

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 582 782**

51 Int. Cl.:

C12N 9/92 (2006.01)

C12N 15/81 (2006.01)

C12P 7/06 (2006.01)

C12P 19/24 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **10.12.2009 E 09787397 (0)**

97 Fecha y número de publicación de la concesión europea: **20.04.2016 EP 2376630**

54 Título: **Microorganismo que expresa xilosa isomerasa**

30 Prioridad:

16.12.2008 GB 0822937
17.12.2008 US 138293 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.09.2016

73 Titular/es:

TERRANOL A/S (100.0%)
Building 223 Søtofts Plads
2800 Lyngby, DK

72 Inventor/es:

RØNNOW, BIRGITTE;
ANDERSEN, THOMAS HVID y
SIBBESEN, OLE

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 582 782 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

Microorganismo que expresa xilosa isomerasa

5 Campo de la invención

La presente invención se refiere a un microorganismo.

10 En particular, la presente invención se refiere a un microorganismo transformado capaz de: una actividad xilosa isomerasa superior a la del microorganismo equivalente antes de la transformación y/o una tasa de crecimiento superior en o sobre un medio de crecimiento que comprende xilosa que la del microorganismo equivalente antes de la transformación y/o un metabolismo de la xilosa más rápido que el del microorganismo equivalente antes de la transformación y/o una producción de etanol superior cuando se cultiva en condiciones anaerobias en xilosa como la fuente de carbono que la del microorganismo equivalente antes de la transformación.

15 La presente invención se refiere además a un método para la preparación de microorganismos transformados capaces de: una actividad de la xilosa isomerasa superior a la del microorganismo antes de la transformación y/o una tasa de crecimiento superior en o sobre un medio de crecimiento que comprende xilosa que la del microorganismo antes de la transformación y/o un metabolismo de la xilosa más rápido que el del microorganismo antes de la transformación y/o una producción de etanol superior cuando se cultiva en condiciones anaerobias en xilosa como la fuente de carbono que la del microorganismo antes de la transformación.

20 Además, la presente invención se refiere a inóculos y medios de cultivo que comprenden un microorganismo de acuerdo con la presente invención o un microorganismo preparado mediante un método de acuerdo con la presente invención.

25 En otro aspecto, la presente invención se refiere a un método de fermentación que comprende cultivar en un medio de cultivo un microorganismo de acuerdo con la presente invención o un microorganismo preparado mediante el método de la presente invención.

30 La presente invención se refiere además a métodos para producir un biocombustible y/o un producto derivado de la xilosa que comprende cultivar en un medio de cultivo un microorganismo de la presente invención o un microorganismo preparado mediante el método de la presente invención.

35 En la presente memoria se describe un biocombustible y/o un producto derivado de la xilosa obtenido mediante un método de la presente invención.

40 Además, la presente invención se refiere al uso de un microorganismo de acuerdo con la presente invención o un microorganismo preparado mediante el método de la presente invención para la producción de un producto derivado de xilosa y/o un biocombustible.

Antecedentes

45 El etanol se considera un combustible para el transporte alternativo atractivo que puede usarse como un aditivo, suplemento o sustitución parcial del combustible fósil gasolina que, en última instancia, es un recurso limitado.

50 El etanol puede ser utilizado en concentraciones más pequeñas en lugar del metil terc-butil éter (MTBE) como un refuerzo de octanos más limpio o como aditivo de la gasolina para la oxigenación o puede ser utilizado para la mezcla en la gasolina en cantidades superiores del 10-85 %, donde las mezclas se pueden utilizar en motores de automóviles sólo ligeramente modificados como los conocemos hoy en día. Por lo tanto, tiene una enorme ventaja sobre otros combustibles alternativos para el transporte. Cualquier sustitución fraccionaria se puede hacer con la máxima flexibilidad, y esto sin cambios drásticos en la tecnología de los motores para transporte tal como la conocemos hoy en día.

55 El etanol tiene la ventaja de ser un recurso renovable, ya que puede ser producido en grandes cantidades a partir de material vegetal. De esto también se deduce la ventaja de tener una contribución neta baja a la liberación de dióxido de carbono a la atmósfera, ya que el dióxido de carbono liberado por el uso de etanol se reabsorbe por la producción necesaria de material vegetal para regenerar el etanol utilizado.

60 En la producción de etanol por fermentación tradicional, el material vegetal utilizado es el azúcar de 6 carbonos, que o bien se extrae directamente de las plantas que contienen monosacáridos o disacáridos libres, o se genera mediante la hidrólisis del almidón a partir de las partes de la planta que contienen almidón. Estos azúcares proceden de partes de plantas que también se utilizan para la alimentación humana o animal. La producción tradicional de etanol está, por lo tanto, compitiendo directamente por el material vegetal, que de otro modo se utiliza para alimento humano o animal. La conversión de los alimentos en combustible plantea preocupaciones éticas en un mundo donde millones de personas están muriendo de hambre; los cálculos muestran que la cantidad de maíz necesaria para

producir un tanque lleno de combustible para un todoterreno común en los EE.UU. es aproximadamente la cantidad que se necesita para alimentar a una persona hambrienta durante un año completo.

Una solución alternativa es el uso de material vegetal lignocelulósico que se considera de otro modo como residuo para la generación de energía y combustible. Es decir, en general, al ser considerado de una manera muy positiva, el desarrollo de formas de utilizar material vegetal lignocelulósico para la producción de etanol ha obtenido en muchos países un gran apoyo, tanto a nivel político como por el público en general. Otra ventaja es la gran cantidad de material lignocelulósico, lo que realmente haría factible la sustitución del 50 % o más del consumo actual de gasolina con etanol.

Existen numerosas fuentes de material lignocelulósico que pueden utilizarse para la producción de etanol, siempre que se pueda establecer un proceso de recolección y conversión eficiente y rentable. Algunos ejemplos de cosechas que se están realizando actualmente incluyen el bagazo de la caña de azúcar, las astillas de madera, los rastrojos de maíz y la paja de trigo.

Los dos principales problemas técnicos encontrados en el uso de material lignocelulósico para la producción de etanol mediante la fermentación son:

1. Las dificultades en la hidrólisis de la celulosa que contiene azúcar de 6 carbonos en glucosa sin generar subproductos que dificulten o impidan una mayor conversión microbiana de la glucosa en etanol y
2. La falta de un organismo que sea capaz de la conversión eficiente de los azúcares de 5 carbonos (por ejemplo, xilosa y arabinosa) presentes en la parte hemicelulosa del material lignocelulósico en etanol.

Desde hace más de 6000 años *Saccharomyces cerevisiae* ha sido el microorganismo preferido para la conversión de los azúcares de 6 carbonos derivados de material vegetal en etanol. Esto es debido a una combinación favorable de un rendimiento de etanol alto, una productividad específica alta y una tolerancia al etanol alta junto con la capacidad de crecer y fermentar a pH bajo, donde otros microorganismos competidores tienen problemas para sobrevivir. Una vez superadas las dificultades en la hidrólisis de la celulosa, entonces lo más probable es que *Saccharomyces cerevisiae* siga siendo el organismo preferido para utilizar en la generación de etanol por fermentación de los monómeros de azúcar de 6 carbonos liberados. Pero desgraciadamente, *S. cerevisiae* es incapaz de metabolizar los azúcares de 5 carbonos que constituyen hasta un tercio del material de azúcar lignocelulósico.

En consecuencia, durante las últimas dos décadas se ha estado trabajando para buscar formas de manipular genéticamente *S. cerevisiae* con el fin de introducir la capacidad de fermentación del azúcar de 5 carbonos. Este trabajo se ha centrado principalmente en la fermentación de la xilosa, ya que la xilosa constituye la parte dominante de los azúcares de 5 carbonos en la lignocelulosa. A pesar de que *S. cerevisiae* no puede metabolizar la xilosa, es capaz de metabolizar el isómero xilulosa (Wang y Schneider, 1980), que es el metabolito que constituye un punto de entrada en la ruta de las pentosas fosfato. Así, en principio, el problema se puede resolver dotando a *S. cerevisiae* de la capacidad de convertir la xilosa en xilulosa.

