.º 11/586315.

La vía preferida para la conversión del piruvato en isobutanol consiste en las etapas enzimáticas "a", "b", "c", "d", y "e" (figura 1) e incluye las siguientes conversiones de sustrato a producto:

- a) piruvato a acetolactato, catalizado, por ejemplo, por acetolactato sintasa,
- b) (S)-acetolactato a 2,3-dihidroxiisovalerato, catalizado, por ejemplo, por acetohidroxiácido isomerorreductasa,
- 5 c) 2,3-dihidroxiisovalerato a α -cetoisovalerato, catalizado, por ejemplo, por acetohidroxiácido deshidratasa,
- d) α -cetoisovalerato a isobutiraldehído, catalizado, por ejemplo, por cetoácido de cadena ramificada descarboxilasa, y
- e) isobutiraldehído a isobutanol, catalizado, por ejemplo, por alcohol de cadena ramificada deshidrogenasa.

Esta vía combina enzimas implicadas en vías bien caracterizadas para la biosíntesis de valina (piruvato a α -cetoisovalerato) y el catabolismo de la valina (α -cetoisovalerato a isobutanol). Puesto que muchas enzimas biosintéticas de la valina también catalizan reacciones análogas en la vía biosintética de la isoleucina, la especificidad de sustrato es una consideración importante a la hora de elegir las fuentes de genes. Por esta razón, los principales genes de interés para la enzima acetolactato sintasa son los procedentes de *Bacillus (alsS)* y *Klebsiella (budB)*. Se sabe que estas acetolactato sintasas concretas participan en la fermentación del butandiol en estos organismos, y muestran mayor afinidad por el piruvato frente al cetobutirato (Gollop *et ál.*, J. Bacteriol., 172, 3444-3449, 1990; y Holtzclaw *et ál.*, J. Bacteriol., 121, 917-922, 1975). La segunda y tercera etapas de la vía son catalizadas por la acetohidroxiácido reductoisomerasa y deshidratasa, respectivamente. Estas enzimas se han caracterizado a partir de una serie de fuentes tales como, por ejemplo, *E. coli* (Chunduru *et ál.*, Biochemistry, 28, 486-493, 1989; y Flint *et ál.*, J. Biol. Chem., 268, 14732-14742, 1993). Se sabe que las últimas dos etapas de la vía del isobutanol preferida se producen en levaduras, que pueden utilizar la valina como fuente de nitrógeno y, en el proceso, segregar isobutanol. El α -cetoisovalerato puede convertirse en isobutiraldehído mediante una serie de enzimas cetoácido descarboxilasas tales como, por ejemplo, piruvato descarboxilasa. Para evitar que el piruvato se dirija hacia fuera de la producción de isobutanol, se desea una descarboxilasa con una menor afinidad por el piruvato. Hasta la fecha, existen dos enzimas de este tipo conocidas en la técnica (Smit *et ál.*, Appl. Environ. Microbiol., 71, 303-311, 2005; y de la Plaza *et ál.*, FEMS Microbiol. Lett., 238, 367-374, 2004). Ambas enzimas proceden de cepas de *Lactococcus lactis* y tienen una preferencia mayor en 50-200 veces por el cetoisovalerato frente al piruvato. Por último, se ha identificado una serie de aldehído reductasas en levaduras, muchas con especificidades de sustrato solapantes. Las que prefieren sustratos de cadena ramificada frente al acetaldehído incluyen, pero no se limitan a alcohol deshidrogenasa VI (ADH6) e Ypr1p (Larroy *et ál.*, Biochem. J., 361, 163-172, 2002; y Ford *et ál.*, Yeast, 19, 1087-1096, 2002), y ambas utilizan el NADPH como donante de electrones. También se ha identificado recientemente en *E. coli* una reductasa dependiente de NADPH, YqhD, que es activa con sustratos de cadena ramificada (Sulzenbacher *et ál.*, J. Mol. Biol., 342, 489-502, 2004).

Dos de las otras vías potenciales para la producción de isobutanol también contienen las tres etapas iniciales de "a", "b" y "c". Una vía consiste en las etapas enzimáticas "a", "b", "c", "f", "g", "e". La etapa "f" contiene una "cetoácido de cadena ramificada deshidrogenasa" con el número EC 1.2.4.4. La etapa "g" contiene una "aldehído acilante deshidrogenasa" con los números EC 1.2.1.10 y 1.2.1.57, además de la etapa "e", que contiene la "alcohol de cadena ramificada deshidrogenasa". Las otras vías potenciales consisten en las etapas "a", "b", "c", "h", "i", "j", "e". El término "transaminasa" (etapa "h") tiene los números EC 2.6.1.42 y 2.6.1.66. La etapa "h" consiste en una "valina deshidrogenasa" con los números EC 1.4.1.8 y 1.4.1.9, o la etapa "i", una "valina descarboxilasa" con el número EC 4.1.1.14. Por último, la etapa "j" emplea una "omega transaminasa" con el número EC 2.6.1.18 para generar isobutiraldehído, que será reducido por la etapa "e" para producir isobutanol. Todas las vías potenciales para la conversión del piruvato en isobutanol se representan en la figura 1.

Además, se sabe que una serie de organismos producen butirato y/o butanol a través de un intermedio de butiril-CoA (Dürre *et ál.*, FEMS Microbiol. Rev., 17, 251-262, 1995; y Abbad-Andaloussi *et ál.*, Microbiology, 142, 1149-1158, 1996). Por tanto, la producción de isobutanol en estos organismos tendrá lugar utilizando las etapas "k", "g" y "e", mostradas en la figura 1. La etapa "k" empleará una "isobutiril-CoA mutasa" con el número EC 5.4.99.13. La siguiente etapa implica el uso de "aldehído acilante deshidrogenasa" con los números EC 1.2.1.10 y 1.2.1.57 para producir isobutiraldehído, seguido de la etapa enzimática "e" para producir isobutanol. Todas estas vías se describen a fondo en la solicitud de patente de EEUU de propiedad de los solicitantes 11/586315.

50 Hospedantes microbianos para la producción de isobutanol

Los hospedantes microbianos para la producción de isobutanol pueden seleccionarse de bacterias, cianobacterias, hongos filamentosos y levaduras. El hospedante microbiano utilizado para la producción de isobutanol debe ser tolerante al isobutanol, de modo que el rendimiento no se vea limitado por la toxicidad del butanol. Los microbios que son metabólicamente activos a altos niveles de titulación de isobutanol no son muy conocidos en la técnica. Aunque se han aislado mutantes tolerantes al butanol a partir de *Clostridium* disolventegénicas, existe poca información disponible acerca de la tolerancia al butanol de otras cepas bacterianas potencialmente útiles. La mayoría de los estudios sobre la comparación de la tolerancia a alcoholes en bacterias sugieren que el butanol es más tóxico que el

etanol (de Cavalho *et ál.*, *Microsc. Res. Tech.*, 64, 215-222, 2004; y Kabelitz *et ál.*, *FEMS Microbiol. Lett.*, 220, 223-227, 2003). Tomas *et ál.*, *J. Bacteriol.*, 186, 2006-2018, 2004 indican que el rendimiento de 1-butanol durante la fermentación en *Clostridium acetobutylicum* puede verse limitado por la toxicidad del 1-butanol. El principal efecto del 1-butanol sobre *Clostridium acetobutylicum* es la alteración de las funciones de membrana (Hermann *et ál.*, *Appl. Environ. Microbiol.*, 50, 1238-1243, 1985).

Los hospedantes microbianos seleccionados para la producción de isobutanol deben ser tolerantes al isobutanol, y deben ser capaces de convertir los carbohidratos en isobutanol. Los criterios para la selección de hospedantes microbianos adecuados incluyen los siguientes: tolerancia intrínseca al isobutanol, alta tasa de utilización de la glucosa, disponibilidad de herramientas genéticas para la manipulación de genes, y capacidad para generar alteraciones cromosómicas estables.

Las cepas de hospedantes adecuadas con tolerancia al isobutanol pueden identificarse mediante una selección basada en la tolerancia intrínseca de la cepa. La tolerancia intrínseca de los microbios al isobutanol puede medirse determinando la concentración de isobutanol que es responsable del 50% de inhibición de la velocidad de crecimiento (CI_{50}) cuando se cultiva en un medio mínimo. Los valores de CI_{50} pueden determinarse utilizando métodos conocidos en la técnica. Por ejemplo, los microbios de interés pueden cultivarse en presencia de diversas cantidades de isobutanol, y controlarse la velocidad de crecimiento midiendo la densidad óptica a 600 nanómetros (DO_{600}). Puede calcularse el tiempo de duplicación a partir de la parte logarítmica de la curva de crecimiento, y emplearse como medición de la velocidad de crecimiento. Puede determinarse la concentración de isobutanol que produce 50% de inhibición del crecimiento a partir de una gráfica del porcentaje de inhibición del crecimiento frente a la concentración de isobutanol. Preferiblemente, la cepa hospedante debe tener una CI_{50} para el isobutanol mayor que aproximadamente 0,5%.

El hospedante microbiano para la producción de isobutanol también debe poder utilizar la glucosa a una alta tasa. La mayoría de los microbios son capaces de utilizar carbohidratos. Sin embargo, ciertos microbios ambientales no pueden utilizar carbohidratos con alta eficacia y, por tanto, no son hospedantes adecuados.

La capacidad para modificar genéticamente el hospedante es fundamental para la producción de cualquier microorganismo recombinante. El modo de la tecnología de transferencia de genes puede ser mediante electroporación, conjugación, transducción o transformación natural. Está disponible una amplia gama de plásmidos que se conjugan con hospedantes y marcadores de resistencia a fármacos. Los vectores de clonación se adaptan a los organismos hospedantes basándose en la naturaleza de los marcadores de resistencia a antibióticos que pueden funcionar en ese hospedante.

El hospedante microbiano también tiene que manipularse para inactivar las vías competitivas por el flujo de carbono mediante la delección de diversos genes. Esto requiere la disponibilidad de transposones para dirigir la inactivación o vectores de integración cromosómica. Además, debe poder hacerse una mutagénesis química en el hospedante productor, de modo que puedan obtenerse mutaciones para mejorar la tolerancia al isobutanol intrínseca.

Basándose en los criterios descritos anteriormente, los hospedantes microbianos adecuados para la producción de isobutanol incluyen, pero no se limitan a miembros de los géneros *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Vibrio*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula* y *Saccharomyces*. Los hospedantes preferidos incluyen *Escherichia coli*, *Alcaligenes eutrophus*, *Bacillus licheniformis*, *Paenibacillus macerans*, *Rhodococcus erythropolis*, *Pseudomonas putida*, *Lactobacillus plantarum*, *Enterococcus faecium*, *Enterococcus gallinarum*, *Enterococcus faecalis*, *Bacillus subtilis* y *Saccharomyces cerevisiae*.

Construcción del hospedante de producción

Los organismos recombinantes que contienen los genes necesarios que codifican la vía enzimática para la conversión de un sustrato de carbono fermentable en isobutanol pueden construirse utilizando técnicas muy conocidas en la técnica. En la presente invención, los genes que codifican las enzimas de una de las vías biosintéticas del isobutanol de la invención, por ejemplo, acetolactato sintasa, acetohidroxiácido isomeroreductasa, acetohidroxiácido deshidratasa, α -cetoácido de cadena ramificada descarboxilasa, y alcohol de cadena ramificada deshidrogenasa, pueden aislarse a partir de diversas fuentes, tal como se describió anteriormente.

Los métodos para obtener genes deseados a partir de un genoma bacteriano son habituales y muy conocidos en la técnica de la biología molecular. Por ejemplo, si se conoce la secuencia del gen, pueden crearse bancos genómicos adecuados mediante digestión con endonucleasas de restricción y pueden seleccionarse con sondas complementarias a la secuencia génica deseada. Tras haber aislado la secuencia, el ADN puede amplificarse utilizando métodos convencionales de amplificación dirigida por cebadores, tales como la reacción en cadena de la polimerasa (documento US 4.683.202), para obtener cantidades de ADN adecuadas para la transformación utilizando vectores apropiados. Las herramientas para la optimización de codones para la expresión en un hospedante heterólogo están disponibles con facilidad. Están disponibles algunas herramientas para la optimización de codones basadas en el contenido en GC del organismo hospedante.

Tras haber identificado y aislado los genes de la vía pertinente, estos pueden transformarse en hospedantes de

expresión adecuados por medios muy conocidos en la técnica. Los vectores o módulos útiles para la transformación de una diversidad de células hospedantes son habituales y están disponibles en el mercado en empresas, tales como EPICENTRE® (Madison, WI), Invitrogen Corp. (Carlsbad, CA), Stratagene (La Jolla, CA), y New England Biolabs, Inc. (Beverly, MA). Generalmente, el vector o módulo contiene secuencias que dirigen la transcripción y la traducción del gen pertinente, un marcador seleccionable, y secuencias que permiten la replicación autónoma o la integración cromosómica. Los vectores adecuados comprenden una región 5' del gen que porta los controles de inicio de la transcripción, y una región 3' del fragmento de ADN que controla la terminación de la transcripción. Ambas regiones de control pueden derivarse de genes homólogos a la célula hospedante transformada, aunque debe entenderse que estas regiones de control también pueden derivarse de genes que no son nativos a las especies específicas elegidas como hospedante de producción.

Los promotores o las regiones de control del inicio, que son útiles para conducir la expresión de las regiones codificadoras de la vía pertinentes en la célula hospedante deseada, son numerosos y familiares para los expertos en la técnica. Casi todos los promotores capaces de dirigir estos elementos genéticos son adecuados para la presente invención e incluyen, pero no se limitan a *CYC1*, *HIS3*, *GAL1*, *GAL10*, *ADH1*, *PGK*, *PHO5*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO*, *TPI* (útiles para la expresión en *Saccharomyces*); *AOX1* (útil para la expresión en *Pichia*); y *lac*, *ara*, *tet*, *trp*, *IP_L*, *IP_R*, *T7*, *tac*, y *trc* (útiles para la expresión en *Escherichia coli*, *Alcaligenes*, y *Pseudomonas*), así como los promotores *amy*, *apr*, *npr* y diversos promotores de fagos útiles para la expresión en *Bacillus subtilis*, *Bacillus licheniformis*, y *Paenibacillus macerans*.

Las regiones de control de la terminación también pueden derivarse de diversos genes nativos a los hospedantes preferidos. Opcionalmente, un sitio de terminación puede ser innecesario, aunque se prefiere que se incluya.

Ciertos vectores son capaces de replicarse en una amplia gama de bacterias hospedantes y pueden ser transferidos mediante conjugación. Está disponible la secuencia completa y comentada de pRK404 y tres vectores relacionados (pRK437, pRK442, y pRK442(H)). Estos derivados han demostrado ser herramientas valiosas para la manipulación genética de bacterias Gram-negativas (Scott *et ál.*, Plasmid, 50, 74-79, 2003). También están disponibles varios derivados de plásmidos del plásmido RSF1010 Inc P4 de amplia gama de hospedantes, con promotores que pueden actuar en una diversidad de bacterias Gram-negativas. Los plásmidos pAYC36 y pAYC37 tienen promotores activos junto con múltiples sitios de clonación para permitir la expresión heteróloga en bacterias Gram-negativas.

Las herramientas de sustitución de genes cromosómicos también están ampliamente disponibles. Por ejemplo, un variante termosensible del replicón de amplia gama de hospedantes pWV101 se ha modificado para construir el plásmido pVE6002, que puede utilizarse para realizar la sustitución de genes en una gama de bacterias Gram-negativas (Maguin *et ál.*, J. Bacteriol., 174, 5633-5638, 1992). Además, están disponibles transposomas *in vitro* para crear mutaciones aleatorias en una diversidad de genomas, a partir de fuentes comerciales, tales como EPICENTRE®.

La expresión de una vía biosintética del isobutanol en diversos hospedantes microbianos preferidos se describe con más detalle a continuación.

Expresión de una vía biosintética del isobutanol en *E. coli*

Los vectores o módulos útiles para la transformación de *E. coli* son habituales y están disponibles en el mercado en las empresas listadas anteriormente. Por ejemplo, los genes de una vía biosintética del isobutanol pueden aislarse a partir de diversas fuentes, clonarse en un vector pUC19 modificado, y transformarse en *E. coli* NM522.

Expresión de una vía biosintética del isobutanol en *Rhodococcus erythropolis*

Está disponible una serie de vectores lanzadera de *E. coli-Rhodococcus* para la expresión en *R. erythropolis* que incluyen, pero no se limitan a pRhBR17 y pDA71 (Kostichka *et ál.*, Appl. Microbiol. Biotechnol., 62, 61-68, 2003). Además, está disponible una serie de promotores para la expresión de genes heterólogos en *R. erythropolis* (Nakashima *et ál.*, Appl. Environ. Microbiol., 70, 5557-5568, 2004; y Tao *et ál.*, Appl. Microbiol. Biotechnol., 68, 346-354, 2005). La alteración de genes cromosómicos dirigida en *R. erythropolis* puede crearse utilizando el método descrito por Tao *et ál.*, *supra*, y Brans *et ál.* (Appl. Environ. Microbiol., 66, 2029-2036, 2000).

Los genes heterólogos necesarios para la producción de isobutanol, tal como se describió anteriormente, pueden clonarse en un principio en pDA71 o pRhBR17, y transformarse en *E. coli*. Los vectores entonces pueden transformarse en *R. erythropolis* mediante electroporación, según se describe en Kostichka *et ál.*, *supra*. Los recombinantes pueden cultivarse en un medio sintético que contenga glucosa, y puede seguirse la producción de isobutanol utilizando métodos conocidos en la técnica.

Expresión de una vía biosintética del isobutanol en *B. subtilis*

Los métodos para la expresión de genes y la creación de mutaciones en *B. subtilis* también son muy conocidos en la técnica. Por ejemplo, los genes de una vía biosintética del isobutanol pueden aislarse a partir de diversas fuentes, clonarse en un vector pUC19 modificado, y transformarse en *Bacillus subtilis* BE1010. Además, los cinco genes de una vía biosintética del isobutanol pueden dividirse entre dos operones para la expresión. Los tres genes de la vía

(*bubB*, *ilvD*, y *kivD*) pueden integrarse en el cromosoma de *Bacillus subtilis* BE1010 (Payne, *et ál.*, J. Bacteriol., 173, 2278-2282, 1991). Los dos genes restantes (*ilvc* y *bdhB*) pueden clonarse en un vector de expresión y transformarse en la cepa de *Bacillus* que porta los genes de isobutanol integrados.

Expresión de una vía biosintética del isobutanol en *B. licheniformis*

- 5 La mayoría de los plásmidos y vectores lanzadera que se replican en *B. subtilis* pueden utilizarse para transformar *B. licheniformis* mediante transformación de protoplastos o electroporación. Los genes necesarios para la producción de isobutanol pueden clonarse en plásmidos derivados de pBE20 o pBE60 (Nagarajan *et ál.*, Gene, 114, 121-126, 1992). Los métodos para transformar *B. licheniformis* son conocidos en la técnica (Fleming *et ál.*, Appl. Environ. Microbiol., 61, 3775-3780, 1995). Los plásmidos construidos para la expresión en *B. subtilis* pueden transformarse en *B. licheniformis* para producir un hospedante microbiano recombinante que produzca isobutanol.

