

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 607 891**

51 Int. Cl.:

**C12N 1/14** (2006.01)

**C12N 1/18** (2006.01)

**C12P 7/00** (2006.01)

**C12P 7/08** (2006.01)

**C12N 9/92** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.07.2005 PCT/NL2005/000516**

87 Fecha y número de publicación internacional: **26.01.2006 WO06009434**

96 Fecha de presentación y número de la solicitud europea: **15.07.2005 E 05763114 (5)**

97 Fecha y número de publicación de la concesión europea: **28.09.2016 EP 1781772**

54 Título: **Ingeniería metabólica de células eucariotas que fermentan xilosa**

30 Prioridad:

**16.07.2004 EP 04077073**

**16.07.2004 US 588381 P**

**22.07.2004 US 589833 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**04.04.2017**

73 Titular/es:

**DSM IP ASSETS B.V. (100.0%)**

**Het Overloon 1**

**6411 TE Heerlen, NL**

72 Inventor/es:

**WINKLER. AARON, ADRIAAN;  
KUYPER, SIPKO, MAARTEN;  
DE LAAT, WILHELMUS, THEODORUS,  
ANTONIUS, MARIA;  
VAN DIJKEN, JOHANNES, PIETER y  
PRONK, JACOBUS, THOMAS**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

ES 2 607 891 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

Ingeniería metabólica de células eucariotas que fermentan xilosa.

Campo de la invención

5 La presente invención se refiere a modificaciones genéticas adicionales en células anfitrionas eucariotas que han sido transformadas para expresar una xilosa isomerasa que confiere a la célula anfitriona la capacidad de isomerizar la xilosa a xilulosa. Las modificaciones genéticas adicionales están dirigidas a mejorar la eficiencia del metabolismo de xilosa e incluyen, p. ej., la disminución de la actividad de la aldosa reductasa inespecífica, aumento de la actividad de xilulosa cinasa y aumento del flujo de la ruta de la pentosa fosfato. Las células anfitrionas modificadas de la invención son adecuadas para la producción de una amplia variedad de productos de fermentación en  
10 procedimientos que comprenden xilosa como fuente de carbono.

Antecedentes de la invención

15 La producción de etanol económicamente viable a partir de la fracción de hemicelulosa de la biomasa vegetal requiere la transformación simultánea tanto de pentosas como de hexosas a velocidades comparables y con altos rendimientos. Las levaduras, en particular *Saccharomyces spp.*, son las candidatas más apropiadas para este procedimiento ya que pueden crecer rápidamente en hexosas, tanto aeróbicamente como anaeróbicamente. Además, son mucho más resistentes al ambiente tóxico de los hidrolizados de lignocelulosa que las bacterias (modificadas genéticamente).

20 En estudios previos se ha proporcionado evidencia de que la ingeniería metabólica de *S. cerevisiae* para la utilización de xilosa, debe basarse en la introducción de xilosa isomerasa (XI, EC 5.3.1.5) Bruinenberg et al. (1983, Eur J. Appl. Microbiol. Biotechnol., 18: 287-292). A diferencia de las cepas basadas en xilosa reductasa (XR, EC 1.1.1.21) y xilitol deshidrogenasa (XD, EC 1.1.1.9), las cepas que expresan actividad XI presentan altos rendimientos en alcohol y apenas producen xilitol como se ha demostrado recientemente en el documento WO 03/0624430 y Kuyper et al. (2004, FEMS Yeast Res. 4: 655-664). Desde un punto de vista teórico, esto no es sorprendente ya que la ruta a través de XR y XD conduce a una obstrucción en el equilibrio del NADH que en ausencia de oxígeno puede ser aliviada, p. ej., a través de la formación de xilitol.  
25

El documento WO 03/0624430 divulga que la introducción de XI funcional de *Piromyces* en *S. cerevisiae* permite el metabolismo lento de xilosa a través de la xilulocinasa endógena (EC 2.7.1.17) codificada por *XKS1* y las enzimas de la parte no oxidativa de la ruta de la pentosa fosfato y confiere a los transformantes de levadura la capacidad de crecer en xilosa.

30 Kuyper et al. (*supra*) describen cepas de *S. cerevisiae* en las que se ha introducido XI de *Piromyces* y que después se someten a evolución dirigida en frascos agitados muestran velocidades mejoradas de fermentación de xilosa, pero todavía requieren oxígeno para el crecimiento. La selección posterior a través de un régimen de limitación extrema del oxígeno con exceso de xilosa, seguida de selección anaerobia dio como resultado una cepa de laboratorio (RWB202-AFX) que cumple al menos uno de los requisitos previos para la utilización de hemicelulosa, a saber, un aceptable rendimiento de etanol en xilosa. Sin embargo, la velocidad específica de producción de etanol en esta cepa es todavía inaceptablemente baja. En particular, la velocidad específica de consumo de azúcar durante el crecimiento en xilosa (345 mg de xilosa/g de biomasa/h) es todavía diez veces menor que en glucosa. Los intentos por mejorar aún más la cepa RWB202-AFX a través de ingeniería evolutiva han fracasado hasta ahora.  
35

40 El documento WO 03/0624430 enumera una serie de modificaciones genéticas alternativas que pueden dar como resultado una mejora adicional de las velocidades específicas de producción de etanol y/o de consumo de azúcar en xilosa en células anfitrionas que expresan el gen XI de *Piromyces* a un nivel que sería necesario para la utilización comercial de hemicelulosa. Estas alternativas incluyen: (a) aumentar el transporte de xilosa a la célula anfitriona; (b) aumento de la actividad de xilulosa cinasa; (c) aumento del flujo de la ruta de la pentosa fosfato; (d) disminución de la sensibilidad a la represión catabólica; (e) aumento de la tolerancia a etanol, a la osmolaridad o a los ácidos orgánicos; y (f) producción reducida de subproductos (tales como, p. ej., xilitol, glicerol y/o ácido acético). Más específicamente, el documento WO 03/0624430 sugiere sobreexpresar uno o más de los genes que codifican un transportador de hexosa o de pentosa, una xilulosa cinasa (tal como la *S. cerevisiae XKS1*) una enzima de la ruta de la pentosa fosfato tal como las enzimas glicolíticas transaldolasa (*TAL1*) o transcetolasa (*TKL1*), enzimas etanológicas tal como alcohol deshidrogenasas, y/o para inactivar un gen de hexosa cinasa, por ejemplo el gen *HXK2* de *S. cerevisiae*, los genes *MIG1* o *MIG2* de *S. cerevisiae*, los genes de la aldosa reductasa (inespecífica) tal como el gen *GRE3* de *S. cerevisiae*, o los genes de enzimas implicadas en el metabolismo de glicerol tales como los genes glicerol-fosfato deshidrogenasa 1 y/o 2 de *S. cerevisiae*. Sin embargo, el documento WO 03/0624430 no divulga cuál de estas muchas alternativas producen realmente una mejora en las velocidades específicas de producción de etanol y/o de consumo de xilosa en células anfitrionas que transportan el gen XI de *Piromyces*.  
45  
50

55 Karhumaa et al. (2004, "Development of a Xylose-growing *Saccharomyces cerevisiae* strain expressing bacterial xilose isomerase", Presentación en póster de la segunda reunión sobre Physiology of Yeasts and Filamentous Fungi; 24-28 de marzo de 2004, Anglet, Francia, página 43; y, 2004, "New Xylose-growing *Saccharomyces cerevisiae* strain for biofuel ethanol production", Presentación oral en el 26º Simposio sobre Biotechnology for fuels and chemicals, 9-

12 de mayo de 2004, Chattanooga (TN), EE. UU, página 19) divulgan una cepa de *S. cerevisiae* que expresa una XI bacteriana de *Thermus Thermophilus*. La cepa contiene además una serie de modificaciones genéticas sugeridas en el documento WO 03/0624430: sobreexpresión de la xilulosa cinasa y las cuatro enzimas de la ruta de la pentosa fosfato no oxidativa así como la inactivación del gen (GRE3) de la aldosa reductasa inespecifica de *S. cerevisiae*.

5 Sin embargo, a pesar de estas modificaciones genéticas esta cepa es incapaz de crecer sobre la xilosa. Sólo después de la adaptación al crecimiento aeróbico en xilosa se obtuvo una cepa TMB3050 que es capaz de crecer en xilosa a baja velocidad ( $\mu = 0,04 \text{ h}^{-1}$ ) y con una baja velocidad específica de consumo de xilosa de 4,3 mg de xilosa/g de células/h. Dado que las modificaciones genéticas no definidas (acumuladas durante la adaptación) son claramente necesarias para el crecimiento en xilosa en primer lugar, no se puede deducir del trabajo de Karhumaa et al., cuales, si las hubiera, de las modificaciones genéticas definidas (como la sobreexpresión de xilulosa cinasa o de cualquiera de las enzimas de la ruta de la pentosa fosfato o la inactivación del gen aldosa reductasa), si acaso, contribuyen en realidad a la capacidad de que la cepa adaptada crezca en xilosa.

10 Por lo tanto, un objeto de la presente invención es proporcionar células anfitrionas eucariotas, tales como células anfitrionas fúngicas, que son transformadas con un gen XI que confiere la capacidad de crecer en xilosa y cuyas células anfitrionas tienen velocidades específicas de consumo de xilosa y/o de formación de producto (etanol) que son compatibles con la aplicación comercial de las células anfitrionas.

Descripción de la invención

Definiciones

Xilosa isomerasa

20 La enzima "xilosa isomerasa" (EC 5.3.1.5) se define en el presente documento como una enzima que cataliza la isomerización directa de D-xilosa en D-xilulosa y viceversa. La enzima también se conoce como una D-xilosa cetoisomerasa. Algunas xilosa isomerasas son también capaces de catalizar la transformación entre D-glucosa y D-fructosa y, por lo tanto, se denominan a veces glucosa isomerasa. Las xilosa isomerasas requieren cationes bivalentes como magnesio o manganeso como cofactor. Las xilosa isomerasas de la invención pueden definirse adicionalmente por su secuencia de aminoácidos como se describe en el presente documento más adelante. Del mismo modo, las xilosa isomerasas pueden definirse por las secuencias de nucleótidos que codifican la enzima, así como por las secuencias de nucleótidos que se hibridan con una secuencia de nucleótidos de referencia que codifican una xilosa isomerasa como se describe en el presente documento más adelante.

25 Una unidad (U) de actividad de xilosa isomerasa se define en el presente documento como la cantidad de enzima que produce 1 nmol de xilulosa por minuto, en las condiciones descritas por Kuyper et al. (2003, FEMS Yeast Res. 4: 69-78).

Xilulosa cinasa

30 La enzima "xilulosa cinasa" (EC 2.7.1.17) se define en el presente documento como una enzima que cataliza la reacción  $\text{ATP} + \text{D-xilulosa} = \text{ADP} + \text{D-xilulosa 5-fosfato}$ . La enzima también se conoce como una xilulocinasa fosforilante, D-xilulocinasa o ATP: D-xilulosa 5-fosfotransferasa. Una xilulosa cinasa de la invención puede definirse además por su secuencia de aminoácidos como se describe en el presente documento más adelante. Del mismo modo, una xilulosa cinasa puede ser definida por las secuencias de nucleótidos que codifican la enzima así como por las secuencias de nucleótidos que se hibridan con una secuencia de nucleótidos de referencia que codifica una xilulosa cinasa como se describe en el presente documento más adelante. Una unidad de actividad de xilulocinasa se define en el Ejemplo 1.13 del presente documento.

Ribulosa 5-fosfato epimerasa

35 La enzima "ribulosa 5-fosfato epimerasa" (5.1.3.1) se define en el presente documento como una enzima que cataliza la epimerización de D-xilulosa 5-fosfato en D-ribulosa 5-fosfato y viceversa. La enzima también se conoce como fosforribulosa epimerasa; eritrosa-4-fosfato isomerasa; fosfocetopentosa 3-epimerasa; xilulosa fosfato 3-epimerasa; fosfocetopentosa epimerasa; ribulosa 5-fosfato 3-epimerasa; D-ribulosa fosfato-3-epimerasa; D-ribulosa 5-fosfato epimerasa; D-ribulosa-5-P 3-epimerasa; D-xilulosa-5-fosfato 3-epimerasa; pentosa-5-fosfato 3-epimerasa; o D-ribulosa-5-fosfato 3-epimerasa. Una ribulosa 5-fosfato epimerasa de la invención se puede definir además por su secuencia de aminoácidos como se describe en el presente documento más adelante. Del mismo modo, una ribulosa 5-fosfato epimerasa se puede definir por las secuencias de nucleótidos que codifican la enzima así como por las secuencias de nucleótidos que se hibridan con una secuencia de nucleótidos de referencia que codifica una ribulosa 5-fosfato epimerasa como se describe en el presente documento más adelante.

Ribulosa 5-fosfato isomerasa

40 La enzima "ribulosa 5-fosfato isomerasa" (EC 5.3.1.6) se define en el presente documento como una enzima que cataliza la isomerización directa de D-ribosa 5-fosfato en D-ribulosa 5-fosfato y viceversa. La enzima también se conoce como fosfopentosa isomerasa; fosforiboisomerasa; ribosa fosfato isomerasa; 5-fosforribosa isomerasa; D-ribosa 5-fosfato isomerasa; D-ribosa-5-fosfato cetol-isomerasa; o D-ribosa-5-fosfato aldosa-cetosa-isomerasa. Una

ribulosa 5-fosfato isomerasa de la invención se puede definir además por su secuencia de aminoácidos como se describe en el presente documento más adelante. Del mismo modo, una ribulosa 5-fosfato isomerasa se puede definir por las secuencias de nucleótidos que codifican la enzima así como por secuencias de nucleótidos que se hibridan con una secuencia de nucleótidos de referencia que codifica una ribulosa 5-fosfato isomerasa como se describe en el presente documento más adelante.

#### Transcetolasa

La enzima "transcetolasa" (EC 2.2.1.1) se define en el presente documento como una enzima que cataliza la reacción:



y viceversa. La enzima también se conoce como glicolaldehído transferasa o sedoheptulosa-7-fosfato:D-gliceraldehído-3-fosfato glicolaldehído transferasa. Una transcetolasa de la invención se puede definir además por su secuencia de aminoácidos como se describe en el presente documento más adelante. De manera similar, una transcetolasa se puede definir por las secuencias de nucleótidos que codifican la enzima así como por secuencias de nucleótidos que se hibridan con una secuencia de nucleótidos de referencia que codifica una transcetolasa como se describe en el presente documento más adelante.

#### Transaldolasa

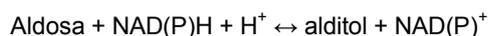
La enzima "transaldolasa" (EC 2.2.1.2) se define en el presente documento como una enzima que cataliza la reacción:



y viceversa. La enzima también se conoce como dihidroxiacetona transferasa; dihidroxiacetona sintasa; formaldehído transcetolasa; o sedoheptulosa-7-fosfato: D-gliceraldehído-3-fosfato glicerona transferasa. Una transaldolasa de la invención se puede definir además por su secuencia de aminoácidos como se describe en el presente documento más adelante. Del mismo modo, una transaldolasa se puede definir por las secuencias de nucleótidos que codifican la enzima, así como por las secuencias de nucleótidos que se hibridan con una secuencia de nucleótidos de referencia que codifica una transaldolasa como se describe en el presente documento más adelante.

#### Aldosa reductasa

La enzima "aldosa reductasa" (EC 1.1.1.21) se define en el presente documento como cualquier enzima que es capaz de reducir xilosa o xilulosa a xilitol. En el contexto de la presente invención, una aldosa reductasa puede ser cualquier aldosa reductasa inespecífica que se encuentre de forma natural (endógena) en una célula anfitriona de la invención y que sea capaz de reducir xilosa o xilulosa a xilitol. Las aldosas reductasas inespecíficas catalizan la reacción:



La enzima tiene una amplia especificidad y se conoce también como aldosa reductasa; poliol deshidrogenasa (NADP<sup>+</sup>); alditol:NADP oxidorreductasa; alditol:NADP<sup>+</sup> 1-oxidorreductasa; NADPH-aldopentosa reductasa; o NADPH-aldosa reductasa. Un ejemplo particular de una aldosa reductasa inespecífica de este tipo que es endógena a *S. cerevisiae* y que está codificada por el gen *GRE3* (Traff et al., 2001, Appl. Environ. Microbiol., 67: 5668-74). De este modo, una aldosa reductasa de la invención se puede definir además por su secuencia de aminoácidos como se describe en el presente documento más adelante. Del mismo modo, una aldosa reductasa se puede definir por las secuencias de nucleótidos que codifican la enzima así como por secuencias de nucleótidos que se hibridan con una secuencia de nucleótidos de referencia que codifica una aldosa reductasa como se describe en el presente documento más adelante.

#### Identidad y similitud de secuencia

La identidad de secuencia se define en el presente documento como una relación entre dos o más secuencias de aminoácidos (polipéptido o proteína) o dos o más secuencias de ácidos nucleicos (polinucleótido), según se determina comparando las secuencias. En la técnica, "identidad" también significa el grado de parentesco de las secuencias entre las secuencias de aminoácidos o de ácidos nucleicos, según sea el caso, determinado por el emparejamiento entre cadenas de tales secuencias. La "similitud" entre dos secuencias de aminoácidos se determina comparando la secuencia de aminoácidos y sus sustitutos de aminoácidos conservados de un polipéptido con la secuencia de un segundo polipéptido. La "identidad" y la "similitud" se pueden calcular fácilmente por procedimientos conocidos, que incluyen pero que no se limitan a los descritos en Computational Molecular Biology, Lesk, A. M., ed., Oxford University Press, Nueva York, 1988, Biocomputing: Informatics and Genome Projects, Smith, D. W., ed., Academic Press, Nueva York, 1993; Computer Analysis of Sequence Data, Parte 1, Griffin, A. M., y Griffin, HG, eds., Humana Press, Nueva Jersey, 1994. Sequence Analysis in Molecular Biology, Von Heine, G.,

Academic Press, 1987, y Sequence Analysis Primer, Gribskov, M. y Devereux, J., eds., M. Stockton Press, Nueva York, 1991, y Carillo, H. y Lipman, D., SIAM J. Applied Math., 48: 1073 (1988).

5 Procedimientos preferidos para determinar la identidad se diseñan para dar el mayor emparejamiento entre las secuencias ensayadas. Procedimientos para determinar la identidad y similitud se codifican en programas informáticos disponibles para el público. Los procedimientos de los programas informáticos preferidos para determinar la identidad y similitud entre dos secuencias incluyen, p. ej., el paquete de programas GCG (Devereux, J. et al., Nucleic Acids Research 12 (1): 387 (1984)), BestFit, BLASTP, BLASTN y FASTA (Altschul, S. F. et al., J. Mol. Biol. 215: 403-410 (1990) El programa BLAST X está disponible para el público a través del NCBI y otras fuentes (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul, S., et al., J. Mol. Biol. 215: 403-410 (1990) El conocido algoritmo de Smith Waterman también puede usarse para determinar la identidad.

10 Los parámetros preferidos para la comparación de secuencias de polipéptidos incluyen los siguientes: Algoritmo: Needleman y Wunsch, J. Mol. Biol. 48: 443-453 (1970); Matriz de comparación: BLOSSUM62 de Hentikoff y Hentikoff, Proc. Natl. Acad. Sci. EE. UU. 89:10915-10919 (1992); Penalización por Hueco: 12; y la penalización por la Longitud del Hueco: 4. Un programa útil con estos parámetros está disponible para el público como programa "Ogap" de Genetics Computer Group, situado en Madison, WI. Los parámetros mencionados anteriormente son parámetros por defecto para las comparaciones de aminoácidos (junto con ninguna penalidad para los huecos terminales).

15 Los parámetros preferidos para la comparación de ácidos nucleicos incluyen los siguientes: Algoritmo: Needleman y Wunsch, J. Mol. Biol. 48: 443-453 (1970); Matriz de comparación: emparejamientos = +10, desemparejamientos = 0; Penalización por Hueco: 50; Penalización por Longitud de Hueco: 3. Disponible como programa "Gap" de Genetics Computer Group, situado en Madison, Wis. Los parámetros por defecto para las comparaciones de ácidos nucleicos se han dado anteriormente.