Se han introducido con éxito dos rutas bioquímicas existentes diferentes en *S. cerevisiae* con el fin de canalizar la xilosa en la ruta de las pentosas fosfato a través de la xilulosa. En los hongos que metabolizan xilosa de forma natural, la xilosa se convierte en xilulosa en un proceso de dos etapas (ver Figura 1). La xilosa se reduce primero en xilitol mediante una xilosa reductasa (XR-EC 1.1.1.21). El xilitol se deshidrogena después en xilulosa por la xilitol deshidrogenasa (XDH - EC 1.1.1.9). La reductasa y la deshidrogenasa no están presentes en *S. cerevisiae*, pero ha sido posible transferir y expresar los genes que codifican estas dos enzimas de *Pichia stipitis* en *S. cerevisiae* y la cepa de *S. cerevisiae* resultante modificada es capaz de metabolizar la xilosa.

La actividad de las dos enzimas requiere NADPH y NAD⁺ y genera NADP⁺ y NADH, de modo que es necesario para el organismo regenerar el equilibrio NAD⁺ - NADH y NADP⁺ - NADPH por procesos redox en otros lugares del metabolismo, de lo contrario el metabolismo de la xilosa dejará de funcionar cuando el NADPH y el NAD⁺ se hayan agotado. La tasa bastante reducida del metabolismo de la xilosa y la generación de grandes cantidades de xilitol por cepas de *S. cerevisiae* modificadas de esta manera se ha atribuido a este problema inherente a la ruta de la xilosa reductasa - xilitol deshidrogenasa

En las bacterias que metabolizan xilosa, la conversión de xilosa en xilulosa es diferente: una sola enzima, la xilosa isomerasa (EC 5.3.1.5), convierte la xilosa directamente en xilulosa (véase la Figura 1). Esto parece ser más simple y no genera el desequilibrio NAD⁺ - NADPH mencionado anteriormente. Por lo tanto, también fue la primera estrategia que se propuso para conferir la capacidad de metabolizar la xilosa en *S. cerevisiae*. Pero ha resultado difícil que esta estrategia funcione. La expresión de la mayoría de los genes de la xilosa isomerasa de bacterias no da lugar a la presencia de una xilosa isomerasa activa en *S. cerevisiae* y todavía se desconoce la razón exacta de ello. Diversos estudios han puesto de manifiesto que a menudo se puede detectar una proteína producida a partir de la expresión de un gen bacteriano, pero la proteína expresada aparentemente no se pliega en una enzima activa en *S. cerevisiae*.

Las enzimas termoestables son a menudo también más estables en otros tipos de condiciones extremas, tales como una alta concentración de sal y valores de pH extremos. Esto puede indicar que las enzimas termoestables son más propensas a mantener (y probablemente también conseguir) un pliegue correcto. Y, de hecho, ha sido posible encontrar ejemplos de genes de la xilosa isomerasa bacteriana aislados de organismos termófilos que expresan xilosa isomerasas activas en *S. cerevisiae* (Ver: Walfridsson et al, 1996; y Bao et al, 1999). Se ha demostrado que estos genes permiten a *S. cerevisiae* metabolizar la xilosa, pero a una tasa muy baja. La baja tasa se ha atribuido al hecho de que las xilosa isomerasas termoestables tienen una actividad óptima a 80 - 90 °C, pero las temperaturas que permiten la supervivencia de la mayoría de las cepas de *S. cerevisiae* son 30 - 35 °C.

10 Declaraciones de la invención

La presente invención se refiere a un microorganismo transformado capaz de:

- 15 (a) una actividad xilosa isomerasa superior a la del microorganismo equivalente antes de la transformación y/o
 (b) una tasa de crecimiento superior en o sobre un medio de crecimiento que comprende xilosa que la del microorganismo equivalente antes de la transformación y/o
 (c) un metabolismo de la xilosa más rápido que el del microorganismo equivalente antes de la transformación y/o
 (d) una producción de etanol superior cuando se cultiva en condiciones anaerobias en xilosa como la fuente de carbono que la del microorganismo equivalente antes de la transformación

20 en el que dicho microorganismo ha sido transformado con una secuencia de nucleótidos que codifica una xilosa isomerasa en el que dicha xilosa isomerasa es una xilosa isomerasa exógena derivada de una bacteria mesófila y en el que dicha xilosa isomerasa es una xilosa isomerasa exógena derivada de una especie de *Lactococcus*; y en el que dicho microorganismo transformado es una levadura transformada en el que dicha levadura transformada es *Saccharomyces*.

25 En otro aspecto, la presente invención se refiere a un inóculo que comprende un microorganismo de acuerdo con la presente invención.

30 La presente invención se refiere a, en otro aspecto, un medio de cultivo que comprende un microorganismo de acuerdo con la presente invención.

35 En un aspecto adicional, la presente invención se refiere a un método para preparar un microorganismo transformado, comprendiendo dicho método la etapa de transformar un microorganismo, de tal modo que dicho microorganismo transformado es capaz de:

- 40 (a) una actividad de la xilosa isomerasa superior a la del microorganismo antes de la transformación y/o
 (b) una tasa de crecimiento superior en o sobre un medio de crecimiento que comprende xilosa que la del microorganismo antes de la transformación y/o
 (c) un metabolismo de la xilosa más rápido que el del microorganismo antes de la transformación y/o
 (d) una producción de etanol superior cuando se cultiva en condiciones anaerobias en xilosa como la fuente de carbono que la del microorganismo antes de la transformación

45 en el que dicho microorganismo ha sido transformado con una secuencia de nucleótidos que codifica una xilosa isomerasa en el que dicha xilosa isomerasa es una xilosa isomerasa exógena derivada de una bacteria mesófila y en el que dicha xilosa isomerasa es una xilosa isomerasa exógena derivada de una especie de *Lactococcus*; y en el que dicho microorganismo transformado es una levadura transformada en el que dicha levadura transformada es *Saccharomyces*.

50 En otro aspecto, la presente invención se refiere a un método de fermentación que comprende cultivar en un medio de cultivo un microorganismo de acuerdo con la presente invención o un microorganismo preparado mediante el método de la presente invención.

55 La presente invención proporciona, en un aspecto adicional, un método para producir un producto derivado de la xilosa que comprende cultivar en un medio de cultivo un microorganismo de acuerdo con la presente invención o un microorganismo preparado mediante el método de la presente invención.

60 La presente invención proporciona, en otro aspecto, un método para producir un biocombustible en el que dicho método comprende la etapa de cultivar en un medio de cultivo un microorganismo de la presente invención o un microorganismo preparado mediante el método de la presente invención.

En la presente memoria se describe un biocombustible obtenido mediante un método de la presente invención.

65 En la presente memoria se describe un producto derivado de la xilosa obtenido mediante un método de la presente invención.

Además, la presente invención, proporciona el uso de un microorganismo de acuerdo con la presente invención o un microorganismo preparado mediante el método de la presente invención para la producción de un producto derivado de la xilosa.

5 Además, la presente invención proporciona el uso de un microorganismo de acuerdo con la presente invención o un microorganismo preparado mediante el método de la presente invención para la producción de un biocombustible.

En otro aspecto, la presente invención proporciona un microorganismo transformado, en el que dicho microorganismo transformado comprende una secuencia de nucleótidos exógena que codifica una xilosa isomerasa.