Expresión de una vía biosintética del isobutanol en *Paenibacillus macerans*

Pueden construirse plásmidos tal como se describió anteriormente para la expresión en *B. subtilis* y emplearse para transformar *Paenibacillus macerans* mediante transformación de protoplastos para producir un hospedante microbiano recombinante que produzca isobutanol.

- 15 Expresión de una vía biosintética del isobutanol en *Alcaligenes (Ralstonia) eutrophus*

Los métodos para la expresión de genes y la creación de mutaciones en *Alcaligenes eutrophus* son conocidos en la técnica (Taghavi *et ál.*, Appl. Environ. Microbiol., 60, 3585-3591, 1994). Los genes para una vía biosintética del isobutanol pueden clonarse en cualquiera de los vectores de amplia gama de hospedantes descritos anteriormente, y electroporarse para generar recombinantes que produzcan isobutanol. La vía del polihidroxibutirato en *Alcaligenes eutrophus*, y estas herramientas pueden aplicarse para modificar una vía biosintética del isobutanol.

Expresión de una vía biosintética del isobutanol en *Pseudomonas putida*

- 25 Los métodos para la expresión de genes en *Pseudomonas putida* son conocidos en la técnica (véase, por ejemplo, Ben-Bassat *et ál.*, patente de EEUU n.º 6.586.229). Los genes de la vía del butanol pueden insertarse en pPCU18 y este ADN acoplado puede electroporarse en células de *Pseudomonas putida* DOT-T1 C5aAR1 electrocompetentes para generar recombinantes que produzcan isobutanol.

Expresión de una vía biosintética del isobutanol en *Saccharomyces cerevisiae*

- 30 Los métodos para la expresión de genes en *Saccharomyces cerevisiae* son conocidos en la técnica (por ejemplo, Methods in Enzymology, volumen 194, Guide to Yeast Genetics and Molecular and Cell Biology, parte A, 2004, Christine Guthrie y Gerald R. Fink, eds., Elsevier Academic Press, San Diego, CA). La expresión de genes en levaduras generalmente requiere un promotor, seguido del gen de interés, y un terminador de la transcripción. Puede utilizarse una serie de promotores de levaduras para construir módulos de expresión para genes que codifican una vía biosintética del isobutanol que incluyen, pero no se limitan a los promotores constitutivos FBA, GPD, ADH1, y GPM, y los promotores inducibles GAL1, GAL10, y CUP1. Los terminadores de la transcripción adecuados incluyen, pero no se limitan a FBA_t, GPD_t, GPM_t, ERG10_t, GAL1_t, CYC1, y ADH 1. Por ejemplo, los promotores, terminadores de la transcripción y genes de una vía biosintética del isobutanol adecuados pueden clonarse en vectores lanzadera de *E. coli*-levaduras.

Expresión de una vía biosintética del isobutanol en *Lactobacillus plantarum*

- 40 El género *Lactobacillus* pertenece a la familia *Lactobacillales*, y muchos plásmidos y vectores que se emplean para la transformación de *Bacillus subtilis* y *Streptococcus* pueden utilizarse para *Lactobacillus*. Los ejemplos no limitantes de vectores adecuados incluyen pABβ1 y sus derivados (Renault *et ál.*, Gene, 183, 175-182, 1996); y (O'Sullivan *et ál.*, Gene, 137, 227-231, 1993); pMBB1 y pHW800, un derivado de pMBB1 (Wyckoff *et ál.*, Appl. Environ. Microbiol., 62, 1481-1486, 1996); pMG1, un plásmido conjugativo (Tanimoto *et ál.*, J. Bacteriol., 184, 5800-5804, 2002); pNZ9520 (Kleerebezem *et ál.*, Appl. Environ. Microbiol., 63, 4581-4584, 1997); pAM401 (Fujimoto *et ál.*, Appl. Environ. Microbiol., 67, 1262-1267, 2001); y pAT392 (Arthur *et ál.*, Antimicrob. Agents Chemother., 38, 1899-1903, 1994). También se han indicado varios plásmidos de *Lactobacillus plantarum* (van Kranenburg R., *et ál.*, Appl. Environ. Microbiol., 71, 1223-1230, 2005).

Expresión de una vía biosintética del isobutanol en diversas especies de *Enterococcus* (*E. faecium*, *E. gallinarum* y *E. faecalis*)

- 50 El género *Enterococcus* pertenece a la familia *Lactobacillales*, y muchos plásmidos y vectores que se emplean para la transformación de *Lactobacillus*, *Bacillus* y *Streptococcus* pueden utilizarse para especies de *Enterococcus*. Los ejemplos no limitantes de vectores adecuados incluyen pABβ1 y sus derivados (Renault *et ál.*, Gene, 183, 175-182, 1996); y (O'Sullivan *et ál.*, Gene, 137, 227-231, 1993); pMBB1 y pHW800, un derivado de pMBB1 (Wyckoff *et ál.*, Appl. Environ. Microbiol., 62, 1481-1486, 1996); pMG1, un plásmido conjugativo (Tanimoto *et ál.*, J. Bacteriol., 184,

5800-5804, 2002); pNZ9520 (Kleerebezem *et ál.*, Appl. Environ. Microbiol., 63, 4581-4584, 1997); pAM401 (Fujimoto *et ál.*, Appl. Environ. Microbiol., 67, 1262-1267, 2001); y pAT392 (Arthur *et ál.*, Antimicrob. Agents Chemother., 38, 1899-1903, 1994). También pueden utilizarse vectores de expresión para *E. faecalis* que emplean el gen *nisA* de *Lactococcus* (Eichenbaum *et ál.*, Appl. Environ. Microbiol., 64, 2763-2769, 1998). Además, pueden utilizarse vectores para la sustitución de genes en el cromosoma de *E. faecium* (Nallaappareddy *et ál.*, Appl. Environ. Microbiol., 72, 334-345, 2006).

Medios de fermentación

Los medios de fermentación en la presente invención deben contener sustratos de carbono adecuados. Los sustratos adecuados pueden incluir, pero no se limitan a monosacáridos, tales como glucosa y fructosa, oligosacáridos, tales como lactosa o sacarosa, polisacáridos, tales como almidón o celulosa o sus mezclas, y mezclas no purificadas de materias primas, tales como permeado de suero de queso, destilado de macerado de maíz, melaza de remolacha azucarera, y malta de cebada. Además, el sustrato de carbono también puede ser un sustrato de un solo carbono, tal como dióxido de carbono, o metanol, para el que se ha demostrado la conversión metabólica en intermedios bioquímicos clave. Además de los sustratos de un solo carbono y de dos carbonos, también se sabe que los organismos metiltróficos utilizan una serie de otros compuestos que contienen carbono, tales como metilamina, glucosamina y una diversidad de aminoácidos para la actividad metabólica. Por ejemplo, se sabe que las levaduras metiltróficas utilizan el carbono de la metilamina para formar trehalosa o glicerol (Bellion *et ál.*, Microb. Growth C1 Compd. [Int. Symp.] 7th (1993), 415-432. eds.: Murrell, J. Collin; Kelly, Don P., edición: Intercept, Andover, Reino Unido). De modo similar, diversas especies de *Candida* metabolizan la alanina o el ácido oleico (Sulter *et ál.*, Arch. Microbiol., 153, 485-489, 1990). Por tanto, se contempla que la fuente de carbono utilizada en la presente invención pueda comprender una amplia gama de sustratos que contienen carbono y que estos solo se verán limitados por la elección del organismo.

Aunque se contempla que todos los sustratos de carbono mencionados anteriormente y sus mezclas son adecuados en la presente invención, los sustratos de carbono preferidos son glucosa, fructosa y sacarosa.

Además de una fuente de carbono adecuada, el medio de fermentación debe contener minerales, sales, cofactores, tampones y otros componentes adecuados, conocidos por los expertos en la técnica, que sean adecuados para el crecimiento de los cultivos y la estimulación de la vía enzimática necesaria para la producción de isobutanol.

Condiciones de cultivo

Generalmente, las células se cultivan a una temperatura en el intervalo de aproximadamente 25 °C a aproximadamente 40 °C en un medio apropiado. Los medios de crecimiento adecuados en la presente invención son medios habituales preparados del mercado, tales como caldo Luria Bertani (LB), caldo Sabouraud dextrosa (SD), o caldo de medio de levadura (YM). También pueden utilizarse otros medios de crecimiento definidos o sintéticos, y el medio apropiado para el crecimiento del microorganismo concreto será conocido por los expertos en la técnica de la microbiología o la ciencia de la fermentación. El uso de agentes conocidos por modular la represión de catabolitos directa o indirectamente, por ejemplo, adenosina 2',3'-monofosfato cíclico (AMPC), también pueden incorporarse en el medio de fermentación.

Los intervalos de pH adecuados para la fermentación son entre pH 5,0 a pH 9,0, siendo preferido un pH 6,0 a pH 8,0 para las condiciones iniciales.

Las fermentaciones pueden realizarse bajo condiciones aerobias o anaerobias, siendo preferidas las condiciones anaerobias o microaerobias.

Fermentaciones continuas y discontinuas industriales

El presente proceso emplea un método discontinuo de fermentación. Una fermentación discontinua clásica es un sistema cerrado en el que la composición del medio se ajusta al comienzo de la fermentación y no se somete a alteraciones artificiales durante la fermentación. Así, al comienzo de la fermentación, el medio se inocula con el organismo u organismos deseados, y se deja que la fermentación se desarrolle sin añadir nada al sistema. Sin embargo, generalmente, una fermentación "discontinua" es discontinua con respecto a la adición de la fuente de carbono, y a menudo se intentan controlar factores, tales como el pH y la concentración de oxígeno. En los sistemas discontinuos, las composiciones de los metabolitos y de la biomasa del sistema cambian constantemente hasta el momento en que se detiene la fermentación. Dentro de los cultivos discontinuos, las células fluctúan desde una fase de retraso estática a una fase logarítmica de alto crecimiento y, por último, a una fase estacionaria en la que la velocidad de crecimiento disminuye o se detiene. Si no se tratan, las células en la fase estacionaria finalmente mueren. Las células en la fase logarítmica en general son responsables de la producción a granel del producto final o del intermedio.

Una variación del sistema discontinuo convencional es el sistema de alimentación discontinua. Los procesos de fermentación de alimentación discontinua también resultan adecuados en la presente invención, y comprenden un sistema discontinuo típico con la excepción de que el sustrato se añade en incrementos a medida que avanza la fermentación. Los sistemas de alimentación discontinua son útiles cuando la represión de catabolitos es apta para

inhibir el metabolismo de las células, y cuando resulta deseable tener cantidades limitadas de sustrato en el medio. Las mediciones de la concentración real del sustrato en los sistemas de alimentación discontinua son difíciles y, por tanto, se calcula sobre la base de los cambios en los factores mensurables, tales como pH, oxígeno disuelto y la presión parcial de gases residuales, tales como CO₂. Las fermentaciones discontinuas y de alimentación discontinua son habituales y muy conocidas en la técnica, y pueden encontrarse ejemplos en Thomas D. Brock, en *Biotechnology: A Textbook of Industrial Microbiology*, 2ª edición (1989), Sinauer Associates, Inc., Sunderland, MA.; o Deshpande, Mukund (Appl. Biochem. Biotechnol., 36, 227, 1992).

Aunque la presente invención se realiza en modo discontinuo, se contempla que el método podría adaptarse a métodos de fermentación continua. La fermentación continua es un sistema abierto en el que un medio de fermentación definido se añade continuamente a un biorreactor, y una cantidad igual de medio acondicionado se retira de modo simultáneo para su procesamiento. La fermentación continua en general mantiene los cultivos a una alta densidad constante, en los que las células están principalmente en fase de crecimiento logarítmico.

La fermentación continua permite la modulación de un factor o de una serie de factores que afectan al crecimiento celular o la concentración del producto final. Por ejemplo, un método mantiene un nutriente limitante, tal como la fuente de carbono o el nivel de nitrógeno, a un nivel fijo, y permite que los otros parámetros fluctúen. En otros sistemas, una serie de factores que afectan al crecimiento pueden alterarse de modo continuo, mientras que la concentración de células, medida a través de la turbidez del medio, se mantiene constante. Los sistemas continuos buscan mantener condiciones de crecimiento en estado constante y, así, la pérdida de células debida a la extracción del medio debe equilibrarse frente a la velocidad de crecimiento celular en la fermentación. Los métodos para modular los nutrientes y los factores del crecimiento para los procesos de fermentación continua, así como las técnicas para maximizar la velocidad de formación del producto, son muy conocidos en la técnica de la microbiología industrial, y una diversidad de métodos se detallan en Brock, *supra*.

Se contempla que la presente invención pueda practicarse utilizando procesos discontinuos, de alimentación discontinua o continuos, y que cualquier modo de fermentación sería adecuado. Además, se contempla que las células puedan inmovilizarse sobre un sustrato como catalizadores de células enteras, y someterse a condiciones de fermentación para la producción de isobutanol.

Métodos para el aislamiento del isobutanol a partir del medio de fermentación

El isobutanol producido biológicamente puede aislarse del medio de fermentación utilizando métodos conocidos en la técnica para las fermentaciones de acetona-butanol-etanol (ABE) (véase, por ejemplo, Durre, Appl. Microbiol. Biotechnol., 49, 639-648, 1998; Groot *et ál.*, Process. Biochem., 27, 61-75, 1992; y las referencias citadas en ellos). Por ejemplo, los sólidos pueden retirarse del medio de fermentación mediante centrifugación, filtración, decantación, y el isobutanol puede aislarse del medio de fermentación utilizando métodos tales como destilación, destilación azeotrópica, extracción de líquido-líquido, adsorción, destilación de gases, evaporación de membrana, o pervaporación.

Ejemplos

La presente invención se define con más detalle en los siguientes ejemplos. Debe entenderse que estos ejemplos, aunque indican realizaciones preferidas de la invención, sólo se ofrecen como ilustración.

Métodos generales

Las técnicas de ADN recombinante y de clonación molecular convencionales utilizadas en los ejemplos son muy conocidas en la técnica, y se describen en Sambrook, J., Fritsch, E. F., y Maniatis, T. (*Molecular Cloning: A Laboratory Manual*; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1989, en la presente mencionado como "Maniatis"), y en Maniatis (*supra*), y en Silhavy *et ál.* (Silhavy, *et ál.*, *Experiments with Gene Fusions*, Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y. 1984), y en Ausubel *et ál.* (Ausubel *et ál.*, *Current Protocols in Molecular Biology*, pub. por Greene Publishing Assoc. and Wiley-Interscience, 1987).

Los materiales y los métodos adecuados para el mantenimiento y el crecimiento de cultivos bacterianos son muy conocidos en la técnica. Las técnicas adecuadas para su uso en los siguientes ejemplos pueden encontrarse en *Manual of Methods for General Bacteriology* (Phillipp *et ál.*, eds., American Society for Microbiology, Washington, DC., 1994), o en Thomas D. Brock, en Brock, *Biotechnology: A Textbook of Industrial Microbiology*, 2ª edición, Sinauer Associates, Inc., Sunderland, MA (1989). Todos los reactivos, enzimas de restricción y materiales utilizados para el crecimiento y el mantenimiento de las células bacterianas se obtuvieron en Aldrich Chemicals (Milwaukee, WI), BD Diagnostic Systems (Sparks, MD), Life Technologies (Rockville, MD), o Sigma Chemical Company (St. Louis, MO), a menos que se indique lo contrario.

Los cebadores oligonucleotídicos utilizados en los siguientes ejemplos se indican en la tabla 3.

Tabla 3

CEBADORES OLIGONUCLEOTÍDICOS UTILIZADOS EN ESTA INVENCIÓN		
SECUENCIA ID NO:	SECUENCIA	Descripción
11	GAGCTCCTTAAGAAGGAGGTAATCACCATGGC TAACTACTTCAA	Cebador para la amplificación de <i>ilvC</i>
12	GGATCCGATCGAGCTAGCGCGGCCGCTTAACC CGCAACAGCAATACGTTTC	Cebador para la amplificación de <i>ilvC</i>
13	GCTAACAGGAGGAAGAGCTCATGGCACCCCTCGCTC	<i>pBAD-SAC1-F</i> directo
14	GAGCGAGGGTGCCATGAGCTCTTCCTCCTGTTAGC	<i>pBAD-SAC1-R</i> inverso
15	ATCACCGAGCTCATGGCTAACTACTTCAATACACT GAATCTGCG	<i>ilvC-trc-Sac1-F</i> directo
16	GGCCGCAAGCTTTTAACCCGCAACAGCAATACGT TTCATATCTGTC	<i>ilvC-trc-HindIII-R</i> inverso
17	CCGTAAAGATATCACCGTAG	<i>ilvC-tro-F3</i>
18	CAGTATGAAGGCAAAATCGG	<i>ilvC-tro-F5</i>
19	CGTACTCAGCGGTATCAGAG	<i>ilvC-tro-R2</i>
20	CAGATTTCACTTCCGCAACG	<i>ilvC-tro-R4</i>
21	CGCAACTCTCTACTGTTTCTCCATACCCG	<i>pBAD-e-F1</i>
22	ACCGCTTCTGCGTTCTGATTTAATC	<i>PALPK-R1</i>
23	CAAACAGCCAAGCTTTTAGTTCTTGCTCTTGTC GACGATCTTG	PAO1-C-F1
24	CAGGAGGAAGAGCTCATGCGGTTTTCTACGAT AAAGACTGTG	PAO1-C-R1
25	CAAACAGCCAAGCTTTTAGTTCTTGCTTTGTC GACGATTTTG	PF5-C-F1
26	CAGGAGGAAGAGCTCATGAAAGTTTTCTACGATA AAGACTGCGAC	PF5-C-R1
27	GATCATGATCGCGCCGAAGG	PF5-S-F2
28	CTGCTCACCGAACAGGTCGG	PF5-S-R2
37	CTGCAGCACATGAAGACTCCATGGCACCC TCGCTCGACTCGATCTCGCACTCGTTTCGC AAACG	PAL-F1 directo
38	TCTCTCATCCGCCAAAACAGAAGCTTCTAA GCGAGCATCT	PAL-R1 inverso
39	GGGCTAACAGGAGGAAGAATTCATGGCAC CCTCGCTCGACTCG	PAL-EcoR1-F1 directo
40	CGAGTCGAGCGAGGGTGCCATGAATTCTT CCTCCTGTAGCCC	PAL-EcoR1-R1 inverso

El significado de las abreviaturas es el siguiente: "seg" significa segundo o segundos, "min" significa minuto o minutos, "h" significa hora u horas, "nm" significa nanómetros, "ul" significa microlitro o microlitros, "ml" significa mililitro o mililitros, "mg/ml" significa miligramos por mililitro, "l" significa litro o litros, "nm" significa nanómetros, "mM" significa milimolar, "M" significa molar, "mmol" significa milimol o milimoles, "mmol" significa micromol o micromoles, "kg" significa kilogramo, "g" significa gramo o gramos, "mg" significa microgramo o microgramos, y "ng" significa nanogramo o nanogramos, "PCR" significa reacción en cadena de la polimerasa, "DO" significa densidad óptica, "DO₆₀₀" significa la densidad óptica medida a una longitud de onda de 600 nm, "kDa" significa kilodaltons, "g" significa la constante de gravitación, "pb" significa par o pares de bases, "kbp" significa kilopares de bases, "kb" significa kilobase, "%" significa porcentaje, "% en p/v" significa porcentaje en peso/volumen, "% en v/v" significa

5

10

porcentaje en volumen/volumen, "HPLC" significa cromatografía líquida de alta resolución, "g/l" significa gramos por litro, "ug/l" significa microgramos por litro, "ng/ul" significa nanogramos por microlitro, "pmol/ul" significa picomoles por microlitro, "rpm" significa rotaciones por minuto, "umol/min/mg" significa micromoles por minuto por miligramo, "p/v" significa peso por volumen, "v/v" significa volumen por volumen.