20 Opcionalmente, para determinar el grado de similitud de aminoácidos, la persona experta puede también tener en cuenta las sustituciones de aminoácidos denominadas "conservadoras", como será evidente para la persona experta. Las sustituciones conservadoras de aminoácidos se refieren a la intercambiabilidad de residuos que tienen cadenas laterales similares. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas es glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tienen cadenas laterales alifático-hidroxilo es serina y treonina; un grupo de aminoácidos que tienen cadenas laterales que contienen amida es asparragina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas es fenilalanina, tirosina y triptófano; un grupo de aminoácidos que tienen cadenas laterales con bases es lisina, arginina e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Los grupos de sustituciones conservadoras de aminoácidos preferidos son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina y asparragina-glutamina. Las variantes de sustitución de la secuencia de aminoácidos divulgada en el presente documento son aquellas en las que se ha eliminado al menos un residuo en las secuencias divulgadas y se ha insertado en su lugar un residuo diferente. Preferiblemente, el cambio de aminoácidos es conservador. Las sustituciones conservadoras preferidas para cada uno de los aminoácidos de origen natural son las siguientes: Ala a ser; Arg a lys; Asn a gln o his; Asp a glu; Cys a ser o ala; Gln a asn; Glu a asp; Gly a pro; His a asn o gln; Ile a leu o val; Leu a ile o val; Lys a arg; gln o glu; Met a leu o ile; Phe a met, leu o tyr; Ser a thr; Thr a ser; Trp a tyr; Tyr a trp o phe; y, Val a ile o leu.

#### 40 Hibridación de secuencias de ácido nucleico

Las secuencias de nucleótidos que codifican las enzimas de la invención también se pueden definir por su capacidad para hibridarse con las secuencias de nucleótidos de las SEQ ID NO. 9-16 y 18, respectivamente, en condiciones de hibridación moderada o preferiblemente rigurosas. Las condiciones de hibridación rigurosas se definen en el presente documento como las condiciones que permiten que una secuencia de ácidos nucleicos de al menos aproximadamente 25, preferiblemente aproximadamente 50 nucleótidos, 75 o 100 y lo más preferiblemente de aproximadamente 200 o más nucleótidos, se hibride a una temperatura de aproximadamente 65 °C en una solución que comprenda aproximadamente sal 1 M, preferiblemente 6 x SSC o cualquier otra solución que tenga una fuerza iónica comparable, y lavado a 65 °C en una solución que comprenda aproximadamente ,sal 0,1 M, o menos, preferiblemente 0,2 x SSC o cualquier otra solución que tenga una fuerza iónica comparable. Preferiblemente, la hibridación se realiza durante la noche, es decir, al menos durante 10 horas y, preferiblemente, el lavado se realiza durante al menos una hora con al menos dos cambios de la solución de lavado. Estas condiciones permitirán normalmente la hibridación específica de secuencias que tienen aproximadamente un 90 % o más de identidad de secuencia.

55 Las condiciones moderadas se definen en el presente documento como las condiciones que permiten que las secuencias de ácidos nucleicos de al menos 50 nucleótidos, preferiblemente de aproximadamente 200 o más nucleótidos, se hibriden a una temperatura de aproximadamente 45 °C en una solución que comprenda aproximadamente sal 1 M, preferiblemente 6 x SSC o cualquier otra solución que tenga una fuerza iónica comparable, y lavado a temperatura ambiente en una solución que comprenda aproximadamente sal 1 M, preferiblemente 6 x SSC o cualquier otra solución que tenga una fuerza iónica comparable. Preferiblemente, la hibridación se realiza durante la noche, es decir, al menos durante 10 horas y, preferiblemente, el lavado se realiza

durante al menos una hora con al menos dos cambios de la solución de lavado. Estas condiciones permitirán normalmente la hibridación específica de secuencias que tengan hasta un 50 % de identidad de secuencia. La persona experta en la técnica podrá modificar estas condiciones de hibridación con el fin de identificar específicamente secuencias que varían en identidad entre 50 % y 90 %.

## 5 Enlazado operativamente

Tal como se utiliza aquí, la expresión "enlazado operativamente" se refiere a un enlace de elementos polinucleótidos en una relación funcional. Un ácido nucleico está "enlazado operativamente" cuando se coloca en una relación funcional con otra secuencia de ácidos nucleicos. Por ejemplo, un promotor o potenciador está enlazado operativamente a una secuencia codificante si afecta a la transcripción de la secuencia codificante. Enlazado operativamente significa que las secuencias de ADN que están enlazadas son típicamente contiguas y, cuando es necesario, unen dos regiones codificantes de proteínas, contiguas y en el marco de lectura.

### Promotor

Tal como se usa en el presente documento, el término "promotor" se refiere a un fragmento de ácido nucleico que funciona para controlar la transcripción de uno o más genes, situados aguas arriba con respecto a la dirección de transcripción del sitio de inicio de la transcripción del gen, y se identifica estructuralmente por la presencia de un sitio de unión para la ARN polimerasa dependiente de ADN, sitios de iniciación de la transcripción y cualquier otra de las secuencias de ADN, que incluya pero que no se limite a los sitios de unión al factor de transcripción, sitios represor y activador de la unión de las proteínas, y cualquier otra de las secuencias de nucleótidos conocidas por el experto en la técnica para actuar directa o indirectamente para regular la cantidad de transcripción desde el promotor. Un promotor "constitutivo" es un promotor que es activo en la mayoría de las condiciones ambientales y de desarrollo. Un promotor "inducible" es un promotor que es activo bajo regulación ambiental o de desarrollo.

### Homólogo

El término "homólogo" cuando se usa para indicar la relación entre una determinada molécula de ácido nucleico (recombinante) o polipéptido y un determinado organismo anfitrión o célula anfitriona, se entiende que significa que en la naturaleza la molécula de ácido nucleico o de polipéptidos se produce por una célula anfitriona u organismos de la misma especie, preferiblemente de la misma variedad o cepa. Si es homóloga a una célula anfitriona, una secuencia de ácidos nucleicos que codifica un polipéptido normalmente estará enlazada operativamente a otra secuencia promotora o, si procede, otra secuencia de señales secretoras y/o secuencia terminadora que esté en su entorno natural. Cuando se usa para indicar el parentesco de dos secuencias de ácidos nucleicos, el término "homólogo" significa que una secuencia de ácido nucleico monocatenario se puede hibridar con una secuencia de ácido nucleico monocatenario complementaria. El grado de hibridación puede depender de una serie de factores que incluyen la cantidad de identidad entre las secuencias y las condiciones de hibridación tales como temperatura y concentración de sal como se discute más adelante. Preferiblemente, la zona de identidad es mayor que aproximadamente 5 pb, más preferiblemente la zona de identidad es mayor que 10 pb.

## 35 Heterólogo

El término "heterólogo" cuando se usa con respecto a un ácido nucleico (de ADN o ARN) o proteína se refiere a un ácido nucleico o proteína que no se da de forma natural como parte del organismo, célula, genoma o secuencia de ADN o ARN en la que está presente, o que se encuentra en una célula o lugar o lugares en el genoma o secuencia de ADN o ARN que difiere de aquella en la que se encuentra en la naturaleza. Los ácidos nucleicos o proteínas heterólogos no son endógenos a la célula en la que se introducen, sino que se han obtenido a partir de otra célula o se han producido de forma sintética o recombinante. En general, aunque no necesariamente, tales ácidos nucleicos codifican proteínas que normalmente no son producidas por la célula en la que el ADN se transcribe o se expresa. De forma similar, el ARN exógeno codifica proteínas que no se expresan normalmente en la célula en la que está presente el ARN exógeno. Los ácidos nucleicos y proteínas heterólogos pueden también denominarse ácidos nucleicos o proteínas foráneos. Cualquier ácido nucleico o proteína que el experto en la técnica reconozca como heteróloga o foránea a la célula en la que se expresa está contemplado en el presente documento por el término ácido nucleico o proteína heterólogo. El término heterólogo también se aplica a combinaciones no naturales de secuencias de ácido nucleico o aminoácidos, es decir, combinaciones donde al menos dos de las secuencias combinadas son foráneas entre sí.

En este documento y en sus reivindicaciones, el verbo "comprender" y sus conjugaciones se utilizan en su sentido no limitativo que significa que los artículos que siguen a la palabra están incluidos, pero los artículos no mencionados específicamente no están excluidos. Además, la referencia a un elemento por el artículo indefinido "un", "una" o "uno" no excluye la posibilidad de que más de uno de los elementos esté presente, a menos que el contexto exija claramente que exista uno y sólo uno de los elementos. El artículo indefinido "un", "una" o "uno" significa, por lo general, "al menos uno".

### Descripción detallada de la invención

La presente invención se refiere a una célula anfitriona de levadura u hongo filamentoso transformada con un

constructo de ácido nucleico que comprende una secuencia de nucleótidos que codifica una xilosa isomerasa, cuya xilosa isomerasa comprende una secuencia de aminoácidos que tiene al menos 95 % de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO. 2, por lo que el constructo de ácido nucleico, tras la transformación de la célula anfitriona, confiere a la célula anfitriona la capacidad de isomerizar directamente la xilosa a xilulosa, por lo que la célula anfitriona comprende una modificación genética que aumenta el flujo de la ruta de la pentosa fosfato, y por lo que la modificación genética comprende la sobreexpresión de al menos un gen de la parte no oxidativa de la ruta de la pentosa fosfato en comparación con una célula anfitriona que es genéticamente idéntica excepto para la modificación genética. La capacidad de la célula anfitriona transformada para isomerizar la xilosa en xilulosa es la isomerización directa de xilosa a xilulosa. Esto se entiende que significa que la xilosa se isomeriza en xilulosa en una única reacción catalizada por una xilosa isomerasa, en oposición a la transformación en dos etapas de xilosa en xilulosa a través de un intermedio de xilitol catalizado por xilosa reductasa y xilitol deshidrogenasa, respectivamente.

De este modo, la expresión de la secuencia de nucleótidos en la célula anfitriona produce una xilosa isomerasa con una actividad específica de al menos 10 U de actividad de xilosa isomerasa por mg de proteína a 30°C, preferiblemente al menos 20, 25, 30, 50, 100, 200, 300 o 500 U por mg a 30 °C. La actividad específica de la xilosa isomerasa expresada en la célula anfitriona transformada se define en el presente documento como la cantidad de unidades de actividad de xilosa isomerasa por mg de proteína de lisado libre de células de la célula anfitriona, p. ej., un lisado libre de células de levadura. La determinación de la actividad de xilosa isomerasa, la cantidad de proteína y la preparación del lisado libre de células son como se describe en el Ejemplo 1.13. Por consiguiente, la expresión de la secuencia de nucleótidos que codifica la xilosa isomerasa en la célula anfitriona produce una xilosa isomerasa con una actividad específica de al menos 50 U de actividad de xilosa isomerasa por mg de proteína a 30°C, preferiblemente al menos 100, 200, 500, 750 o 1000 U por mg a 30 °C.

Preferiblemente, la expresión de la secuencia de nucleótidos que codifica la xilosa isomerasa en la célula anfitriona produce una xilosa isomerasa con un Km para xilosa que es menor que 50, 40, 30 o 25 mM, más preferiblemente, el Km para xilosa es de aproximadamente 20 mM o menos. Una secuencia de nucleótidos preferida que codifica la xilosa isomerasa puede seleccionarse del grupo que consiste en:

- (a) secuencias de nucleótidos que comprenden una secuencia de nucleótidos que tiene al menos una identidad de secuencia de 80, 90, 95, 97, 98 o 99 % con la secuencia de nucleótidos de la SEQ ID NO. 10.

La secuencia de nucleótidos que codifica la xilosa isomerasa puede codificar una xilosa isomerasa procariótica o una eucariota, es decir, una xilosa isomerasa con una secuencia de aminoácidos que es idéntica a la de una xilosa isomerasa que se da de forma natural en el organismo procariótico o eucariótico. Los autores de la presente invención han encontrado que la capacidad de una determinada xilosa isomerasa para conferir a una célula anfitriona eucariota la capacidad de isomerizar la xilosa en xilulosa no depende tanto de si la isomerasa es de origen procariótico o eucariótico. Más bien esto depende del parentesco de la secuencia de aminoácidos de la isomerasa con la de la secuencia de *Piromyces* (SEQ ID NO. 1). Sorprendentemente, la isomerasa eucariota *Piromyces* está más relacionada con las isomerases procarióticas que con otras isomerases eucariotas conocidas. La isomerasa *Piromyces* comparte 61 % de identidad de aminoácidos con una enzima *Xanthomonas* y 82 % con una enzima *Bacteroides* (SEQ ID NO. 2), mientras que sólo comparte una identidad del 49-52 % con varias xilosa isomerases vegetales. No se han publicado informes de una xilosa isomerasa vegetal que se exprese activamente en levadura. Por el contrario, en el Ejemplo 3 del presente documento se describe que una xilosa isomerasa *Bacteroides* confiere a una célula anfitriona eucariota la capacidad para isomerizar la xilosa en xilulosa y crecer en xilosa como única fuente de carbono. Por lo tanto, una secuencia de nucleótidos preferida codifica una xilosa isomerasa que tiene una secuencia de aminoácidos que está relacionada con la secuencia de *Piromyces* como se definió anteriormente. Una secuencia de nucleótidos preferida codifica una xilosa isomerasa fúngica (p. ej., a partir de un *Basidiomycete*), más preferiblemente una xilosa isomerasa de un hongo anaeróbico, p. ej., una xilosa isomerasa de un hongo anaerobio que pertenezca a las familias *Neocallimastix*, *Caecomyces*, *Piromyces*, *Orpinomyces* o *Ruminomyces*. Como alternativa, una secuencia de nucleótidos preferida codifica una xilosa isomerasa bacteriana, preferiblemente una bacteria Gram-negativa, más preferiblemente una isomerasa de la clase *Bacteroides*, o del género *Bacteroides*, lo más preferiblemente de *B. thetaiotaomicron* (SEQ ID NO. 2).

Para aumentar la probabilidad de que la xilosa isomerasa se exprese en forma activa en una célula anfitriona eucariota tal como levadura, la secuencia de nucleótidos que codifica la xilosa isomerasa puede adaptarse para optimizar su uso de codones al de la célula anfitriona eucariota. La adaptabilidad de una secuencia de nucleótidos que codifica la xilosa isomerasa (u otras enzimas de la invención, véase más adelante) para el uso de codones de la célula anfitriona puede expresarse como índice de adaptación de codones (CAI). El índice de adaptación de codones se define en el presente documento como una medida de la adaptabilidad relativa del uso de codones de un gen hacia el uso de codones de genes muy expresados. La adaptabilidad relativa (w) de cada codón es la relación entre la utilización de cada codón y la del codón más abundante para el mismo aminoácido. El índice CAI se define como la media geométrica de estos valores de adaptabilidad relativa. Se excluyen los codones no sinónimos y los codones de terminación (dependientes del código genético). Los valores de CAI varían de 0 a 1, con los valores más altos indicando una mayor proporción de los codones más abundantes (véase Sharp y Li, 1987, *Nucleic Acids Research* 15: 1281-1295; véase también: Jansen et al., 2003, *Nucleic Acids Res.* 31(8): 2242-51). Una secuencia de nucleótidos adaptada tiene preferiblemente un CAI de al menos 0,2, 0,3, 0,4, 0,5, 0,6 o 0,7.

Una célula anfitriona para la transformación con la secuencia de nucleótidos que codifica la xilosa isomerasa como

se describió anteriormente, es preferiblemente una anfitriona capaz del transporte activo o pasivo de xilosa dentro de la célula. La célula anfitriona contiene preferiblemente glucólisis activa. La célula anfitriona puede contener además una ruta de la pentosa fosfato endógena y puede contener actividad de xilulosa cinasa endógena de manera que la xilulosa isomerizada a partir de xilosa puede ser metabolizada a piruvato. La anfitriona contiene además, preferiblemente, enzimas para la transformación de piruvato en un producto de fermentación deseado tal como etanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, aminoácidos, 1,3-propano- diol, etileno, glicerol, antibióticos  $\beta$ -lactámicos y cefalosporinas.

Una célula anfitriona preferida es una célula anfitriona que es capaz de fermentación alcohólica de forma natural, preferiblemente fermentación alcohólica anaerobia. La célula anfitriona tiene además, preferiblemente, una gran tolerancia a etanol, una gran tolerancia al pH bajo (es decir, capaz de crecer a un pH inferior a 5, 4, 3, o 2,5) y a los ácidos orgánicos como ácido láctico, ácido acético o ácido fórmico y productos de degradación del azúcar tal como furfural e hidroximetilfurfural, y una gran tolerancia a temperaturas elevadas. Cualquiera de estas características o actividades de la célula anfitriona puede estar presente de forma natural en la célula anfitriona o puede ser introducida o modificada por modificación genética. Una célula anfitriona adecuada es un microorganismo eucariótico como p. ej., un hongo, sin embargo, más adecuadas como célula anfitriona son las levaduras o los hongos filamentosos.

Las levaduras se definen en el presente documento como microorganismos eucarióticos e incluyen todas las especies de la subdivisión Eumycotina (Alexopoulos, C. J., 1962, en: Introductory Mycology, John Wiley & Sons, Inc., Nueva York) que crecen, predominantemente, en forma unicelular. Las levaduras pueden crecer bien brotando de un talo unicelular o bien pueden crecer por fisión del organismo. Las levaduras preferidas como células anfitrionas pertenecen a los géneros *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces* y *Yarrowia*. Preferiblemente, la levadura es capaz de fermentación anaeróbica, más preferiblemente de fermentación alcohólica anaeróbica.

Los hongos filamentosos se definen en el presente documento como microorganismos eucarióticos que incluyen todas las formas filamentosas de la subdivisión Eumycotina. Estos hongos se caracterizan por un micelio vegetativo compuesto de quitina, celulosa y otros polisacáridos complejos. Los hongos filamentosos de la presente invención son morfológicamente, fisiológicamente y genéticamente distintos de las levaduras. El crecimiento vegetativo por hongos filamentosos es por alargamiento hifal, y el catabolismo de carbono de la mayoría de los hongos filamentosos es obligatoriamente aeróbico. Los hongos filamentosos preferidos como células anfitrionas pertenecen a los géneros *Aspergillus*, *Trichoderma*, *Humicola*, *Acremonium*, *Fusarium* y *Penicillium*.

A lo largo de los años se han hecho sugerencias para la introducción de diversos organismos para la producción de bioetanol a partir de azúcares de cultivo. En la práctica, sin embargo, todos los principales procesos de producción de bioetanol han seguido utilizando las levaduras del género *Saccharomyces* como productor de etanol. Esto se debe a las muchas características atractivas de la especie *Saccharomyces* para los procesos industriales, es decir, una gran tolerancia a ácidos, a etanol y osmotolerancia, capacidad de crecimiento anaeróbico y, por supuesto, su gran capacidad de fermentación alcohólica. Las especies de levadura preferidas como células anfitrionas incluyen *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*; *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus* y *K. fragilis*.

La célula anfitriona de la invención es, de este modo, una célula anfitriona que se transforma con un constructo de ácido nucleico que comprende la secuencia de nucleótidos que codifica la xilosa isomerasa como se ha definido anteriormente. El constructo de ácido nucleico que comprende la secuencia que codifica xilosa isomerasa es, preferiblemente, capaz de la expresión de la xilosa isomerasa en la célula anfitriona. Con este fin, el constructo de ácido nucleico puede construirse como se describe en, p. ej., el documento WO 03/0624430. La célula anfitriona puede comprender una sola copia pero preferiblemente comprende múltiples copias del constructo de ácido nucleico. El constructo de ácido nucleico puede mantenerse de forma episómica y, con ello, comprender una secuencia para la replicación autónoma, tal como una secuencia ARS. Los constructos de ácido nucleico episómico adecuados pueden, p. ej. basarse en los plásmidos de levadura  $2\mu$  o pKD1 (Fleer et al., 1991, Biotechnology 9: 968-975). Preferiblemente, sin embargo, el constructo de ácido nucleico está integrado en una o más copias en el genoma de la célula anfitriona. La integración en el genoma de la célula anfitriona puede ocurrir al azar por recombinación ilegítima, pero el constructo de ácido nucleico se integra en el genoma de la célula anfitriona, preferiblemente, por recombinación homóloga como es bien sabido en la técnica de genética molecular fúngica (véanse, p. ej., los documentos WO 90/14423, EP-A-0 481 008, EP-A-0 635 574 y US 6.265.186).