10 Discusión

Los esfuerzos anteriores para encontrar genes de xilosa isomerasa bacterianos de origen mesófilo con alta actividad cuando se expresan en, por ejemplo, *Saccharomyces cerevisiae* no han tenido éxito. Inesperadamente, como se divulga en la presente memoria, ahora los inventores han tenido éxito en la identificación de bacterias, en particular bacterias mesófilas, que contienen secuencias de nucleótidos que codifican la xilosa isomerasa, la cual tras la expresión en, por ejemplo, *S. cerevisiae*, da lugar a una alta actividad intracelular de xilosa isomerasa en el organismo. Tales secuencias se pueden utilizar, por ejemplo, para diseñar cepas de la familia *Saccharomyces* para permitir que la levadura convierta xilosa en xilulosa y de ese modo facilitar la producción y/o el metabolismo eficiente de la xilosa en etanol. La figura 4 detalla el metabolismo de la xilosa en etanol; la xilulosa-5-fosfato - un intermedio que se puede derivar a partir de la D-xilosa - entra en la ruta de las pentosas fosfato y se metaboliza adicionalmente en etanol en condiciones anaerobias. Pero además de eso, dichas secuencias pueden utilizarse también para diseñar cepas de, por ejemplo, *S. cerevisiae* que sean productoras eficientes de otros compuestos derivados de xilosa través de la ruta de las pentosas fosfato. Ejemplos de tales compuestos incluyen, pero no se limitan a, aminoácidos aromáticos, ácido láctico, ácido succínico, ácido acético, acetaldehído, furfural, ácido itacónico, ácido glutámico, ácido cítrico, cresol, lisina, ácido 3-hidroxiisovalérico, poli-3-hidroxiacanoatos, ácido protocatéquico, pirocatecol, guayacol, veratrol, resveratrol, vainillina, ácido vainillínico, alcohol vainillínico, ácido mucónico, ácido adípico, ácido 4-hidroxibenzoico, 4-hidroxibenzaldehído, ácido 4-metoxibenzoico, 4-aminobenzoato, 4-hidroxianilina, 4-metoxianilina, quinol, anisol, fenol, ácido antranílico, 3-hidroxiantranilato, ácido 2,3-dihidroxibenzoico, 2-aminofenol, 1,4-ciclohexanodiona, isopreno y estireno.

Algunas ventajas

Ventajosamente, la presente invención se refiere a la transformación (ingeniería) de microorganismos, tales como los de la familia *Saccharomyces*, para permitir que los microorganismos metabolicen xilosa. Esto, a su vez, aumenta ventajosamente la eficiencia de los microorganismos, en particular cepas de *Saccharomyces*, para producir etanol a partir de materia prima que contiene xilosa (tales como material lignocelulósico). Además, ventajosamente aumenta la eficacia de los microorganismos para producir otros metabolitos naturales o modificados derivados de los intermedios en la ruta de las pentosas fosfato, tales como aminoácidos aromáticos, ácido láctico, ácido succínico, ácido acético, acetaldehído, furfural, ácido itacónico, ácido glutámico, ácido cítrico, cresol, lisina, ácido 3-hidroxiisovalérico, poli-3-hidroxiacanoatos, ácido protocatéquico, pirocatecol, guayacol, veratrol, resveratrol, vainillina, ácido vainillínico, alcohol vainillínico, ácido mucónico, ácido adípico, ácido 4-hidroxibenzoico, 4-hidroxibenzaldehído, ácido 4-metoxibenzoico, 4-aminobenzoato, 4-hidroxianilina, 4-metoxianilina, quinol, anisol, fenol, ácido antranílico, 3-hidroxiantranilato, ácido 2,3-dihidroxibenzoico, 2-aminofenol, 1,4-ciclohexanodiona, isopreno y estireno. Además, facilita el uso general de material que contiene xilosa como materia prima para el crecimiento de las cepas de la familia *Saccharomyces* para fines industriales.

Una ventaja adicional es que la secuencia de nucleótidos que codifica la xilosa isomerasa se deriva de un organismo mesófilo y/o es capaz de una alta actividad a temperaturas mesófilas.

Otra ventaja es que los microorganismos de acuerdo con la presente invención producen una xilosa isomerasa activa.

Ventajosamente, mediante el uso de los microorganismos de acuerdo con la presente invención, se pueden producir biocombustibles (tales como etanol).

De forma más ventajosa, mediante el uso de los microorganismos de acuerdo con la presente invención, los biocombustibles pueden ser producidos a partir de materiales de desecho, tales como los residuos agrícolas (incluyendo paja de cereales, tales como paja de trigo, pulpa de remolacha, bagazo de caña de azúcar; rastrojos, tales como rastrojos de sorgo, soja, rastrojos de maíz y astillas de madera). Con la presente invención no hay necesidad (o hay una necesidad reducida) de utilizar los materiales (tales como extracto de caña de azúcar, extracto de remolacha azucarera, almidón de sorgo, almidón de maíz o almidón de trigo) que de otro modo podrían ser utilizados como un alimento fuente para los seres humanos y/o como pienso para los animales.

Ventajosamente, los microorganismos de acuerdo con la presente invención permiten el uso óptimo de los azúcares liberados por hidrólisis del material lignocelulósico mediante la fermentación de azúcares pentosa, en particular xilosa.

5 La presente invención permite la producción de un biocombustible que es más neutro en términos de CO₂ en comparación con el transporte de combustible a base de petróleo. Con la presente invención, las emisiones de CO₂ serán bajas (incluso más bajas) durante la producción del biocombustible de acuerdo con la presente invención que en comparación con la producción de un combustible para transporte basado en petróleo típico (combustibles fósiles)

10

Figuras

15 Figura 1. Una representación esquemática de las rutas metabólicas que detallan el metabolismo de las dos aldopentosas, D-xilosa y L-arabinosa. Ambas aldopentosas se convierten en una cetopentosa y además en D-xilulosa 5-fosfato. Sin desear estar limitado por la teoría, como se indica en la figura, un tipo de ruta (tipo aldosa reductasa) se encuentra en los hongos, mientras que el otro tipo de ruta (tipo isomerasa) se encuentra en las bacterias. En el tipo de ruta fúngica, las primeras enzimas pueden denominarse "D-xilosa reductasa" y "L-arabinosa reductasa", pero a menudo la misma enzima puede reducir tanto la D-xilosa como la L-arabinosa y ambas rutas pueden servir y ser denominadas con el nombre menos específico de "aldosa reductasa".

20

Figura 2. Sin desear estar limitado por la teoría, la Figura 2 muestra una representación esquemática de los tipos de rutas metabólicas fúngicas y bacterianas de algunas pentosas menos abundantes (D-lixosa y L-lixosa, D-ribosa) detallando el metabolismo inicial hasta la entrada en la ruta de las pentosas fosfato.

25 Figura 3. Una representación esquemática de la parte no oxidativa de la ruta de las pentosas fosfato (RPF).

Figura 4. Una representación esquemática de la conversión metabólica de xilosa en etanol. Aquí, la cetopentosa xilulosa-5-fosfato, que se puede derivar a partir de D-xilosa, se metaboliza adicionalmente en etanol en condiciones anaerobias. XI es xilosa isomerasa y XK es xiluloquinasa. Se muestra la entrada neta de xilosa en el proceso y la producción neta de etanol a partir del proceso (proceso neto).

30

Figura 5. Actividad de la xilosa isomerasa (XI) en *Saccharomyces cerevisiae* transformada con una secuencia de nucleótidos que codifica una xilosa isomerasa bacteriana. La xilosa isomerasa bacteriana se deriva de (i) *Pseudomonas syringae*, (ii) *Thermoanaerobacter thermohydrosulfuricus*, (iii) *Thermoanaerobacter thermohydrosulfurigenes* o (iv) *Lactococcus lactis* subsp. *lactis* (*Lactobacillus xylosus*).

35

Figura 6. El crecimiento en xilosa de *Saccharomyces cerevisiae* transformada con una secuencia de nucleótidos que codifica y expresa una xilosa isomerasa bacteriana y una secuencia de nucleótidos que codifica y expresa una xilulosa quinasa clonada a partir de *Pichia stipitis*.

40

Figura 7. La actividad xilosa isomerasa específica de *Saccharomyces cerevisiae* transformada con una secuencia de nucleótidos que codifica y expresa una variante de la xilosa isomerasa de *Lactococcus*. La variante xilosa isomerasa de *Lactococcus* (XI) tiene la secuencia de aminoácidos SEQ ID No 14, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o la SEQ ID No 20.