5 Ejemplo 1 (comparativo)

Análisis de la actividad de la enzima KARI

Este ejemplo describe la preparación de construcciones de sobreexpresión del gen *ilvC* y la medición de la actividad enzimática utilizando la oxidación dependiente de acetolactato de NADPH por la enzima KARI codificada por el gen *ilvC* de *E. coli*.

10 Construcción del plásmido de expresión pBAD-*ilvC* y aislamiento del gen *ilvC* de *E. coli*

La región codificadora del gen *ilvC* se amplificó a partir del ADN genómico de *E. coli* cepa FM5 (ATCC 53911) utilizando PCR. Las células se cultivaron durante la noche (37 °C, mientras se agita a 300 rpm) en tubos de cultivo de 50 ml que contenían 4 ml de medio Luria Bertani (LB) (Mediatech Inc., Herndon, VA). Después se recolectaron mediante centrifugación a 1000 x g durante 3 min, y el ADN genómico de las células se preparó utilizando el kit

15 Gentra Puregene (Gentra Systems, Inc., Minneapolis, MN; n.º de catálogo D-5000A) según las instrucciones del fabricante. Se preparó un fragmento de ADN de la región codificadora de *ilvC* mediante PCR utilizando ADN de *E. coli* como molde y los cebadores SEQ ID NO:11 y 12.

La PCR se realizó utilizando Finnzymes Phusion™ High-Fidelity PCR Master Mix (New England Biolabs Inc., Beverly, MA, n.º de catálogo F-531) según el protocolo del fabricante. La amplificación se realizó en un termociclador de ADN GeneAmp 9700 (PE Applied Biosystems, Foster city, CA). El producto de la PCR (0,5 ul), sin más purificación, se acopló en pCR4Blunt TOPO (Invitrogen, Carlsbad, CA, n.º de catálogo 45-0031) y se transformó en células TOP10 químicamente competentes (Invitrogen 44-0301). El producto del acoplamiento se sembró en estrías sobre una placa que contenía el medio LB más ampicilina 100 ug/ml (Teknova Inc, Hollister, CA, n.º de catálogo L1004). Los clones que contenía la inserción de *ilvC* fueron confirmados mediante digestión de restricción con

20 Sacl/BamHI. Tres de los cuatro plásmidos digeridos tenían la banda de 1,5 kpb esperada. El clon resultante se denominó pCR4Blunt TOPO-*ilvC*.

El fragmento *ilvC* del vector de clonación pCR4Blunt TOPO-*ilvC* fue liberado mediante digestión con Sacl/BamHI, y se acopló en pTrc99A digerido con Sacl/BamHI (Amann, *et ál.*, Gene, 69, 301-315, 1988) utilizando ADN ligasa de T4 (New England Biolabs, Beverly, MA). Esta construcción se electroporó en células de *E. coli* TOP10 electrocompetentes (Invitrogen 44-0035), y estas se sembraron en estrías sobre una placa de LB/ampicilina tal como se describió anteriormente. El vector que contiene la inserción de 1,5 kb se denominó pTrc99A-*ilvC*.

30

Preparación del vector pBAD para la clonación

Se construyó un derivado del vector pBAD.HisA (Invitrogen) que contenía un sitio Sacl en el extremo 5' del gen, para la clonación del gen *ilvC* en pBAD utilizando sitios de restricción Sacl/HindIII. Esta construcción se creó en tres etapas. En primer lugar, la región codificadora de fenilalanina amoniaco liasa (PAL, EC 4.3.1.5) de *Rhodotorula glutinis* se clonó en el vector pBAD-HisA para fabricar pBAD-PAL. En segundo lugar, se añadió el sitio EcoRI en el extremo 5' del gen inmediatamente antes del codón de inicio sobre la construcción pBAD-PAL para fabricar pBAD-PAL-EcoRI. En tercer lugar, el sitio EcoRI fue reemplazado por un sitio Sacl, y el vector resultante se digirió con Sacl/HindIII para fabricar un vector pBAD-Sacl para la clonación del gen *ilvC*. El gen PAL primero se amplificó mediante PCR a partir del vector pKK223-PAL (patente de EEUU n.º 6521748) utilizando un cebador directo (PAL-F1) (SEQ ID NO: 37) y un cebador inverso (PAL-R1) (SEQ ID NO:38).

35

La PCR se realizó en un termociclador Perkin Elmer PCR9700 (PE Applied Biosystems, Foster City, CA) utilizando la premezcla de ADN polimerasa Taq TaKaRa (TAKARA Bio USA, Madison, WI, n.º de catálogo TAK_R004A) según el protocolo del fabricante. El producto de la PCR fue parcialmente purificado utilizando el kit de purificación de PCR QIAuik (Qiagen, n.º de catálogo 28106) y se digirió con BbsI y HindIII. Esto produjo un fragmento que contenía una proyección NcoI en el extremo 5'. El producto de la digestión después se acopló en un vector pBAD.HisA (Invitrogen) que se había digerido con NcoI/HindIII. La reacción de acoplamiento se realizó utilizando ADN ligasa de T4 (Promega) siguiendo el protocolo convencional proporcionado por el fabricante. Se emplaron 2 ul del producto de acoplamiento para transformar células electrocompetentes TOP10 (Invitrogen) empleando un Bio-RAD Gene Pulser II (Bio-Rad Laboratories Inc., Hercules, CA) siguiendo las indicaciones del fabricante. Las células transformadas se sembraron en estrías sobre placas de agar que contenían medio LB más ampicilina 100 ug/ml (Teknova Inc., Hollister, CA, n.º de catálogo L1004) y se incubaron durante la noche a 37 °C. Los clones que contenían la inserción PAL se confirmaron mediante digestión de restricción con NcoI/HindIII. Esta construcción se denominó pBAD-PAL. Entonces se añadió el sitio EcoRI al extremo 5' del gen PAL en la anterior construcción utilizando un kit de mutagénesis específica dirigida a sitio QuikChange II (Stratagene, La Jolla, CA, n.º de catálogo 200524). Se diseñaron el cebador directo (PAL-EcoRI-F1) (SEQ ID NO:39) y el cebador inverso (PAL-EcoRI-R1) (SEQ ID NO:40), y las mezclas de reacción se prepararon siguiendo las indicaciones del fabricante. Se empleó la construcción pBAD-PAL preparada anteriormente como molde en la siguiente reacción.

45

50

55

Los 50 ul de la mezcla de reacción contenían 1,0 ul de 50 ng/ul del plásmido molde, 1,0 ul de 10 pmol/ul de cada cebador, 5 ul de 10x tampón de reacción, 1,0 ul de mezcla de dNTP, y 3 ul de disolución Quik, 30 ul de agua y 1,0 ul de ADN polimerasa de alta fidelidad pfu-ultra en un tubo de 200 ul de pared fina. Todos los reactivos y la polimerasa utilizados en esta reacción fueron proporcionados por el anterior kit QuikChange II XL. La reacción se realizó en un termociclador de ADN GeneAmp 2400 (PE Applied Biosystems, Foster City, CA) utilizando las siguientes condiciones. La temperatura de inicio fue de 96 °C durante 2 min, tras lo cual se realizaron 18 ciclos de calentamiento/enfriamiento. Cada ciclo consistió en 96 °C durante 30 seg, seguido de 60 °C durante 30 seg, y 72 °C durante 160 seg. Cuando finalizaron los ciclos de temperatura, las muestras se mantuvieron a 72 °C durante 600 seg más, y después se mantuvieron a la espera de la recuperación de la muestra a 4 °C.

Tras completar la reacción, se añadió 1,0 ul de la enzima de restricción DpnI (del anterior kit) a la reacción, seguido de una incubación a 37 °C durante 3 h para digerir los plásmidos molde en la reacción.

Entonces se transformaron 2,0 ul del producto de la reacción digerido con DpnI en 50 ul de células electrocompetentes *E. coli* TOP10 (Invitrogen) utilizando un Bio-RAD Gene Pulser II (Bio-Rad Laboratories Inc., Hercules, CA) siguiendo las indicaciones del fabricante. Se sembraron en estrías diferentes volúmenes (2,0 ul, 5,0 ul y 20 ul) de las células transformadas sobre placas de agar de 10 cm que contenían medio LB y ampicilina 100 ug/ml, y las placas se incubaron a 37 °C durante la noche. Se escogieron tres clones de la placa que contenía las colonias bien separadas. Los plásmidos de los tres clones se purificaron utilizando un kit de minipreparación Spin Qiaprep (Qiagen, Valencia, CA, n.º de catálogo 27106) siguiendo las instrucciones del fabricante. Los clones positivos fueron confirmados mediante un análisis de digestión de restricción empleando las enzimas de restricción EcoRI y HindIII (Promega, Madison, WI) colocando 1,0 ul de 10x tampón de reacción (tampón Promega), 1,0 ul del plásmido purificado y 1,0 ul de cada enzima de restricción en 6,0 ul de agua desionizada. La mezcla de reacción se incubó a 37 °C durante 60 min. El producto digerido de cada clon se separó en un gel de agarosa E al 0,8% (Invitrogen, n.º de catálogo G5018-08). Se detectó un fragmento de ADN de 2,1 kpb y otro de 4,0 kpb sobre el gel en las muestras con ambos sitios de restricción EcoRI y HindIII en la construcción. El sitio EcoRI en esta construcción entonces se reemplazó por un sitio SacI utilizando el mismo protocolo descrito anteriormente con el plásmido molde pBAD-PAL-EcoRI y los cebadores de SEQ ID NO:13 y 14.

Los clones positivos fueron confirmados por análisis de digestión de restricción utilizando las enzimas de restricción SacI y HindIII (Promega, Madison, WI). Tras haber identificado los clones positivos, la anterior reacción de digestión de restricción se ajustó a una escala mayor (50 ul). El fragmento de 4 kpb que contenía el vector digerido se purificó en gel de la mezcla utilizando un gel de agarosa al 1% y un kit de extracción de gel QIAquick (Qiagen, Valencia, CA, n.º de catálogo 28704) siguiendo el protocolo del fabricante. Esta construcción se denominó pBAD-SacI.

Cepas hospedantes utilizadas para sobreexpresar KARI

La cepa hospedante *E. coli* Bw25113 ($\Delta ilvC$), una cepa con el gen *ilvC* inactivado, se empleó para fabricar construcciones que sobreexpresan la enzima KARI. En esta cepa, el gen *ilvC* completo sobre el cromosoma de *E. coli* ha sido reemplazado por un módulo de kanamicina empleando la tecnología de recombinación de homología de rojo lambda (Datsenko y Wanner, Proc. Natl. Acad. Sci. USA, 97, 6640-6645, 2000). Todas las cepas y los vectores necesarios para la creación de la cepa inactivada utilizando esta tecnología fueron obtenidos del profesor Barry Wanner (Purdue University, West Lafayette, IN).

Preparación de la región codificadora de *ilvC* para la clonación

La región codificadora para *ilvC* se amplificó utilizando la polimerasa de alta fidelidad pfu-ultra (Stratagene, La Jolla, CA) con la adición de un sitio SacI en el extremo 5' del cebador directo justo antes de ATG, y se añadió un sitio HindIII al extremo 5' del cebador inverso justo después del codón de fin. Se empleó el cebador con SEQ ID NO:15 (directo: *ilvc-trc-SacI-F*) y el cebador con SEQ ID NO:16 (inverso: *ilvc-trc-HindIII-R*) para esta reacción. El molde utilizado en la reacción de PCR fue la construcción pTrc99A-*ilvC* descrita anteriormente.

Los 50 ul de la mezcla de reacción contenían 5,0 ul de 10x tampón de reacción suministrado con la polimerasa pfu-ultra (Stratagene), 1,0 ul de 50 ng/ul del plásmido molde, 1,0 ul de 10 pmol/ul de cada uno de los cebadores directo e inverso, 1,0 ul de mezcla de dNTP 40 mM (Clontech, Mountain View, CA), 1,0 ul de ADN polimerasa de pfu-ultra (Stratagene) y 39 ul de agua. Esta mezcla de reacción se colocó en un tubo de 200 ul de pared fina para la reacción de PCR en un termociclador de ADN GeneAmp 2400 (PE Applied Biosystems, Foster City, CA). Se emplearon las siguientes condiciones para realizar la reacción de PCR. La temperatura de inicio fue de 94 °C durante 2 min. Después se realizaron 30 ciclos de calentamiento/enfriamiento. Cada ciclo consiste en 94 °C durante 30 seg, 58 °C durante 30 seg, y 68 °C durante 1 min y 40 seg. Cuando finalizaron los ciclos de temperatura, las muestras se mantuvieron a 60 °C durante 10 min más, y después se mantuvieron a la espera de la recuperación de la muestra a 4 °C.

El producto de la PCR fue parcialmente purificado utilizando el kit de purificación de PCR QIAquick (Qiagen, n.º de catálogo 28106) y se digirió con HindIII y SacI, y después se purificó en gel utilizando el protocolo descrito anteriormente. El fragmento de la PCR digerido se acopló en el vector pBAD-SacI digerido con el mismo conjunto de enzimas. Los 20 ul de la mezcla de acoplamiento contenían 1,0 ul de ADN ligasa de T4 (Promega), 2,0 ul de 10x

tampón de reacción que contiene la ADN ligasa de T4, 45 ng de vector y 45 ng de la inserción y agua desionizada. La reacción se incubó a 16 °C durante la noche en un termociclador Eppendorf (Eppendorf North America, Westbury, NY).

- 5 Se transformaron 2 μ l del producto de acoplamiento en células electrocompetentes *E. coli* TOP10 (Invitrogen) empleando un Bio-RAD Gene Pulser II (Bio-Rad Laboratories Inc., Hercules, CA). Los clones transformados se seleccionaron en placas de agar que contenían medio LB y ampicilina 100 μ g/ml. La presencia de la inserción del gen *ilvC* de *E. coli* en el clon se confirmó mediante digestión con SacI y secuenciación de ADN utilizando los cebadores de SEQ ID NO:17-22. La construcción con la inserción del gen *ilvC* se denominó pBAD-K12-*ilvC*.

Preparación de cepas para el análisis de la expresión de KARI

- 10 Los plásmidos de la anterior construcción pBAD-K12-*ilvC* y pTrc99A-*ilvC*, ambos en la cepa hospedante TOP10, se prepararon a partir de 3 ml del cultivo durante la noche en medio LB que contenía ampicilina 100 μ g/ml utilizando el kit de minipreparación Spin Qiaprep (Qiagen, Valencia, CA, n.º de catálogo 27106) siguiendo las instrucciones del fabricante. Se transformó 1 μ l de pBAD-K12-*ilvC* y 1 μ l de pTrc99A-*ilvC* por separado en células electrocompetentes *E. coli* Bw25113 (Δ *ilvC*) utilizando un Bio-RAD Gene Pulser II (Bio-Rad Laboratories Inc., Hercules, CA) siguiendo las indicaciones del fabricante. Las células transformadas se sembraron en estrías sobre placas de agar que contenían medio LB más ampicilina 100 μ g/ml y se incubaron durante la noche a 37 °C. Las colonias de estas placas se utilizaron para la preparación de extractos sin células.

Preparación del extracto sin células

- 20 Las células que contenían pBAD-K12-*ilvC* y pTrc99A-*ilvC* se cultivaron en 3,0 ml de medio LB que contenía ampicilina 100 μ g/ml y el inductor arabinosa al 0,02% (en p/v) e isopropil β -D-1-tiogalactopiranosido (IPTG) 1 mM, respectivamente, a 37 °C mientras se agitaba a 250 rpm. Las células se recolectaron mediante centrifugación a 6000 x g durante 5 min a 22,5 °C, los sedimentos celulares se resuspendieron en 300 μ l de tampón HEPES 100 mM (pH 7,5) en un tubo de microcentrífuga de 1,5 ml, se colocaron en un baño de agua relleno con 40% de agua y 60% de hielo (en volumen), y se sonicaron durante 2-3 min (activaciones de 3,0 seg a fuerza 1,0, seguido de 3,0 seg de reposo) utilizando un sonicador Misonix 300 (Misonix, Farmingdale, NY). Los restos celulares se retiraron mediante centrifugación (microcentrífuga Eppendorf, modelo 5415D, a 9300 x g durante 5 min a 22,5 °C).

- 25 Como alternativa, los extractos celulares se prepararon utilizando la mezcla maestra BugBuster de reactivo de extracción de proteínas basado en detergentes (Novagen, n.º de catálogo 71456). Los sedimentos celulares de 3,0 ml de los cultivos se resuspendieron en 300 μ l de mezcla maestra BugBuster y se incubaron a temperatura ambiente durante 20 min. Los restos celulares se retiraron mediante centrifugación (microcentrífuga Eppendorf, modelo 5415D) a 9300 x g durante 5 min a 22,5 °C.

Cuantificación de las proteínas

- 35 Se midió la concentración de proteínas totales en las muestras mediante el ensayo de Coomassie de Bradford (BCA) empleando Coomassie Plus (Pierce n.º 23238, Rockford, IL). Las muestras y los patrones de proteínas (albúmina de suero bovino, BSA) se dispusieron en una microplaca de 96 pocillos siguiendo el protocolo del fabricante. La concentración de proteínas se midió tras leer la absorbancia a 595 nm utilizando un lector de placas SpectraMax (Molecular Devices Corporation, Sunnyvale, CA).

Protocolo de ensayo de la enzima KARI

- 40 El sustrato de ensayo, (R,S)-acetolactato, se sintetizó como se describe en Aulabaugh y Schloss (Aulabaugh y Schloss, *Biochemistry*, 29, 2824-2830, 1990): se mezclaron 1,0 g de éster etílico del ácido 2-acetoxi-2-metil-3-oxibutírico (Aldrich, Milwaukee, WI) con 10 ml de NaOH 1,0 M y se agitó a temperatura ambiente. Cuando el pH de la disolución llegó a la neutralidad, se añadió lentamente más NaOH para mantener el pH a aproximadamente 8,0. Todos los demás productos químicos utilizados en el ensayo se obtuvieron en Sigma.