En un primer aspecto de la invención, la célula anfitriona de la invención comprende una modificación genética que aumenta el flujo de la ruta de la pentosa fosfato. En particular, la modificación genética provoca un aumento del flujo de la parte no oxidativa de la ruta de la pentosa fosfato. Una modificación genética que provoca un aumento del flujo de la parte no oxidativa de la ruta de la pentosa fosfato se entiende, en el presente documento, que significa una modificación que aumenta el flujo en al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 20 en comparación con el flujo en una cepa que sea genéticamente idéntica excepto por la modificación genética que provoca el aumento del flujo. El flujo de la parte no oxidativa de la ruta de la pentosa fosfato puede medirse cultivando la anfitriona modificada en xilosa como única fuente de carbono, determinando la velocidad específica del consumo de xilosa y substrayendo la velocidad específica de producción de xilitol de la velocidad específica del consumo de xilosa, en el caso de que se produzca algo de xilitol. Sin embargo, el flujo de la parte no oxidativa de la ruta de la pentosa fosfato es proporcional

a la velocidad de crecimiento de la xilosa como única fuente de carbono, preferiblemente con la velocidad de crecimiento anaeróbico en xilosa como única fuente de carbono. Existe una relación lineal entre la velocidad de crecimiento en xilosa como única fuente de carbono ( $\mu_{m\acute{a}x}$ ) y el flujo de la parte no oxidativa de la ruta de la pentosa fosfato. La velocidad específica del consumo de xilosa ( $Q_s$ ) es igual a la velocidad de crecimiento ( $\mu$ ) dividida por el rendimiento de biomasa en azúcar ( $Y_{xs}$ ) porque el rendimiento de la biomasa en el azúcar es constante (en un conjunto dado de condiciones: anaeróbico, medio de cultivo, pH, antecedentes genéticos de la cepa, etc., es decir,  $Q_s = \mu/Y_{xs}$ ). Por lo tanto, el aumento del flujo de la parte no oxidativa de la ruta de la pentosa fosfato puede deducirse del aumento de la velocidad de crecimiento máximo en estas condiciones.

Las modificaciones genéticas que aumentan el flujo de la ruta de la pentosa fosfato pueden ser introducidas en la célula anfitriona de varias maneras. Estas incluyen, p. ej., alcanzar mayores niveles de actividad en estado estacionario de la xilulosa cinasa y/o una o más de las enzimas de la parte no oxidativa de la ruta de la pentosa fosfato y/o un nivel reducido en estado estacionario de la actividad de la aldosa reductasa inespecífica. Estos cambios en los niveles de actividad en estado estacionario pueden efectuarse mediante la selección de mutantes (espontáneos o inducidos por productos químicos o radiación) y/o por tecnología de ADN recombinante, p. ej., por sobreexpresión o inactivación, respectivamente, de genes que codifican las enzimas o factores que regulan estos genes.

En una célula anfitriona preferida, la modificación genética comprende la sobreexpresión de al menos una enzima de la ruta de la pentosa fosfato (la parte no oxidativa). Preferiblemente, la enzima se selecciona del grupo que consiste en las enzimas que codifican ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa, transcetolasa y transaldolasa. Pueden sobreexpresarse varias combinaciones de enzimas de la ruta de la pentosa fosfato (la parte no oxidativa). P. ej., las enzimas que son sobreexpresadas pueden ser al menos las enzimas ribulosa-5-fosfato isomerasa y ribulosa-5-fosfato epimerasa; o al menos las enzimas ribulosa-5-fosfato isomerasa y transcetolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato epimerasa y transcetolasa; o al menos las enzimas ribulosa-5-fosfato epimerasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato epimerasa, transcetolasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa, transcetolasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa y transcetolasa. En una realización de la invención, cada una de las enzimas ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa, transcetolasa y transaldolasa son sobreexpresadas en la célula anfitriona. Más preferida es una célula anfitriona en la que la modificación genética comprende al menos la sobreexpresión tanto de las enzimas transcetolasa como de la transaldolasa, ya que dicha célula anfitriona es ya capaz de crecimiento anaeróbico en xilosa. De hecho, en algunas condiciones los autores de la invención han encontrado que las células anfitrionas que sobreexpresan sólo la transcetolasa y la transaldolasa tienen ya la misma velocidad de crecimiento anaeróbico en xilosa que las células anfitrionas que sobreexpresan las cuatro enzimas, es decir, ribulosa-5-fosfato isomerasa, la ribulosa-5-fosfato epimerasa, transcetolasa y transaldolasa. Además, las células anfitrionas que sobreexpresan ambas enzimas ribulosa-5-fosfato isomerasa y ribulosa-5-fosfato epimerasa son preferibles a las células anfitrionas que sobreexpresan sólo la isomerasa o sólo la epimerasa ya que la sobreexpresión de sólo una de estas enzimas puede producir desequilibrios metabólicos.

Existen diversos medios disponibles en la técnica para la sobreexpresión de enzimas en las células anfitrionas de la invención. En particular, una enzima puede ser sobreexpresada aumentando el número de copias del gen que codifica la enzima en la célula anfitriona, p. ej., integrando copias adicionales del gen en el genoma de la célula anfitriona, expresando el gen de un vector de expresión multicopia episómico o introduciendo un vector de expresión episómico que comprende múltiples copias del gen.

Como alternativa, se puede lograr la sobreexpresión de enzimas en las células anfitrionas de la invención utilizando un promotor que no sea nativo para la secuencia que codifica la enzima a sobreexpresar, es decir, un promotor que sea heterólogo con la secuencia codificante a la que está enlazado operativamente. Aunque el promotor sea, preferiblemente, heterólogo con la secuencia codificante a la que está enlazado operativamente, también se prefiere que el promotor sea homólogo, es decir, endógeno con la célula anfitriona. Preferiblemente, el promotor heterólogo es capaz de producir un mayor nivel en estado estacionario del transcrito que comprende la secuencia codificante (o es capaz de producir más moléculas de transcrito, es decir, moléculas de ARNm, por unidad de tiempo) que el promotor que es nativo de la secuencia codificante, preferiblemente en condiciones donde xilosa, o xilosa y glucosa, están disponibles como fuentes de carbono, más preferiblemente como principales fuentes de carbono (es decir, más del 50 % de la fuente de carbono disponible consiste en xilosa, o en xilosa y glucosa), lo más preferiblemente como únicas fuentes de carbono. Los promotores adecuados en este contexto incluyen tanto promotores naturales constitutivos como inducibles, así como promotores diseñados por ingeniería. Un promotor preferido para uso en la presente invención será además insensible a la represión del catabolito (glucosa) y/o no requerirá, preferiblemente, xilosa para la inducción.

Los promotores que tienen estas características están ampliamente disponibles y son conocidos por el experto en la materia. Ejemplos adecuados de tales promotores incluyen p. ej., promotores de genes glicolíticos, tales como los promotores fosfofructocinasa (PPK), triosa fosfato isomerasa (TPI), gliceraldehído-3-fosfato deshidrogenasa (GPD, TDH3 o GAPDH), piruvato cinasa (PYK), fosfoglicerato cinasa (PGK) de levaduras u hongos filamentosos; más detalles sobre tales promotores de levadura se pueden encontrar en el documento WO 93/03159. Otros promotores

útiles son los promotores de genes que codifican la proteína ribosómica, el promotor del gen de la lactasa (LAC4), promotores de alcohol deshidrogenasa (ADH1, ADH4 y similares) y el promotor de enolasa (ENO). Otros promotores, tanto constitutivos como inducibles, y potenciadores o secuencias activadoras aguas arriba serán conocidos por los expertos en la técnica. Los promotores utilizados en las células anfitrionas de la invención pueden ser modificados, si se desea, para afectar sus características de control.

La secuencia de codificación utilizada para la sobreexpresión de las enzimas es preferiblemente homóloga a la célula anfitriona de la invención. Sin embargo, del mismo modo pueden aplicarse secuencias codificantes que sean heterólogas para la célula anfitriona de la invención.

Una secuencia de nucleótidos utilizada para la sobreexpresión de ribulosa-5-fosfato isomerasa en la célula anfitriona de la invención es una secuencia de nucleótidos que codifica un polipéptido con actividad de ribulosa-5-fosfato isomerasa, por lo que preferiblemente el polipéptido tiene una secuencia de aminoácidos que tiene al menos 50, 60, 70, 80, 90 o 95 % de identidad con la SEQ ID NO. 4 o por lo que la secuencia de nucleótidos es capaz de hibridarse con la secuencia de nucleótidos de la SEQ ID NO. 12, en condiciones moderadas, preferiblemente en condiciones rigurosas.

Una secuencia de nucleótidos utilizada para la sobreexpresión de la ribulosa-5-fosfato epimerasa en la célula anfitriona de la invención es una secuencia de nucleótidos que codifica un polipéptido con actividad de ribulosa-5-fosfato epimerasa, por lo que preferiblemente el polipéptido tiene una secuencia de aminoácidos que tiene al menos 50, 60, 70, 80, 90 o 95 % de identidad con la SEQ ID NO. 5 o por lo que la secuencia de nucleótidos es capaz de hibridarse con la secuencia de nucleótidos de la SEQ ID NO. 13, en condiciones moderadas, preferiblemente en condiciones rigurosas.

Una secuencia de nucleótidos utilizada para la sobreexpresión de transcetolasa en la célula anfitriona de la invención es una secuencia de nucleótidos que codifica un polipéptido con actividad de transcetolasa, por lo que preferiblemente el polipéptido tiene una secuencia de aminoácidos que tiene al menos 50, 60, 70, 80, 90 o 95 % de identidad con la SEQ ID NO. 6 o por lo que la secuencia de nucleótidos es capaz de hibridarse con la secuencia de nucleótidos de la SEQ ID NO. 14, en condiciones moderadas, preferiblemente en condiciones rigurosas.

Una secuencia de nucleótidos utilizada para la sobreexpresión de transaldolasa en la célula anfitriona de la invención es una secuencia de nucleótidos que codifica un polipéptido con actividad de transaldolasa, por lo que preferiblemente el polipéptido tiene una secuencia de aminoácidos que tiene al menos 50, 60, 70, 80, 90 o 95 % de identidad con la SEQ ID NO. 7 o por lo que la secuencia de nucleótidos es capaz de hibridarse con la secuencia de nucleótidos de la SEQ ID NO. 15, en condiciones moderadas, preferiblemente en condiciones rigurosas.

La sobreexpresión de una enzima, cuando se refiere a la producción de la enzima en una célula anfitriona genéticamente modificada, significa que la enzima se produce a un nivel más alto de actividad enzimática específica en comparación con la célula anfitriona no modificada en idénticas condiciones. Generalmente esto significa que la proteína (o proteínas en el caso de enzimas multi-subunidad) enzimáticamente activa se produce en mayores cantidades, o más bien a un nivel de estado estacionario más elevado en comparación con la célula anfitriona no modificada en idénticas condiciones. De forma similar, esto generalmente significa que el ARNm que codifica la proteína enzimáticamente activa se produce en mayores cantidades o más bien, de nuevo, a un nivel de estado estacionario más elevado en comparación con la célula anfitriona no modificada en idénticas condiciones. La sobreexpresión de una enzima se determina preferiblemente midiendo el nivel de la actividad específica de la enzima en la célula anfitriona usando ensayos enzimáticos apropiados como se describe en el presente documento. Como alternativa, la sobreexpresión de la enzima puede determinarse indirectamente cuantificando el nivel específico en estado estacionario de la proteína de la enzima, p. ej., utilizando anticuerpos específicos de la enzima, o cuantificando el nivel estacionario específico del ARNm que codifica la enzima. Este último puede ser particularmente adecuado para enzimas de la ruta de la pentosa fosfato para las que los ensayos enzimáticos no son fácilmente viables ya que los sustratos para las enzimas no están comercialmente disponibles. Preferiblemente, en las células anfitrionas de la invención, una enzima a sobreexpresar se sobreexpresa mediante al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 20 en comparación con una cepa que sea genéticamente idéntica excepto por la modificación genética que causa la sobreexpresión. Debe entenderse que estos niveles de sobreexpresión pueden aplicarse al nivel de estado estacionario de la actividad de la enzima, al nivel de estado estacionario de la proteína de la enzima, así como al nivel de estado estacionario del transcrito que codifica la enzima.

En un segundo aspecto de la invención, la célula anfitriona de la invención comprende una modificación genética que aumenta la actividad específica de xilulosa cinasa. Preferiblemente, la modificación genética provoca la sobreexpresión de una xilulosa cinasa, p. ej., por sobreexpresión de una secuencia de nucleótidos que codifica una xilulosa cinasa. El gen que codifica la xilulosa cinasa puede ser endógeno a la célula anfitriona o puede ser una xilulosa cinasa que sea heteróloga con la célula anfitriona. Una secuencia de nucleótidos utilizada para la sobreexpresión de la xilulosa cinasa en la célula anfitriona de la invención es una secuencia de nucleótidos que codifica un polipéptido con actividad de xilulosa cinasa, por lo que preferiblemente el polipéptido tiene una secuencia de aminoácidos que tiene al menos 50, 60, 70, 80, 90 o 95 % de identidad con la SEQ ID NO. 3 o por lo que la secuencia de nucleótidos es capaz de hibridarse con la secuencia de nucleótidos de la SEQ ID NO. 11, en condiciones moderadas, preferiblemente en condiciones rigurosas.

Una xilulosa cinasa particularmente preferida es una xilosa cinasa que está relacionada con la xilulosa cinasa de *Piromyces* (*xy1B*; véase el documento WO 03/0624430). Esta xilulosa cinasa *Piromyces* está en realidad más relacionada con la cinasa procariótica que con todas las cinasas eucariotas conocidas tales como la cinasa de levadura (SEQ ID NO. 3). Las xilulosa cinasas eucariotas se han indicado como cinasas de azúcares no específicas, que tienen un amplio intervalo de sustratos que incluye xilulosa. Por el contrario, las xilulosa cinasas procarióticas, con las que está relacionada más estrechamente la cinasa de *Piromyces*, han sido indicadas como cinasas más específicas para xilulosa, es decir, que tienen un intervalo de sustrato más estrecho. Por lo tanto, una secuencia de nucleótidos más preferida para uso en la sobreexpresión de xilulosa cinasa en la célula anfitriona de la invención es una secuencia de nucleótidos que codifica un polipéptido con actividad xilulosa cinasa, por lo que preferiblemente el polipéptido tiene una secuencia de aminoácidos que tiene al menos 45, 50, 60, 65, 70, 80, 90 o 95 % de identidad con la SEQ ID NO. 17 o por lo que la secuencia de nucleótidos es capaz de hibridar con la secuencia de nucleótidos de la SEQ ID NO. 18, en condiciones moderadas, preferiblemente en condiciones rigurosas.

En las células anfitrionas de la invención, la modificación genética que aumenta la actividad específica de xilulosa cinasa puede combinarse con cualquiera de las modificaciones que aumentan el flujo de la ruta de la pentosa fosfato como se ha descrito anteriormente, pero esta combinación no es esencial para la invención. De este modo, se incluye específicamente en la invención una célula anfitriona de la invención que comprende sólo una modificación genética que aumenta la actividad específica de xilulosa cinasa. Los diversos medios disponibles en la técnica para conseguir y analizar la sobreexpresión de una xilulosa cinasa en las células anfitrionas de la invención son los mismos que se han descrito anteriormente para las enzimas de la ruta de la pentosa fosfato. Preferiblemente, en las células anfitrionas de la invención, una xilulosa cinasa a sobreexpresar se sobreexpresa mediante al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 20 en comparación con una cepa que seas genéticamente idéntica excepto por la modificación genética que causa la sobreexpresión. Debe entenderse que estos niveles de sobreexpresión pueden aplicarse al nivel de estado estacionario de la actividad de la enzima, al nivel de estado estacionario de la proteína de la enzima, así como al nivel de estado estacionario del transcrito que codifica la enzima.

En un tercer aspecto de la invención, la célula anfitriona de la invención comprende una modificación genética que reduce la actividad de la aldosa reductasa inespecífica en la célula anfitriona. Preferiblemente, la actividad de la aldosa reductasa inespecífica se reduce en la célula anfitriona mediante una o más modificaciones genéticas que reducen la expresión de o inactivan un gen que codifica una aldosa reductasa inespecífica. Preferiblemente, las modificaciones genéticas reducen o inactivan la expresión de cada copia endógena de un gen que codifica una aldosa reductasa inespecífica en la célula anfitriona. Las células anfitrionas pueden comprender múltiples copias de genes que codifican aldosa reductasas inespecíficas como resultado de di-, poli- o aneu-ploidías, y/o la célula anfitriona puede contener varias (iso)enzimas diferentes con actividad de la aldosa reductasa que difieren en la secuencia de aminoácidos y que están codificadas por un gen diferente. También en tales casos, preferiblemente, la expresión de cada gen que codifica una aldosa reductasa inespecífica se reduce o se inactiva. Preferiblemente, el gen se inactiva por eliminación de al menos parte del gen o por interrupción del gen, por lo que en este contexto el término gen incluye también cualquier secuencia no codificante aguas arriba o aguas abajo de la secuencia codificante, cuya eliminación (parcial) o inactivación da como resultado una reducción de la expresión de la actividad de la aldosa reductasa inespecífica en la célula anfitriona. Una secuencia de nucleótidos que codifica una aldosa reductasa cuya actividad debe reducirse en la célula anfitriona de la invención es una secuencia de nucleótidos que codifica un polipéptido con actividad de aldosa reductasa, por la que preferiblemente el polipéptido tiene una secuencia de aminoácidos que tiene al menos 50, 60, 70, 80, 90 o 95 % de identidad con la SEQ ID NO. 8 o por la que la secuencia de nucleótidos es capaz de hibridar con la secuencia de nucleótidos de la SEQ ID NO. 16 en condiciones moderadas, preferiblemente en condiciones rigurosas.

En las células anfitrionas de la invención, la modificación genética que reduce la actividad de la aldosa reductasa inespecífica en la célula anfitriona puede combinarse con cualquiera de las modificaciones que aumentan el flujo de la ruta de la pentosa fosfato y/o con cualquiera de las modificaciones que aumentan la actividad de la xilulosa cinasa específica en las células anfitrionas como se ha descrito anteriormente, pero estas combinaciones no son esenciales para la invención. De este modo, se incluye específicamente en la invención una célula anfitriona de la invención que comprende sólo una modificación genética que reduce la actividad de la aldosa reductasa inespecífica en la célula anfitriona.

En un aspecto adicional, la invención se refiere a células anfitrionas modificadas que están adaptadas además a la utilización de xilosa mediante la selección de mutantes, bien espontáneos o bien inducidos (por ejemplo por radiación o productos químicos), para el crecimiento en xilosa, preferiblemente en xilosa como única fuente de carbono, y más preferiblemente en condiciones anaerobias. La selección de mutantes puede realizarse mediante el paso en serie de cultivos como p. ej., el descrito por Kuyper et al. (2004, FEMS Yeast Res. 4: 655-664) o mediante cultivo bajo presión selectiva en un cultivo quimiostático como se describe en el Ejemplo 4 del presente documento.

En una célula anfitriona preferida de la invención al menos una de las modificaciones genéticas descritas anteriormente, incluidas las modificaciones obtenidas mediante la selección de mutantes, confieren a la célula anfitriona la capacidad de crecer en xilosa como fuente de carbono, preferiblemente como única fuente de carbono, y preferiblemente bajo condiciones anaerobias. Preferiblemente, la célula anfitriona modificada no produce esencialmente xilitol, p. ej., el xilitol producido está por debajo del límite de detección o, p. ej., menor que 5, 2, 1, 0,5, o 0,3 % del carbono consumido en una base molar.

Preferiblemente, la célula anfitriona modificada tiene la capacidad de crecer en xilosa como única fuente de carbono a una velocidad de al menos 0,05, 0,1, 0,2, 0,25 o 0,3 h<sup>-1</sup> en condiciones aerobias o, si procede, a una velocidad de al menos 0,03, 0,05, 0,07, 0,08, 0,09, 0,1, 0,12, 0,15 o 0,2 h<sup>-1</sup> en condiciones anaerobias. Preferiblemente, la célula anfitriona modificada tiene la capacidad de crecer en una mezcla de glucosa y xilosa (en una relación en peso de 1:1) como única fuente de carbono a una velocidad de al menos 0,05, 0,1, 0,2, 0,25 o 0,3 h<sup>-1</sup> en condiciones aerobias o, si procede, a una velocidad de al menos 0,03, 0,05, 0,1, 0,12, 0,15 o 0,2 h<sup>-1</sup> en condiciones anaerobias.

Preferiblemente, la célula anfitriona modificada tiene una velocidad específica de consumo de xilosa de al menos 346, 350, 400, 500, 600, 750 o 1000 mg de xilosa/g de células/h. Preferiblemente, la célula anfitriona modificada tiene un rendimiento de producto de fermentación (tal como etanol) en xilosa que tiene al menos un 55, 60, 70, 80, 85, 90, 95 o 98 % de rendimiento del producto de fermentación (tal como etanol) de la célula anfitriona en glucosa. Más preferiblemente, el rendimiento de la célula anfitriona modificada de producto de fermentación (tal como etanol) en xilosa es igual al rendimiento de la célula anfitriona del producto de fermentación (tal como etanol) sobre la glucosa. Igualmente, el rendimiento de biomasa de la célula anfitriona modificada en xilosa tiene preferiblemente al menos un 55, 60, 70, 80, 85, 90, 95 o 98 % de rendimiento de biomasa de la célula anfitriona en glucosa. Más preferiblemente, el rendimiento de biomasa de la célula anfitriona modificada en xilosa es igual al rendimiento de biomasa de la célula anfitriona en glucosa. Se entiende que en la comparación de los rendimientos en glucosa y xilosa ambos rendimientos se comparan en condiciones aerobias o ambos en condiciones anaerobias.