45

Descripción de la invención

50 En la presente memoria las expresiones "una actividad xilosa isomerasa superior a la del microorganismo equivalente antes de la transformación" y "una actividad xilosa isomerasa superior a la del microorganismo antes de la transformación" se refieren a un microorganismo transformado que tiene una actividad xilosa isomerasa de al menos 0,16, 0,2, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9 o 1,0 unidades de xilosa isomerasa por mg de proteína del microorganismo, cuando se cultiva en condiciones de cultivo adecuadas que permiten al menos el mantenimiento de dicho microorganismo; por ejemplo, dicho medio de cultivo puede comprender uno o más sustratos (por ejemplo, xilosa) capaces de activar la expresión de la secuencia de nucleótidos que codifica la xilosa isomerasa pero dicho medio de cultivo no comprende ningún sustrato que sea capaz de inhibir la expresión de la secuencia de nucleótidos que codifica la xilosa isomerasa. Un ejemplo de condiciones de cultivo adecuadas se describe en el Ejemplo 5. Dicha actividad xilosa isomerasa se mide utilizando el ensayo de cisteína-carbazol descrito por Dische y Borenfreund (1951). La actividad xilosa isomerasa del microorganismo transformado es superior a la del microorganismo equivalente antes de la transformación o a la del microorganismo antes de la transformación, cuando se cultiva en las mismas condiciones de cultivo que el microorganismo transformado.

60

65 En la presente memoria las expresiones "una tasa de crecimiento superior en o sobre un medio de crecimiento que comprende xilosa que la del microorganismo equivalente antes de la transformación" y "una tasa de crecimiento superior en o sobre un medio de crecimiento que comprende xilosa que la del microorganismo antes de la transformación" se refieren a un microorganismo transformado capaz de una mayor tasa de crecimiento de tal manera que el tiempo necesario para una duplicación en el número de microorganismos es por lo menos 10 %, 15

%, 20 % o 25 % menor que la del microorganismo equivalente antes de la transformación o el microorganismo antes de la transformación, cuando se cultivan en las mismas condiciones.

Las expresiones “un metabolismo de la xilosa más rápido que el del microorganismo equivalente antes de la transformación” y un “o un metabolismo de la xilosa más rápido que el del microorganismo antes de la transformación”, como se n en la presente memoria, se refieren a un microorganismo transformado que es capaz de metabolizar xilosa, de tal forma que la cantidad consumida de xilosa en el medio de cultivo es de al menos 10 %, 20 %, 25 %, 30 % o 35 % superior por célula que la del microorganismo equivalente antes de la transformación o el microorganismo antes de la transformación, cuando se cultiva en las mismas condiciones durante un determinado período de tiempo dentro de la fase de crecimiento exponencial. Dicho microorganismo transformado se cultiva en condiciones de cultivo adecuadas que permiten al menos el mantenimiento de dicho microorganismo; por ejemplo, dicho medio de cultivo puede comprender uno o más sustratos (por ejemplo, xilosa) capaces de activar la expresión de la secuencia de nucleótidos que codifica la xilosa isomerasa pero dicho medio de cultivo no comprende ningún sustrato que sea capaz de inhibir la expresión de la secuencia de nucleótidos que codifica la xilosa isomerasa.

En la presente memoria las expresiones “una producción de etanol superior cuando se cultiva en condiciones anaerobias en xilosa como la fuente de carbono que la del microorganismo equivalente antes de la transformación” y “una producción de etanol superior cuando se cultiva en condiciones anaerobias en xilosa como la fuente de carbono que la del microorganismo antes de la transformación” se refieren a un microorganismo transformado que es capaz de producir al menos 10 %, 20 % o 30 % más de etanol por célula que el microorganismo equivalente antes de la transformación cuando se cultiva en condiciones anaerobias, con xilosa como la fuente de carbono, durante un determinado período de tiempo.

La expresión “microorganismo equivalente antes de la transformación” como se usa en la presente memoria incluye referencias al microorganismo antes de la transformación con una secuencia de nucleótidos que codifica la xilosa isomerasa o antes de la transformación con una secuencia de nucleótidos que provoca la regulación positiva (por ejemplo, la sobreexpresión) de la xilosa isomerasa.

En una realización, el microorganismo de acuerdo con la presente invención se ha transformado con una secuencia de nucleótidos que hace que el microorganismo sobreexpresa la xilosa isomerasa. Por ejemplo, se inserta un promotor en el genoma de un microorganismo que permite que el microorganismo sobreexpresa una secuencia de nucleótidos que codifica la xilosa isomerasa endógena.

En otra realización, el microorganismo de acuerdo con la presente invención se ha transformado con una secuencia de nucleótidos que codifica la xilosa isomerasa. Por ejemplo, el microorganismo se transforma con un vector de expresión que comprende una secuencia de nucleótidos que codifica la xilosa isomerasa unida operativamente a una secuencia reguladora.

Preferiblemente, el microorganismo de acuerdo con la presente invención se ha transformado con una secuencia de nucleótidos que codifica una xilosa isomerasa.

Preferiblemente, el microorganismo de acuerdo con la presente invención se ha transformado con una secuencia de nucleótidos que codifica una xilosa isomerasa exógena.

Preferiblemente, el microorganismo de acuerdo con la presente invención se ha transformado con una secuencia de nucleótidos que codifica una xilosa isomerasa exógena derivada de una especie de *Lactococcus*.

En una realización preferida, el microorganismo de acuerdo con la presente invención se ha transformado con una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada como SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20, o una variante, homóloga o derivada de la misma.

En una realización más preferida, el microorganismo de acuerdo con la presente invención se ha transformado con una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada como SEQ ID No 19 o SEQ ID No 20, o una variante, homóloga o derivada de la misma.

En una realización preferida, el microorganismo de acuerdo con la presente invención se ha transformado con la secuencia de nucleótidos mostrada como SEC ID No 1, SEQ ID No 10, SEQ ID No 17, SEQ ID No 12, SEQ ID No 27 o SEQ ID No 28, o una variante, homóloga o derivada de la misma.

En una realización más preferida, el microorganismo de acuerdo con la presente invención se ha transformado con la secuencia de nucleótidos mostrada como SEC ID No 27 o SEQ ID No 28, o una variante, homóloga o derivada de la misma.

En un aspecto, el microorganismo de acuerdo con la presente invención se ha transformado con dos o más secuencias de nucleótidos que codifican la xilosa isomerasa.

Preferiblemente, la secuencia de nucleótidos que codifica la xilosa isomerasa mencionado en la presente memoria está en un vector de expresión que codifica la misma.

5 Preferiblemente, el vector de expresión mencionado en la presente memoria comprende un promotor capaz de sobreexpresar la secuencia de nucleótidos que codifica la xilosa isomerasa. Ejemplos de tales promotores incluyen el promotor GPD, el promotor TEF y el promotor ADP. Los promotores preferidos que pueden ser usados para sobreexpresar la xilosa isomerasa pueden ser cualquiera de los elementos reguladores que controlan la expresión de secuencias de nucleótidos que codifican proteínas implicadas en la glicolisis y la fermentación de la glucosa.

10 En una realización, en el método de acuerdo con la presente invención, el microorganismo se transforma con una secuencia de nucleótidos que hace que el microorganismo sobreexpresa la xilosa isomerasa. Por ejemplo, se inserta un promotor inserta en el genoma de un microorganismo que permite que el microorganismo sobreexpresa una secuencia de nucleótidos que codifica la xilosa isomerasa endógena. En un ejemplo adicional, se inserta un promotor en el genoma de un microorganismo que permite que el microorganismo exprese constitutivamente una
15 secuencia de nucleótidos que codifica la xilosa isomerasa endógena.

En otra realización en el método de acuerdo con la presente invención, el microorganismo se transforma con una secuencia de nucleótidos que codifica la xilosa isomerasa. Por ejemplo, el microorganismo se transforma con un vector de expresión que comprende una secuencia de nucleótidos que codifica la xilosa isomerasa unida operativamente a una secuencia reguladora. En un ejemplo adicional, el microorganismo se transforma con un vector de expresión que comprende una secuencia de nucleótidos que codifica la xilosa isomerasa unida operativamente a una secuencia reguladora en la que dicha secuencia reguladora permite la expresión constitutiva de la secuencia de nucleótidos que codifica la xilosa isomerasa.

25 Preferiblemente, en el método de acuerdo con la presente invención, el microorganismo se transforma con una secuencia de nucleótidos que codifica una xilosa isomerasa.

Preferiblemente, en el método de acuerdo con la presente invención, el microorganismo se transforma con una secuencia de nucleótidos que codifica una xilosa isomerasa exógena.