- 45 La conversión enzimática de acetolactato a 2,3-dihidroxiisovalerato mediante KARI fue seguida midiendo la desaparición del cofactor, NADPH, de la reacción a 340 nm utilizando un espectrofotómetro (Agilent Technologies, Santa Clara, CA). La actividad se calculó utilizando el coeficiente de extinción molar de 6220 M⁻¹ cm⁻¹ para NADPH. Las disoluciones madre utilizadas fueron: HEPES-sal de potasio 100 mM, ajustado mediante HCl/KOH a pH 7,5, MgCl₂ 1,0 M, NADPH 20 mM y acetolactato 90 mM. Los 40 ml de la disolución madre de mezcla de reacción contenían disolución madre de HEPES 100 mM y 400 μ l de disolución madre de MgCl₂.

- 50 El tampón de reacción (194 μ l) se mezcló con disolución madre de NADPH (2,0 μ l) y extracto celular (2,0 μ l) en una cubeta de plástico desechable (Eppendorf UVette, Eppendorf AG, Hamburgo, Alemania) y se registró la absorbancia a 340 nm a 22,5 °C durante 20 segundos. La A₃₄₀ inicial normalmente fue de aproximadamente 0,9-1,0. Después se añadió acetolactato (2,0 μ l) a la cubeta para iniciar la reacción. La concentración final de los ingredientes en el ensayo fue de: HEPES-potasio 100 mM a pH 7,5, MgCl₂ 1,0 mM, NADPH 200 μ M y acetolactato 900 μ M. Esta disolución se mezcló a fondo y se registró su absorbancia a 340 nm durante 80 seg más. La actividad KARI indicada en la presente se define como μ moles de NADPH consumidos por min por mg de proteína total en los extractos

celulares. Los resultados de las concentraciones de proteínas y actividades KARI en los extractos celulares preparados a partir de células *E. coli* Bw25113 (Δ ilvC) transformadas con plásmidos pBAD-K12-ilvC y plásmidos pTrc99A-ilvC se muestran en la tabla 4. Se prepararon dos muestras de extracto celular para la construcción pBAD-K12-ilvC, una por sonicación y la otra utilizando BugBuster. La muestra de extracto celular para la construcción pTrc99A-ilvC se preparó utilizando BugBuster. Estos análisis demuestran que la proteína KARI se expresa a un nivel mayor en las células que contenían los plásmidos pBAD-K12-ilvC que en las que contenían pTrc99A-ilvC, aunque las actividades específicas de la enzima en las muestras de extractos celulares preparadas mediante dos métodos diferentes no fueron significativamente diferentes. La cepa Bw25113 de *E. coli* transformada con pBAD-HisB (Invitrogen) se empleó como control negativo. La velocidad de consumo de NADPH en el control negativo fue extremadamente baja (aproximadamente 1% al 2% de la velocidad de consumo medida en las células que contienen el gen pBAD-K12-ilvC).

Tabla 4

Concentración de proteínas totales y KARI en clones que contienen el gen <i>ilvC</i>		
Clones	Proteína total (μ g/ml)	Actividad KARI umol/min/mg de proteína total
BW25113(Δ ilvC)-ptrc99A-ilvC-Bugbuster	8007	0,16
BW25113(Δ ilvC)-pBAD-K12-ilvC-sonicación	9707	0,83
BW25113(Δ ilvC)-pBAD-K12-ilvC-BugBuster	4595	0,78

Ejemplo 2

Identificación de KARI con alta actividad enzimática específica a partir de diversos microorganismos

El objetivo de este ejemplo es describir el modo de identificar microorganismos que contienen enzimas KARI con alta actividad específica.

Se estableció la hipótesis de que los organismos que contienen KARI con un tiempo de duplicación más rápido que *E. coli*, durante el crecimiento en un medio mínimo, contendrían enzimas KARI muy activas. Se identificaron tres microorganismos, *Pseudomonas aeruginosa* (PAO1), *Pseudomonas fluorescens* (PF5), y *Vibrio cholerae* (N16961), con unos tiempos de duplicación más rápidos que *E. coli* cuando se cultivan en medio mínimo M9 (véase a continuación). Las preparaciones de ADN genómico de estos organismos están disponibles en el mercado. La tabla 5 muestra los tiempos de duplicación de estos organismos comparado con *E. coli* tras el crecimiento en medio M9 mínimo.

Tabla 5

Tiempo de duplicación de las cepas ensayadas durante el crecimiento en medio M9		
Organismo	Tiempo de duplicación en medio M9	Referencia bibliográfica
<i>E. coli</i>	55-60 min	1
<i>V. cholerae</i> (N16961)	45 min	3
<i>P. aeruginosa</i> (PAO1)	42 min	2
<i>P. fluorescens</i> (PF5)	38 min	2

25 Referencias bibliográficas

1. Neidhardt, F. C., et ál., J. Bacteriol., 119, 736-747, 1974.
2. Brinkman, F. S. L., et ál., J. Bacteriol., 181, 4746-4754, 1999.
3. Silva, A. J. y Benítez, J. A., J. Bacteriol., 188, 794-800, 2006.

Tal como se indicó anteriormente, las enzimas KARI se han agrupado en diferentes clases. Las enzimas de *Pseudomonas* PF5 y PAO1 pertenecen al grupo de KARI de clase I, que es el grupo más grande de la familia, mientras que las enzimas de *E. coli* y *V. cholerae* pertenecen al grupo de KARI bacteriana de clase II.

Los ADN genómicos purificados de *P. aeruginosa* (PAO1, ATCC 47085) y *P. fluorescens* (PF5, ATCC BAA-477) se obtuvieron en ATCC (American Type Culture Collection, 10801 University Blvd., Manassas, VA). El ADN genómico de cada organismo (10 μ g cada uno) se rehidrató en 100 μ l de Tris-HCl 10 mM, pH 8,5 para su uso en una reacción

de PCR. Se emplearon los siguientes pares de cebadores, con un sitio SacI unido a los cebadores directos (SEQ ID NO:23 y 25) y un sitio HindIII unido a los cebadores inversos (SEQ ID NO:24 y 26), para amplificar las regiones codificadoras del gen *ilvC* a partir de los ADN genómicos de PAO1 y PF5 mediante PCR utilizando ADN polimerasa de alta fidelidad pfu-ultra (Stratagene). Los cebadores se diseñaron basándose en las secuencias públicas (GeneBank) de los genes *ilvC* de PF5 y PAO1 para estos organismos.

Cada 50 ul de la reacción de PCR contiene 1,0 ul de ADN genómico y 1,0 ul de cada uno de 10 pmol/ul de cebadores directos e inversos para los respectivos genes.

Las reacciones de PCR se realizaron en un gradiente de cicladores maestros Eppendorf (Eppendorf North America, Westbury, NY) utilizando las siguientes condiciones de reacción. La temperatura de inicio fue de 95 °C durante 2 min. Después se realizaron 5 ciclos de calentamiento/enfriamiento. Cada ciclo consiste en 95 °C durante 30 seg, 55 °C durante 30 seg, y 72 °C durante 1 min y 30 seg. Después se realizaron 25 ciclos más de calentamiento/enfriamiento. Cada uno de estos ciclos consiste en 95 °C durante 30 seg, 65 °C durante 30 seg, y 72 °C durante 1,0 min y 30 seg. Después de finalizar estos ciclos de temperatura, las muestras se mantuvieron a 72 °C durante 10 min más, y después se mantuvieron a la espera de la recuperación de la muestra a 4 °C.

Los fragmentos de PCR resultantes se digirieron con HindIII y SacI, se clonaron en el vector de expresión pBAD-SacI, y se transformaron en la cepa con *ilvC* inactivado BW25113($\Delta ilvC$) utilizando los procedimientos descritos en el ejemplo 1. Los clones positivos se identificaron mediante digestión con enzimas de restricción y se validaron mediante secuenciación de ADN de longitud completa empleando los cebadores SEQ ID NO:21 (pBAD-eF1), SEQ ID NO:22 (PALPK-R1), SEQ ID NO: 27 (PF5-S-F2), y SEQ ID NO: 28 (PF5-S-R2). Las cepas resultantes se denominaron BW25113($\Delta ilvC$)-PAO1-*ilvC* y BW25113($\Delta ilvC$)-PF5-*ilvC*.

La región codificadora del gen VC0162 de *V. cholerae* fue sometida a una optimización de codones para la expresión de *E. coli*, basándose en la secuencia conocida de la proteína (n.º de registro NP-229819.1) y se preparó mediante síntesis de genes personalizada sintética (DNA 2.0, Inc., Menlo Park, CA). Se preparó con sitios SacI y HindIII unidos a los extremos del gen. Este fragmento de ADN también se clonó en el vector de expresión pBAD-SacI utilizando sitios de restricción SacI y HindIII, y se transformó en la cepa con *ilvC* inactivado BW25113($\Delta ilvC$). La cepa resultante se denominó BW25113($\Delta ilvC$)-VCopt-VC0162. La secuencia de VC0162 con codones optimizados se indica en SEQ ID NO:30.

Ensayos de actividad KARI y de proteínas de las cepas K12, PAO1, PF5 y VC

Se prepararon extractos sin células de las cepas BW25113($\Delta ilvC$)-K12-*ilvC*, BW25113($\Delta ilvC$)-PAO1-*ilvC*, BW25113($\Delta ilvC$)-PF5-*ilvC* y BW25113($\Delta ilvC$)-VCopt-VC0162, que expresan la enzima KARI, utilizando BugBuster según se describe en el ejemplo 1. El ensayo KARI se realizó utilizando 188 ul de tampón de reacción, 2,0 ul de disolución madre de NADPH 20 mM, 5,0 ul de extracto de células al 20% diluido en tampón de ensayo, y 5,0 ul de acetolactato 90 mM. Por tanto, la disolución de ensayo final utilizada en este ejemplo contenía enzima, HEPES-potasio 100 mM, MgCl₂ 10 mM, NADPH 200 uM y acetolactato 2,25 mM.

La tabla 6 muestra las actividades específicas de KARI de cuatro organismos diferentes cultivados durante la noche en presencia de arabinosa al 0,02% (en p/v) como inductor. La cantidad de proteína total en el extracto celular y la actividad KARI se midieron según se describió anteriormente. Tal como se resume en la tabla 6, las enzimas KARI de los organismos identificados con unos tiempos de duplicación más rápidos cuando se cultivan en medio mínimo (tabla 5) tienen una actividad específica mayor que la KARI de *E. coli*. Cada uno de los extractos tiene unos niveles aproximadamente iguales de expresión de la proteína KARI, según se calcula mediante SDS-PAGE (los datos no se muestran). Estos resultados apoyan la hipótesis de que el tiempo de duplicación durante el crecimiento en medio mínimo puede utilizarse como medio para identificar enzimas KARI con mayor actividad específica.

Tabla 6

Comparación de las actividades específicas de KARI de diferentes organismos				
Cepa	PM (KDa)	Clase de KARI	Proteína total en el extracto celular, ug/ml	Actividad específica de KARI, umol/min/mg de proteína total
BW25113($\Delta ilvC$)-K12- <i>ilvC</i>	54	II	6693	0,72
BW25113($\Delta ilvC$)-VCopt-VC0162	54	II	6730	1,1
BW25113($\Delta ilvC$)-PAO1- <i>ilvC</i>	36	I	4988	1,2
BW25113($\Delta ilvC$)-PF5- <i>ilvC</i>	36	I	7671	1,8

Ejemplo 3

Análisis de la actividad específica de K12-KARI y PF5-KARI purificadas

Para resolver mejor los aumentos en la actividad específica de KARI observados con los extractos de células brutos en el ejemplo 2, K12-KARI y PF5-KARI se purificaron hasta la homogeneidad para permitir la cuantificación precisa de la concentración de las proteínas individuales y determinar la actividad específica de las enzimas KARI purificadas.

Purificación de K12-KARI y PF5-KARI

Se purificaron K12-KARI y PF5-KARI utilizando una columna Spin de intercambio aniónico débil Vivapure IEX D, miniH (Vivascience AG, Hannover, Alemania), seguido de una concentración en un dispositivo Microcon con un límite de exclusión molecular de 100 KDa (YM100, Millipore, Bedford, MA). El procedimiento de purificación se realizó a temperatura ambiente (22,5 °C).

Las disoluciones madre utilizadas en la columna Spin de intercambio aniónico fueron: HEPES-potasio 100 mM a pH 7,0, MgCl₂ 1,0 mM, EDTA 250 mM, Brij35 al 10% y KCl 2 M. El tampón de lavado (tampón A) se preparó añadiendo 5,0 ml de disolución madre de HEPES 100 mM a 15 ml de agua con la adición de 50 ul de disolución madre de MgCl₂, 20 ul de disolución madre de EDTA y 10 ul de Brij35 al 10%. El tampón de elución n.º 1 (tampón B) se preparó añadiendo 5,0 ml de HEPES 100 mM, 2,0 ml de disolución madre de KCl a 13 ml de agua con la adición de 50 ul de disolución madre de MgCl₂, 20 ul de disolución madre de EDTA y 10 ul de Brij35 al 10%. El tampón de elución n.º 2 (tampón C) se preparó añadiendo 5 ml de disolución madre de HEPES 100 mM, 5,0 ml de disolución madre de KCl a 10 ml de agua con la adición de 50 ul de disolución madre de MgCl₂, 20 ul de disolución madre de EDTA y 10 ul de Brij35 al 10%. La concentración final de KCl en el tampón B es de aproximadamente 200 mM y de aproximadamente 500 mM en el tampón C.

Se prepararon extractos sin células de las cepas BW25113(Δ ilvC)-K12-ilvC y BW25113(Δ ilvC)-PF5-ilvC utilizando BugBuster según se describe en el ejemplo 1. Para preparar el extracto celular diluido para cargarlo en las columnas Vivapure IEX D, se añadieron 600 ul agua doblemente ionizada a 200 ul del extracto.

Las columnas Vivapure IEX D primero se lavaron con 400 ul de tampón A mediante centrifugación (microcentrifuga Eppendorf, modelo 5415D) a 2000 x g durante 5 min. Se empleó un equipo y un proceso idénticos en el procedimiento de purificación de Vivapure IEX D completo. El extracto celular diluido (descrito anteriormente) se cargó sobre la columna y se centrifugó en dos lotes de 400 ul cada uno. La columna entonces se lavó (x2) con 400 ul de tampón A. Para la muestra de PF5-KARI, se cargaron 400 ul del tampón B para eluir la enzima de la columna hacia un tubo de recolección. Para la muestra de K12-KARI, se emplearon, en su lugar, 400 ul del tampón C.

Los dispositivos Microcon YM100 primero se lavaron con 400 ul de agua desionizada mediante centrifugación (microcentrifuga Eppendorf, modelo 5415D) a 13800 x g durante 5 min. La muestra recogida de la purificación de Vivapure IEX D después se cargó y se centrifugó a 13800 x g durante 4 min. Se rechazó la corriente y se añadieron 400 ul del tampón B a la cámara de muestras y se centrifugó a 13800 x g durante 4 min. El procedimiento de lavado se repitió (x2) antes de añadir 200 ul del tampón B a la cámara de muestras. La cámara de muestras se invirtió hacia un tubo de recolección limpio y se centrifugó a 5000 x g durante 2 min para recoger la muestra purificada.

La pureza de cada muestra de KARI purificada se validó mediante electroforesis capilar (Agilent 2100 Bioanalyzer, Agilent Technology, Santa Clara, CA). Las muestras se prepararon utilizando el kit de reactivo Protein 230 y se aplicaron a un Protein Labchip (suministrado con los kits de reactivo) siguiendo las instrucciones del fabricante, y se analizó con el Bioanalyzer. Se observó un único pico con poco fondo en el electrograma para cada muestra purificada.

Cuantificación de proteínas de las muestras KARI purificadas

Se realizó la medición de la absorción de UV de las muestras KARI purificada a 280 nm utilizando un espectrofotómetro (Agilent Technology, Santa Clara, CA) y cubetas de plástico desechables de 1 cm de longitud de paso (UVette, Eppendorf, Hamburgo, Alemania) para cuantificar la cantidad de KARI en las muestras purificadas. Los coeficientes de extinción a 280 nm para PF5-KARI (0,73 para 1 mg/ml) y K12-KARI (0,98 para 1 mg/ml) fueron predichos por el programa Protparam, disponible en el sitio web ExpAsy (Pace, C. N., *et ál.*, Protein Sci., 11, 2411-2423, 1995). La muestra purificada fue diluida hasta 20% (en v/v) en tampón B para la medición de la absorción de UV. La A₂₈₀ para la muestra de PF5-KARI diluida fue de 0,41 y para la K12-KARI diluida fue de 0,36.

Ensayo de actividad para KARI purificada

Las condiciones de ensayo utilizadas en este ejemplo fueron las mismas que en el ejemplo 2, excepto que se emplearon 5 ul de muestra purificada al 20% (en v/v) en lugar de extracto celular. Las concentraciones de proteína de las muestras purificadas y sus actividades específicas se muestran en la tabla 7. La actividad específica de PF5-KARI purificada, de aquel que crece con más rapidez de los ensayados, fue dos veces la actividad específica de K12-KARI. Estos resultados son coherentes con los datos obtenidos utilizando las preparaciones brutas de estas dos enzimas en el ejemplo 2, por lo que esto proporciona aún más apoyo a la hipótesis de que el tiempo de duplicación durante el crecimiento en medio mínimo puede utilizarse como medio para identificar enzimas KARI con

mayor actividad específica comparada con la enzima de *E. coli*.