En un aspecto preferido, la célula anfitriona modificada de la invención es una célula anfitriona para la producción de etanol. En otro aspecto, la invención se refiere a una célula anfitriona transformada para la producción de productos de fermentación distintos del etanol. Tales productos de fermentación no etanólicos incluyen en principio cualquier producto químico masivo o fino que pueda obtenerse mediante un microorganismo eucariótico tal como una levadura o un hongo filamentoso. Tales productos de fermentación incluyen p. ej., ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, aminoácidos, 1,3-propano-diol, etileno, glicerol, antibióticos β-lactámicos y cefalosporinas. Una célula anfitriona modificada preferida de la invención para la producción de productos de fermentación no etanólicos es una célula anfitriona que contiene una modificación genética que da como resultado una actividad disminuida de alcohol deshidrogenasa.

En un aspecto adicional, la invención se refiere a procesos de fermentación en los que las células anfitrionas modificadas de la invención se utilizan para la fermentación de una fuente de carbono que comprende una fuente de xilosa, tal como xilosa. Además de una fuente de xilosa, la fuente de carbono en el medio de fermentación puede comprender también una fuente de glucosa. La fuente de xilosa o glucosa puede ser xilosa o glucosa como tal o puede ser cualquier hidrato de carbono oligo- o polímero que comprende unidades de xilosa o glucosa, tal como p. ej., lignocelulosa, xilanos, celulosa, almidón y similares. Para la liberación de unidades de xilosa o glucosa a partir de tales hidratos de carbono, pueden añadirse carbohidrasas apropiadas (tal como xilanasas, glucanasas, amilasas y similares) al medio de fermentación o pueden ser producidas por la célula anfitriona modificada. En este último caso, la célula anfitriona modificada puede ser diseñada por ingeniería genética para producir y excretar tales carbohidrasas. Una ventaja adicional de usar fuentes oligo- o poliméricas de glucosa es que permite mantener una baja (o más baja) concentración de glucosa libre durante la fermentación, p. ej., utilizando cantidades de las carbohidrasas que limitan la velocidad. Esto, a su vez, evitará la represión de los sistemas requeridos para el metabolismo y transporte de azúcares que no son glucosa como la xilosa. En un procedimiento preferido, la célula anfitriona modificada fermenta tanto la xilosa como la glucosa, preferiblemente simultáneamente, en cuyo caso se utiliza preferiblemente una célula anfitriona modificada que sea insensible a la represión de la glucosa para prevenir el crecimiento diáuxico. Además de una fuente de xilosa (y glucosa) como fuente de carbono, el medio de fermentación comprenderá además el ingrediente apropiado requerido para el crecimiento de la célula anfitriona modificada. Las composiciones de medios de fermentación para el crecimiento de microorganismos tales como las levaduras son bien conocidas en la técnica.

El procedimiento de fermentación es un procedimiento para la producción de un producto de fermentación tal como p. ej., etanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, aminoácidos, 1,3-propanodiol, etileno, glicerol, antibióticos β-lactámicos tales como la Penicilina G o la Penicilina V y derivados fermentativos de los mismos, y cefalosporinas. El procedimiento de fermentación puede ser un procedimiento de fermentación aeróbico o anaeróbico. Un procedimiento de la fermentación anaeróbico se define en el presente documento como un procedimiento de fermentación que se lleva a cabo en ausencia de oxígeno o en el que no se consume sustancialmente oxígeno, preferiblemente menos que 5, 2,5 o 1 mmol//h, más preferiblemente 0 mmol//h (es decir, el consumo de oxígeno no es detectable), y en el que las moléculas orgánicas sirven como dadores de electrones y como aceptores de electrones. En ausencia de oxígeno, el NADH producido en la glucólisis y formación de biomasa, no puede ser oxidado por fosforilación oxidativa. Para resolver este problema, muchos microorganismos usan piruvato o uno de sus derivados como un electrón y un aceptor de hidrógeno, regenerando de ese modo NAD<sup>+</sup>. De este modo, en un procedimiento de fermentación anaeróbica preferido, el piruvato se usa como un electrón (y aceptor de hidrógeno) y se reduce a productos de fermentación tales como etanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, aminoácidos, 1,3-propanodiol, etileno, glicerol, antibióticos β-lactámicos y cefalosporinas.

El procedimiento de fermentación se lleva a cabo preferiblemente a una temperatura que es óptima para la célula anfitriona modificada. De este modo, para la mayoría de células anfitrionas levaduras u hongos, el procedimiento de

fermentación se realiza a una temperatura que es inferior a 42 °C, preferiblemente inferior a 38 °C. Para las células anfitrionas levaduras u hongos filamentosos, el procedimiento de fermentación se realiza, preferiblemente, a una temperatura que es menor que 35, 33, 30 o 28 °C y a una temperatura que es mayor que 20, 22 o 25 °C.

5 Un procedimiento preferido es un procedimiento para la producción de etanol, por lo que el procedimiento comprende las etapas de: (a) fermentar un medio que contiene una fuente de xilosa con una célula anfitriona modificada como se ha definido anteriormente, por el que la célula anfitriona fermenta xilosa a etanol; y  
 10 opcionalmente, (b) recuperación del etanol. El medio de fermentación puede comprender también una fuente de glucosa que también se fermenta a etanol. En el procedimiento, la productividad volumétrica de etanol es preferiblemente al menos 0,5, 1,0, 1,5, 2,0, 2,5, 3,0, 5,0 o 10,0 g de etanol por litro y hora. El rendimiento de etanol en xilosa y/o glucosa en el procedimiento es preferiblemente de al menos 50, 60, 70, 80, 90, 95 o 98 %. El rendimiento de etanol se define en el presente documento como un porcentaje del rendimiento máximo teórico, que, para glucosa y xilosa es de 0,51 g. de etanol por g. de glucosa o xilosa.

15 En un aspecto adicional, la invención se refiere a un procedimiento para producir un producto de fermentación seleccionado del grupo que consiste en ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, aminoácidos, 1,3-propanodiol, etileno, glicerol, antibióticos  $\beta$ -lactámicos y cefalosporinas. El procedimiento comprende, preferiblemente, las etapas de (a) fermentar un medio que contiene una fuente de xilosa con una célula anfitriona modificada como se ha definido anteriormente, por el que la célula anfitriona fermenta xilosa al producto de fermentación y, opcionalmente, (b) recuperación del producto de fermentación. En un procedimiento preferido, el medio también contiene una fuente de glucosa.

## 20 Modificaciones genéticas

Para la sobreexpresión de enzimas en las células anfitrionas de las invenciones como se ha descrito anteriormente, así como para la modificación genética adicional de las células anfitrionas, preferiblemente levaduras, las células anfitrionas se transforman con los diversos constructos de ácido nucleico de la invención por procedimientos bien conocidos en la técnica. Tales procedimientos son p. ej., conocidos a través de los manuales de normas, tales como  
 25 Sambrook y Russel (2001) *Molecular Cloning: A Laboratory Manual* (3ª edición), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, o F. Ausubel et al, eds. "Current protocols in molecular biology", Green Publishing and Wiley Interscience, Nueva York (1987). Procedimientos para la transformación y modificación genética de células anfitrionas fúngicas se conocen, p. ej., a partir de los documentos EP-A-0 635 574, WO 98/46772, WO 99/60102 y WO 00/37671.

30 Anteriormente se han descrito los promotores para uso en los constructos de ácido nucleico para la sobreexpresión de enzimas en las células anfitrionas de la invención. En los constructos de ácido nucleico para la sobreexpresión, el extremo 3' de la secuencia de ácido nucleotídico que codifica la enzima (o enzimas) está preferiblemente enlazado operativamente a una secuencia terminadora de la transcripción. Preferiblemente, la secuencia terminadora es operable en una célula anfitriona de elección, tal como p. ej., la especie de la levadura de elección. En cualquier  
 35 caso, la elección del terminador no es crítica; puede, p. ej., ser de cualquier gen de levadura, aunque los terminadores a veces pueden funcionar si proceden de un gen eucariótico, no levadura. La secuencia de terminación de la transcripción comprende además, preferiblemente, una señal de poliadenilación.

Opcionalmente, un marcador seleccionable puede estar presente en el constructo de ácido nucleico. Tal como se utiliza en el presente documento, el término marcador se refiere a un gen que codifica un rasgo o un fenotipo que  
 40 permite la selección, o el cribado, de una célula anfitriona que contiene el marcador. El gen marcador puede ser un gen de resistencia a los antibióticos por el que el antibiótico apropiado puede usarse para seleccionar las células transformadas de entre células que no se transforman. Ejemplos de marcadores adecuados de resistencia a los antibióticos incluyen p. ej., dihidrofolato reductasa, higromicin-B-fosfotransferasa, 3'-O-fosfotransferasa II (resistencia a kanamicina, neomicina y a G418). Aunque los marcadores de resistencia a los antibióticos pueden ser más convenientes para la transformación de células anfitrionas poliploides, preferiblemente, sin embargo, se utilizan marcadores no resistentes a los antibióticos, tales como marcadores auxotróficos (URA3, TRP1, LEU2) o el gen TPI de *S. pombe* (descrito por Russell PR, 1985, Gene 40: 125-130). En una realización preferida, las células anfitrionas transformadas con los constructos de ácido nucleico están libres de genes marcadores. En el documento EP-A-0 635 574 se divulgan procedimientos para construir células anfitrionas microbianas libres de genes marcadores  
 45 recombinantes y se basan en el uso de marcadores bidireccionales tales como el gen *amdS* (acetamidasa) de *A. nidulans* o los genes URA3 y LYS2 de levadura. Como alternativa, un marcador cribable como la Proteína Fluorescente Verde, *lacZ*, luciferasa, cloranfenicol acetiltransferasa, beta-glucuronidasa se puede incorporar en los constructos de ácido nucleico de la invención que permiten cribar las células transformadas.

Los elementos opcionales adicionales que pueden estar presentes en los constructos de ácido nucleico de la invención incluyen, pero no se limitan a, una o más secuencias líderes, potenciadores, factores de integración y/o genes reporteros, secuencias de intrones, centrómeros, telómeros y/o secuencias de unión a la matriz (MAR). Los constructos de ácido nucleico de la invención pueden comprender además una secuencia para la replicación autónoma, tal como una secuencia ARS. Los constructos de ácido nucleico episómico adecuados pueden, p. ej., basarse en los plásmidos de levadura  $2\mu$  o pKD1 (Fleer et al., 1991, Biotechnology 9: 968-975). Como alternativa, el  
 60 constructo de ácido nucleico puede comprender secuencias para la integración, preferiblemente mediante

recombinación homóloga. Dichas secuencias pueden ser así secuencias homólogas al sitio diana para la integración en el genoma de la célula anfitriona. Los constructos de ácido nucleico de la invención pueden proporcionarse de una manera conocida per se, que generalmente implica técnicas tales como restringir y enlazar secuencias de ácidos nucleicos/ácido nucleico, para lo cual se hace referencia a los manuales estándar, tales como Sambrook y Russel (2001) "Molecular Cloning: A Laboratory Manual (3ª edición), Cold Spring Harbor Laboratory, Cold Spring Harbour Laboratory Press, o F. Ausubel et al, eds., "Current protocols in molecular biology ", Green Publishing y Wiley Interscience, Nueva York (1987).

Los procedimientos para la inactivación y la interrupción génica en levadura u hongos filamentosos son bien conocidos en la técnica (véase, p. ej., Fincham, 1989, Microbiol Rev. 53(1): 148-70 y EP-A-0 635 574).

10 Descripción de las figuras

Figura 1. Gráfico típico del crecimiento anaeróbico de la cepa RWB 212 en fermentadores en medio sintético con 2 % (p/v) de xilosa como fuente de carbono, los experimentos duplicados difirieron en menos del 5 %. Panel A: Xilosa (●), etanol (○), glicerol (■) y CO<sub>2</sub> acumulado producido por litro deducido del análisis de gas (-). Panel B: peso en seco (●), acetato (○), xilitol (□), succinato (▲), lactato (Δ)

15 Figura 2. Gráfico típico del crecimiento anaeróbico de la cepa RWB 212 en fermentadores en medio sintético con glucosa al 2 % (p/v) y xilosa al 2 % (p/v) como fuente de carbono, los experimentos duplicados difirieron en menos del 5 %. Panel A: Glucosa (●), xilosa (○), etanol (■), glicerol (□) y CO<sub>2</sub> acumulado producido por litro deducido del análisis de gas (-). Panel B: peso en seco (●), acetato (○), xilitol (■), lactato (□) succinato (▲).

**Ejemplos**

20 1. Materiales y procedimientos

1.1. Construcción de plásmidos

Con el fin de integrar el promotor *TPI1* delante de varios de los genes diana se construyeron plásmidos. Primero se cortó el promotor *TPI1* como un fragmento *XhoI-EcoRV* de pYX012-Aat (A. A. Winkler, derivado de pYX012 (R & D Systems, Minneapolis, MN, EE. UU.)) y se ligó al corte de pUG6 [3] con *SalI-PvuII*. Esto dio pUG6P<sub>TPI</sub>, que después pudo utilizarse una casete de integración Kanlox-P<sub>TPI</sub> para la PCR.

25 En los casos en que los ORF supuestos se localizaban muy cerca del ATG de los genes diana, se clonaron esos genes en pUG6P<sub>TPI</sub>. Se aislaron un fragmento *RPE1* de 0,8 kb y un fragmento *TKL1* de 2,3 kb a partir del gel y se cortaron con *EcoRI* y *XhoI* (presentes en los cebadores, véase la Tabla 3) y se ligaron en pUG6P<sub>TPI</sub> digerido con *EcoRI-SalI*, dando como resultado pUG6P<sub>TPI</sub>-RPE1 y pUG6P<sub>TPI</sub>-TKL1.

30 Con el fin de aumentar la actividad de la xilulocinasa, el gen *XKS1* se amplificó por PCR como un fragmento *SpeI-SalI* (sitios en los cebadores, véase la Tabla 3) y se clonó en corte de p415ADH [4] con *XbaI-XhoI*, dando como resultado p415ADHXKS.

Se usaron endonucleasas de restricción (New England Biolabs, Beverly, MA, EE. UU. y Roche, Basilea, Suiza) y ADN ligasa (Roche) según las especificaciones de los fabricantes. El aislamiento de plásmido de *E. coli* se realizó con el kit Qiaprep Spin Miniprep (Qiagen, Hilden, Alemania). Los fragmentos de ADN se separaron en un 1 % de gel agarosa (Sigma, St. Louis, MO, EE. UU.) en 1xTBE [5]. El aislamiento de los fragmentos del gel se llevó a cabo con el kit Qiaquick Gel Extraction (Qiagen). La amplificación de *RPE1*, *TKL1* y *XKS1* se realizó con ADN polimerasa de Vent® (New England Biolabs) según la especificación del fabricante. El molde era ADN cromosómico de CEN.PK113-7D (wt, tipo silvestre). La reacción en cadena de la polimerasa (PCR) se realizó en un Biometra TGradient Thermocycler (Biometra, Göttingen, Alemania) con los siguientes ajustes: 30 ciclos de 1 min de recocado a 60 °C, 3 min de extensión a 75 °C y 1 min de desnaturalización a 94 °C.

1.2. Cepas y medios

La cepa de *Saccharomyces cerevisiae* utilizada en este estudio es RWB212 (*MATA ura3-52 leu2-112 loxP-P<sub>TPI</sub>::(-266,-1) TAL1 gre3::hphMX pUGP<sub>TPI</sub>-TKL1 pUGP<sub>TPI</sub>-RPE1 KanloxP-P<sub>TPI</sub>::(-?,-1) RKI1*), que se deriva de CEN.PK102-3A (*MATA ura3-52 leu2-112*).

50 Durante la construcción, las cepas se mantuvieron en complejos (YP: 10 g l<sup>-1</sup> de extracto de levadura (BD Difco), 20 g l<sup>-1</sup> de peptona (BD Difco)) o medio sintético (MY) [6] complementado con glucosa (2 %) como fuente de carbono (YPD o MYD) y 1,5 % de agar en el caso de placas. Después de la transformación, los integrantes se seleccionaron en placas sobre YPD que contenía genética (G418) (Invitrogen/GIBCO) a 200 µg/ml o higromicina B (Roche Diagnostics GmbH, Mannheim, Alemania) a 300 µg/ml. Después de la transformación con plásmidos, las cepas se cultivaron en MYD. Las transformaciones de la levadura se realizaron según Gietz y Woods [7].

Los plásmidos se amplificaron en cepa XL-1 azul de *Escherichia coli* (Stratagene, La Jolla, CA, EE. UU.). La transformación se realizó según Inoue et al. [8]. Se cultivó *E. coli* en placas LB (Luria-Bertani) o en medio líquido TB (Terrific Broth) para el aislamiento de los plásmidos [5].

### 1.3. Construcción de la cepa

Para *TAL1* y *RK11*, la integración del promotor 5' *TPI1* del marco abierto de lectura se realizó amplificando un fragmento de la PCR con el marcador *KanMX* y el promotor *TPI1* y dirigiéndolo a la posición diana a través de terminales homólogos. La PCR se realizó con ADN polimerasa de *Taq* (Amersham Pharmacia, Piscataway, EE. UU.) según las especificaciones del fabricante. El molde era pUG6P<sub>TPI</sub> con P<sub>TALdisA</sub> y P<sub>TALdisB</sub> o P<sub>RK1disA</sub> y P<sub>RK1disB</sub> (Tabla 3) como cebadores.

En el caso de *TKL1* y *RPE1*, los plásmidos pUG6P<sub>TPI</sub>-*TKL1* y pUG6P<sub>TPI</sub>-*RPE1* se linealizaron con *PvuII* y *SalI*, respectivamente, y se integraron en el genoma. La correcta integración de los constructos se verificó en la PCR de colonias con *TAL1* interno + *KanA* para *TAL1* y cebador P<sub>TPI</sub> + "interno" para *TKL1*, *RPE1* y *RPI1*. Los cebadores "internos" se recuecen aguas abajo de los constructos integrados, mientras que el cebador P<sub>TPI</sub> se recuece en el terminal 3' del promotor *TPI1*.

Después de la integración de un constructo, se eliminó el marcador *KanMX* con la cre recombinasa. Con este fin, las cepas se transformaron con pSH47 [3]. Las colonias con el plásmido se resuspendieron en YP con galactosa al 1 % y se incubaron durante 1 hora a 30 °C. Después se depositaron alrededor de 200 células en placas sobre YPD. Las colonias resultantes se comprobaron en cuanto a la pérdida del marcador *KanMX* y pSH47 (*URA3*).

Además, el gen *GRE3* se reemplazó por la casete *hphMX* de pAG32, aportando resistencia a higromicina [9]. El casete *hphMX* se amplificó utilizando 5'gre3::*Kanlox* y 3'gre3::*Kanlox* de oligo. La integración correcta se verificó utilizando PCR con 5'*GRE3* + *KanA* y 3'*GRE3* + *KanB* (Tabla 3). *KanA* y *KanB* se recuecen con el promotor y terminador *TEF* de *A. gossipi*, respectivamente, mientras que los otros cebadores se recuecen fuera del marco abierto de lectura de *GRE3*.

La PCR de colonias se realizó con ADN polimerasa *Taq* (Amersham Pharmacia, Piscataway, EE. UU.) según las especificaciones del fabricante. Como células de plantilla se resuspendieron en 2,5 µl de NaOH 0,02 M a los que se añadió la mezcla de reacción de la PCR. La PCR se realizó en un Biometra TGradient Thermocycler (Biometra, Göttingen, Alemania) con los siguientes ajustes: 30 ciclos de 1 min de recocido a 60 °C, 3 min de extensión a 72 °C y 1 min de desnaturalización a 94 °C.

La cepa resultante *RWB212* (*MATA ura3-52 leu2-112 loxP-P<sub>TPI</sub>::(-266,-1)TAL1 gre3::hphMX pUGP<sub>TPI</sub>-TKL1 pUGP<sub>TPI</sub>-RPE1 KanloxP-P<sub>TPI</sub>::(-?,-1)RK11*), se transformó después con pAKX002, un vector multicopia que contiene el *E2 XylA* de *Piromyces* sp. detrás del promotor *TPI1*, así como p415ADHXKS. Lo cual da *RWB217* (*MATA ura3-52 leu2-112 loxP-P<sub>TPI</sub>::(-266,-1) TAL1 gre3::hphMX pUGP<sub>TPI</sub>-TKL1 pUGP<sub>TPI</sub>-RPE1 KanloxP-P<sub>TPI</sub>::(-?,-1) RK11 {pAKX002, p415ADHXKS}*).

### 1.4. Mantenimiento de la cepa

Los cultivos de reserva se hicieron crecer a 30 °C en frascos agitados en medio sintético [6] complementado con 20 g·l<sup>-1</sup> de glucosa. Cuando se alcanzó la fase estacionaria, se añadió glicerol estéril al 30 % (vol/vol) y se almacenaron alícuotas de 2 ml en viales estériles. a -80 °C.