30 En una realización preferida, en el método de acuerdo con la presente invención, el microorganismo se transforma con una secuencia de nucleótidos, que codifica una xilosa isomerasa, que comprende la secuencia de nucleótidos mostrada como SEC ID No 1, SEQ ID No 10, SEQ ID No 17, SEQ ID No 12, SEQ ID No 27 o SEQ ID No 28, o una variante, homóloga o derivada de la misma.

35 En una realización preferida, en el método de acuerdo con la presente invención, el microorganismo se transforma con una secuencia de nucleótidos que codifica una xilosa isomerasa que comprende la secuencia de aminoácidos mostrada como SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o la SEQ ID No 20 o una variante, homóloga o derivada de la misma.

40 En un aspecto, en el método de acuerdo con la presente invención, el microorganismo se transforma con dos o más secuencias de nucleótidos que codifican la xilosa isomerasa.

45 En el método de acuerdo con la presente invención, preferiblemente la secuencia de nucleótidos que codifica una xilosa isomerasa está en un vector de expresión que codifica la misma.

Tal como se utiliza en la presente memoria, la expresión "método de fermentación" se refiere al cultivo de un microorganismo o microorganismos en condiciones aerobias y anaerobias.

50 En una realización, el medio de cultivo comprende xilosa y/o una fuente de xilosa.

En una realización, el medio de cultivo comprende un azúcar pentosa y/o una fuente de azúcar pentosa. Preferiblemente, el azúcar pentosa es xilosa. En una realización, el azúcar pentosa se deriva y/o es derivable de material lignocelulósico.

55 Como alternativa, o además, el medio de cultivo comprende material derivado de material lignocelulósico.

Preferiblemente, el método de acuerdo con la presente invención comprende además la etapa de obtener el biocombustible del medio de cultivo.

60 En una realización, el producto derivado de la xilosa se selecciona del grupo que consiste en xilulosa, xilulosa-5-fosfato, etanol, aminoácidos aromáticos, ácido láctico, ácido succínico, ácido acético, acetaldehído, furfural, ácido itacónico, ácido glutámico, ácido cítrico, cresol, lisina, ácido 3-hidroxiisovalérico, poli-3-hidroxiacanoatos, ácido protocatéquico, pirocatecol, guayacol, veratrol, resveratrol, vainillina, ácido vainílico, alcohol vainílico, ácido mucónico, ácido adípico, ácido 4-hidroxibenzoico, 4-hidroxibenzaldehído, 4-metoxibenzoico, 4-aminobenzoato, 4-
65

hidroxianilina, 4-metoxianilina, quinol, anisol, fenol, ácido antranílico, 3-hidroxiantranilato, ácido 2,3-dihidroxibenzoico, 2-aminofenol, 1,4-ciclohexanodiona, isopreno y estireno.

5 Preferiblemente, el producto derivado de la xilosa se selecciona del grupo que consiste en xilulosa, xilulosa-5-fosfato, etanol, aminoácidos aromáticos, ácido láctico, ácido succínico, ácido acético, acetaldehído, furfural, ácido itacónico, ácido glutámico, ácido cítrico, cresol, lisina, ácido 3-hidroxipropiónico y poli-3-hidroxialcanoatos.

En una realización muy preferida, el producto derivado de la xilosa es etanol.

10 Tal como se utiliza en la presente memoria, el término “sobreexpresé” en la frase “una secuencia de nucleótidos que hace que el microorganismo sobreexpresé la xilosa isomerasa” y “un promotor capaz de sobreexpresar la secuencia de nucleótidos que codifica la xilosa isomerasa” se refiere a un aumento en la expresión desde cero hasta un nivel de expresión o que va desde un nivel más bajo de expresión hasta un nivel más alto de expresión (por ejemplo, regulación positiva) cuando el microorganismo transformado se compara con el microorganismo equivalente antes de la transformación. Los microorganismos que sobreexpresan xilosa isomerasa tienen una mayor capacidad para catalizar la conversión de xilosa en xilulosa. La capacidad de convertir xilosa en xilulosa puede ser confirmada mediante la determinación de la capacidad del lisado de células para producir xilulosa a partir de xilosa ensayando la xilulosa producida usando, por ejemplo, un ensayo de cetopentosa como se describe por Dische y Borenfreund (Dische y Borenfreund, 1951), también descrito en la sección de Ejemplos de la presente memoria.

20 Preferiblemente, el microorganismo transformado que sobreexpresa xilosa isomerasa es capaz de:

(a) una actividad xilosa isomerasa superior a la del microorganismo equivalente antes de la transformación y/o
 (b) una tasa de crecimiento superior en o sobre un medio de crecimiento que comprende xilosa que la del microorganismo equivalente antes de la transformación y/o
 (c) un metabolismo de la xilosa más rápido que el del microorganismo equivalente antes de la transformación y/o
 (d) una producción de etanol superior cuando se cultiva en condiciones anaerobias en xilosa como la fuente de carbono que la del microorganismo equivalente antes de la transformación.

30 Ejemplos de microorganismos que sobreexpresan xilosa isomerasa incluyen: (i) microorganismos transformados con un vector de expresión que codifica la xilosa isomerasa (antes de la transformación dicho microorganismo no era capaz de expresar la xilosa isomerasa) y (ii) microorganismos transformados para regular positivamente la expresión de una xilosa isomerasa endógena (antes de la transformación dicho microorganismo era capaz de expresar dicha xilosa isomerasa en una serie determinada de condiciones de cultivo durante el crecimiento exponencial, pero después de la transformación dicho microorganismo es capaz de expresar dicha xilosa isomerasa en un nivel superior, en las mismas condiciones de cultivo, durante el crecimiento exponencial).

40 La expresión “una secuencia de nucleótidos que codifica la xilosa isomerasa” como se usa en la presente memoria incluye secuencias de nucleótidos que comprenden secuencias reguladoras que permiten la expresión de la secuencia de nucleótidos que codifica la xilosa isomerasa tales como promotores y potenciadores, los cuales pueden estar asociados de forma nativa o no nativa a la secuencia de nucleótidos que codifica la xilosa isomerasa.

45 En un aspecto, la presente invención proporciona una célula de levadura modificada genéticamente que tiene un gen de la xilosa isomerasa exógeno funcional derivado de un microorganismo de la familia *Lactococcus*, en el que el gen de la xilosa isomerasa exógeno se une operativamente a secuencias de promotor y terminador que son funcionales en la levadura.

50 En otro aspecto, la presente invención proporciona un proceso de fermentación en el cual un microorganismo como se describe en la presente memoria se cultiva en condiciones de fermentación en un caldo de fermentación que incluye el material derivado del material lignocelulósico.

55 En aún otro aspecto, la presente invención proporciona un proceso de fermentación utilizando un microorganismo de acuerdo con la presente invención en el que el caldo de fermentación (es decir, el medio de cultivo) comprende un azúcar pentosa.

Microorganismo transformado

60 Como se ha mencionado en la presente memoria, la expresión “microorganismo transformado” se refiere a un microorganismo que ha sido genéticamente alterado mediante tecnología de ADN recombinante. El término “transformado” como se utiliza en la presente memoria es sinónimo de términos tales como “transfectado”, “recombinante”, “manipulado genéticamente” y “modificado genéticamente”.

65 La expresión “microorganismo transformado” en relación con la presente invención incluye cualquier microorganismo que comprende un vector(es) de expresión que comprende la secuencia(s) de nucleótidos mencionada en la presente memoria y/o un promotor(es) que es capaz de permitir la expresión (en particular la sobreexpresión, es decir, la regulación positiva) de la secuencia(s) de nucleótidos mencionada en la presente memoria. En una

realización, la secuencia(s) de nucleótidos se incorpora en el genoma del microorganismo. En otra realización, el promotor se incorpora en el genoma del microorganismo. Estas características permiten que el microorganismo transformado (en comparación con un microorganismo equivalente antes de la transformación) tenga (a) una actividad xilosa isomerasa superior y/o (b) una tasa de crecimiento superior en o sobre un medio de crecimiento que comprende xilosa y/o (c) un metabolismo de la xilosa más rápido y/o (d) una producción de etanol superior cuando se cultiva en condiciones anaerobias en xilosa como la fuente de carbono.