Tabla 7

Concentración y actividad específica de KARI en cepas de <i>E.coli</i> y <i>Pseudomonas</i>		
Muestra	Concentración de KARI (mg/ml)	Actividad específica, umol/min/mg de KARI
K12-KARI	1,85	1,1
P5F-KARI	2,80	2,2

LISTADO DE SECUENCIAS

<110> E. I du Pont de Nemours and Co.
 <120> Producción fermentativa de isobutanol utilizando enzimas cetol-ácido reductoisomerasas muy activas
 <130> CL3761
 5 <160> 40
 <170> Patent In version 3. 5
 <210> 1
 <211> 1680
 <212> ADN
 10 <213> K. pneumoniae
 <400> 1

at ggacaaac agt at ccggt acgccagt gg ggcacggcg ccgat ct cgt cgt cagt cag 60
 ct ggaagct c agggagt acg ccaggt gt t c ggcat ccccg ggcctaaaat cgacaaggt c 120
 t t t gat t cac t gct ggat t c ct ccat t cgc at t at t ccgg t acgccacga agccaacgcc 180
 gcat t t at gg ccgccgccgt cggacgcat t accggcaaag cgggct ggc gct ggt cacc 240
 t ccggt ccgg gct gt t ccaa cct gat cacc ggcat ggcca ccgcaaacag cgaaggcgac 300
 ccggt ggt gg ccct gggcgg cgcggt aaaa cgcgccgat a aagcgaagca ggt ccaccag 360
 agt at ggat a cggg ggcat gt t cagcccgt gt caccctaat acgcat cga ggt gacggcg 420
 ccggt ggc t ggcggaagt ggt ct ccaac gcct t ccgcg ccgccgagca gggccggccg 480
 ggcagcgcgt t cgt t agcct gccgcaggat gt ggt cgat g gcccggt cag cggcaaagt g 540
 ct gccggcca gcggggcccc gcagat gggc gccgcgccgg at gat gccat cgaccaggt g 600
 gcaagct t a t cggccaggc gaagaacccg at ct t cct gc t cggcct gat ggccagccag 660
 ccgaaaaaca gcaaggcgct gcgccgt t t g ct ggagacca gccat at t cc agt caccagc 720
 acct at cagg ccgccggagc ggt gaat cag gat aact t ct ct cgct t cgc cggccgggt t 780
 gggct gt t t a acaaccaggc cggggaccgt ct gct gcagc t cggcgacct ggt gat ct gc 840
 at cggct aca gcccggt gga at acgaaccg gcgat gt gga acagcggcaa cgcgacgct g 900
 gt gcacat cg acgt gct gcc cgcct at gaa gagcgcaact acaccccgga t gt cgagct g 960
 gt gggcgat a t cgcggcac t ct caacaag ct ggcgcaaa at at cgat ca t cggct ggt g 1020
 ct ct ccccg c aggcggcgga gat cct ccgc gaccgccagc accagcgcga gct gct ggac 1080
 cggcgggcg cgcagct caa ccagt t t gcc ct gcac cccc t gcgcat cgt t cgcgcat g 1140
 caggat at cg t caacagcga cgt cacgt t g accgt ggaca t gggcagct t ccat at ct gg 1200
 at t gcccgct acct gt acac gt t ccgccc cgt cagggt ga t gat ct ccaa cggccagcag 1260
 accat gggcg t cgcct gcc ct gggct at c ggcgct ggc t ggt caat cc t gagcgcaaa 1320
 gt ggt ct ccg t ct ccggcga cggcggct t c ct gcagt cga gcat ggagct ggagaccgcc 1380
 gt ccgct ga aagcaaacgt gct gcat ct t at ct gggct cg at aacggct a caacat ggt c 1440
 gct at ccagg aagagaaaaa at at cagcgc ct gt ccggcg t cgagt t t gg gccgat ggat 1500
 t t t aaagcct at gccgaat c ct t cggcgcg aaagggt t t g ccgt ggaaag cggcgaggcg 1560
 ct ggagccga ccct gcgcgc ggcgat ggac gt cgacggcc cggcgggt agt ggccat cccg 1620
 gt ggat t at c gcgat aacc gct gct gat g ggccagct gc at ct gagt ca gat t ct gt aa 1680

ES 2 426 070 T3

<210> 2
 <211> 559
 <212> PRT
 <213> K. pneumoni ae

5 <400> 2

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

ES 2 426 070 T3

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

ES 2 426 070 T3

485 490 495

Gly Pro Met Asp Phe Lys Ala Tyr Ala Glu Ser Phe Gly Ala Lys Gly
500 505 510

Phe Ala Val Glu Ser Ala Glu Ala Leu Glu Pro Thr Leu Arg Ala Ala
515 520 525

Met Asp Val Asp Gly Pro Ala Val Val Ala Ile Pro Val Asp Tyr Arg
530 535 540

Asp Asn Pro Leu Leu Met Gly Gln Leu His Leu Ser Gln Ile Leu
545 550 555

<210> 3
<211> 1476
<212> ADN
<213> E. coli

5

<400> 3
at ggct aact act t caat ac act gaat ct g cgccagcagc t ggcacagct gggcaaat gt 60
cgct t t at gg gccgcgat ga at t cgccgat ggcgcgagct acct t caggg t aaaaaagt a 120
gt cat cgt cg gct gt ggcgc acagggt ct g aaccagggcc t gaacat gcg t gat t ct ggt 180
ct cgat at ct cct acgct ct gcgt aaagaa gcgat t gccg agaagcgcgc gt cct ggcgt 240
aaagcgaccg aaaat ggt t t t aaagt ggg t act t acgaag aact gat ccc acaggcggat 300
ct ggt gat t a acct gacgcc ggacaagcag cact ct gat g t agt gcgcac cgt acagcca 360
ct gat gaaag acggcgcggc gct gggct ac t cgcacggt t t caacat cgt cgaagt gggc 420
gagcagat cc gt aaagat at caccgt agt g at ggt t gcgc cgaaat gcc aggcaccgaa 480
gt gcgt gaag agt acaaacg t ggg t t cggc gt accgacgc t gat t gccgt t caccggaa 540
aacgat ccga aaggcgaagg cat ggcgat t gccaaagcct gggcggct gc aaccggt ggt 600
caccgt gcgg gt gt gct gga at cgt cct t c gt t gcggaag t gaaat ct ga cct gat gggc 660
gagcaaacca t cct gt gcgg t at gt t gcag gct ggct ct c t gct gt gct t cgacaagct g 720
gt ggaagaag gt accgat cc agcat acgca gaaaaact ga t t cagt t cgg t t gggaaacc 780
at caccgaag cact gaaaca gggcggcat c accct gat ga t ggaccgt ct ct ct aaccg 840
gcgaaact gc gt gct t at gc gct t t ct gaa cagct gaaag agat cat ggc acccct gt t c 900
cagaaacat a t ggacgacat cat ct ccggc gaat t ct ct t ccggt at gat ggcggact gg 960
gccaacgat g at aagaaact gct gacct gg cgt gaagaga ccggcaaac cgcgt t t gaa 1020
accgcgccgc agt at gaagg caaaat cggc gagcaggagt act t cgat aa aggcgt act g 1080
at gat t gcga t ggt gaaagc gggcgt t gaa ct ggcgt t cg aaacct ggt cgat t ccggc 1140
at cat t gaag agt ct gcat a t t at gaat ca ct gcacgagc t gccgct gat t gccaacacc 1200
at cgcccgt a agcgt ct gt a cgaaat gaac gt ggt t at ct ct gat accgc t gagt acggt 1260
aact at ct gt t ct ct t acgc t t gt gt gccg t t gct gaaac cgt t t at ggc agagct gaa 1320
ccgggcgacc t ggg t aaagc t at t ccggaa ggcgcggt ag at aacgggcā act gcgt gat 1380
gt gaacgaag cgat t cgcag ccat gcgat t gagcagg t ag gt aagaaact gcgcggct at 1440
at gacagat a t gaaacgt at t gct gt t gcg ggt t aa 1476

ES 2 426 070 T3

<210> 4
 <211> 491
 <212> PRT
 <213> E. coli

5 <400> 4

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

ES 2 426 070 T3

Leu 225 Oys Gly Met Leu 230 Gn Ala Gly Ser Leu 235 Leu Cys Phe Asp Lys Leu 240
 Val Gu Gu Gy Thr 245 Asp Pro Ala Tyr Ala 250 Gu Lys Leu Ile Gn Phe 255
 Gy Trp Gu Thr 260 Ile Thr Gu Ala Leu 265 Lys Gn Gy Gy Ile Thr Leu 270
 Met Met Asp 275 Arg Leu Ser Asn Pro Ala Lys Leu Arg Ala 285 Tyr Ala Leu
 Ser Gu Gn 290 Leu Lys Gu Ile 295 Met Ala Pro Leu Phe Gn Lys His Met
 Asp 305 Asp Ile Ile Ser Gy 310 Gu Phe Ser Ser Gy 315 Met Met Ala Asp Trp 320
 Ala Asn Asp Asp Lys 325 Lys Leu Leu Thr Trp 330 Arg Gu Gu Thr Gy 335 Lys
 Thr Ala Phe Gu 340 Thr Ala Pro Gn Tyr 345 Gu Gy Lys Ile Gy Gu Gn 350
 Gu Tyr Phe 355 Asp Lys Gy Val Leu Met Ile Ala Met Val 365 Lys Ala Gy
 Val Gu 370 Leu Ala Phe Gu Thr 375 Met Val Asp Ser Gy 380 Ile Ile Gu Gu
 Ser 385 Ala Tyr Tyr Gu 390 Ser Leu His Gu Leu Pro 395 Leu Ile Ala Asn Thr 400
 Ile Ala Arg Lys Arg 405 Leu Tyr Gu Met Asn 410 Val Val Ile Ser Asp Thr 415
 Ala Gu Tyr Gy 420 Asn Tyr Leu Phe Ser 425 Tyr Ala Cys Val Pro 430 Leu Leu
 Lys Pro Phe 435 Met Ala Gu Leu Gn Pro Gy Asp Leu Gy 445 Lys Ala Ile
 Pro Gu Gy Ala Val Asp Asn 455 Gy Gn Leu Arg Asp 460 Val Asn Gu Ala
 Ile Arg Ser His Ala 470 Ile Gu Gn Val Gy Lys 475 Lys Leu Arg Gy Tyr 480
 Met Thr Asp Met Lys 485 Arg Ile Ala Val Ala Gy 490

<210> 5
 <211> 1851
 <212> ADN

ES 2 426 070 T3

<213> E. coli

```

<400> 5
at gcct aagt accgt t ccgc caccaccact cat ggt cgt a at at ggccggg t gct cgt gcg 60
ct gt ggccgc ccaccggaat gaccgacgcc gat t t ccgt a agccgat t at cgccgt t gt g 120
aact cgt t ca cccaat t t gt accgggt cac gt ccat ct gc gcgat ct cgg t aaact ggt c 180
gccgaacaaa t t gaagcggc t ggccggcgt t gccaaagagt t caacacat t gcggg ggat 240
gat gggat t g ccat gggcca cgggggggat g ct t t at t cac t gccat ct cg cgaact gat c 300
gct gat t ccg t t gagt at at ggt caacgcc cact gcgccg acgccat ggt ct gcat ct ct 360
aact gcgaca aaat cacccc ggggat gct g at ggct t ccc t gcgcct gaa t at t ccggg t 420
at ct t t gt t t ccggcggccc gat ggaggcc gggaaaacca aact t t ccga t cagat cat c 480
aagct cgat c t ggt t gat gc gat gat ccag ggcgcagacc cgaaagt at c t gact cccag 540
agcgat cagg t t gaacgt t c cgct gt ccg acct gcggg t cct gct ccgg gat gt t t acc 600
gct aact caa t gaact gcct gaccgaagcg ct gggcct gt cgcagccggg caacggct cg 660
ct gct ggcaa cccacgccga ccgt aagcag ct gt t cct t a at gct ggt aa acgcat t gt t 720
gaat t gacca aacgt t at t a cgagcaaaac gacgaaagt g cact gccgcg t aat at cgcc 780
agt aaggcgg cgt t t gaaaa cgccat gacg ct ggat at cg cgat gggg gg at cgact aac 840
accgt act t c acct gct gcc ggcggcgcag gaagcggaaa t cgact t cac cat gagt gat 900
at cgat aagc t t t cccgcaa ggt t ccacag ct gt gt aaag t t gcgccgag caccagaaa 960
t accat at gg aagat gt t ca ccgt gct ggt ggt gt t at cg gt at t ct cgg cgaact ggat 1020
cgccgggggt t act gaaccg t gat gt gaaa aacgt act t g gcct gacgt t gccgcaaacg 1080
ct ggaacaat acgacgt t at gct gaccag gat gacgcgg t aaaaaat at gt t ccgcgca 1140
ggt cct gcag gcat t cgt ac cacacaggca t t ct cgcaag at t gccgt t g ggat acgct g 1200
gacgacgat c gcgccaat gg ct gt at ccgc t cgct ggaac acgcct acag caaagacggc 1260
ggcct ggcgg t gct ct acgg t aact t t gcg gaaaacggct gcat cgt gaa aacggcaggc 1320
gt cgat gaca gcat cct caa at t caccggc ccggcgaaag t gt acgaaag ccaggacgat 1380
gcggg agaag cgat t ct cgg ccgt aaagt t gt cgccggag at gt ggt agt aat t cgct at 1440
gaaggcccga aaggcggg cc ggggat gcag gaaat gct ct acccaaccag ct t cct gaaa 1500
t caat gggg c t cggcaaagc ct gt gcgct g at caccgacg gt cgt t t ct c t ggt ggcacc 1560
t ct ggt ct t t ccat cggcca cgt ct caccg gaagcggcaa gcggcggcag cat t ggct g 1620
at t gaagat g gt gacct gat cgct at cgac at cccgaacc gt ggcat t ca gt t acaggt a 1680
agcgat gccg aact ggcggc gcgt cgt gaa gcgcaggacg ct cgaggt ga caaagcct gg 1740
acgccgaaaa at cgt gaacg t caggt ct cc t t t gccct gc gt gct t at gc cagcct ggca 1800
accagcggc acaaaggcgc ggt gcgcat aaat cgaaac t ggggggt t a a 1851

```

<210> 6
 <211> 616
 <212> PRT
 <213> E. coli

5

<400> 6

ES 2 426 070 T3

Mèt Pro Lys Tyr Arg Ser Ala Thr Thr Thr His Gly Arg Asn Mèt Ala
 1 5 10 15
 Gly Ala Arg Ala Leu Trp Arg Ala Thr Gly Mèt Thr Asp Ala Asp Phe
 20 25 30
 Gly Lys Pro Ile Ile Ala Val Val Asn Ser Phe Thr Gn Phe Val Pro
 35 40 45
 Gly His Val His Leu Arg Asp Leu Gly Lys Leu Val Ala Gu Gn Ile
 50 55 60
 Gu Ala Ala Gly Gly Val Ala Lys Gu Phe Asn Thr Ile Ala Val Asp
 65 70 75 80
 Asp Gly Ile Ala Mèt Gly His Gly Gly Mèt Leu Tyr Ser Leu Pro Ser
 85 90 95
 Arg Gu Leu Ile Ala Asp Ser Val Gu Tyr Mèt Val Asn Ala His Cys
 100 105 110
 Ala Asp Ala Mèt Val Cys Ile Ser Asn Cys Asp Lys Ile Thr Pro Gly
 115 120 125
 Mèt Leu Mèt Ala Ser Leu Arg Leu Asn Ile Pro Val Ile Phe Val Ser
 130 135 140
 Gly Gly Pro Mèt Gu Ala Gly Lys Thr Lys Leu Ser Asp Gn Ile Ile
 145 150 155 160
 Lys Leu Asp Leu Val Asp Ala Mèt Ile Gn Gly Ala Asp Pro Lys Val
 165 170 175
 Ser Asp Ser Gn Ser Asp Gn Val Gu Arg Ser Ala Cys Pro Thr Cys
 180 185 190
 Gly Ser Cys Ser Gly Mèt Phe Thr Ala Asn Ser Mèt Asn Cys Leu Thr
 195 200 205
 Gu Ala Leu Gly Leu Ser Gn Pro Gly Asn Gly Ser Leu Leu Ala Thr
 210 215 220
 His Ala Asp Arg Lys Gn Leu Phe Leu Asn Ala Gly Lys Arg Ile Val
 225 230 235 240
 Gu Leu Thr Lys Arg Tyr Tyr Gu Gn Asn Asp Gu Ser Ala Leu Pro
 245 250 255

ES 2 426 070 T3

Arg Asn Ile Ala Ser Lys Ala Ala Phe Gu Asn Ala Met Thr Leu Asp
 260 265 270
 Ile Ala Met Gly Gly Ser Thr Asn Thr Val Leu His Leu Leu Ala Ala
 275 280 285
 Ala Gn Gu Ala Gu Ile Asp Phe Thr Met Ser Asp Ile Asp Lys Leu
 290 295 300
 Ser Arg Lys Val Pro Gn Leu Cys Lys Val Ala Pro Ser Thr Gn Lys
 305 310 315 320
 Tyr His Met Gu Asp Val His Arg Ala Gly Gly Val Ile Gly Ile Leu
 325 330 335
 Gly Gu Leu Asp Arg Ala Gly Leu Leu Asn Arg Asp Val Lys Asn Val
 340 345 350
 Leu Gly Leu Thr Leu Pro Gn Thr Leu Gu Gn Tyr Asp Val Met Leu
 355 360 365
 Thr Gn Asp Asp Ala Val Lys Asn Met Phe Arg Ala Gly Pro Ala Gly
 370 375 380
 Ile Arg Thr Thr Gn Ala Phe Ser Gn Asp Cys Arg Trp Asp Thr Leu
 385 390 395 400
 Asp Asp Asp Arg Ala Asn Gly Cys Ile Arg Ser Leu Gu His Ala Tyr
 405 410 415
 Ser Lys Asp Gly Gly Leu Ala Val Leu Tyr Gly Asn Phe Ala Gu Asn
 420 425 430
 Gly Cys Ile Val Lys Thr Ala Gly Val Asp Asp Ser Ile Leu Lys Phe
 435 440 445
 Thr Gly Pro Ala Lys Val Tyr Gu Ser Gn Asp Asp Ala Val Gu Ala
 450 455 460
 Ile Leu Gly Gly Lys Val Val Ala Gly Asp Val Val Val Ile Arg Tyr
 465 470 475 480
 Gu Gly Pro Lys Gly Gly Pro Gly Met Gn Gu Met Leu Tyr Pro Thr
 485 490 495
 Ser Phe Leu Lys Ser Met Gly Leu Gly Lys Ala Cys Ala Leu Ile Thr
 500 505 510
 Asp Gly Arg Phe Ser Gly Gly Thr Ser Gly Leu Ser Ile Gly His Val
 515 520 525