### 1.5. Cultivo y medios

El cultivo en frascos agitados se realizó a 30 °C en un medio sintético [6]. El pH del medio se ajustó a 6,0 con KOH 2 M antes de la esterilización. Se prepararon precultivos inoculando 100 ml de medio que contenía 20 g l<sup>-1</sup> de xilosa en un frasco agitado de 500 ml con un cultivo de material congelado. Después de 24 a 48 h de incubación a 30 °C en un agitador orbital (200 rpm), este cultivo se utilizó para inocular cultivos en frascos agitados o cultivos en fermentadores. El medio sintético para el cultivo anaeróbico se complementó con 0,01 g l<sup>-1</sup> de ergosterol y 0,42 g l<sup>-1</sup> de Tween 80 disuelto en etanol [10, 11], esto dio como resultado etanol 11-13 mM en el medio de cultivo.

### 1.6. Cultivo discontinuo anaerobio en fermentadores

Los cultivos discontinuos anaerobios se llevaron a cabo en fermentadores de laboratorio de 2 litros (Applikon, Schiedam, Holanda) equipados con tubería de Norprene, con un volumen de trabajo de 1,5 litros, a 30 °C y a pH 5,0. Los cultivos se agitaron a 800 rpm y se roció con 0,5 l min<sup>-1</sup> de nitrógeno de alta calidad (<5 ppm de oxígeno). El medio sintético se complementó con los factores de crecimiento anaeróbicos ergosterol y Tween 80 (0,01 y 0,42 g l<sup>-1</sup>, respectivamente), así como 100 µl l<sup>-1</sup> de antiespumante de sílica (BDH, Poole, Reino Unido).

### 1.7. Determinación del peso en seco del cultivo

Se filtraron muestras de cultivo (10,0 ml) sobre filtros de nitrocelulosa previamente pesados (tamaño de poro de 0,45 µm, laboratorio Gelman, Ann Arbor, EE. UU.). Después de retirar el medio, los filtros se lavaron con agua desmineralizada y se secaron en un horno de microondas (Bosch, Stuttgart, Alemania) durante 20 minutos a 360 W y se pesaron. Las determinaciones por duplicado variaban en menos del 1 %.

### 1.8. Análisis de gas

El gas de escape se enfrió en un condensador (2 °C) y se secó con un secador Permapure de tipo MD-110-48P-4 (Permapure, Toms River, EE. UU.). Las concentraciones de O<sub>2</sub> y CO<sub>2</sub> se determinaron con un analizador NGA 2000 (Rosemount Analytical, Orrville, EE. UU.). El caudal del gas de escape y las velocidades del consumo específico de oxígeno y de la producción de dióxido de carbono se determinaron como se ha descrito anteriormente [12, 13]. Al calcular estas velocidades específicas de biomasa, se hizo una corrección para los cambios de volumen causados por la retirada de muestras de cultivo.

#### 1.9. Análisis de metabolitos

Glucosa, xilosa, xilitol, ácidos orgánicos, glicerol y etanol se detectaron por análisis con HPLC en una Waters Alliance 2690 HPLC (Waters, Milford, EE. UU.) que contenía una columna Biorad HPX 87H (Biorad, Hercules, EE. UU.). La columna se eluyó a 60 °C con 0,5 g l<sup>-1</sup> de H<sub>2</sub>SO<sub>4</sub> a un caudal de 0,6 ml min<sup>-1</sup>. La detección se realizó por medio de un detector de índice de refracción Waters 2410 y un detector por UV Waters 2487. La xilulosa se determinó enzimáticamente de la manera siguiente. La mezcla de reacción consistió en tampón Tris-HCl 100 mM (pH 7,5) con MgCl<sub>2</sub> 10 mM, NADH 0,30 mM y una cantidad adecuada de muestra (1 ml de volumen total), el ensayo se inició mediante la adición de 0,2 U de sorbitol deshidrogenasa (Sigma, St Louis, EE. UU.). La concentración de xilulosa se calculó usando un coeficiente de absorción de 6,3 mM<sup>-1</sup> cm<sup>-1</sup> para NADH.

#### 1.10. Recuperaciones de carbono y evaporación de etanol

Las recuperaciones de carbono se calcularon como carbono en productos formados dividido entre la cantidad total del carbono de azúcares consumido y se basaron en un contenido de carbono de biomasa de 48 %. Para corregir la evaporación de etanol durante las fermentaciones, se supuso que la cantidad de etanol producida era igual a la producción acumulada medida de CO<sub>2</sub> menos la producción de CO<sub>2</sub> producida por la síntesis de biomasa (5,85 mmol de CO<sub>2</sub> por gramo de biomasa [14]) y el CO<sub>2</sub> asociado con la formación de acetato como se ha descrito anteriormente [2].

#### 1.11 Análisis por micromatrices

El muestreo de células de quimiostatos, la preparación de la sonda y la hibridación con micromatrices Genechip<sup>®</sup> de Affymetrix se realizaron como se ha descrito anteriormente [15]. Los resultados para cada estado de crecimiento se obtuvieron a partir de tres réplicas cultivadas independientemente.

#### 1.12. Adquisición y análisis de los datos

La adquisición y cuantificación de las imágenes de matrices y el filtrado de los datos se realizaron utilizando los paquetes de software Affymetrix: Microarray Suite v5.0, MicroDB v3.0 y Data Mining Tool v3.0.

Antes de la comparación, todas las matrices se escalaron globalmente a un valor diana de 150 utilizando la señal media de todas las características del gen utilizando Microarray Suite v5.0. De las 9.335 características del transcrito en las matrices YG-S98, se aplicó un filtro para extraer 6.383 marcos de lectura abiertos de levadura de los que había 6.084 genes diferentes. Esta discrepancia se debía a varios genes que se representan más de una vez cuando se utilizaban conjuntos de sondas sub-óptimas en el diseño de la matriz.

Para representar la variación en las mediciones por triplicado, se calculó el coeficiente de variación (C.V., desviación estándar dividida por la media) como se ha descrito previamente por Boer et al. [16].

Para análisis estadísticos adicionales, se utilizó Microsoft Excel que llevó a cabo el análisis significativo de micromatrices (SAM v1.12) en adición [17] para posibles pares de comparaciones adecuadas de los ocho conjuntos de datos. Se consideró que los genes se modificaban en expresión si se les llamaba significativamente cambiados usando SAM (mediana esperada de la tasa de falso descubrimiento (FDR) de 1 %) al menos 2 veces de cada condición. El agrupamiento jerárquico de los conjuntos obtenidos de los niveles de expresión significativamente cambiados se realizó posteriormente por Genespring (Silicon Genetics).

#### 1.13 Ensayos de enzimas

La actividad de la xilosa isomerasa se ensayó a 37 °C en una mezcla de reacción que contenía tampón fosfato 50 mM (pH 7,0), xilosa 10 mM, MgCl<sub>2</sub> 10 mM y una cantidad adecuada de extracto libre de células. La cantidad de xilulosa formada se determinó mediante el procedimiento de cisteína-carbazol (9). Como alternativa, se ensayó la actividad de xilosa isomerasa en un ensayo de enzimas a 30°C desarrollado por Kersters-Hildersson et al. (Kinetic characterization of D-xylose isomerases by enzymatic assays using D-sorbitol dehydrogenase. *Enz. Microb. Technol.* 9 (1987) 145-148). La actividad in vitro de xilosa isomerasa en los extractos libres de células de cepas de *S. cerevisiae* transformadas depende de cationes bivalentes (Mg<sup>2+</sup> o Co<sup>2+</sup>).

Las actividades de xilulosa cinasa y xilosa reductasa se ensayaron como se describió por Witteveen et al. (28). Una unidad de actividad se define como la cantidad de enzima que produce 1 nmol de xilulosa por minuto en las condiciones de ensayo. La xilulosa formada se determinó por el procedimiento de Dische y Borenfreund (Dische y Borenfreund, 1951, *J. Biol. Chem.* 192: 583-587) o mediante HPLC usando una columna Biorad HPX-87N operada a

80 °C y eluida a 0,6 ml/min usando Na<sub>2</sub>HPO<sub>4</sub> 0,01 M como eluyente. La xilosa y la xilulosa se detectaron mediante un detector del Índice de Refracción a una temperatura interna de 60°C.

La actividad específica se expresa como unidades por mg de proteína. La proteína se determinó con el reactivo de proteínas Bio-Rad (Bio-Rad Laboratories, Richmond, CA, EE. UU.) con  $\gamma$ -globulina bovina como patrón.

## 5 2. Resultados

### 2.1 Sobreexpresión de los genes de la ruta de la pentosa fosfato (PPP)

Previamente los autores de la invención han demostrado que expresar una xilosa isomerasa fúngica en *Saccharomyces cerevisiae* es en principio suficiente para permitir el crecimiento anaeróbico de esta levadura en xilosa como única fuente de carbono siempre que se aplique una presión selectiva suficiente [2]. Sin embargo, la cepa seleccionada todavía no cumplía con los requisitos industriales (Tabla 1).

Con el fin de investigar la posibilidad de etapas de limitación de la velocidad en el metabolismo de la pentosa fosfato se decidió construir una cepa que sobreprodujera todas las enzimas requeridas para transformar xilosa en fructosa-6-fosfato y gliceraldehído-3-fosfato. Las enzimas sobreexpresadas fueron: xilosa isomerasa (XI), xilulocinasa (XKS), ribulosa-5-fosfato isomerasa (R5PI), ribulosa-5-fosfato epimerasa (R5PE), transcetolasa (TKT) y transaldolasa (TAL). Adicionalmente, se eliminó la aldosa reductasa no específica codificada por GRE3, que media la producción no deseada de xilitol [18]. Dado que algunos de los sustratos de las enzimas en la PPP no están disponibles comercialmente, se decidió comprobar la sobreexpresión a través de matrices de ADN en lugar de a través de mediciones de la actividad enzimática. Los resultados enumerados en la Tabla 1 confirmaban además que la transcripción de los genes diana se modificaba con éxito en la cepa RWB 212.

Tabla 1. Niveles de ARNm de los genes estructurales que codifican las enzimas de la ruta xilulocinasa y pentosa-fosfato en la cepa CEN.PK113-7D de *S. cerevisiae* de referencia y en la cepa RWB212 de *S. cerevisiae* de ingeniería que expresa la xilosa-isomerasa. Ambas cepas se desarrollaron en cultivos aeróbicos de quimiostatos, limitados en glucosa ( $D = 0,10 \text{ h}^{-1}$ ). Los niveles de transcritos se determinaron con micromatrices GeneChip de Affymetrix. Los datos son la media  $\pm$  desviación media de la media de los análisis en tres cultivos independientes para cada cepa. ACT1 (Ng y Abelson, 1980, Proc. Nat. Acad. Sci. EE. UU. 77: 3912-3916) y PDA1 (Wenzel et al., 1995, Nucleic Acids Res. 23: 883-884) se incluyen como patrones internos.

Gen	Nombre sistemático	Nombre de la enzima	Nivel del transcrito		Veces de cambio Mutante frente a WT (tipo silvestre)
			CEN.PK113-7D	RWB212	
XylA	-	Xilosa isomerasa	n.d.	n.d.	
XKS1	YGR194C	Xilulocinasa	91 $\pm$ 7	321 $\pm$ 54	+ 3,5
TAL1	YLR354C	Transaldolasa	574 $\pm$ 49	959 $\pm$ 91	+ 1,7
TKL1	YPR074C	Transcetolasa 1	450 $\pm$ 37	1.982 $\pm$ 79	+ 4,4
RPE1	YJL121C	D-ribulosa-5-fosfato epimerasa	299 $\pm$ 24	2.551 $\pm$ 385	+ 8,5
RKI1	YOR095C	D-ribosa-5-fosfato cetol-isomerasa	96 $\pm$ 8	483 $\pm$ 64	+ 5,0
GRE3	YHR104w	Aldosa reductasa	322 $\pm$ 6	12 $\pm$ 0	-26,8
ACT1	YFL039C	Actina	408 $\pm$ 32	366 $\pm$ 56	NC <sup>a</sup>
PDA1	YER178W	Subunidad E1 $\alpha$ de complejo piruvato deshidrogenasa	2.901 $\pm$ 142	3.217 $\pm$ 182	NC

n.d. = no determinado (no representado en las micromatrices de Affymetrix); <sup>a</sup>NC = no cambiado.

### 2.2 Caracterización fisiológica de la cepa de ingeniería

Una de las propiedades sorprendentes de la cepa de ingeniería fue su capacidad para crecer de forma anaerobia en xilosa (Fig. 1) sin que se requiera ninguna presión selectiva. El crecimiento anaerobio en xilosa en medio mineral prosiguió con una velocidad de crecimiento tan alta como  $0,09 \text{ h}^{-1}$ . Xilulosa no se acumuló, pero la formación de xilitol, aunque extremadamente pequeña, era detectable (Fig. 1), los rendimientos de biomasa, etanol y glicerol de la cepa RWB 212 en xilosa eran comparables a los de RWB 202-AFX que se obtenían a través de ingeniería evolutiva (Tabla 2). A partir de la Tabla 2 se puede calcular una velocidad específica del consumo de xilosa de más de  $1,0 \text{ g}$  de xilosa/g de biomasa/h ( $Q_s = 0,09/0,085 = 1,059 \text{ g Xyl/g X/h}$ ), comparado con  $345 \text{ mg}$  de xilosa/g de biomasa/h para RWB 202-AFX mientras que se obtuvo un rendimiento al menos similar al rendimiento en glucosa.

### 2.3 Utilización de sustrato mixto

Como se señala en la introducción: la transformación económica de los hidrolizados de hemicelulosa en etanol requiere la fermentación tanto de glucosa como de xilosa, preferiblemente simultáneamente. Con el fin de verificar las propiedades de la cepa RWB 212 con respecto a la utilización de azúcar mezclado, la levadura se cultivó en una mezcla de glucosa y xilosa ( $20 \text{ g l}^{-1}$  de cada una). Los resultados representados en la Fig. 2 muestran que ambos azúcares se consumieron completamente, pero la glucosa era el sustrato preferido. El consumo de xilosa comenzaba después de que se consumiera aproximadamente el 80 % de la glucosa.

### 3. Expresión funcional de la xilosa isomerasa de *B. thetaiotaomicron* en levadura

Se clonó la secuencia de nucleótidos que codifica la xilosa isomerasa VPI-5482 de *B. thetaiotaomicron* (números de acceso AAO75900 o NP 809706; SEQ ID NO. 10) en un vector multicopia de expresión de levadura para dar p426GPDBtXI. Este plásmido se utilizó para transformar RWB215 (*MAT $\alpha$  ura3-52 leu2-112 loxP-PTPI::(-266,-1) TAL1 gre3::hphMX pUGP<sub>TP1</sub>-TKL1 pUGP<sub>TP1</sub>-RPE1 KanloxP-P<sub>TP1</sub>::(-?, -1) RKI1*), que se transformó además con p415ADHXKS para sobreexpresión de xilulocinasa. Se recogieron dos transformantes independientes y ambos fueron capaces de crecer en medio mínimo con xilosa como única fuente de carbono y en lisados de los transformantes se midió una actividad de xilosa isomerasa específica de  $140 \pm 20 \text{ U}$  por mg de proteína, en comparación con aproximadamente  $1.300 \text{ U}$  por mg de proteína para las cepas que expresan la xilosa isomerasa *Piromyces*.

Tabla 3: cebadores utilizados en este estudio

Nombre del Oligo	
P <sub>TALdisA</sub>	CCTTTCCAACGAATCGTATATACTAACATGCGCGCGCTTCTATGCATAGGCCACTAGTGGATCTG
P <sub>TALdisB</sub>	AGAGAGTTGTTAGCAACCTTTTGTCTTTTGTGAGCTGGTTCAGACATGGTGAATTCCTGTATGTG
5'TAL1	CTGACTGAGCCATATTGAGG
TAL1 interno	CACCAGTGTCCGCAACAACG
P <sub>RKIdisA</sub>	TCTTGTAGAAAATTAACAACATCGTGTTACATAAACTTGGTTACGCATAGGCCACTAGTGGATCTG
P <sub>RKIdisB</sub>	TTGCCCAAAGATTCTAACGCATCAATTTTTGGGACACCGGCAGCCATGGTGAATTCCTGTATGTG
RKI1 interno	CAGCTCTCTTGGCATCCTCC
EcoRI-5'TKL1	GGAATTCATGACTCAATTCAGTACATTG
3'TKI1-XhoI	GGCCTCGAGCTTGAATGGTGTGATTGTCT
TKL1 interno	CCGCCATTGGTGTATGTACAG
EcoRI-5'RPEI	GGAATTCATGGTCAAACCAATTATAGC
3'RPEI-XhoI	CCGCTCGAGTTAGGCACTTACGTATCTTG
RPE1 interno	GGAAGCCTTATGGAGTGTCA
cebador P <sub>TP11</sub>	TGGCATGTGAGATTCTCCGA

Nombre del Oligo	
KanA	CGCACGTCAAGACTGTCAAG
KanB	TCGTATGTGAATGCTGGTCG
5'gre3::Kanlox	AAAATACTGTAATATAAATCGTAAAGGAAAATTGGAAATTTTTTCAGCTGAAGCTTCGTACGC
3'gre3::Kanlox	TGGATTTTACTGGCTGGATCAGGCCAAAAGTGGGGAATTTACCGCATAGGCCACTAGTGGATCTG
5'GRE3	CCTGGTGGAAACATCCTAGAA
3'GRE3	GGATGACACCACAGGCAGAA
SpeI-5'XKS1	GACTAGTATGTTGTGTTTCAGTAATTCAG
3'XKS1-SalI	TGCAGTCGACATTTTAGATGAGAGTCTTTTCC

Tabla 4: plásmidos utilizados en este documento

pUG6	casete loxP-KanMX-loxP	Guldener et al. [3]
pUG6P <sub>TPI1</sub>	pUG6 con el promotor TPI1	este trabajo
pUG6P <sub>TPI1</sub> -RPE1	pUG6 con RPE1 detrás del promotor TPI1	este trabajo
PUG6P <sub>TPI1</sub> -TKL1	pUG6 con TKL1 detrás del promotor TPI1	este trabajo
pAG32	casete loxP-hphMX-loxP	Goldstein y McCusker [9]
PAKX002	2 $\mu$ , <i>URA3</i> , <i>Piromyces XylA</i> detrás del promotor TPI1	Kuyper et al. [20]
P415ADH	<i>CEN</i> , <i>LEU2</i> , <i>promotor ADH1</i>	Mumberg et al. [21]
p415ADHXKS1	<i>CEN</i> , <i>LEU2</i> , P <sub>ADH1</sub> - <i>XKS1</i>	este trabajo
PSH47	<i>CEN</i> , <i>URA3</i> , Cre recombinasa detrás de P <sub>GAL1</sub>	Guldener et al. [3]