La expresión "microorganismo transformado" no abarca secuencias codificadoras de nucleótidos nativas en su medio natural cuando están bajo el control de su promotor nativo, el cual también está en su entorno natural.

Por lo tanto, el microorganismo transformado de la presente invención incluye un microorganismo que comprende una cualquiera de, o combinaciones de, las secuencias de nucleótidos que codifican las enzimas mencionadas en la presente memoria, construcciones que comprenden dichas secuencias de nucleótidos, vectores que comprenden dichas secuencias de nucleótidos, plásmidos que comprenden dichas secuencias de nucleótidos y vectores de expresión que comprenden dichas secuencias de nucleótidos.

Por lo tanto, una realización adicional de la presente invención proporciona microorganismos transformados o transfectados con una secuencia(s) de nucleótidos que expresa la enzima(s) mencionada en la presente memoria. El microorganismo se elige de forma que sea compatible con el vector y puede ser, por ejemplo, células bacterianas, fúngicas o de levaduras.

Ejemplos de organismos hospedadores bacterianos adecuados son las especies bacterianas grampositivas o gramnegativas.

Dependiendo de la naturaleza de la secuencia(s) de nucleótidos que codifica la enzima(s) mencionada en la presente memoria, pueden preferirse hospedadores eucariotas, tales como levaduras u otros hongos. En general, se prefieren células de levadura a las células fúngicas porque son más fáciles de manipular.

El uso de microorganismos adecuados, tales como células hospedadoras de levaduras y hongos, puede proporcionar modificaciones post-traduccionales (por ejemplo, miristoilación, glicosilación, truncamiento, lipidación y fosforilación de tirosina, serina o treonina) según se requiera para conferir actividad biológica óptima a los productos de expresión recombinantes mencionados en la presente memoria.

Los microorganismos adecuados incluyen bacterias, hongos y levaduras. Preferiblemente, el microorganismo es una levadura.

Preferiblemente, dicho microorganismo transformado es una levadura transformada. Preferiblemente, dicha levadura transformada se deriva del género *Saccharomyces*. Más preferiblemente dicha levadura transformada es *Saccharomyces cerevisiae*.

En una realización, el microorganismo transformado descrito en la presente memoria es capaz de una actividad xilosa isomerasa superior que el microorganismo equivalente antes de la transformación.

En otro aspecto, el microorganismo transformado descrito en la presente memoria es capaz de una tasa de crecimiento superior en o sobre un medio de crecimiento que comprende xilosa que el microorganismo equivalente antes de la transformación.

En un otro aspecto adicional, el microorganismo transformado descrito en la presente memoria es capaz de un metabolismo de la xilosa más rápido que el microorganismo equivalente antes de la transformación.

En otro aspecto, el microorganismo transformado descrito en la presente memoria es capaz de una producción de etanol superior cuando se cultiva en condiciones anaerobias en xilosa como la fuente de carbono que el microorganismo equivalente antes de la transformación.

El microorganismo se transforma usando técnicas que son de rutina en la técnica tales como electroporación (Sambrook et al 1989). Además, la presencia de una secuencia en un microorganismo transformado puede ser determinada por la selección de crecimiento en un medio adecuado, el cual selecciona para el crecimiento del microorganismo transformado. Como alternativa, o además, la presencia de, secuencias de ADN heterólogo insertado puede ser determinada por PCR directa en la colonia usando cebadores diseñados específicamente para la secuencia insertada. Tales técnicas son bien conocidas y de rutina en la técnica (véase, por ejemplo, Sambrook et al 1989 y Ausubel et al 1995).

Los microorganismos transformados de acuerdo con la presente invención se pueden usar en combinación con uno o más microorganismos. Por ejemplo, uno o más microorganismos transformados de acuerdo con la presente invención pueden cultivarse en combinación con al menos un microorganismo capaz de producir, en ciertas condiciones de cultivo, uno o más componentes seleccionados de la lista que consiste en: etanol, aminoácidos

aromáticos, ácido láctico, ácido succínico, ácido acético, acetaldehído, furfural, ácido itacónico, ácido glutámico, ácido cítrico, cresol, lisina, ácido 3-hidroxipropiónico, poli-3-hidroxialcanoatos, ácido protocatéquico, pirocatecol, guayacol, veratrol, resveratrol, vainillina, ácido vainílico, alcohol vainílico, ácido mucónico, ácido adípico, ácido 4-hidroxibenzoico, 4-hidroxibenzaldehído, ácido 4-metoxibenzoico, 4-aminobenzoato, 4-hidroxianilina, 4-metoxianilina, quinol, anisol, fenol, ácido antranílico, 3-hidroxiantranilato, ácido 2,3-dihidroxibenzoico, 2-aminofenol, 1,4-ciclohexanodiona, isopreno y estireno.

En un aspecto adicional, se proporciona una combinación de (i) uno o más microorganismos transformadas de acuerdo con la presente invención y (ii) al menos otro microorganismo capaz además de producir, en ciertas condiciones de cultivo, uno o más componentes seleccionados de la lista que consiste en: etanol, aminoácidos aromáticos, ácido láctico, ácido succínico, ácido acético, acetaldehído, furfural, ácido itacónico, ácido glutámico, ácido cítrico, cresol, lisina, ácido 3-hidroxipropiónico, poli-3-hidroxialcanoatos, ácido protocatéquico, pirocatecol, guayacol, veratrol, resveratrol, vainillina, ácido vainílico, alcohol vainílico, ácido mucónico, ácido adípico, ácido 4-hidroxibenzoico, 4-hidroxibenzaldehído, ácido 4-metoxibenzoico, 4-aminobenzoato, 4-hidroxianilina, 4-metoxianilina, quinol, anisol, fenol, ácido antranílico, 3-hidroxiantranilato, ácido 2,3-dihidroxibenzoico, 2-aminofenol, 1,4-ciclohexanodiona, isopreno y estireno.

Adicionalmente, la presente invención proporciona un inóculo que comprende una combinación de un microorganismo transformado de acuerdo con la presente invención y uno o más microorganismos.

Además, se proporciona un medio de cultivo que comprende una combinación de un microorganismo transformado de acuerdo con la presente invención y uno o más microorganismos.

Además, la presente invención proporciona un kit que comprende un inóculo que comprende uno o más microorganismos de acuerdo con la presente invención.

Además, la presente invención proporciona un kit que comprende (i) un inóculo que comprende uno o más microorganismos de acuerdo con la presente invención y (ii) un inóculo que comprende uno o más microorganismos.

LEVADURA TRANSFORMADA

En una realización preferida, el microorganismo es una levadura transgénica.

Una revisión de los principios de la expresión génica heteróloga en levaduras se proporcionan en, por ejemplo, *Methods Mol Biol* (1995), 49: 341-54 y *Curr Opin Biotechnol* (1997) Oct; 8(5):554-60

A este respecto, la levadura, tal como la especie *Saccharomyces cerevisiae* o *Pichia pastoris* (ver *FEMS Microbiol Rev* (2000 24(1):45-66)), se puede utilizar como un vehículo para la expresión génica heteróloga.

Una revisión de los principios de la expresión génica heteróloga en *Saccharomyces cerevisiae* y la secreción de productos génicos viene dada por E Hinchcliffe E Kenny (1993, "Yeast as a vehicle for the expression of heterologous genes", *Yeasts*, Vol 5, Anthony H Rose y J Stuart Harrison, eds, segunda edición, Academic Press Ltd.).

Para la transformación de la levadura, se han desarrollado varios protocolos de transformación. Por ejemplo, una *Saccharomyces* transgénica de acuerdo con la presente invención se puede preparar siguiendo las enseñanzas de Hinnen et al., (1978, *Proceedings of the National Academy of Sciences of the USA* 75, 1929); Beggs, J D (1978, *Nature*, Londres, 275, 104) e Ito, H et al (1983, *J Bacteriology* 153, 163-168).

Las células de levadura transformadas pueden seleccionarse utilizando varios marcadores selectivos, tales como marcadores auxotróficos y marcadores de resistencia a antibióticos dominantes.

Vector de expresión

La expresión "vector de expresión" significa una construcción capaz de expresión *in vivo* o *in vitro*.

En un aspecto, el vector de expresión se incorpora en el genoma de un microorganismo adecuado. El término "incorporado" abarca preferiblemente la incorporación estable en el genoma.