ES 2 426 070 T3

Ser Pro Glu Ala Ala Ser Gly Gly Ser Ile Gly Leu Ile Glu Asp Gly
530 535 540

Asp Leu Ile Ala Ile Asp Ile Pro Asn Arg Gly Ile Gn Leu Gn Val
545 550 555 560

Ser Asp Ala Glu Leu Ala Ala Arg Arg Glu Ala Gn Asp Ala Arg Gly
565 570 575

Asp Lys Ala Trp Thr Pro Lys Asn Arg Glu Arg Gn Val Ser Phe Ala
580 585 590

Leu Arg Ala Tyr Ala Ser Leu Ala Thr Ser Ala Asp Lys Gly Ala Val
595 600 605

Arg Asp Lys Ser Lys Leu Gly Gly
610 615

- <210> 7
- <211> 1662
- <212> ADN
- <213> Lactococcus lactis

5

<400> 7
t ct agacat a t gt at act gt gggggat t ac ct gct ggat c gcct gcacga act ggggat t 60
gaagaaat t t t cggt gt gcc aggcgat t at aacct gcagt t cct ggacca gat t at ct cg 120
caciaagat a t gaagt ggggt cggt aacgcc aacgaact ga acgcgagct a t at ggcagat 180
ggg t at gccc gt accaaaa agct gct gcg t t t ct gacga cct t t ggcgt t ggcgaact g 240
agcgccgt ca acggact ggc aggaagct ac gccgagaacc t gccagt t gt cgaaat t gt t 300
gggt cgcct a ct t ct aaggt t cagaat gaa ggcaaat t t g t gcacat ac t ct ggct gat 360
ggggat t t t a aacat t t t at gaaaat gcat gaaccggt t a ct gcggcccg cacgct gct g 420
acagcagaga at gct acggt t gagat cgac cgcgt cct gt ct gcgct gct gaaagagcgc 480
aagccggt at at at caat ct gcct gt cgat gt t gccgcag cgaaagccga aaagccgt cg 540
ct gccact ga aaaaagaaaa cagcacct cc aat acat cgg accaggaaat t ct gaat aaa 600
at ccaggaat cact gaagaa t gcgaagaaa ccgat cgt ca t caccggaca t gagat cat c 660
t ct t t t ggcc t ggaaaaaac ggt cacgcag t t cat t t ct a agaccaaact gcct at cacc 720
accct gaact t cggcaaat c t agcgt cgat gaagcgt gc cgagt t t t ct gggat at ct at 780
aat ggt accc t gt ccgaacc gaacct gaaa gaat t cgt cg aaagcgcgga ct t t at cct g 840
at gct gggcg t gaaact gac ggat agct cc acaggcgcgt t t accacca t ct gaacgag 900
aat aaaat ga t t t ccct gaa t at cgacgaa ggcaaat ct t t aacgagcg cat ccagaac 960
t t cgat t t t g aat ct ct gat t agt t cgct g ct ggat ct gt ccgaaat t ga gt at aaaggt 1020
aaat at at t g at aaaaaaca ggaggat t t t gt gccgt ct a at gcgct gct gagt caggat 1080
cgt ct gt ggc aagccgt aga aaacct gaca cagt ct aat g aaacgat t gt t gcggaacag 1140

ES 2 426 070 T3

ggaacttcat tttt cggcgc ctcatccatt tttctgaaat ccaaaagccā tttcattggc 1200
 caaccgctgt gggggagt at tggttat acc tttccggcgg cgctgggttc acagattgca 1260
 gat aaggaat cacgccat ct gctgtttat t ggtgacggca gcctgcagct gactgtccag 1320
 gaactggggc tggcgatccg tgaaaaaatc aatccgat t t gctttatcat caat aacgac 1380
 ggctacaccg t cgaacgcga aat tcatgga ccgaatcaaa gttacaatga catcccgatg 1440
 t ggaact at a gaaactgcc ggaatccttt ggcgcgacag aggatcgcgt ggtgagt aaa 1500
 at t g t g c g t a c g g a a a c g a a t t t g t g t c g g t t a t g a a a g a a g c g c a g g c t g a c c c g a a t 1560
 c g c a t g t a t t g g a t t g a a c t g a t c c t g g c a a a g a a g g c g c a c c g a a a g t t c t g a a a a a g 1620
 a t g g g g a a a c t g t t t g c g g a g c a a a a t a a a a g c t a a g g a t c c 1662

<210> 8
 <211> 548
 <212> PRT

5 <213> Lactococcus lactis

<400> 8

M e t T y r T h r V a l G l y A s p T y r L e u L e u A s p A r g L e u H i s G u L e u G l y
 1 5 10 15
 I l e G u G u I l e P h e G l y V a l P r o G l y A s p T y r A s n L e u G n P h e L e u
 20 25 30
 A s p G n I l e I l e S e r H i s L y s A s p M e t L y s T r p V a l G l y A s n A l a A s n
 35 40 45
 G u L e u A s n A l a S e r T y r M e t A l a A s p G l y T y r A l a A r g T h r L y s L y s
 50 55 60
 A l a A l a A l a P h e L e u T h r T h r P h e G l y V a l G l y G u L e u S e r A l a V a l
 65 70 75 80
 A s n G l y L e u A l a G l y S e r T y r A l a G u A s n L e u P r o V a l V a l G u I l e
 85 90 95
 V a l G l y S e r P r o T h r S e r L y s V a l G n A s n G u G l y L y s P h e V a l H i s
 100 105 110
 H i s T h r L e u A l a A s p G l y A s p P h e L y s H i s P h e M e t L y s M e t H i s G u
 115 120 125
 P r o V a l T h r A l a A l a A r g T h r L e u L e u T h r A l a G u A s n A l a T h r V a l
 130 135 140
 G u I l e A s p A r g V a l L e u S e r A l a L e u L e u L y s G u A r g L y s P r o V a l
 145 150 155 160
 T y r I l e A s n L e u P r o V a l A s p V a l A l a A l a A l a L y s A l a G u L y s P r o
 165 170 175

ES 2 426 070 T3

Ser Leu Pro Leu Lys Lys G u Asn Ser Thr Ser Asn Thr Ser Asp G n
180 185 190

G u Ile Leu Asn Lys Ile G n G u Ser Leu Lys Asn Ala Lys Lys Pro
195 200 205

Ile Val Ile Thr G y His G u Ile Ile Ser Phe G y Leu G u Lys Thr
210 215 220

Val Thr G n Phe Ile Ser Lys Thr Lys Leu Pro Ile Thr Thr Leu Asn
225 230 235 240

Phe G y Lys Ser Ser Val Asp G u Ala Leu Pro Ser Phe Leu G y Ile
245 250 255

Tyr Asn G y Thr Leu Ser G u Pro Asn Leu Lys G u Phe Val G u Ser
260 265 270

Ala Asp Phe Ile Leu Met Leu G y Val Lys Leu Thr Asp Ser Ser Thr
275 280 285

G y Ala Phe Thr His His Leu Asn G u Asn Lys Met Ile Ser Leu Asn
290 295 300

Ile Asp G u G y Lys Ile Phe Asn G u Arg Ile G n Asn Phe Asp Phe
305 310 315 320

G u Ser Leu Ile Ser Ser Leu Leu Asp Leu Ser G u Ile G u Tyr Lys
325 330 335

G y Lys Tyr Ile Asp Lys Lys G n G u Asp Phe Val Pro Ser Asn Ala
340 345 350

Leu Leu Ser G n Asp Arg Leu Trp G n Ala Val G u Asn Leu Thr G n
355 360 365

Ser Asn G u Thr Ile Val Ala G u G n G y Thr Ser Phe Phe G y Ala
370 375 380

Ser Ser Ile Phe Leu Lys Ser Lys Ser His Phe Ile G y G n Pro Leu
385 390 395 400

Trp G y Ser Ile G y Tyr Thr Phe Pro Ala Ala Leu G y Ser G n Ile
405 410 415

Ala Asp Lys G u Ser Arg His Leu Leu Phe Ile G y Asp G y Ser Leu
420 425 430

G n Leu Thr Val G n G u Leu G y Leu Ala Ile Arg G u Lys Ile Asn
435 440 445

ES 2 426 070 T3

Pro Ile Cys Phe Ile Ile Asn Asn Asp Gly Tyr Thr Val Glu Arg Glu
450 455 460

Ile His Gly Pro Asn Gln Ser Tyr Asn Asp Ile Pro Met Trp Asn Tyr
465 470 475 480

Ser Lys Leu Pro Glu Ser Phe Gly Ala Thr Glu Asp Arg Val Val Ser
485 490 495

Lys Ile Val Arg Thr Glu Asn Glu Phe Val Ser Val Met Lys Glu Ala
500 505 510

Gln Ala Asp Pro Asn Arg Met Tyr Trp Ile Glu Leu Ile Leu Ala Lys
515 520 525

Glu Gly Ala Pro Lys Val Leu Lys Lys Met Gly Lys Leu Phe Ala Glu
530 535 540

Gln Asn Lys Ser
545

<210> 9
<211> 1164
<212> ADN
<213> E. coli

5

<400> 9
at gaacaact t t aat ct gca caccccaacc cgcattctgt ttggt aaagg cgcaat cgct 60
ggtttacgcg aacaaattcc tcacgatgct cgcgtat t ga ttacct acgg cggcggcagc 120
gt gaaaaaaaa ccggcgt t ct cgat caagt t ct ggat gcc t gaaaggcat ggacgt gct g 180
gaat ttggcg gt at t gagcc aaaccggct t at gaaacgc t gat gaacgc cgt gaaact g 240
gt t cgcgaac agaaagt gac t t t cct gct g gcggt t ggcg gcggt t ct gt act ggacggc 300
accaaatt t a t cgccgcagc ggct aact at ccggaaaat a t cgat ccgt g gcacat t ct g 360
caaacgggcg gt aaagagat t aaaagcgcc at cccgat gg gct gt gt gct gacgct gcca 420
gcaaccgggt t cagaat ccaa cgcaggcgcg gt gat ct ccc gt aaaaccac aggcgacaag 480
caggcgt t cc at t ct gccca t gt t cagccg gt at t t gccg t gct cgat cc ggt t t at acc 540
t acaccct gc cgccgcgt ca ggt ggct aac ggcgt agt gg acgcct t t gt acacaccgt g 600
gaacagt at g t t accaaacc ggt t gat gcc aaaat t cagg accgt t t cgc agaaggcat t 660
t t gct gacgc t aat cgaaga t ggt ccgaaa gccct gaaag agccagaaaa ct acgat gt g 720
cgcgccaacg t cat gt gggc ggcgact cag gcgct gaacg gt t t gat t gg cgct ggcgt a 780
ccgcaggact gggcaacgca t at gct gggc cacgaact ga ct gcgat gca cggt ct ggat 840
cacgcgcaaa cact ggct at cgt cct gcct gcact gt gga at gaaaaacg cgat accaag 900
cgcgct aagc t gct gcaat a t gct gaacgc gt ct ggaaca t cact gaagg t t ccgat gat 960
gagcgt at t g acgccgcgat t gccgcaacc cgcaat t t ct t t gagcaat t aggcgt gccg 1020
accacct ct ccgact acgg t ct ggacggc agct ccat cc cggct t t gct gaaaaaact g 1080
gaagagcacg gcat gacca act gggcgaa aat cat gaca t t acgt t gga t gt cagccgc 1140
cgt at at acg aagccgcccg ct aa 1164

ES 2 426 070 T3

<210> 10
 <211> 387
 <212> PRT
 <213> E. coli

5 <400> 10
 Met Asn Asn Phe Asn Leu Hi s Thr Pro Thr Arg Ile Leu Phe Gly Lys
 1 5 10 15
 Gly Ala Ile Ala Gly Leu Arg Gu Gn Ile Pro Hi s Asp Ala Arg Val
 20 25 30
 Leu Ile Thr Tyr Gly Gly Gly Ser Val Lys Lys Thr Gly Val Leu Asp
 35 40 45
 Gn Val Leu Asp Ala Leu Lys Gly Met Asp Val Leu Gu Phe Gly Gly
 50 55 60
 Ile Gu Pro Asn Pro Ala Tyr Gu Thr Leu Met Asn Ala Val Lys Leu
 65 70 75 80
 Val Arg Gu Gn Lys Val Thr Phe Leu Leu Ala Val Gly Gly Gly Ser
 85 90 95
 Val Leu Asp Gly Thr Lys Phe Ile Ala Ala Ala Ala Asn Tyr Pro Gu
 100 105 110
 Asn Ile Asp Pro Trp Hi s Ile Leu Gn Thr Gly Gly Lys Gu Ile Lys
 115 120 125
 Ser Ala Ile Pro Met Gly Cys Val Leu Thr Leu Pro Ala Thr Gly Ser
 130 135 140
 Gu Ser Asn Ala Gly Ala Val Ile Ser Arg Lys Thr Thr Gly Asp Lys
 145 150 155 160
 Gn Ala Phe Hi s Ser Ala Hi s Val Gn Pro Val Phe Ala Val Leu Asp
 165 170 175
 Pro Val Tyr Thr Tyr Thr Leu Pro Pro Arg Gn Val Ala Asn Gly Val
 180 185 190
 Val Asp Ala Phe Val Hi s Thr Val Gu Gn Tyr Val Thr Lys Pro Val
 195 200 205
 Asp Ala Lys Ile Gn Asp Arg Phe Ala Gu Gly Ile Leu Leu Thr Leu
 210 215 220

ES 2 426 070 T3

Ile Gu Asp Gly Pro Lys Ala Leu Lys Gu Pro Gu Asn Tyr Asp Val
225 230 235 240

Arg Ala Asn Val Met Trp Ala Ala Thr Gn Ala Leu Asn Gly Leu Ile
245 250 255

Gly Ala Gly Val Pro Gn Asp Trp Ala Thr His Met Leu Gly His Gu
260 265 270

Leu Thr Ala Met His Gly Leu Asp His Ala Gn Thr Leu Ala Ile Val
275 280 285

Leu Pro Ala Leu Trp Asn Gu Lys Arg Asp Thr Lys Arg Ala Lys Leu
290 295 300

Leu Gn Tyr Ala Gu Arg Val Trp Asn Ile Thr Gu Gly Ser Asp Asp
305 310 315 320

Gu Arg Ile Asp Ala Ala Ile Ala Ala Thr Arg Asn Phe Phe Gu Gn
325 330 335

Leu Gly Val Pro Thr His Leu Ser Asp Tyr Gly Leu Asp Gly Ser Ser
340 345 350

Ile Pro Ala Leu Leu Lys Lys Leu Gu Gu His Gly Met Thr Gn Leu
355 360 365

Gly Gu Asn His Asp Ile Thr Leu Asp Val Ser Arg Arg Ile Tyr Gu
370 375 380

Ala Ala Arg
385

<210> 11

<211> 44

<212> ADN

5 <213> secuencia artificial

<220>

<223> cebador

<400> 11

gagct cct t a agaaggaggt aat caccat g gctaactact tcaa 44

10 <210> 12

<211> 51

<212> ADN

<213> secuencia artificial

<220>

15 <223> cebador

<400> 12

ggatccgat c gagct agcgc ggccgct t aa cccgcaacag caatacgtttc 51

<210> 13

<211> 35

20 <212> ADN

<213> secuencia artificial

ES 2 426 070 T3

<220>
 <223> cebador

 <400> 13
 gct aacagga ggaagagct c at ggcaccct cgctc 35

 5 <210> 14
 <211> 35
 <212> ADN
 <213> secuencia artificial

 <220>
 10 <223> cebador

 <400> 14
 gagcgagggt gccat gagct cttcctcctgtagc 35

 <210> 15
 <211> 44
 15 <212> ADN
 <213> secuencia artificial

 <220>
 <223> cebador

 <400> 15
 20 at caccgagc t cat ggct aa ctactcaat aactgaat c tgcg 44

 <210> 16
 <211> 46
 <212> ADN
 <213> secuencia artificial

 25 <220>
 <223> cebador

 <400> 16
 ggccgcaagc ttttaaccg caacagcaat acgttccatatctgctc 46

 <210> 17
 30 <211> 20
 <212> ADN
 <213> secuencia artificial

 <220>
 <223> cebador

 35 <400> 17
 ccgt aaagat at caccgt ag 20

 <210> 18
 <211> 20
 <212> ADN
 40 <213> secuencia artificial

 <220>
 <223> cebador

 <400> 18
 cagtatgaag gcaaaatcgg 20

 45 <210> 19
 <211> 20
 <212> ADN
 <213> secuencia artificial

 <220>
 50 <223> cebador

ES 2 426 070 T3

	<400> 19 cgt act cagc ggt at cagag	20
5	<210> 20 <211> 20 <212> ADN <213> secuencia artificial <220> <223> cebador	
10	<400> 20 cagattcac ttccgcaacg	20
15	<210> 21 <211> 29 <212> ADN <213> secuencia artificial <220> <223> cebador	
	<400> 21 cgcaact ct tactgttct ccatacccg	29
20	<210> 22 <211> 25 <212> ADN <213> secuencia artificial <220> <223> cebador	
25	<400> 22 accgct t ct g cgt tctgatt t aat c	25
30	<210> 23 <211> 44 <212> ADN <213> secuencia artificial <220> <223> cebador	
35	<400> 23 caaaacagcc aagcttttag t t ct t gct ct tgcgacgat ctg	44
40	<210> 24 <211> 43 <212> ADN <213> secuencia artificial <220> <223> cebador	
45	<400> 24 caggaggaag agct cat gcg cgtttctac gat aaagact gtg	43
50	<210> 25 <211> 44 <212> ADN <213> secuencia artificial <220> <223> cebador	
	<400> 25 caaaacagcc aagcttttag ttctggctt tgcgacgat ttg	44

ES 2 426 070 T3

<210> 26
 <211> 45
 <212> ADN
 <213> secuencia artificial

5 <220>
 <223> cebador

<400> 26
 caggaggaag agct cat gaa agt t t t ct ac gat aaagact gcgac 45

10 <210> 27
 <211> 20
 <212> ADN
 <213> secuencia artificial

<220>
 <223> cebador

15 <400> 27
 gat cat gat gcgccgaagg 20

<210> 28
 <211> 20
 <212> ADN

20 <213> secuencia artificial

<220>
 <223> cebador

<400> 28
 ct gct caccg aacaggt cgg 20

25 <210> 29
 <211> 1476
 <212> ADN
 <213> secuencia genoma c

<400> 29

at ggct aact act t caat ac act gaat ct g cgccagcagc tggcacagct gggcaaat gt 60

cgct t t at gg gccgcgat ga at t cgccgat ggcgcgagct acct t caggg t aaaaaagt a 120

gt cat cgt cg gct gt ggcgc acagggt ct g aaccagggcc t gaacat gcg t gat t ct ggt 180

ct cgat at ct cct acgct ct gcgt aaagaa gcgat t gccg agaagcgcgc gt cct ggcgt 240

30 aaagcgaccg aaaat ggt t t t aaagt gggg act t acgaag aact gat ccc acaggcggat 300

ES 2 426 070 T3

ct ggt gat t a acct gacgcc ggacaagcag cact ct gat g t agt gcgca^c cgt acagcca 360
ct gat gaaag acggcgcggc gct gggct ac t cgcacggt t t caacat cgt cgaagt gggc 420
gagcagat cc gt aaagat at caccgt agt g at ggt t gcgc cgaaat gccc aggaccgaa 480
gt gcgt gaag agt acaaacg t ggggt t cggc gt accgacgc t gat t gccgt t caccggaa 540
aacgat ccga aaggcgaagg cat ggcgat t gccaaagcct gggcggct gc aaccggt ggt 600
caccgt gcgg gt gt gct gga at cgt cct t c gt t gcggaag t gaaat ct ga cct gat gggc 660
gagcaaacca t cct gt gcgg t at gt t gcag gct ggct ct c t gct gt gct t cgacaagct g 720
gt ggaagaag gt accgat cc agcat acgca gaaaaact ga t t cagt t cgg t t gggaaacc 780
at caccgaag cact gaaaca gggcggcat c accct gat ga t ggaccgt ct ct ct aaccg 840
gcgaaact gc gt gct t at gc gct t t ct gaa cagct gaaag agat cat ggc acccct gt t c 900
cagaaacat a t ggacgacat cat ct ccggc gaat t ct ct t ccggt at gat ggcggact gg 960
gccaacgat g at aagaaact gct gacct gg cgt gaagaga ccggcaaac cgcgt t t gaa 1020
accgcgccgc agt at gaagg caaaat cggc gagcaggagt act t cgat aa aggct act g 1080
at gat t gcga t ggt gaaagc gggcgt t gaa ct ggcgt t cg aaacct ggt cgat t ccggc 1140
at cat t gaag agt ct gcat a t t at gaat ca ct gcacgagc t gccgct gat t gccaacacc 1200
at cgcccgt a agcgt ct gt a cgaaat gaac gt ggt t at ct ct gat accgc t gagt accgt 1260
aact at ct gt t ct ct t acgc t t gt gt gccg t t gct gaaac cgt t t at ggc agagct gcaa 1320
ccgggcgacc t ggggt aaagc t at t ccggaa ggcgcggt ag at aacgggca act gcgt gat 1380
gt gaacgaag cgat t cgcag ccat gcgat t gagcaggt ag gt aagaaact gcgcggct at 1440
at gacagat a t gaaacgt at t gct gt t gcg ggt t aa 1476