## Referencias

1. Bruinenberg, P. M., De Bot, P. H. M, Van Dijken, J. P. y Scheffers, W. A. (1983). The role of the redox balance in the anaerobic fermentation of xylose by yeasts. *Eur. J. Appl. Microbiol. Biotechnol.* 18, 287-292.
- 5 2. Kuyper, M., Winkler, A. A., Van Dijken, J. P. y Pronk, J. T. (2004). Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle. *FEMS Yeast Res.* 4, 655-664.
3. Guldener, U., Heck, S., Fiedler, T., Beinhauer, J. y Hegemann, J. H. (1996). A new efficient gene disruption cassette for repeated use in budding yeast. *Nucleic Acids Res.* 24, 2519-2524.
- 10 4. Mumberg, D., Muller, R. y Funk, M. (1995). Yeast Vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156,119-122.
5. Sambrook, K., Fritsch, E. F., y Maniatis, I. (1989). *Molecular cloning: A laboratory manual*, 2º ed., Cold spring harbour, Nueva York.
6. Verduyn, C., Postma, E., Scheffers, W. A. y Van Dijken, J. P. (1992). Effect of benzoic acid on metabolic fluxes in yeasts: a continuous-culture study on the regulation of respiration and alcoholic fermentation. *Yeast* 8, 501-517.
- 15 7. Gietz, R. D. y Woods, R. A. (2002). Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method. *Methods Enzymol.* 350, 87-96.
8. Inoue, H., Nojima, H. y Okayama, H. (1990). High efficiency transformation of *Escherichia coli* with plasmids. *Gene* 96, 23-28.
- 20 9. Goldstein, A. L. y McCusker, J. H. (1999). Three new dominant drug resistance cassettes for gene disruption in *Saccharomyces cerevisiae*. *Yeast* 15, 1541-1553.
10. Andreasen, A. A. y Stier, T. J. (1953). Anaerobic nutrition of *Saccharomyces cerevisiae*. I. Ergosterol requirement for growth in a defined medium. *J. Cell Physiol.* 41, 23-36.
11. Andreasen, A. A. y Stier, T. J. (1954). Anaerobic nutrition of *Saccharomyces cerevisiae*. II. Unsaturated fatty acid requirement for growth in a defined medium. *J. Cell Physiol.* 43, 271-281.
- 25 12. Van Urk, H., Mak, P. R., Scheffers, W. A. y Van Dijken, J. P. (1988). Metabolic responses of *Saccharomyces cerevisiae* CBS 8066 y *Candida utilis* CBS 621 upon transition from glucose limitation to glucose excess. *Yeast* 4, 283-291.
13. Weusthuis, R. A., Luttkik, M. A., Scheffers, W. A., Van Dijken, J. P. y Pronk, J. T. (1994). Is the Kluyver effect in yeasts caused by product inhibition? *Microbiology* 140, 1723-1729.
- 30 14. Verduyn, C., Postma, E., Scheffers, W. A. y Van Dijken, J. P. (1990). Physiology of *Saccharomyces cerevisiae* in anaerobic glucose-limited chemostat cultures. *J. Gen. Microbiol.* 136, 395-403.
15. Piper, M. D., Daran-Lapujade, P., Bro, C., Regenber, B., Knudsen, S., Nielsen, J. y Pronk, J. T. (2002). Reproducibility of oligonucleotide microarray transcriptome analyses. An interlaboratory comparison using chemostat cultures of *Saccharomyces cerevisiae*. *J. Biol. Chem.* 277, 37001-37008.
- 35 16. Boer, V. M., de Winde, J. H., Pronk, J. T. y Piper, M. D. (2003). The genome-wide transcriptional responses of *Saccharomyces cerevisiae* grown on glucose in aerobic chemostat cultures limited for carbon, nitrogen, phosphorus, or sulfur. *J. Biol. Chem.* 278, 3265-3274.
17. Tusher, V. G., Tibshirani, R. y Chu, G. (2001). Significance analysis of microarrays applied to the ionizing radiation response. *Proc. Natl. Acad. Sci. EE. UU.* 98, 5116-5121.
- 40 18. Träff, K. L., Otero Cordero, R. R., Van Zyl, W. H. y Hahri-Hägerdal, B. (2001). Deletion of the GRE3 aldosa reductasa gene and its influence on xylose metabolism in recombinant strains of *Saccharomyces cerevisiae* expressing the xylA and XKS1 genes. *Appl. Environ. Microbiol.* 67, 5668-5674.
19. Jeffries, T. W. y Jin, Y. S. (2004). Metabolic engineering for improved fermentation of pentoses by yeasts. *Applied Microbiology and Biotechnology* 63, 495-509.
- 45 20. Kuyper, M., Harhangi, H. R., Stave, A. K., Winkler, A. A., Jetten, M. S., De Laat, W. T., D Ridder, J. J., Op den Camp, H. J., Van Dijken, J. P. y Pronk, J. T. (2003). High-level functional expression of a fungal xilosa isomerasa: the key to efficient ethanolic fermentation of xylose by *Saccharomyces cerevisiae*? *FEMS Yeast Res.* 4, 69-78.
21. Mumberg, D., Muller, R. y Funk, M. (1995). Yeast Vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156,119-122.

**LISTA DE SECUENCIAS**

- <110> Technische Universiteit Delft
- <120> Ingeniería metabólica de células eucariotas que fermentan xilosa
- <130> P214817EP
- 5 <140> P214817EP
- <141> 2004-07-07
- <160> 18
- <170> PatentIn version 3.1
- <210> 1
- 10 <211> 437
- <212> PRT
- <213> *Piromyces* sp.
- <220>
- <221> misc feature (características diversas)
- 15 <223> xilosa isomerasa
- <400> 1

```

Met Ala Lys Glu Tyr Phe Pro Gln Ile Gln Lys Ile Lys Phe Glu Gly
1          5          10          15
Lys Asp Ser Lys Asn Pro Leu Ala Phe His Tyr Tyr Asp Ala Glu Lys
20          25          30
Glu Val Met Gly Lys Lys Met Lys Asp Trp Leu Arg Phe Ala Met Ala
35          40          45
    
```

ES 2 607 891 T3

Trp Trp His Thr Leu Cys Ala Glu Gly Ala Asp Gln Phe Gly Gly Gly  
50 55 60

Thr Lys Ser Phe Pro Trp Asn Glu Gly Thr Asp Ala Ile Glu Ile Ala  
65 70 75 80

Lys Gln Lys Val Asp Ala Gly Phe Glu Ile Met Gln Lys Leu Gly Ile  
85 90 95

Pro Tyr Tyr Cys Phe His Asp Val Asp Leu Val Ser Glu Gly Asn Ser  
100 105 110

Ile Glu Glu Tyr Glu Ser Asn Leu Lys Ala Val Val Ala Tyr Leu Lys  
115 120 125

Glu Lys Gln Lys Glu Thr Gly Ile Lys Leu Leu Trp Ser Thr Ala Asn  
130 135 140

Val Phe Gly His Lys Arg Tyr Met Asn Gly Ala Ser Thr Asn Pro Asp  
145 150 155 160

Phe Asp Val Val Ala Arg Ala Ile Val Gln Ile Lys Asn Ala Ile Asp  
165 170 175

Ala Gly Ile Glu Leu Gly Ala Glu Asn Tyr Val Phe Trp Gly Gly Arg  
180 185 190

Glu Gly Tyr Met Ser Leu Leu Asn Thr Asp Gln Lys Arg Glu Lys Glu  
195 200 205

His Met Ala Thr Met Leu Thr Met Ala Arg Asp Tyr Ala Arg Ser Lys  
210 215 220

Gly Phe Lys Gly Thr Phe Leu Ile Glu Pro Lys Pro Met Glu Pro Thr  
225 230 235 240

Lys His Gln Tyr Asp Val Asp Thr Glu Thr Ala Ile Gly Phe Leu Lys  
245 250 255

Ala His Asn Leu Asp Lys Asp Phe Lys Val Asn Ile Glu Val Asn His  
260 265 270

Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Cys Ala Val  
275 280 285

ES 2 607 891 T3

Asp Ala Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Tyr Gln  
 290 295 300

Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Gln Tyr Glu Leu Val  
 305 310 315 320

Gln Ala Trp Met Glu Ile Ile Arg Gly Gly Gly Phe Val Thr Gly Gly  
 325 330 335

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

Ile Ile Ile Ala His Val Ser Gly Met Asp Ala Met Ala Arg Ala Leu  
 355 360 365

Glu Asn Ala Ala Lys Leu Leu Gln Glu Ser Pro Tyr Thr Lys Met Lys  
 370 375 380

Lys Glu Arg Tyr Ala Ser Phe Asp Ser Gly Ile Gly Lys Asp Phe Glu  
 385 390 395 400

Asp Gly Lys Leu Thr Leu Glu Gln Val Tyr Glu Tyr Gly Lys Lys Asn  
 405 410 415

Gly Glu Pro Lys Gln Thr Ser Gly Lys Gln Glu Leu Tyr Glu Ala Ile  
 420 425 430

Val Ala Met Tyr Gln  
 435

<210> 2

< 211> 438

5 < 212> PRT

< 213> Bacteroides thetaiotaomicron

<220>

< 221> misc\_feature

< 223> xilosa isomerasa

10 <400> 2

Met Ala Thr Lys Glu Phe Phe Pro Gly Ile Glu Lys Ile Lys Phe Glu



ES 2 607 891 T3

Lys Ala His Gly Leu Asp Lys Asp Phe Lys Val Asn Ile Glu Val Asn  
 260 265 270

His Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Val Ala  
 275 280 285

Val Asp Asn Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Tyr  
 290 295 300

Gln Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Asn Tyr Glu Leu  
 305 310 315 320

Thr Gln Ala Met Met Gln Ile Ile Arg Asn Gly Gly Leu Gly Thr Gly  
 325 330 335

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

Asp Ile Phe Ile Ala His Ile Ala Gly Met Asp Ala Met Ala Arg Ala  
 355 360 365

Leu Glu Ser Ala Ala Ala Leu Leu Asp Glu Ser Pro Tyr Lys Lys Met  
 370 375 380

Leu Ala Asp Arg Tyr Ala Ser Phe Asp Gly Gly Lys Gly Lys Glu Phe  
 385 390 395 400

Glu Asp Gly Lys Leu Thr Leu Glu Asp Val Val Ala Tyr Ala Lys Thr  
 405 410 415

Lys Gly Glu Pro Lys Gln Thr Ser Gly Lys Gln Glu Leu Tyr Glu Ala  
 420 425 430

Ile Leu Asn Met Tyr Cys  
 435

<210> 3

< 211> 600

5 < 212> PRT

< 213> Saccharomyces cerevisiae

<220>

< 221> misc\_feature

< 223> xilulocinasa

10 <400> 3

ES 2 607 891 T3

Met Leu Cys Ser Val Ile Gln Arg Gln Thr Arg Glu Val Ser Asn Thr  
 1 5 10 15

Met Ser Leu Asp Ser Tyr Tyr Leu Gly Phe Asp Leu Ser Thr Gln Gln  
 20 25 30

Leu Lys Cys Leu Ala Ile Asn Gln Asp Leu Lys Ile Val His Ser Glu  
 35 40 45

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

Val Tyr Ile His Gly Asp Thr Ile Glu Cys Pro Val Ala Met Trp Leu  
 65 70 75 80

Glu Ala Leu Asp Leu Val Leu Ser Lys Tyr Arg Glu Ala Lys Phe Pro  
 85 90 95

Leu Asn Lys Val Met Ala Val Ser Gly Ser Cys Gln Gln His Gly Ser  
 100 105 110

Val Tyr Trp Ser Ser Gln Ala Glu Ser Leu Leu Glu Gln Leu Asn Lys  
 115 120 125

Lys Pro Glu Lys Asp Leu Leu His Tyr Val Ser Ser Val Ala Phe Ala  
 130 135 140

Arg Gln Thr Ala Pro Asn Trp Gln Asp His Ser Thr Ala Lys Gln Cys  
 145 150 155 160

Gln Glu Phe Glu Glu Cys Ile Gly Gly Pro Glu Lys Met Ala Gln Leu  
 165 170 175

Thr Gly Ser Arg Ala His Phe Arg Phe Thr Gly Pro Gln Ile Leu Lys  
 180 185 190

Ile Ala Gln Leu Glu Pro Glu Ala Tyr Glu Lys Thr Lys Thr Ile Ser  
 195 200 205

ES 2 607 891 T3

Leu Val Ser Asn Phe Leu Thr Ser Ile Leu Val Gly His Leu Val Glu  
 210 215 220

Leu Glu Glu Ala Asp Ala Cys Gly Met Asn Leu Tyr Asp Ile Arg Glu  
 225 230 235 240

Arg Lys Phe Ser Asp Glu Leu Leu His Leu Ile Asp Ser Ser Ser Lys  
 245 250 255

Asp Lys Thr Ile Arg Gln Lys Leu Met Arg Ala Pro Met Lys Asn Leu  
 260 265 270

Ile Ala Gly Thr Ile Cys Lys Tyr Phe Ile Glu Lys Tyr Gly Phe Asn  
 275 280 285

Thr Asn Cys Lys Val Ser Pro Met Thr Gly Asp Asn Leu Ala Thr Ile  
 290 295 300

Cys Ser Leu Pro Leu Arg Lys Asn Asp Val Leu Val Ser Leu Gly Thr  
 305 310 315 320

Ser Thr Thr Val Leu Leu Val Thr Asp Lys Tyr His Pro Ser Pro Asn  
 325 330 335

Tyr His Leu Phe Ile His Pro Thr Leu Pro Asn His Tyr Met Gly Met  
 340 345 350

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

Leu Asn Lys Glu Arg Glu Asn Asn Tyr Glu Lys Thr Asn Asp Trp Thr  
 370 375 380

Leu Phe Asn Gln Ala Val Leu Asp Asp Ser Glu Ser Ser Glu Asn Glu  
 385 390 395 400

Leu Gly Val Tyr Phe Pro Leu Gly Glu Ile Val Pro Ser Val Lys Ala  
 405 410 415

Ile Asn Lys Arg Val Ile Phe Asn Pro Lys Thr Gly Met Ile Glu Arg  
 420 425 430

Glu Val Ala Lys Phe Lys Asp Lys Arg His Asp Ala Lys Asn Ile Val  
 435 440 445

ES 2 607 891 T3

Glu Ser Gln Ala Leu Ser Cys Arg Val Arg Ile Ser Pro Leu Leu Ser  
 450 455 460

Asp Ser Asn Ala Ser Ser Gln Gln Arg Leu Asn Glu Asp Thr Ile Val  
 465 470 475 480

Lys Phe Asp Tyr Asp Glu Ser Pro Leu Arg Asp Tyr Leu Asn Lys Arg  
 485 490 495

Pro Glu Arg Thr Phe Phe Val Gly Gly Ala Ser Lys Asn Asp Ala Ile  
 500 505 510

Val Lys Lys Phe Ala Gln Val Ile Gly Ala Thr Lys Gly Asn Phe Arg  
 515 520 525

Leu Glu Thr Pro Asn Ser Cys Ala Leu Gly Gly Cys Tyr Lys Ala Met  
 530 535 540

Trp Ser Leu Leu Tyr Asp Ser Asn Lys Ile Ala Val Pro Phe Asp Lys  
 545 550 555 560

Phe Leu Asn Asp Asn Phe Pro Trp His Val Met Glu Ser Ile Ser Asp  
 565 570 575

Val Asp Asn Glu Asn Trp Asp Arg Tyr Asn Ser Lys Ile Val Pro Leu  
 580 585 590

Ser Glu Leu Glu Lys Thr Leu Ile  
 595 600

<210> 4

< 211> 258

5 < 212> PRT

< 213> Saccharomyces cerevisiae

<220>

< 221> misc\_feature

< 223> ribulosa 5-fosfato isomerasa

10 <400> 4

Met Ala Ala Gly Val Pro Lys Ile Asp Ala Leu Glu Ser Leu Gly Asn



ES 2 607 891 T3

<220>

< 221> misc\_feature

< 223> ribulosa 5-fosfato epimerasa

<400> 5

5

Met Val Lys Pro Ile Ile Ala Pro Ser Ile Leu Ala Ser Asp Phe Ala  
 1 5 10 15

Asn Leu Gly Cys Glu Cys His Lys Val Ile Asn Ala Gly Ala Asp Trp  
 20 25 30

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

Gly Gln Pro Ile Val Thr Ser Leu Arg Arg Ser Val Pro Arg Pro Gly  
 50 55 60

Asp Ala Ser Asn Thr Glu Lys Lys Pro Thr Ala Phe Phe Asp Cys His  
 65 70 75 80

Met Met Val Glu Asn Pro Glu Lys Trp Val Asp Asp Phe Ala Lys Cys  
 85 90 95

Gly Ala Asp Gln Phe Thr Phe His Tyr Glu Ala Thr Gln Asp Pro Leu  
 100 105 110

His Leu Val Lys Leu Ile Lys Ser Lys Gly Ile Lys Ala Ala Cys Ala  
 115 120 125

Ile Lys Pro Gly Thr Ser Val Asp Val Leu Phe Glu Leu Ala Pro His  
 130 135 140

Leu Asp Met Ala Leu Val Met Thr Val Glu Pro Gly Phe Gly Gly Gln  
 145 150 155 160

Lys Phe Met Glu Asp Met Met Pro Lys Val Glu Thr Leu Arg Ala Lys  
 165 170 175

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

Ile Pro Lys Ala Ala Lys Ala Gly Ala Asn Val Ile Val Ala Gly Thr  
 195 200 205

Ser Val Phe Thr Ala Ala Asp Pro His Asp Val Ile Ser Phe Met Lys  
 210 215 220

Glu Glu Val Ser Lys Glu Leu Arg Ser Arg Asp Leu Leu Asp  
 225 230 235



ES 2 607 891 T3

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

ES 2 607 891 T3

Gly Val Glu Ala Asn Asn Lys Trp Asn Lys Leu Phe Ser Glu Tyr Gln  
305 310 315 320

Lys Lys Phe Pro Glu Leu Gly Ala Glu Leu Ala Arg Arg Leu Ser Gly  
325 330 335

Gln Leu Pro Ala Asn Trp Glu Ser Lys Leu Pro Thr Tyr Thr Ala Lys  
340 345 350

Asp Ser Ala Val Ala Thr Arg Lys Leu Ser Glu Thr Val Leu Glu Asp  
355 360 365

Val Tyr Asn Gln Leu Pro Glu Leu Ile Gly Gly Ser Ala Asp Leu Thr  
370 375 380

Pro Ser Asn Leu Thr Arg Trp Lys Glu Ala Leu Asp Phe Gln Pro Pro  
385 390 395 400

Ser Ser Gly Ser Gly Asn Tyr Ser Gly Arg Tyr Ile Arg Tyr Gly Ile  
405 410 415

Arg Glu His Ala Met Gly Ala Ile Met Asn Gly Ile Ser Ala Phe Gly  
420 425 430

Ala Asn Tyr Lys Pro Tyr Gly Gly Thr Phe Leu Asn Phe Val Ser Tyr  
435 440 445

Ala Ala Gly Ala Val Arg Leu Ser Ala Leu Ser Gly His Pro Val Ile  
450 455 460

Trp Val Ala Thr His Asp Ser Ile Gly Val Gly Glu Asp Gly Pro Thr  
465 470 475 480

His Gln Pro Ile Glu Thr Leu Ala His Phe Arg Ser Leu Pro Asn Ile  
485 490 495

Gln Val Trp Arg Pro Ala Asp Gly Asn Glu Val Ser Ala Ala Tyr Lys  
500 505 510

Asn Ser Leu Glu Ser Lys His Thr Pro Ser Ile Ile Ala Leu Ser Arg  
515 520 525

Gln Asn Leu Pro Gln Leu Glu Gly Ser Ser Ile Glu Ser Ala Ser Lys  
530 535 540

Gly Gly Tyr Val Leu Gln Asp Val Ala Asn Pro Asp Ile Ile Leu Val



ES 2 607 891 T3

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

ES 2 607 891 T3

260 265 270

Arg Val Leu Asp Pro Val Ser Ala Lys Lys Glu Ala Gly Asp Lys Ile  
 275 280 285

Ser Tyr Ile Ser Asp Glu Ser Lys Phe Arg Phe Asp Leu Asn Glu Asp  
 290 295 300

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

Asp Ile Val Thr Leu Phe Asp Leu Ile Glu Lys Lys Val Thr Ala  
 325 330 335

<210> 8

< 211> 327

5 < 212> PRT

< 213> *Saccharomyces cerevisiae*

<220>

< 221> misc\_feature

< 223> aldosa reductasa

10 <400> 8

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

Gly Leu Gly Cys Trp Lys Ile Asp Lys Lys Val Cys Ala Asn Gln Ile  
 20 25 30

Tyr Glu Ala Ile Lys Leu Gly Tyr Arg Leu Phe Asp Gly Ala Cys Asp  
 35 40 45

Tyr Gly Asn Glu Lys Glu Val Gly Glu Gly Ile Arg Lys Ala Ile Ser  
 50 55 60

Glu Gly Leu Val Ser Arg Lys Asp Ile Phe Val Val Ser Lys Leu Trp  
 65 70 75 80

Asn Asn Phe His His Pro Asp His Val Lys Leu Ala Leu Lys Lys Thr  
 85 90 95

ES 2 607 891 T3

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

ES 2 607 891 T3

<210> 9

< 211> 1669

< 212> ADN

< 213> Piromyces sp.