Las secuencias de nucleótidos mencionadas en la presente memoria pueden estar presentes en un vector en el cual la secuencia de nucleótidos está unida operativamente a las secuencias reguladoras capaces de proporcionar la expresión de la secuencia de nucleótidos por un microorganismo hospedador adecuado.

Los vectores se transforman en un microorganismo hospedador adecuado tal como se describe en la presente memoria.

La elección del vector, por ejemplo, un plásmido, cósmido o vector de fago con frecuencia dependerá del microorganismo en el que se va a introducir.

5 Los vectores para uso en la presente memoria pueden contener una o más secuencias de nucleótidos de marcador seleccionables, tal como una secuencia de nucleótidos que confiere resistencia a antibióticos, por ejemplo, resistencia a ampicilina, kanamicina, cloranfenicol o tetraciclina. Como alternativa, la selección puede realizarse por cotransformación (como se describe en el documento WO91/17243).

10 Los vectores pueden ser utilizados *in vitro*, por ejemplo, para transfectar, transformar, transducir o infectar un microorganismo hospedador.

15 El vector puede comprender además una secuencia de nucleótidos que permite al vector replicarse en el microorganismo hospedador en cuestión. Ejemplos de tales secuencias son los orígenes de replicación de los plásmidos pUC19, pACYC177, pUB110, pE194, pAMB1 y pIJ702.

En un aspecto preferido, el microorganismo capaz de convertir la xilosa en xilulosa como se ha mencionado en la presente memoria comprende una secuencia de nucleótidos que codifica la xilosa isomerasa.

20 Preferiblemente un vector de expresión como se ha mencionado en la presente memoria comprende una secuencia de nucleótidos que codifica la xilosa isomerasa.

En un aspecto adicional, preferiblemente el microorganismo capaz de convertir la xilosa en xilulosa como se ha mencionado en la presente memoria comprende al menos un vector de expresión que codifica la xilosa isomerasa.

25 Preferiblemente, en otro aspecto, el microorganismo capaz de convertir la xilosa en xilulosa como se ha mencionado en la presente memoria puede comprender además al menos un vector de expresión que codifica una o más enzimas seleccionadas del grupo que consiste en xiluloquinasa, D-ribuloquinasa, ribosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa, transaldolasa, transcetolasa y cualquier otra enzima de la ruta de las pentosas fosfato. Más preferiblemente, dicho microorganismo capaz de convertir una aldopentosa en una cetopentosa como se ha mencionado en la presente memoria comprende además al menos un vector de expresión que codifica la xiluloquinasa.

35 En un aspecto, un vector de expresión como se ha mencionado en la presente memoria, puede codificar además una o más enzimas seleccionadas del grupo que consiste en una aldosa-1-epimerasa, xilosa reductasa, D-xilulosa reductasa, arabinosa reductasa, L-arabitol-4-deshidrogenasa, L-xilulosa reductasa, L-arabinosa isomerasa, ribuloquinasa, ribulosa fosfato-4-epimerasa, D-lixosa isomerasa, D-ribosa isomerasa, xiluloquinasa, D-ribuloquinasa, ribulosa-5-fosfato epimerasa, ribosa-5-fosfato isomerasa, transaldolasa y transcetolasa.

40 En un aspecto, un vector de expresión como se ha mencionado en la presente memoria, puede codificar además uno o más enzimas seleccionadas del grupo que consiste en aldosa-1-epimerasa, xiluloquinasa, D-ribuloquinasa, ribosa-5-fosfato isomerasa, D-ribulosa-5-fosfato epimerasa, transaldolasa, transcetolasa y cualquier otra enzima de la ruta de las pentosas fosfato. Preferiblemente, dicho vector de expresión como se ha mencionado en la presente memoria además codifica xiluloquinasa.

45 En un aspecto preferido, el microorganismo capaz de convertir la xilosa en xilulosa como se ha mencionado en la presente memoria comprende además al menos un vector de expresión que codifica una aldosa-1-epimerasa.

50 Preferiblemente un vector de expresión como se ha mencionado en la presente memoria comprende además una secuencia de nucleótidos que codifica una aldosa-1-epimerasa.

Secuencias reguladoras

55 En algunas aplicaciones, la secuencia(s) de nucleótidos mencionadas en la presente memoria está unida operativamente a una secuencia reguladora que es capaz de proporcionar la expresión de la secuencia de nucleótidos, tal como por el microorganismo elegido. A modo de ejemplo, la presente invención abarca el uso de un vector que comprende la secuencia(s) de nucleótidos mencionada en la presente memoria unida operativamente a una secuencia reguladora de este tipo, es decir, el vector es un vector de expresión.

60 La expresión "unida operativamente" se refiere a una yuxtaposición en la que los componentes descritos están en una relación que les permite funcionar de la manera pretendida. Una secuencia reguladora "operativamente unida" a una secuencia codificante se liga de tal manera que la expresión de la secuencia codificante se consigue en condiciones compatibles con las secuencias de control.

65 La expresión "secuencias reguladoras" incluye promotores y potenciadores y otras señales de regulación de la expresión.

En una realización, la secuencia reguladora permite a la secuencia(s) de nucleótidos operativamente unida a la secuencia reguladora expresarse constitutivamente en una célula hospedadora.

5 Tal como se utiliza en la presente memoria, las expresiones “expresarse constitutivamente” y “expresión constitutiva” se refieren a la transcripción continua de una secuencia de nucleótidos unida operativamente a la secuencia reguladora (tal como, un promotor constitutivo). Así, por ejemplo, el medio de cultivo no necesita comprender un sustrato con el fin de activar la expresión de la secuencia de nucleótidos que codifica la xilosa isomerasa. En otro ejemplo, el medio de cultivo no podría comprender un sustrato que inhibe la secuencia reguladora y, por lo tanto, inhibe la expresión de la secuencia de nucleótidos que codifica la xilosa isomerasa.

10 El término “promotor” se utiliza en el sentido normal de la técnica, por ejemplo, un sitio de unión a la ARN polimerasa.

15 La expresión potenciada de la secuencia(s) de nucleótidos que codifica la enzima(s) mencionada en la presente memoria también se puede conseguir mediante la selección de regiones reguladoras heterólogas, por ejemplo, las regiones del promotor, del líder de secreción y de terminación.

20 Preferiblemente, la secuencia(s) de nucleótidos mencionada en la presente memoria está unida operativamente a al menos un promotor.

Pueden utilizarse incluso otros promotores para dirigir la expresión del polipéptido(s) mencionado en la presente memoria.

25 Ejemplos de promotores adecuados para dirigir la transcripción de la secuencia de nucleótidos en una célula bacteriana, fúngica o de levadura son bien conocidos en la técnica.

30 El promotor puede incluir adicionalmente características para asegurar o aumentar la expresión en un hospedador adecuado. Por ejemplo, las características pueden ser regiones conservadas tales como una caja Pribnow o una caja TATA.

En una realización, el promotor permite a la secuencia(s) de nucleótidos operativamente unida al promotor expresarse de forma constitutiva en una célula hospedadora.

35 Construcciones

El término “construcción”, que es sinónimo de términos tales como “conjugado”, “casete” e “híbrido” incluye una secuencia de nucleótidos mencionada en la presente memoria, directa o indirectamente unida a un promotor.

40 Un ejemplo de una unión indirecta es la provisión de un grupo espaciador adecuado, tal como una secuencia de intrón, tal como el intrón Sh1 o el intrón ADH, intermedia entre el promotor y la secuencia(s) de nucleótidos mencionada en la presente memoria. Lo mismo se aplica al término “fusionado” en relación con la presente invención que incluye la unión directa o indirecta. En algunos casos, los términos no abarcan la combinación natural de la secuencia de nucleótidos que codifica para la proteína asociada normalmente con el promotor del gen de tipo silvestre y cuando ambos están en su entorno natural.

45 La construcción puede incluso contener o expresar un marcador, que permite la selección de la construcción genética.

50 Para algunas aplicaciones, preferiblemente la construcción comprende al menos la secuencia(s) de nucleótidos mencionada en la presente memoria unida operativamente a un promotor.