<210> 30
<211> 1485
<212> ADN
<213> vibrio

5

<400> 30
at ggct aact at t t t aacac t ct gaat ct g cgcaacagc t ggat cagct ggggt cgt t gt 60
cgct t cat gg cgcgcgagga at t cgcaact gaagcggact acct gaaggg t aagaaagt g 120
gt cat cgt ag gct gt ggcgc gcagggcct g aaccagggt c t gaat at gcg t gat agcggc 180
ct ggacgt gt ct t acgcact gcgt caagcg gccat cgat g aacagcgt ca gt ct t t caaa 240
aat gct aaga acaacggt t t caat gt aggt t ct t acgaac agct gat t cc gacggcggac 300
ct ggt t at ca at ct gacccc ggat aaacag cacaccagcg t ggt t aacgc ggt t at gccg 360
ct gat gaaac agggcgct gc t ct gggct at agccat ggt t t t aacat cgt ggaagagggc 420
at gcagat t c gt aaagacat cacggt agt a at ggt ggcgc cgaaat gccc aggt act gag 480
gt t cgt gagg aat acaagcg t ggt t t t gcc gt t cct accc t gat t gcagt gcaccggaa 540
aacgat ccgc agggcgaagg t t gggaaat c gct aaggcat gggct gcagc cact ggt ggc 600
caccgcgct g gt t gcct ggc ct ct agct t c gt cgcagaag t aaaaagcga cct gat gggc 660

ES 2 426 070 T3

gaacagacca t cct gt gt gg t at gct gcag gccggg agca t cgt t t gct ā cgaaaagat g 720
 gt t gct gat g gt at cgaccc gggct acgcg ggcaact gc t gcagt t t gg t t gggaaact 780
 at t accgagg ct ct gaaat t cggg ggcat c acccacat ga t ggaccgt ct gagcaaccg 840
 gcaaagat ca aagcat t cga act gt ct gaa gagct gaaag acct gat gcg t ccgct gt ac 900
 aacaagcaca t ggacgacat t at t t ccggc cact t t t ct t ct accat gat ggcggact gg 960
 gccaacgcg acaaggat ct gt t cggct gg cgt gcggaga ccgct gaaac cgcat t t gaa 1020
 aact at ccga ccact gacgt t aagat cgcg gagcaggaat act t t gacaa cggg at cct g 1080
 at gat t gcaa t ggt t cgcgc gggcgt t gaa ct ggcgt t t g aagcgat gac cgct t ct ggt 1140
 at cat cgat g aat ct gcat a ct at gagt cc ct gcacgagc t gccgct gat cgct aacacg 1200
 gt ggcccgc aacgt ct gt a t gaaat gaat gt cgt aat ct ct gacacggc t gaat at ggt 1260
 aat t at ct gt t cgccaat gt ggcagt t cct ct gct gcgt g agaaat t cat gccgaaagt g 1320
 ggt accgat g t t at t ggt aa aggt ct gggg gt cgt t t ct a accaggt aga t aacgcgact 1380
 ct gat cgaag t gaact ccat cat ccgt aac caccggg gg aat acat t gg t gaagaact g 1440
 cgt ggt t at a t gaaggacat gaagcgt at c gccgt t ggt g at t aa 1485

<210> 31
 <211> 1014
 <212> ADN
 <213> Pseudomonas

5

<400> 31
 at gcgcgt t t t ct acgat aa agact gt gac ct ct cgat ca t ccagggcaa gaaagt t gcc 60
 at cat cggct acggct ccca gggccacgcc cat gcct gca acct gaagga ct ccggcgt c 120
 gacgt caccg t gggcct gcg t agcggct cc gccaccgt gg ccaaggccga agcgcacggg 180
 ct gaagggt t g ccgacgt gaa gaccgccgt c gccgcagccg acgt ggt cat gat cct cacc 240
 ccggacgagt t ccagggccg cct gt acaag gaagagat cg agccgaacct gaagaagggc 300
 gccaccct gg cct t cgct ca cggct t ct cc at ccaact aca accaggt cgt cccgcgcgcc 360
 gacct cgacg t gat cat gat cgcgccgaag gcaccgggt c acaccgt gcg t t ccgagt t c 420
 gt caagggcg gt ggcat ccc t gacct gat c gccat ct acc aggacgt t c cggcaacgcc 480
 aagaacgt cg ccct gt cct a cgcct gcggc gt cggcggcg gt cgt accgg t at cat cgaa 540
 accacct t ca aggacgagac cgaaaccgac ct gt t cggg g agcaggccgt t ct ct gcggt 600
 ggt t gcgt cg agct ggt caa ggccggg t t c gaaaccct gg t cgaagccgg t t acgcgccg 660
 gaaat ggcct act t cgagt g cct gcacgag ct gaagct ga t cgt cgacct gat gt acgaa 720
 ggcggcat cg ccaacat gaa ct act ccat c t ccaacaat g ccgaat acgg t gagt acgt a 780
 accggg ccgg aggt gat caa cgccgagt cc cgt gct gccat t gcgcaacgc cct gaagcgc 840
 at ccaggacg gcgagt acgc gaaaat gt t c at t accgaag gt gcggccaa ct acccgt cg 900
 at gact gcct accgccgcaa caacgccgt caccggat cg agcagat cgg cgagaagct g 960
 cgcgcgat ga t gccgt ggat cgcagccaac aagat cgt cg acaagagcaa gaac 1014

<210> 32
 <211> 1014
 <212> ADN
 <213> Pseudomonas fluorescens

10

ES 2 426 070 T3

<400> 32
at gaaagt t t t ct acgat aa agact gcgac ct gt cgat ca t ccaaggt aa gaaagt t gcc 60
at cat cggct acggt t ccca gggccacgct caagcat gca acct gaagga t t cggcgt a 120
gacgt gact g t t ggcct gcg t aaaggct cg gct accgt t g ccaaggct ga agcccacggc 180
t t gaaagt ga cgcacgt t gc t gcagccgt t gccggt gccg act t ggt cat gat cct gacc 240
ccggacgagt t ccagt ccca gct gt acaag aacgaaat cg agccgaacat caagaagggc 300
gccact ct gg cct t ct ccca cggct t cgcg at ccaact aca accaggt t gt gcct cgt gcc 360
gacct cgacg t gat cat gat cgcgccgaag gct ccaggcc acaccgt acg t t ccgagt t c 420
gt caagggcg gt ggt at t cc t gacct gat c gcgat ct acc aggacgct t c cggcaacgcc 480
aagaacgt t g ccct gt cct a cgccgcaggc gt gggcggcg gccgt accgg cat cat cgaa 540
accacct t ca aggacgagac t gaaccgac ct gt t cggt g agcaggct gt t ct gt gt ggc 600
ggt accgt cg agct ggt caa agccggt t t c gaaaccct gg t t gaagct gg ct acgct cca 660
gaaat ggcct act t cgagt g cct gcacgaa ct gaagct ga t cgt t gacct cat gt acgaa 720
ggcgggt at cg ccaacat gaa ct act cgat c t ccaacaacg ct gaat acgg cgagt acgt g 780
act ggt ccag aagt cat caa cgccgaat cc cgt caggcca t gcgcaat gc t ct gaagcgc 840
at ccaggacg gcgaat acgc gaagat gt t c at cagcgaag gcgct accgg ct accat cg 900
at gaccgcca agcgt cgt aa caacgct gct cacggt at cg aaat cat cgg cgagcaact g 960
cgct cgat ga t gcct t ggat cggg gccaac aaaat cgt cg acaaagccaa gaac 1014

<210> 33
<211> 491
5 <212> PRT
<213> E. coli

<400> 33
Mèt Al a Asn Tyr Phe Asn Thr Leu Asn Leu Arg G n G n Leu Al a G n
1 5 10 15
Leu G y Lys Cys Arg Phe Mèt G y Arg Asp G u Phe Al a Asp G y Al a
20 25 30
Ser Tyr Leu G n G y Lys Lys Val Val Il e Val G y Cys G y Al a G n
35 40 45
G y Leu Asn G n G y Leu Asn Mèt Arg Asp Ser G y Leu Asp Il e Ser
50 55 60
Tyr Al a Leu Arg Lys G u Al a Il e Al a G u Lys Arg Al a Ser Trp Arg
65 70 75 80

ES 2 426 070 T3

Lys Ala Thr Gu Asn Gly Phe Lys Val Gly Thr Tyr Gu Gu Leu Ile
 85 90 95
 Pro Gn Ala Asp Leu Val Ile Asn Leu Thr Pro Asp Lys Gn His Ser
 100 105 110
 Asp Val Val Arg Thr Val Gn Pro Leu Met Lys Asp Gly Ala Ala Leu
 115 120 125
 Gly Tyr Ser His Gly Phe Asn Ile Val Gu Val Gly Gu Gn Ile Arg
 130 135 140
 Lys Asp Ile Thr Val Val Met Val Ala Pro Lys Cys Pro Gly Thr Gu
 145 150 155 160
 Val Arg Gu Gu Tyr Lys Arg Gly Phe Gly Val Pro Thr Leu Ile Ala
 165 170 175
 Val His Pro Gu Asn Asp Pro Lys Gly Gu Gly Met Ala Ile Ala Lys
 180 185 190
 Ala Trp Ala Ala Ala Thr Gly Gly His Arg Ala Gly Val Leu Gu Ser
 195 200 205
 Ser Phe Val Ala Gu Val Lys Ser Asp Leu Met Gly Gu Gn Thr Ile
 210 215 220
 Leu Cys Gly Met Leu Gn Ala Gly Ser Leu Leu Cys Phe Asp Lys Leu
 225 230 235 240
 Val Gu Gu Gly Thr Asp Pro Ala Tyr Ala Gu Lys Leu Ile Gn Phe
 245 250 255
 Gly Trp Gu Thr Ile Thr Gu Ala Leu Lys Gn Gly Gly Ile Thr Leu
 260 265 270
 Met Met Asp Arg Leu Ser Asn Pro Ala Lys Leu Arg Ala Tyr Ala Leu
 275 280 285
 Ser Gu Gn Leu Lys Gu Ile Met Ala Pro Leu Phe Gn Lys His Met
 290 295 300
 Asp Asp Ile Ile Ser Gly Gu Phe Ser Ser Gly Met Met Ala Asp Trp
 305 310 315 320
 Ala Asn Asp Asp Lys Lys Leu Leu Thr Trp Arg Gu Gu Thr Gly Lys
 325 330 335
 Thr Ala Phe Gu Thr Ala Pro Gn Tyr Gu Gly Lys Ile Gly Gu Gn
 340 345 350
 Gu Tyr Phe Asp Lys Gly Val Leu Met Ile Ala Met Val Lys Ala Gly

ES 2 426 070 T3

355 360 365

Val Glu Leu Ala Phe Glu Thr Met Val Asp Ser Gly Ile Ile Glu Glu
 370 375 380

Ser Ala Tyr Tyr Glu Ser Leu His Glu Leu Pro Leu Ile Ala Asn Thr
 385 390 395 400

Ile Ala Arg Lys Arg Leu Tyr Glu Met Asn Val Val Ile Ser Asp Thr
 405 410 415

Ala Glu Tyr Gly Asn Tyr Leu Phe Ser Tyr Ala Cys Val Pro Leu Leu
 420 425 430

Lys Pro Phe Met Ala Glu Leu Gln Pro Gly Asp Leu Gly Lys Ala Ile
 435 440 445

Pro Glu Gly Ala Val Asp Asn Gly Gln Leu Arg Asp Val Asn Glu Ala
 450 455 460

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

Met Thr Asp Met Lys Arg Ile Ala Val Ala Gly
 485 490

<210> 34
 <211> 494
 <212> PRT
 5 <213> Vibri on

<400> 34
 Met Ala Asn Tyr Phe Asn Thr Leu Asn Leu Arg Glu Gln Leu Asp Gln
 1 5 10 15

Leu Gly Arg Cys Arg Phe Met Ala Arg Glu Glu Phe Ala Thr Glu Ala
 20 25 30

Asp Tyr Leu Lys Gly Lys Lys Val Val Ile Val Gly Cys Gly Ala Gln
 35 40 45

Gly Leu Asn Gln Gly Leu Asn Met Arg Asp Ser Gly Leu Asp Val Ser
 50 55 60

Tyr Ala Leu Arg Gln Ala Ala Ile Asp Glu Gln Arg Gln Ser Phe Lys
 65 70 75 80

Asn Ala Lys Asn Asn Gly Phe Asn Val Gly Ser Tyr Glu Gln Leu Ile
 85 90 95

Pro Thr Ala Asp Leu Val Ile Asn Leu Thr Pro Asp Lys Gln His Thr
 100 105 110

ES 2 426 070 T3

Ser Val Val Asn Ala Val Met Pro Leu Met Lys Gln Gly Ala Ala Leu
115 120 125

Gly Tyr Ser His Gly Phe Asn Ile Val Glu Glu Gly Met Gln Ile Arg
130 135 140

Lys Asp Ile Thr Val Val Met Val Ala Pro Lys Cys Pro Gly Thr Glu
145 150 155 160

Val Arg Glu Glu Tyr Lys Arg Gly Phe Gly Val Pro Thr Leu Ile Ala
165 170 175

Val His Pro Glu Asn Asp Pro Gln Gly Glu Gly Trp Glu Ile Ala Lys
180 185 190

Ala Trp Ala Ala Ala Thr Gly Gly His Arg Ala Gly Cys Leu Ala Ser
195 200 205

Ser Phe Val Ala Glu Val Lys Ser Asp Leu Met Gly Glu Gln Thr Ile
210 215 220

Leu Cys Gly Met Leu Gln Ala Gly Ser Ile Val Cys Tyr Glu Lys Met
225 230 235 240

Val Ala Asp Gly Ile Asp Pro Gly Tyr Ala Gly Lys Leu Leu Gln Phe
245 250 255

Gly Trp Glu Thr Ile Thr Glu Ala Leu Lys Phe Gly Gly Ile Thr His
260 265 270

Met Met Asp Arg Leu Ser Asn Pro Ala Lys Ile Lys Ala Phe Glu Leu
275 280 285

Ser Glu Glu Leu Lys Asp Leu Met Arg Pro Leu Tyr Asn Lys His Met
290 295 300

Asp Asp Ile Ile Ser Gly His Phe Ser Ser Thr Met Met Ala Asp Trp
305 310 315 320

Ala Asn Asp Asp Lys Asp Leu Phe Gly Trp Arg Ala Glu Thr Ala Glu
325 330 335

Thr Ala Phe Glu Asn Tyr Pro Thr Thr Asp Val Lys Ile Ala Glu Gln
340 345 350

Glu Tyr Phe Asp Asn Gly Ile Leu Met Ile Ala Met Val Arg Ala Gly
355 360 365

Val Glu Leu Ala Phe Glu Ala Met Thr Ala Ser Gly Ile Ile Asp Glu
370 375 380

ES 2 426 070 T3

Ser Ala Tyr Tyr Glu Ser Leu His Glu Leu Pro Leu Ile Ala Asn Thr
385 390 395 400

Val Ala Arg Lys Arg Leu Tyr Glu Met Asn Val Val Ile Ser Asp Thr
405 410 415

Ala Glu Tyr Gly Asn Tyr Leu Phe Ala Asn Val Ala Val Pro Leu Leu
420 425 430 435

Arg Glu Lys Phe Met Pro Lys Val Gly Thr Asp Val Ile Gly Lys Gly
435 440 445

Leu Gly Val Val Ser Asn Glu Val Asp Asn Ala Thr Leu Ile Glu Val
450 455 460

Asn Ser Ile Ile Arg Asn His Pro Val Glu Tyr Ile Gly Glu Glu Leu
465 470 475 480

Arg Gly Tyr Met Lys Asp Met Lys Arg Ile Ala Val Gly Asp
485 490

<210> 35
<211> 338
<212> PRT
5 <213> Pseudomonas

<400> 35
Met Arg Val Phe Tyr Asp Lys Asp Cys Asp Leu Ser Ile Ile Glu Gly
1 5 10 15

Lys Lys Val Ala Ile Ile Gly Tyr Gly Ser Glu Gly His Ala His Ala
20 25 30

Cys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Ser
35 40 45

Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Ala
50 55 60

Asp Val Lys Thr Ala Val Ala Ala Ala Asp Val Val Met Ile Leu Thr
65 70 75 80

Pro Asp Glu Phe Glu Gly Arg Leu Tyr Lys Glu Glu Ile Glu Pro Asn
85 90 95

Leu Lys Lys Gly Ala Thr Leu Ala Phe Ala His Gly Phe Ser Ile His
100 105 110

Tyr Asn Glu Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
115 120 125

Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
130 135 140

ES 2 426 070 T3

G y I l e P r o A s p L e u I l e A l a I l e T y r G n A s p A l a S e r G y A s n A l a
145 150 155 160

L y s A s n V a l A l a L e u S e r T y r A l a C y s G y V a l G y G y G y A r g T h r
165 170 175

G y I l e I l e G u T h r T h r P h e L y s A s p G u T h r G u T h r A s p L e u P h e
180 185 190

G y G u G n A l a V a l L e u C y s G y G y C y s V a l G u L e u V a l L y s A l a
195 200 205

G y P h e G u T h r L e u V a l G u A l a G y T y r A l a P r o G u M e t A l a T y r
210 215 220

P h e G u C y s L e u H i s G u L e u L y s L e u I l e V a l A s p L e u M e t T y r G u
225 230 235 240

G y G y I l e A l a A s n M e t A s n T y r S e r I l e S e r A s n A s n A l a G u T y r
245 250 255

G y G u T y r V a l T h r G y P r o G u V a l I l e A s n A l a G u S e r A r g A l a
260 265 270

A l a M e t A r g A s n A l a L e u L y s A r g I l e G n A s p G y G u T y r A l a L y s
275 280 285

M e t P h e I l e T h r G u G y A l a A l a A s n T y r P r o S e r M e t T h r A l a T y r
290 295 300

A r g A r g A s n A s n A l a A l a H i s P r o I l e G u G n I l e G y G u L y s L e u
305 310 315 320