5 <220>

< 221> misc\_feature

< 223> xilosa isomerasa

<400> 9

gtaaatggct aaggaatatt tcccacaaat tcaaaagatt aagttcgaag gtaaggattc  
60

taagaateca ttagccttcc actactaoga tgctgaaaag gaagtcatgg gtaagaaaat  
120

gaaggattgg ttacgtttcg ccattggcctg gtggcacact ctttgcgccg aagggtctga  
180

ccaattcggg ggaggtacaa agtctttccc atggaacgaa ggtactgatg ctattgaaat  
240

tgccaagcaa aaggttgatg ctggtttcga aatcatgcaa aagcttggtg ttccatacta  
300

ctggtttccac gatggtgatc ttgtttccga aggtaactct attgaagaat acgaatccaa  
360

ccttaaggct gtcgttgctt acctcaagga aaagcaaaag gaaaccggtg ttaagcttct  
420

ctggagtact gctaaagctt tcggtcacaa gcgttacatg aacggtgctt ccaactaaccc  
480

agactttgat gttgtcgccc gtgctattgt tcaaattaag aacgccatag acgcccgtat  
540

tgaacttggg gctgaaaact acgtcttctg gggtggtcgt gaaggttaca tgagtctcct  
600

taacactgac caaaagcgtg aaaaggaaca catggcact atgcttacca tggtcgtgga  
660

ctacgctcgt tccaagggat tcaagggtac tttcctcatt gaaccaaacg caatggaacc  
720

aaccaagcac caatacgatg ttgacactga aaccgctatt ggtttcctta aggccacaaa  
780

cttagacaag gacttcaagg tcaacattga agttaaccac gctactcttg ctggtcacac  
840

10

ES 2 607 891 T3

tttcgaacac gaacttgccct gtgctgttga tgctggatg ctcggtcca ttgatgctaa  
900  
ccgtggtagc taccaaacg gttgggatac tgatcaattc ccaattgatc aatagcaact  
960  
cgtccaagct tggatgaaa tcacccgtgg tgggtggttc gttactgggtg gtaccaactt  
1020  
cgatccaag actcgtcgta actctactga cctcgaagac atcatcattg cccacgtttc  
1080  
tggtaggat gctatggctc gtgctcttga aaacgctgcc aagctcctcc aagaatctcc  
1140  
atacacciaag atgaagaagg aacgttacgc ttccttcgac agtgggtattg gtaaggactt  
1200  
tgaagatggt aagctcacc tcgaacaagt ttacgaatac gtaagaaga acggtgaacc  
1260  
aaagcaact tctggtaagc aagaactcta cgaagctatt gttgccatgt accaataagt  
1320  
taatcgtagt taaattggta aaataattgt aaaatcaata aacttgtaa tcctccaatc  
1380  
aagtttaaaa gatcctatct ctgtactaat taaatatagt acaaaaaaaaa atgtataaac  
1440  
aaaaaaaaagt ctaaaagacg gaagaattta atttaggaa aaaataaaaa taataataaa  
1500  
caatagataa atcctttata ttaggaaaat gtcccattgt attatittca tttctactaa  
1560  
aaaagaaagt aaataaaaca caagaggaaa tttcccttt ttttttttt tgtaataaat  
1620  
tttatgcaaa tataaatata aataaataa taaaaaaaa aaaaaaaaa  
1669

<210> 10

< 211> 1317

5 < 212> ADN

< 213> Bacteroides thetaiotaomicron

<220>

< 221> misc\_feature

< 223> xilosa isomerasa

10 <400> 10

ES 2 607 891 T3

atggcaacaa aagaatTTTT tccgggaatt gaaaagatta aatttgaagg taaagatagt  
60  
aagaaccgga tggcattccg ttattacgat gcagagaagg tgattaatgg taaaaagatg  
120  
aaggattggc tgagattcgc tatggcatgg tggcacacat tgtgcgctga aggtggtgat  
180  
cagttcgggtg gcggaacaaa gcaattccca tggaatggta atgcagatgc tatacaggca  
240  
gcaaaagata agatggatgc aggatttgaa ttcatgcaga agatgggtat cgaatactat  
300  
tgcttccatg acgtagactt ggtttcggaa ggtgccagtg tagaagaata cgaagctaac  
360  
ctgaaagaaa tcgtagctta tgcaaaacag aaacaggcag aaaccgggat caaactactg  
420  
tggggactctg ctaatgtatt cggtcacgcc cgctatatga acggtgcagc taccaatcct  
480  
gacttcgatg tagtagctcg tctgctggtt cagatcaaaa atgogattga tgcaacgatt  
540  
gaacttggcg gagagaatta tgtgttttgg ggtggctcgtg aaggctatat gtctcttctg  
600  
aacacagatc agaaacgtga aaaagaacac cttgcacaga tgttgacgat tgctcgtgac  
660  
tatgcccgtg cccgtggttt caaaggctact ttccctgatcg aaccgaaacc gatggaaccg  
720  
actaaacatc aatatgacgt agatacggaa actgtaatcg gcttcctgaa agctcatggt  
780  
ctggataagg atttcaaagt aatatcggag gtgaatcacg caactttggc aggtcacact  
840  
ttcgagcatg aattggctgt agctgtagac aatgggtatgt tgggctcaat tgacgccaat  
900  
cgtggtgact atcagaatgg ctgggataca gaccaattcc cgatcgacaa ttatgaactg  
960  
actcaggcta tgatgcagat tatccgtaat ggtggctcgtg gtaccggtgg tacgaacttt  
1020  
gatgctaaaa cccgtcgtaa ttotactgat ctggaagata tctttattgc tcacatcgca  
1080  
ggtatggacg ctatggcccg tgcaactcgaa agtgcagcgg ctctgctcga cgaatctccc  
1140  
  
tataagaaga tgctggctga ccgttatgct tcatttgatg ggggcaaagg taaagaattt  
1200  
gaagacggca agctgactct ggaggatgtg gttgcttatg caaaaacaaa aggccaaccg  
1260  
aaacagacta gcggcaagca agaactttat gaggcaattc tgaatatgta ttgctaa  
1317

<210> 11

< 211> 2467

< 212> ADN

< 213> *Saccharomyces cerevisiae*

5 <220>

< 221> misc\_feature

< 223> xilulocinasa

<400> 11

ggatccaaga ccattattcc atcagaatgg aaaaaagttt aaaagatcac ggagatthttg  
60

ttcttctgag cttctgctgt ccttgaaaac aaattattcc gctggccgcc ccaaacaata  
120

acaaccccga ttttaataaca ttgtcacagt attagaaatt ttctttttac aaattaccat  
180

ttccagctta ctacttctca taatctctca tottcagcaa gcgacgcagg gaatagccgc  
240

tgaggtgcat aactgtcact tttcaattcg gccaatgcaa tctcaggcgg acgaataagg  
300

gggccctctc gagaaaaaca aaaggaggat gagattagta cttaaatggt gtgttcagta  
360

attcagagac agacaagaga ggtttccaac acaatgtctt tagactcata ctatcttggg  
420

tttgatcttt cgacccaaca actgaaatgt ctgcaccatta accaggacct aaaaattgtc  
480

cattcagaaa cagtgaatt tgaaaaggat cttccgcatt atcacacaaa gaagggtgtc  
540

tatatacacg gcgacactat cgaatgtccc gtagccatgt ggtaggggc tctagatctg  
600

gttctctcga aatatcgaga ggctaaattt ccattgaaca aagttatggc cgtctcaggg  
660

10

ES 2 607 891 T3

tcctgccagc agcacgggtc tgtctactgg tcctccaag ccgaatctct gttagagcaa  
720

ttgaataaga aaccggaaaa agatttattg cactacgtga gctctgtagc atttgcaagg  
780

caaaccgccc ccaattggca agaccacagt actgcaaagc aatgtcaaga gtttgaagag  
840

tgcataaggc ggctgaaaa aatggctcaa ttaacagggt ccagagccca ttttagattt  
900

actggctctc aaattctgaa aattgcacaa ttagaaccag aagcttacga aaaaacaaag  
960

accatttctt tagtgtctaa ttttttgact tctatcttag tggggcatct tgttgaatta  
1020

gaggaggcag atgcctgtgg tatgaacctt tatgatatac gtgaaagaaa attcatgtat  
1080

gagctactac atctaattga tagttcttct aaggataaaa ctatcagaca aaaattaatg  
1140

agagcaocca tgaaaaatth gatagcgggt accatctgta aatattttat tgagaagtac  
1200

ggtttcaata caaactgcaa ggtctctccc atgactgggg ataatttagc cactatatgt  
1260

tctttacccc tgcggaagaa tgacgttctc gtttccctag gaacaagtac tacagttctt  
1320

ctggtcaccg ataagtatca cccctctccg aactatcctc ttttcattca tccaactctg  
1380

ccaaaccatt atatgggtat gatttgttat tgtaatgggt ctttggcaag ggagaggata  
1440

agagacgagt taaacaaaga acgggaaaat aattatgaga agactaacga ttggactctt  
1500

tttaatcaag ctgtgctaga tgactcagaa agtagtgaaa atgaattagg tgtatatttt  
1560

cctctggggg agatcgttcc tagogtaaaa gccataaaca aaagggttat cttcaatcca  
1620

aaaacgggta tgattgaaag agaggtggcc aagttcaaag acaagaggca cgatgccaaa  
1680

aatattgtag aatcacaggc tttaagttgc aggtaagaa tatctcccct gctttcggat  
1740

tcaaacgcaa gctcacaaca gagactgaac gaagatacaa tcgtgaagtt tgattacgat  
1800

gaatctccgc tgcgggacta cctaaataaa aggccagaaa ggactttttt tgtaggtggg  
1860

ES 2 607 891 T3

gcttctaaaa acgatgctat tgtgaagaag tttgctcaag tcattggtgc taaaaaggt  
1920

aattttaggc tagaaacacc aaactcatgt gcccttggtg gttgttataa ggccatgtgg  
1980

tcattggtat atgactctaa taaaattgca gttccttttg ataaatttct gaatgacaat  
2040

tttccatggc atgtaatgga aagcatatcc gatgtggata atgaaaattg gatcgctata  
2100

attccaagat tgtcccctta agcgaactgg aaaagactct catctaaaat atgtttgaat  
2160

aatttatcat gccctgacaa gtacacacaa acacagacac ataataatac tacatatata  
2220

tatataccg ttattatgcg tgcacatgac aatgcccttg tatgtttcgt atactgtagc  
2280

aagtagtoat cttttgttc ccggttcgga aaatgacaaa aagtaaaatc aataaatgaa  
2340

gagtaaaaaa caatttatga aagggtgagc gaccagcaac gagagagaca aatcaaatta  
2400

gcgctttcca gtgagaatat aagagagcat tgaagagct aggttattgt taaatcatct  
2460

cgagctc  
2467

<210> 12

< 211> 777

5 < 212> ADN

< 213> *Saccharomyces cerevisiae*

<220>

< 221> misc\_feature

< 223> ribulosa 5-fosfato isomerasa

10 <400> 12

atggctgccg gtgtcccaaa aattgatgcg ttagaatctt tgggcaatcc tttggaggat  
60

gccaagagag ctgcagcata cagagcagtt gatgaaaatt taaaatttga tgatcacaaa  
120

attattggaa ttggtagtgg tagcacagtg gtttatgttg ccgaaagaat tggacaatat  
180

ES 2 607 891 T3

ttgcatgacc ctaaatttta tgaagtagcg tctaaattca tttgcattcc aacaggattc  
240

caatcaagaa acttgatttt ggataacaag ttgcaattag gctccattga acagtatcct  
300

cgcattgata tagcgtttga cggtgctgat gaagtggatg agaatttaca attaattaaa  
360

ggtggtggtg cttgtctatt tcaagaaaaa ttggttagta ctagtgctaa aaccttcatt  
420

gtcgttgctg attcaagaaa aaagtcacca aaacatttag gtaagaactg gaggcaaggt  
480

gttcccattg aaattgtacc ttctcatac gtgagggtca agaatgatct attagaacaa  
540

ttgcatgctg aaaaagttga catcagacaa ggaggttctg ctaaagcagg tctgtttgta  
600

actgacaata ataacttcat tatcgatgcg gatctcggtg aaatttccga tccaagaaaa  
660

ttgcatagag aaatcaaact gttagtgggc gtgggtgaaa caggtttatt catcgacaac  
720

gcttcaaaaag cctacttcgg taattctgac ggtagtgttg aagttaccga aaagtga  
777

<210> 13

< 211> 1328

5 < 212> ADN

< 213> *Saccharomyces cerevisiae*

<220>

< 221> misc\_feature

< 223> ribulosa 5-fosfato epimerasa

10 <400> 13

gtaggcact tacgtatctt gtatagtagg aatggctcgg tttatgtata ttaggagatc  
60

aaaacgagaa aaaaatacca tatcgtatag tatagagagt ataaatataa gaaatgccgc  
120

atatgtacaa ctaatctagc aaatctctag aacgcaattc cttcgagact tcttctttca  
180

ES 2 607 891 T3

tgaaggagat aacatcgtgc gggtcagctg cagtgaaaac actggtacca ggcacaataa  
240

cgttggcacc ggctttggcg gctttcggga tggctcctt gcccaaacca ccacgactt  
300

ggaalalcaa atgggggaac ttggctctca aagtttccac ttttggcacc atgtcttcca  
360

tgaatttttg gcctccaaac ccaggttcca cagtcataac aagagccata tccaaatgag  
420

gagctagttc aaataaaaac tcaacagaag taccaggttt gatggcgcat gcagctttga  
480

tgcccttaga cttaatcaac ttaactaaat gcaaagggtc ttgtgtggcc tcgtagtgga  
540

acgtaaattg gtcagcacca catttagoaa aatcgtcgac ccatttttca ggattttcaa  
600

ccatcatgtg acaatcgaag aacgcagtg gcttcttttc tgtgttgcta gcacgcccag  
660

ggcgtggcac agaacgacgt agggaggtaa caattggttg gccagagta atgtttggaa  
720

caaatggcc gtccatgaca tcgatalgta accaatctgc gccggcgttg atgaccttat  
780

gacattcgca acccaagttg gogaagtcag aagcaaggat actgggagct ataattggtt  
840

tgaccatttt ttcttgtgtg tttacctcgc tcttgaatt agcaaatgyc ctctctgcat  
900

gaaattgtat cgagtttgc tttttttct ttttacgggc ggattctttc tattctggtt  
960

ttctataac agagatcatg aaagaagttc cagcttacgg atcaagaaag tacctataca  
1020

tatacaaaaa tctgattact ttcccagctc gacttggata gctgttcttg tttctcttg  
1080

ggcacacatt ttttgtttct gaagccacgt cctgctttat aagaggacat ttaaagttgc  
1140

aggacttgaa tgcaattacc ggaagaagca accaaccggc atggttcagc atacaatata  
1200

catttgatta gaaaagcaga gaataaatag acatgatacc tctcttttta tctctgcag  
1260

cgtattattg tttattccac gcaggcatcg gtcgttggct gttgttatgt ctacagataag  
1320

cgcgtttg  
1328

<210> 14

< 211> 2046

< 212> ADN

5 < 213> Saccharomyces cerevisiae

<220>

< 221> misc\_feature

ES 2 607 891 T3

< 223> transcetolasa

<400> 14

atggcacagt tctccgacat tgataaactt gcggtttcca cttaagatt actttccggt  
60  
gaccaggtgg aaagcgcaca atctggccac ccaggtgcac cactaggatt ggcaccagtt  
120  
gcccattgaa ttttcaagca actgcgctgt aaccctaaca atgaacattg gatcaataga  
180  
gacaggtttg ttctgtcgaa cggtcactca tgcgctcttc tgtactcaat gctccatcta  
240  
ttaggatacg attactctat cgaggacttg agacaattta gacaagtaa ctcaaggaca  
300  
ccgggtcacc cagaattcca ctacgcggga gtggaaatca ctcccggtcc gctaggccag  
360  
ggtatctcaa atgctgttgg tatggcaata gcgcaggcca actttgccgc cacttataac  
420  
gaggatggct ttcccatttc cgactcatat acgtttgcta ttgtagggga tggttgctta  
480  
caagagggtg tttcttcgga gacctcttcc ttagcgggac atctgcaatt gggtacttg  
540  
attacgtttt atgacagtaa tagcatttcc attgacggta aaacctgta ctcgttcgac  
600  
gaagatgttt tgaagcgata cgaggcatat ggttgggaag tcatggaagt cgataaagga  
660  
gacgacgata tggaatccat ttctagcgtt ttggaaaagg caaaactatc gaaggacaag  
720  
ccaaccataa tcaaggtaac tactacaatt ggatttgggt cctacaaca gggactgct  
780  
ggtgttcatg ggtccgcttt gaaggcagat gatgttaaac agttgaagaa gaggtggggg  
840

ES 2 607 891 T3

tttgaoccaa ataatcatt tgtagtaacct caagaggtgt acgattatta taagaagact  
900

gttgtggaac ccggtcaaaa acttaatgag gaatgggata ggatgtttga agaatacaaa  
960

accaaatctc ccgagaaggg taaagaattg caaagaagat tgaatggtga gttaccgga  
1020

ggttgggaaa agcatttacc gaagtttact ccggaacgacg atgctctggc aacaagaaaag  
1080

acatcccagc aggtgctgac gaacatggtc caagttttgc ctgaattgat cgggtggtct  
1140

gccgatttga caocttogaa tctgacaagg tgggaaggcg cggtagattt ccaaoctccc  
1200

attaccaac taggtaacta tgcaggaagg tacattagat acggtgtgag ggaacacgga  
1260

atgggtgcca ttatgaacgg tatctctgcc tttggtgcaa actacaagcc ttacggtggt  
1320

acctttttga acttctgtctc ttatgctgca ggagccgta ggtagccgc cttgtctggt  
1380

aatccagtca tttgggttgc aacacatgac tctatcgggc ttggtgagga tgtccaacg  
1440

caccaaccta ttgaaactct ggctcacttg agggctattc caaacatgca tgtatggaga  
1500

octgtgatg gtaacgaaac ttctgctgcg tattattctg ctatcaaate tggatgaaca  
1560

ccatctgttg tggctttatc acgacagaat cttcctcaat tggagcattc ctcttttgaa  
1620

aaagccttga aggggtggcta tgtgatccat gacgtggaga atcctgatat tatcctggtg  
1680

tcaacaggat cagaagtctc catttctata gatgcagcca aaaaattgta cgatactaaa  
1740

aaaatcaaag caagagttgt ttcoctgcca gacttttata cttttgacag gcaaagtga  
1800

gaatacagat tctctgttct accagacggt gttccgatca tgtcctttga agtattggct  
1860

acttcaagct ggggtaagta tgcctcatcaa tcgttcggac tgcagaatt tggctggtca  
1920

ggcaaggggc ctgaaattta caaattgttc gatttcacag cggacgggtg tgcgtcaagg  
1980

gctgaaaaga caatcaatta ctacaaagga aagcagttgc tttctcctat ggaagagct  
2040

ttctaa  
2046

ES 2 607 891 T3

<210> 15

< 211> 1008

< 212> ADN

< 213> *Saccharomyces cerevisiae*

5 <220>

< 221> misc\_feature

< 223> transaldolasa

<400> 15

atgtctgaac cagctcaaaa gaaacaaaag gttgctaaca actctctaga acaattgaaa  
60

gctccggca ctgtcgttgt tgccgacct ggtgatttcg gctctattgc caagtttcaa  
120

cctcaagact ccacaactaa cccatcattg atcttggtcg ctgccaaagca accaacttac  
180

gccaaagtga togatgttgc cgtggaatac ggtaagaagc atggtaagac caccgaagaa  
240

caagtcgaaa atgctgtgga cagattgtta gtcgaattcg gtaaggagat cttaaagatt  
300

gttccaggca gagtctccac cgaagttgat gctagattgt cttttgacac tcaagctacc  
360

attgaaaagg ctagacatat cattaattg tttgaacaag aagggtgtctc caaggaaaga  
420

gtccttatta aaattgcttc cacttgggaa ggtattcaag ctgccaaaga attggaagaa  
480

aaggacggta tccactgtaa tttgactcta ttattctcct tcgttcaage agttgcctgt  
540

gccgaggccc aagttacttt gatttcccca tttgttggtg gaattctaga ctggtacaaa  
600

tccagcactg gtaaagatta caagggtgaa gccgaccag gtgttatttc cgtcaagaaa  
660

atctacaact actacaagaa gtacggttac aagactattg ttatgggtgc ttctttcaga  
720

agcactgacg aaatcaaaaa cttggctggt gttgactatc taacaatttc tccagcttta  
780

10

ttggacaagt tgatgaacag tactgaacct ttcccaagag ttttgaccc tgtctccgct  
840

aagaaggaag ccggcgacaa gatctttac atcagcgacg aatctaaatt cagattcgac  
900

ttgaatgaag acgctatggc cactgaaaaa ttgtccgaag gtatcagaaa attctctgcc  
960

gatattgtta ctctattcga cttgattgaa aagaaagtta ccgcttaa  
1008

ES 2 607 891 T3

<210> 16

< 211> 984

< 212> ADN

< 213> *Saccharomyces cerevisiae*

5 <220>

< 221> misc\_feature

< 223> aldosa reductasa

<400> 16

atgtcttcac tggttactct taataacggt ctgaaaatgc cctagtcgg cttagggtgc  
60

tggaaaattg acaaaaaagt ctgtgcgaat caaatttatg aagctatcaa attaggctac  
120

cgtttatcog atggtgcttg cgactacggc aacgaaaagg aagttggtga aggtatcagg  
180

aaagccatct ccgaaggctc tgtttctaga aaggatatat ttgttgtttc aaagttatgg  
240

aacaattttc accatcctga tcatgtaaaa ttagctttta agaagacctt aagcgatatg  
300

ggacttgatt atttagacct gtattatatt cacttcccaa tcgccttcaa atatggtcca  
360

tttgaagaga aataccctcc aggattctat acgggcgcag atgacgagaa gaaaggtcac  
420

atcaccgaag cacatgtacc aatcatagat acgtaccggg ctctggaaga atgtggtgat  
480

10 gaaggcttga ttaagtctat tgggtgttcc aactttcagg gaagcttgat tcaagattta  
540

ES 2 607 891 T3

ttacgtgggt gtagaatcaa gcccgtaggt ttgcaaattg aacaccatcc ttatttgact  
600

caagaacacc tagttgagtt ttgtaaatta cacgatatcc aagtagttgc ttaactcctcc  
660

ttcggctctc aatcattcat tgagatggac ttacagttgg caaaaaccac gccaaactctg  
720

ttcgagaatg atgtaatcaa gaaggtctca caaaaccatc caggcagtac cacttcccaa  
780

gtattgctta gatgggcaac tcagagaggg attgccgtca ttccaaaatc ttccaagaag  
840

gaaaggttac ttggcaacct agaaatcgaa aaaaagttca cttaacgga gcaagaattg  
900

aaggatattt ctgcactaaa tgccaacatc agatttaatg atccatggac ctggttgat  
960

ggtaaattcc ccaactttgc ctga  
984

<210> 17

< 211> 494

5 < 212> PRT

< 213> Piromyces sp.

<220>

< 221> misc\_feature

< 223> xilulosa cinasa

10 <400> 17

Met Lys Thr Val Ala Gly Ile Asp Leu Gly Thr Gln Ser Met Lys Val  
1 5 10 15

Val Ile Tyr Asp Tyr Glu Lys Lys Glu Ile Ile Glu Ser Ala Ser Cys  
20 25 30

Pro Met Glu Leu Ile Ser Glu Ser Asp Gly Thr Arg Glu Gln Thr Thr  
35 40 45

Glu Trp Phe Asp Lys Gly Leu Glu Val Cys Phe Gly Lys Leu Ser Ala  
50 55 60

ES 2 607 891 T3

Asp Asn Lys Lys Thr Ile Glu Ala Ile Gly Ile Ser Gly Gln Leu His  
65 70 75 80

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

Leu Trp Cys Asp Thr Ala Thr Val Glu Glu Cys Lys Ile Ile Thr Asp  
100 105 110

Ala Ala Gly Gly Asp Lys Ala Val Ile Asp Ala Leu Gly Asn Leu Met  
115 120 125

Leu Thr Gly Phe Thr Ala Pro Lys Ile Leu Trp Leu Lys Arg Asn Lys  
130 135 140

Pro Glu Ala Phe Ala Asn Leu Lys Tyr Ile Met Leu Pro His Asp Tyr  
145 150 155 160

Leu Asn Trp Lys Leu Thr Gly Asp Tyr Val Met Glu Tyr Gly Asp Ala  
165 170 175

Ser Gly Thr Ala Leu Phe Asp Ser Lys Asn Arg Cys Trp Ser Lys Lys  
180 185 190

Ile Cys Asp Ile Ile Asp Pro Lys Leu Leu Asp Leu Leu Pro Lys Leu  
195 200 205

Ile Glu Pro Ser Ala Pro Ala Gly Lys Val Asn Asp Glu Ala Ala Lys  
210 215 220

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

Asn Met Met Gly Ala Val Gly Thr Gly Thr Val Ala Asp Gly Phe Leu  
245 250 255

Thr Met Ser Met Gly Thr Ser Gly Thr Leu Tyr Gly Tyr Ser Asp Lys  
260 265 270

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

Gly Gly Trp Leu Pro Leu Leu Cys Thr Met Asn Cys Thr Val Ala Thr  
290 295 300

ES 2 607 891 T3

Glu Phe Val Arg Asn Leu Phe Gln Met Asp Ile Lys Glu Leu Asn Val  
 305 310 315 320

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

Phe Phe Asn Gly Glu Arg Thr Pro Asn Leu Pro Asn Gly Arg Ala Ser  
 340 345 350

Ile Thr Gly Leu Thr Ser Ala Asn Thr Ser Arg Ala Asn Ile Ala Arg  
 355 360 365

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

Phe Arg Lys Leu Gly Phe Gln Pro Lys Glu Ile Arg Leu Ile Gly Gly  
 385 390 395 400

Gly Ser Lys Ser Asp Leu Trp Arg Gln Ile Ala Ala Asp Ile Met Asn  
 405 410 415

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

Ala Val Gln Ala Leu Trp Cys Leu Lys Asn Gln Ser Gly Lys Cys Asp  
 435 440 445

Ile Val Glu Leu Cys Lys Glu His Ile Lys Ile Asp Glu Ser Lys Asn  
 450 455 460

Ala Asn Pro Ile Ala Glu Asn Val Ala Val Tyr Asp Lys Ala Tyr Asp  
 465 470 475 480

Glu Tyr Cys Lys Val Val Asn Thr Leu Ser Pro Leu Tyr Ala  
 485 490

<210> 18

<211> 2041

<212> ADN

5 <213> Piromyces sp.

<220>

<221> misc\_feature

<223> xilulosa cinasa

<400> 18

ES 2 607 891 T3

attatataaa ataactttaa ataaaaaat tttatattgt ttatttaatt attcaaaaa  
60  
aattaaagta aaagaaaaat aatacagtag aacaatagta ataatatcaa aatgaagact  
120  
gttgctggta ttgatcttgg aactcaaagt atgaaagtcg ttatttacga ctatgaaaag  
180  
aaagaaatta ttgaaagtgc tagctgtcca atggaattga tttccgaaag tgacgggtacc  
240  
cgtgaacaaa ccaactgaatg gttgacaag ggtcttgaag tttgtttgg taagcttagt  
300  
gctgataaca aaaagactat tgaagctatt ggtatttctg gtcaattaca cggttttgtt  
360  
cctcttgatg ctaacggtaa ggctttatac aacatcaaac tttgggtgga tactgctacc  
420  
gttgaagaat gtaagattat cactgatgct gccgggtggtg acaaggctgt tattgatgcc  
480  
cttggttaacc ttatgctcac cggtttcacc gctccaaaga tcctctggct caagcgcaac  
540  
aagccagaag ctttcgctaa cttaaagtac attatgcttc cacacgatta cttaaactgg  
600  
aagcttactg gtgattacgt tatggaatac ggtgatgcct ctggtaccgc tctcttcgat  
660  
tctaagaacc gttgctggtc taagaagatt tgcgatatca ttgacccaaa acttttagat  
720  
ttacttcaa agttaattga accaagcgcct ccagctggta aggttaatga tgaagccgct  
780  
aaggcttacg gtattccagc cggattcca gtttccgctg gtggtggtga taacatgatg  
840  
gggtctggtg gtactggtac tgttctgat ggtttcotta ccatgtotat gggttacttct  
900  
ggtactcttt acggttacag tgacaagcca attatgacc cagctaattg ttttaagtgg  
960  
ttctgttctt ctactggtgg atggcttcca ttactttgta ctatgaactg tactgttgcc  
1020  
actgaattcg ttcgtaacct cttccaaatg gatattaagg aacttaatgt tgaagctgcc  
1080

ES 2 607 891 T3

aagtctccat gtggtagtga aggtgtttta gttattccat tcttcaatgg tgaagaact  
1140

ccaaacttac caaacggctg tgctagtatt actggtctta cttctgctaa caccagccgt  
1200

gctaacattg ctctgtctag tttcgaatcc gccgtttctg ctatgcgtgg tggtttagat  
1260

gctttccgta agttaggttt ccaaccaaag gaaattcgtc ttattggtgg tggttctaag  
1320

tctgatctct ggagacaaat tgccgctgat atcatgaacc ttccaatcag agttccactt  
1380

ttagaagaag ctgctgctct tgggtggtct gttcaagctt tatggtgtct taagaaccaa  
1440

tctggaagt gtgatattgt tgaactttgc aaagaacaca ttaagattga tgaatctaag  
1500

aatgctaacc caattgccga aaatgttgct gtttacgaca aggcttacga tgaatactgc  
1560

aaggttgtaa atactcttct tccattatat gcttaaattg ccaatgtaa aaaaaatata  
1620

atgccatata attgccttgt caatacactg ttcattgtca tataatcata ggacattgaa  
1680

tttacaaggt ttatacaatt aatatctatt atcatattat tatacagcat ttcattttct  
1740

aagattagac gaaacaattc ttggttcctt gcaatatata aaatttacat gaatttttag  
1800

aatagtctcg tatttatgcc caataatcag gaaaattacc taatgctgga ttcttgtaa  
1860

taaaaacaaa ataaataaat taaataaaca aataaaaatt ataagtaaat ataaatata  
1920

aagtaatata aaaaaaaaaagt aataaataa ataaataaat aaaaattttt tgcaaatata  
1980

taataaata aataaaatat aaaaataatt tagcaataa attaaaaaaaa aaaaaaaaaa  
2040

a  
2041

## REIVINDICACIONES

1. Una célula anfitriona de levadura u hongos filamentosos transformada con un constructo de ácido nucleico que comprende una secuencia de nucleótidos que codifica una xilosa isomerasa, cuya xilosa isomerasa comprende una secuencia de aminoácidos que tiene al menos 95 % de identidad de secuencia con la secuencia de aminoácidos de SEQ ID NO. 2, por la que el constructo de ácido nucleico, tras la transformación de la célula anfitriona, confiere a la célula anfitriona la capacidad de isomerizar directamente xilosa a xilulosa, por la que la célula anfitriona comprende una modificación genética que aumenta el flujo de la ruta de la pentosa fosfato y la modificación genética comprende la sobreexpresión de al menos un gen de la parte no oxidativa de la ruta de la pentosa fosfato en comparación con una célula anfitriona que es genéticamente idéntica excepto por la modificación genética.
2. Una célula anfitriona según la reivindicación 1, en donde la secuencia de nucleótidos codifica una xilosa isomerasa obtenible a partir de una bacteria de la clase *Bacteroides*.
3. Una célula anfitriona según la reivindicación 1 o 2, por la que el gen de la parte no oxidativa de la ruta de la pentosa fosfato se selecciona del grupo que consiste en los genes que codifican ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa, transcetolasa y transaldolasa.
4. Una célula anfitriona según la reivindicación 3, por la que la modificación genética comprende la sobreexpresión de al menos los genes que codifican una transcetolasa y una transaldolasa.
5. Una célula anfitriona según cualquiera de las reivindicaciones 1-4, por la que la célula anfitriona comprende además una modificación genética que aumenta la actividad de la xilulosa cinasa específica.
6. Una célula anfitriona según la reivindicación 5, por la que la modificación genética comprende la sobreexpresión de un gen que codifica una xilulosa cinasa que es sobreexpresada.
7. Una célula anfitriona según una cualquiera de las reivindicaciones 1-6, por la que el gen que está sobreexpresado es endógeno a la célula anfitriona.
8. Una célula anfitriona según una cualquiera de las reivindicaciones 1-7, por la que la célula anfitriona comprende una modificación genética que reduce la actividad de la aldosa reductasa inespecífica en la célula anfitriona.
9. Una célula anfitriona según la reivindicación 8, por la que la modificación genética reduce la expresión de, o inactiva un gen que codifica una aldosa reductasa inespecífica.
10. Una célula anfitriona según la reivindicación 9, por la que el gen es inactivado por la eliminación de al menos parte del gen o por interrupción del gen.
11. Una célula anfitriona según las reivindicaciones 8 o 9, por la que la expresión de cada gen en la célula anfitriona que codifica una aldosa reductasa inespecífica se reduce o se inactiva.
12. Una célula anfitriona según una cualquiera de las reivindicaciones precedentes, en donde la célula anfitriona es una levadura, preferiblemente una levadura que pertenece a uno de los géneros: *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces* y *Yarrowia*.
13. Una célula anfitriona según la reivindicación 12, en donde la levadura pertenece a una de las especies: *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus*, y *K. fragilis*.
14. Una célula anfitriona según una cualquiera de las reivindicaciones 1 a 11, en donde la célula anfitriona es un hongo filamentosos, preferiblemente un hongo filamentosos que pertenece a uno de los géneros: *Aspergillus*, *Trichoderma*, *Humicola*, *Acremonium*, *Fusarium* y *Penicillium*.
15. Una célula anfitriona según una cualquiera de las reivindicaciones precedentes, por la que la célula anfitriona expresa una o más enzimas que confieren a la célula anfitriona la capacidad de producir al menos un producto de fermentación seleccionado del grupo que consiste en etanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, aminoácidos, 1,3-propanodiol, etileno, glicerol, antibióticos  $\beta$ -lactámicos y cefalosporinas.
16. Un procedimiento para producir etanol, por el que el procedimiento comprende las etapas de:  
 (a) fermentar un medio que contiene una fuente de xilosa con una célula anfitriona modificada como se define en una cualquiera de las reivindicaciones 1-14, por el que la célula anfitriona fermenta xilosa a etanol, y opcionalmente,  
 (b) recuperación del etanol.
17. Un procedimiento según la reivindicación 16, por el que el medio contiene también una fuente de glucosa.
18. Un procedimiento según las reivindicaciones 16 o 17, por el que la productividad volumétrica de etanol es al menos de 0,5 g de etanol por litro y hora.

19. Un procedimiento según una cualquiera de las reivindicaciones 16-18, por el que el rendimiento de etanol es al menos del 50 %.

5 20. Un procedimiento para producir un producto de fermentación seleccionado del grupo que consiste en ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, aminoácidos, 1,3-propanodiol, etileno, glicerol, antibióticos  $\beta$ -lactámicos y cefalosporinas, por el que el procedimiento comprende las etapas de: (a) fermentar un medio que contiene una fuente de xilosa con una célula anfitriona modificada como se define en la reivindicación 15, por el que la célula anfitriona fermenta la xilosa hasta el producto de fermentación, y opcionalmente, (b) recuperación del producto de fermentación.

21. Un procedimiento según la reivindicación 20, por el que el medio contiene también una fuente de glucosa.

10

Fig 1a

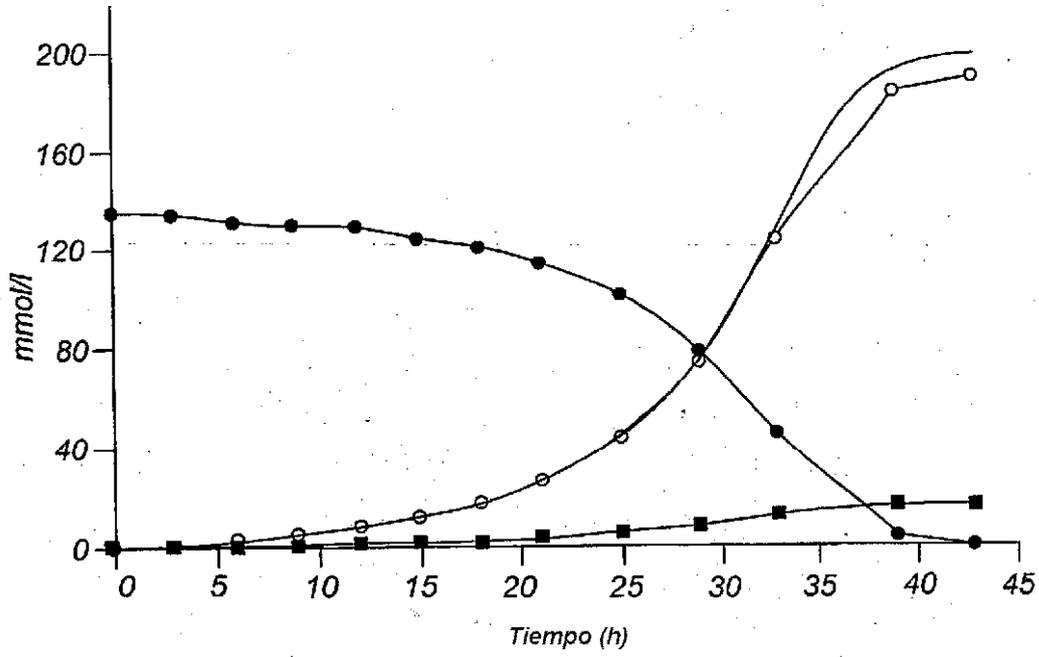


Fig 1b

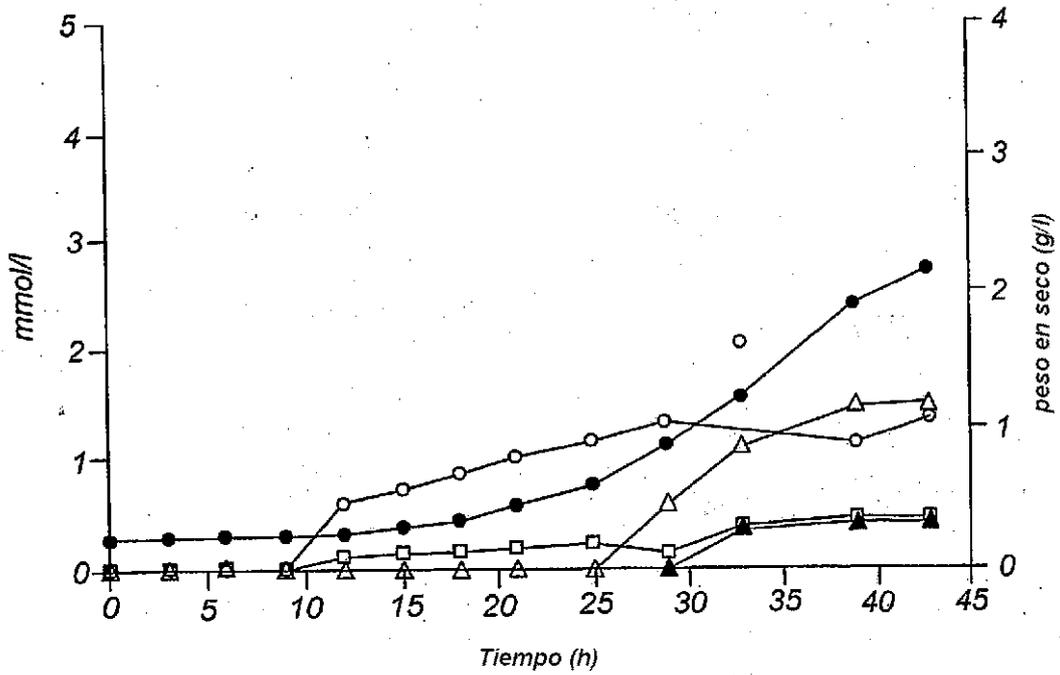


Fig 2a

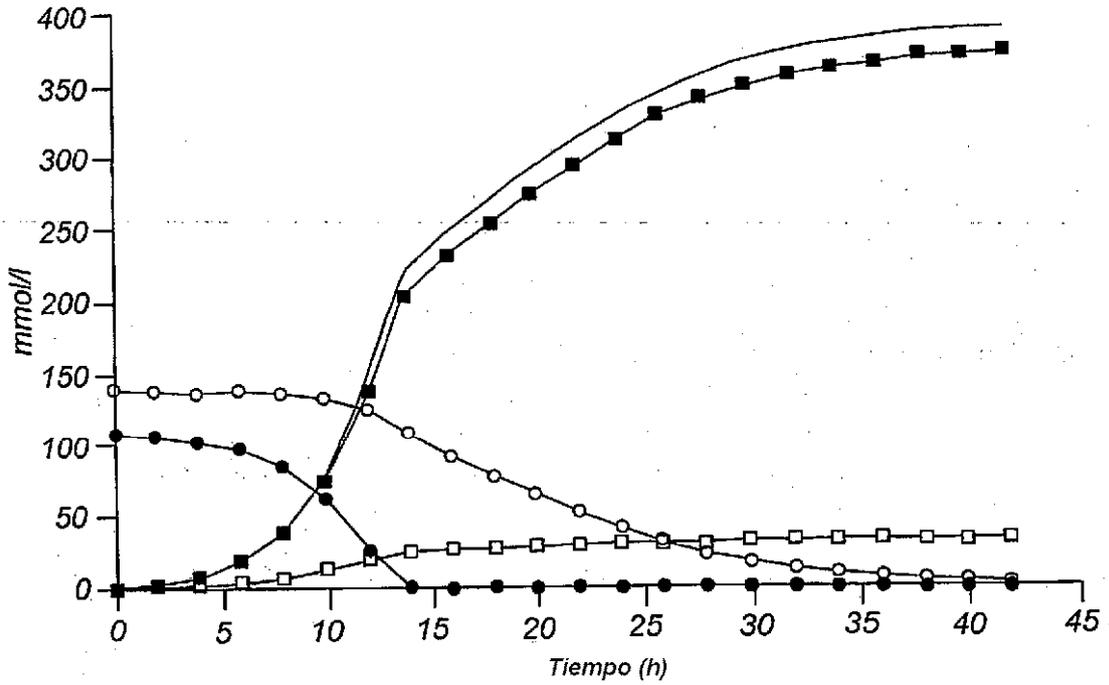


Fig 2b