A r g A l a M e t M e t P r o T r p I l e A l a A l a A s n L y s I l e V a l A s p L y s S e r
325 330 335

Lys Asn

<210> 36

<211> 338

<212> PRT

5 <213> Pseudomonas

<400> 36

M e t L y s V a l P h e T y r A s p L y s A s p C y s A s p L e u S e r I l e I l e G n G y
1 5 10 15

L y s L y s V a l A l a I l e I l e G y T y r G y S e r G n G y H i s A l a G n A l a
20 25 30

ES 2 426 070 T3

Oys Asn Leu Lys Asp Ser Gly Val Asp Val Thr Val Gly Leu Arg Lys
 35 40 45
 Gly Ser Ala Thr Val Ala Lys Ala Glu Ala His Gly Leu Lys Val Thr
 50 55 60
 Asp Val Ala Ala Ala Val Ala Gly Ala Asp Leu Val Met Ile Leu Thr
 65 70 75 80
 Pro Asp Glu Phe Gn Ser Gn Leu Tyr Lys Asn Glu Ile Glu Pro Asn
 85 90 95
 Ile Lys Lys Gly Ala Thr Leu Ala Phe Ser His Gly Phe Ala Ile His
 100 105 110
 Tyr Asn Gn Val Val Pro Arg Ala Asp Leu Asp Val Ile Met Ile Ala
 115 120 125
 Pro Lys Ala Pro Gly His Thr Val Arg Ser Glu Phe Val Lys Gly Gly
 130 135 140
 Gly Ile Pro Asp Leu Ile Ala Ile Tyr Gn Asp Ala Ser Gly Asn Ala
 145 150 155 160
 Lys Asn Val Ala Leu Ser Tyr Ala Ala Gly Val Gly Gly Gly Arg Thr
 165 170 175
 Gly Ile Ile Gu Thr Thr Phe Lys Asp Glu Thr Gu Thr Asp Leu Phe
 180 185 190
 Gly Gu Gn Ala Val Leu Oys Gly Gly Thr Val Gu Leu Val Lys Ala
 195 200 205
 Gly Phe Gu Thr Leu Val Gu Ala Gly Tyr Ala Pro Gu Met Ala Tyr
 210 215 220
 Phe Gu Oys Leu His Gu Leu Lys Leu Ile Val Asp Leu Met Tyr Gu
 225 230 235 240
 Gly Gly Ile Ala Asn Met Asn Tyr Ser Ile Ser Asn Asn Ala Gu Tyr
 245 250 255
 Gly Gu Tyr Val Thr Gly Pro Gu Val Ile Asn Ala Gu Ser Arg Gn
 260 265 270
 Ala Met Arg Asn Ala Leu Lys Arg Ile Gn Asp Gly Gu Tyr Ala Lys
 275 280 285
 Met Phe Ile Ser Gu Gly Ala Thr Gly Tyr Pro Ser Met Thr Ala Lys
 290 295 300
 Arg Arg Asn Asn Ala Ala His Gly Ile Gu Ile Ile Gly Gu Gn Leu

ES 2 426 070 T3

305 310 315 320

Arg Ser Met Met Pro Trp Ile Gly Ala Asn Lys Ile Val Asp Lys Ala
 325 330

Lys Asn

5 <210> 37
 <211> 63
 <212> ADN
 <213> secuencia artificial
 <220>
 <223> cebador directo PAL- F1
 <400> 37
 ct gcagcaca t gaagact cc at ggcaccct cgct cgact c gat ct cgcac t cgt t cgcaa 60
 acg 63

10 <210> 38
 <211> 40
 <212> ADN
 <213> secuencia artificial
 <220>
 15 <223> cebador inverso PAL- R1
 <400> 38
 t ct ct cat cc gccaaaacag aagct t ct aa gcgagcat ct 40

20 <210> 39
 <211> 43
 <212> ADN
 <213> secuencia artificial
 <220>
 <223> cebador PAL- EcoRI - F1
 <400> 39
 25 gggct aacag gaggaagaat t cat ggcacc ct cgct cgac t cg 43

30 <210> 40
 <211> 42
 <212> ADN
 <213> secuencia artificial
 <220>
 <223> cebador inverso PAL- EcoRI - R1
 <400> 40
 cgagt cgagc gagggg gccca tgaattctc ctctgtagc cc 42

REIVINDICACIONES

- 1.- Un método para la conversión de acetolactato a dihidroxiisovalerato, que comprende:
- 5 a) proporcionar una célula hospedante microbiana que comprende una construcción genética que codifica un polipéptido que tienen una actividad específica de cetol-ácido reductoisomerasa mayor que la actividad específica de una cetol-ácido reductoisomerasa de *E. coli*, en el que dicho polipéptido tiene una actividad específica de cetol-ácido reductoisomerasa mayor que 1,1 $\mu\text{moles/min/mg}$ basada en la proteína purificada, medida mediante el ensayo de consumo de NADPH, ejecutado bajo las siguientes condiciones:
- i) pH de aproximadamente 7,5;
- ii) una temperatura de aproximadamente 22,5 °C; y
- 10 iii) potasio mayor que aproximadamente 10 mM; y
- b) poner en contacto la célula hospedante de a) con acetolactato, en el que se produce 2,3-dihidroxiisovalerato.
- 2.- Un método para la producción de isobutanol, que comprende:
- a) proporcionar una célula hospedante microbiana que comprende las siguientes construcciones genéticas:
- 15 1) al menos una construcción genética que codifica una enzima acetolactato sintasa para la conversión del piruvato a acetolactato;
- 2) al menos una construcción genética que codifica una enzima cetol-ácido reductoisomerasa que tiene una actividad específica mayor que 1,1 $\mu\text{mol/min/mg}$ basada en la proteína purificada, medida mediante el ensayo del consumo de NADPH, ejecutado bajo las siguientes condiciones:
- i) pH de aproximadamente 7,5;
- 20 ii) una temperatura de aproximadamente 22,5 °C; y
- iii) potasio mayor que aproximadamente 10 mM,
- para la conversión de (S)-acetolactato a 2,3-dihidroxiisovalerato;
- 3) al menos una construcción genética que codifica una acetohidroxiácido deshidratasa para la conversión de 2,3-dihidroxiisovalerato a α -cetoisovalerato;
- 25 4) al menos una construcción genética que codifica una cetoácido de cadena ramificada descarboxilasa para la conversión de α -cetoisovalerato a isobutiraldehído;
- 5) al menos una construcción genética que codifica una alcohol de cadena ramificada deshidrogenasa para la conversión de isobutiraldehído a isobutanol; y
- b) cultivar la célula hospedante de a) bajo condiciones en las que se produce isobutanol.
- 30 3.- El método según la reivindicación 1 o la reivindicación 2, en el que al menos una construcción genética que codifica un polipéptido que tiene actividad cetol-ácido reductoisomerasa se aísla de *Pseudomonas*.
- 4.- El método según la reivindicación 1 o la reivindicación 2, en el que el polipéptido que tiene actividad cetol-ácido reductoisomerasa se aísla de *Pseudomonas aeruginosa* (PAO1), *Pseudomonas fluorescens* (PF5), o *Vibrio cholerae* (N16961).
- 35 5.- El método según una cualquiera de las reivindicaciones 1-4, en el que la célula hospedante es un miembro de un género seleccionado del grupo que consiste en *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, *Brevibacterium*, *Pichia*, *Candida*, *Hansenula*, *Vibrio* y *Saccharomyces*.
- 40 6.- El método según la reivindicación 5, en el que la célula hospedante es *Escherichia coli*, *Lactobacillus plantarum*, o *Saccharomyces cerevisiae*.
- 7.- El método según una cualquiera de las reivindicaciones 2-6, en el que la acetolactato sintasa tiene una secuencia de aminoácidos según se indica en SEQ ID NO:2.
- 8.- El método según una cualquiera de las reivindicaciones 1-7, en el que el polipéptido que tiene actividad cetol-ácido reductoisomerasa tiene una secuencia de aminoácidos seleccionada del grupo que consiste en SEQ ID NO:34, SEQ ID NO:35 y SEQ ID NO:36.
- 45

- 9.- El método según una cualquiera de las reivindicaciones 2-8, en el que la actividad cetohidroxiácido deshidratasa tiene una secuencia de aminoácidos según se indica en SEQ ID NO:6.
- 10.- El método según una cualquiera de las reivindicaciones 2-9, en el que la alcohol de cadena ramificada deshidrogenasa tiene una secuencia de aminoácidos según se indica en SEQ ID NO:10.
- 5 11.- El método según una cualquiera de las reivindicaciones 2-10, en el que la α -cetoácido de cadena ramificada descarboxilasa tiene una secuencia de aminoácidos según se indica en SEQ ID NO:8.
- 12.- Una célula hospedante recombinante que comprende una construcción genética que codifica una enzima cetol-ácido reductoisomerasa que tiene una actividad específica mayor que la actividad específica de una cetol-ácido reductoisomerasa de *E. coli*, en la que la enzima cetol-ácido reductoisomerasa que tiene una actividad específica mayor que la actividad específica de una cetol-ácido reductoisomerasa de *E. coli* tiene una actividad específica mayor que 1,1 μ moles/min/mg basada en la proteína purificada, medida mediante el ensayo de consumo de NADPH, ejecutado bajo las siguientes condiciones:
- 10 a) pH de aproximadamente 7,5;
- b) una temperatura de aproximadamente 22,5 °C; y
- 15 c) potasio mayor que aproximadamente 10 mM.
- 13.- Un método para la identificación y el aislamiento de una construcción genética, que codifica una enzima cetol-ácido reductoisomerasa que tiene una actividad específica mayor que 1,1 μ mol/min/mg basada en la proteína purificada, medida mediante el ensayo del consumo de NADPH, ejecutado bajo las siguientes condiciones:
- i) pH de aproximadamente 7,5;
- 20 ii) una temperatura de aproximadamente 22,5 °C; y
- iii) potasio mayor que aproximadamente 10 mM;
- que comprende las etapas de:
- a) identificar especies bacterianas que tengan un tiempo de duplicación más corto que *E. coli* cuando se cultivan en medio mínimo M9;
- 25 b) seleccionar las especies bacterianas de (a) para la actividad cetol-ácido reductoisomerasa para identificar las especies bacterianas activas;
- c) sondear el ADN genómico de las especies bacterianas activas de b) con secuencias de ácidos nucleicos que tengan homología con construcciones genéticas conocidas por codificar una cetol-ácido reductoisomerasa, para identificar y aislar las construcciones genéticas que codifican una cetol-ácido reductoisomerasa a partir de dichas especies bacterianas activas; y
- 30 d) amplificar y expresar las construcciones genéticas que codifican una cetol-ácido reductoisomerasa a partir de dichas especies bacterianas activas; y
- e) seleccionar las construcciones genéticas expresadas de la etapa d) para detectar las que tengan una actividad específica mayor que 1,1 μ mol/min/mg basada en la proteína purificada, medida mediante el ensayo del consumo de NADPH, ejecutado bajo las siguientes condiciones:
- 35 i) pH de aproximadamente 7,5;
- ii) una temperatura de aproximadamente 22,5 °C; y
- iii) potasio mayor que aproximadamente 10 mM.
- 14.- El método de la reivindicación 13, en el que dicha especie bacteriana activa se selecciona del grupo que consiste en *Clostridium*, *Zymomonas*, *Escherichia*, *Salmonella*, *Rhodococcus*, *Pseudomonas*, *Bacillus*, *Vibrio*, *Lactobacillus*, *Enterococcus*, *Alcaligenes*, *Klebsiella*, *Paenibacillus*, *Arthrobacter*, *Corynebacterium*, y *Brevibacterium*.
- 40 15.- El método según la reivindicación 13 o la reivindicación 14, en el que el tiempo de duplicación de la etapa a) es igual o menor que 80% del tiempo de duplicación de *E. coli* cuando se cultiva en medio mínimo M9.

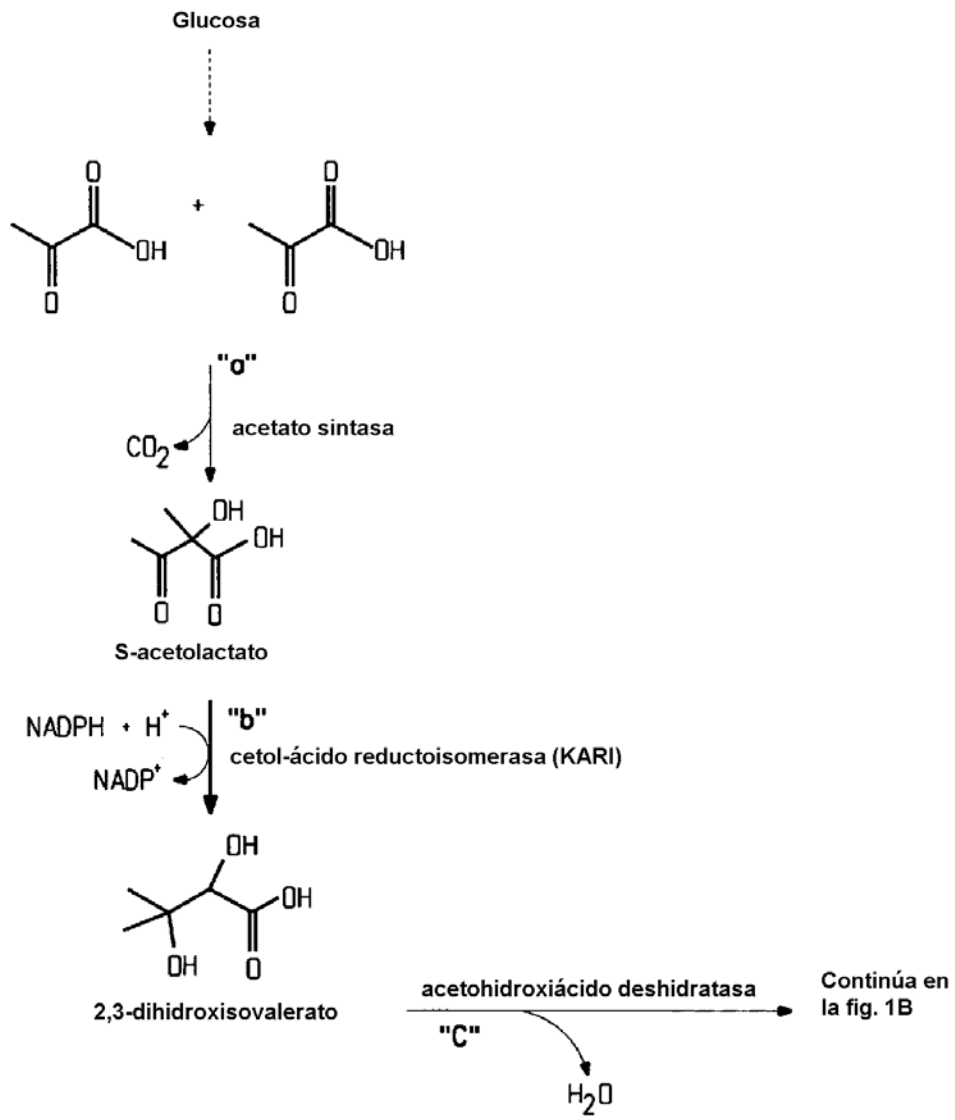


FIG. 1A

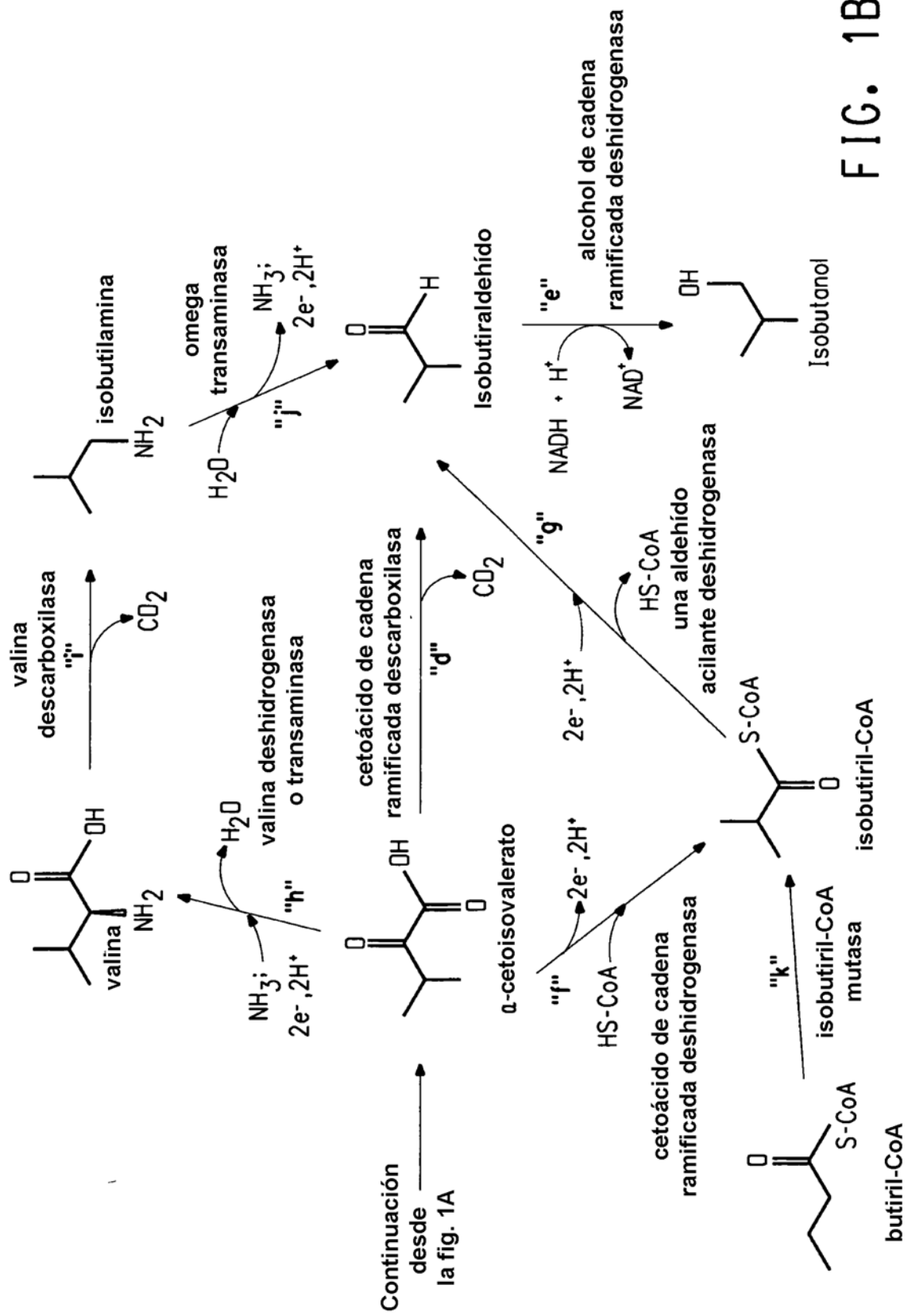


FIG. 1B