Promotores

55 Como se ha mencionado en la presente memoria, en un aspecto la presente invención se refiere a un microorganismo que ha sido transformada con una secuencia de nucleótidos, tal como un promotor, que hace que el microorganismo sobreexpresa la xilosa isomerasa.

60 Por ejemplo, el promotor se inserta en el genoma de un microorganismo que permite que el microorganismo sobreexpresa (por ejemplo, por regulación positiva) una secuencia de nucleótidos que codifica la xilosa isomerasa endógena.

En una realización, el promotor se inserta en el genoma de un microorganismo que permite que el microorganismo exprese constitutivamente una secuencia de nucleótidos que codifica la xilosa isomerasa endógena.

65 En otro aspecto, el promotor está unido operativamente a una secuencia de nucleótidos en, por ejemplo, un vector de expresión.

En una realización, el promotor permite que la secuencia(s) de nucleótidos operativamente unida al promotor se exprese de forma constitutiva en una célula hospedadora.

En otro aspecto, el promotor no está reprimido por la presencia de glucosa.

5 Ejemplos de promotores adecuados que podrían utilizarse en los microorganismos de acuerdo con la presente invención, tales como *Saccharomyces cerevisiae*, incluyen: el promotor del gen gliceraldehído-3-fosfato deshidrogenasa (GPD); el promotor del gen de la alcohol deshidrogenasa (ADH) y el promotor del gen del factor embrionario tirotrópico (TEF).

10 Los promotores preferidos que pueden ser usados para sobreexpresar la xilosa isomerasa pueden ser cualquiera de los elementos reguladores que controlan la expresión de secuencias de nucleótidos que codifican proteínas implicadas en la glicolisis y la fermentación de la glucosa en la levadura, en particular *Saccharomyces* tales como *S. cerevisiae*. Ejemplos de estos son el promotor de la glucoquinasa (GLK1), el promotor de la fosfoglucoasa isomerasa (PGI1), el promotor de la fosfofructoquinasa (PFK1) y el promotor de la gliceraldehído-3-fosfato deshidrogenasa (TDH3).

Biocombustible

20 Tal como se utiliza en la presente memoria, el término "biocombustible" se refiere a un combustible (por ejemplo, un combustible líquido) adecuado para su uso (por ejemplo) en motores de combustión. Dicho biocombustible se deriva de materia biológica que comprende azúcares pentosa y/o de azúcares pentosa que se pueden derivar por hidrólisis por medios enzimáticos y/o por tratamiento ácido. Preferiblemente, dicho azúcar pentosa es la xilosa aldopentosa.

25 Los materiales vegetales, incluyendo los residuos de plantas que comprenden material lignocelulósico (por ejemplo: paja de cereales, tales como paja de trigo, pulpa de remolacha, bagazo de caña de azúcar; rastrojo de sorgo; rastrojo de soja; rastrojo de maíz; astillas de madera y pasta de papel) y plantas enteras (como las que se cultivan con fines energéticos, por ejemplo, pasto varilla) son fuentes adecuadas para los azúcares pentosa, en particular azúcares aldopentosa (tales como xilosa), para la presente invención. Otras fuentes adecuadas de material vegetal incluyen productos que no son residuos (en otras palabras, fuentes de alimentos y piensos) como el extracto de caña de azúcar, extracto de remolacha azucarera, sorgo, soja, almidón de trigo y almidón de maíz.

Preferiblemente, el biocombustible mencionado en la presente memoria comprende al menos un alcohol.

35 En un aspecto preferido, el alcohol se selecciona del grupo que consiste en metanol, etanol, propanol y butanol. Más preferiblemente, el biocombustible comprende etanol.

40 Preferiblemente, dicho biocombustible se obtiene (o se puede obtener), en otras palabras, se extrae (o se puede extraer) del medio de cultivo en el que uno o más microorganismos transformados de acuerdo con la presente invención se han cultivado en condiciones adecuadas. Dicho biocombustible se obtiene (o se puede obtener) del medio de cultivo utilizando técnicas que son de rutina en la técnica tales como la eliminación de microorganismo por centrifugación, el aislamiento del sobrenadante, seguido por destilación y una etapa adicional para producir un alcohol 99,5 % puro tal como etanol.

45 El biocombustible puede comprender uno o más componentes de biocombustible adicionales tales como butanol.

El uno o más componentes adicionales del biocombustible se pueden mezclar con el biocombustible antes y/o después de que se obtiene o extrae (obtenible o extraíble) el biocombustible de un cultivo.

50 Como alternativa, o además, uno o más componentes adicionales del biocombustible se pueden producir por cultivo de un microorganismo en un medio de cultivo antes y/o después y/o al mismo tiempo que un microorganismo transformado de acuerdo con la presente invención es/se ha cultivado en un medio de cultivo con el fin de producir el biocombustible.

55 En la presente memoria se describe un transporte de combustible que comprende un biocombustible producido utilizando los microorganismos de acuerdo con la presente invención.

El etanol utilizado como combustible de transporte puede servir a dos propósitos diferentes:

- 60 (i) puede actuar como un aditivo oxigenado que aumenta el índice de octano y reduce las emisiones en la gasolina reformulada (GR) (sustituto del tetraetilo de plomo o del MTBE);
- (ii) puede actuar como un sustituto parcial o total de la gasolina común (GC) para reducir la dependencia del suministro de gasolina.

65 El etanol anhidro tiene un índice de octano de 130 y se puede añadir en concentraciones de 5-10 % (dependiendo de la temporada) a la gasolina común obtenida directamente de las refinerías. Tradicionalmente, el tetraetilo de plomo se ha utilizado para aumentar el octanaje, sin embargo, debido a problemas de salud, el uso del plomo se ha

prohibido en casi todo el mundo. La adición de compuestos oxigenados a la gasolina común reduce las emisiones de monóxido de carbono, así como de otras partículas que contribuyen a la contaminación atmosférica. El metil terc-butil éter (MTBE) se utilizó inicialmente como un aditivo oxigenado, sin embargo, la ocurrencia de contaminación de MTBE en los acuíferos de agua potable ha llevado a algunos estados a prohibir el uso de este compuesto oxigenado. El etanol se utiliza cada vez más en todo el mundo como un sustituto del MTBE como aditivo oxigenado para la fabricación de GR.

Además de servir como un aditivo oxigenado en la producción de la gasolina reformulada, el etanol puede ser utilizado como un sustituto general para la gasolina común. Los coches pueden utilizar mezclas E10 (10 % etanol añadido) sin ninguna modificación del motor.

Además, los vehículos que se han fabricado pueden funcionar con etanol al 100 %, en otras palabras, no hay ninguna necesidad de un combustible fósil.

Los microorganismos transformados de acuerdo con la presente invención o de los microorganismos preparados por un método de acuerdo con la presente invención son capaces de producir un biocombustible a un ritmo mayor que el microorganismo equivalente antes de la transformación. Tal como se utiliza en la presente memoria, la expresión "ritmo mayor" se refiere a un microorganismo transformado que es capaz de producir en el medio de cultivo al menos 5 %, 10 %, 20 %, 25 %, 30 % o 35 % más de biocombustible (como bioetanol) por célula que el microorganismo equivalente antes de la transformación cuando se cultiva en las mismas condiciones de cultivo durante un determinado período de tiempo dentro de la fase de crecimiento exponencial.

Producto derivado de la xilosa

Tal como se utiliza en la presente memoria, la expresión "producto derivado de la xilosa" se refiere a cualquier compuesto derivado de la xilosa.

Ejemplos de productos derivados de la xilosa incluyen, pero no se limitan a: etanol, aminoácidos aromáticos, ácido láctico, ácido succínico, ácido acético, acetaldehído, furfural, ácido itacónico, ácido glutámico, ácido cítrico, cresol, lisina, ácido 3-hidroxi propiónico, poli-3-hidroxi alcanoatos, ácido protocatéutico, pirocatecol, guayacol, veratrol, resveratrol, vainillina, ácido vainílico, alcohol vainílico, ácido mucónico, ácido adípico, ácido 4-hidroxibenzoico, 4-hidroxibenzaldehído, ácido 4-metoxibenzoico, 4-aminobenzoato, 4-hidroxianilina, 4-metoxianilina, quinol, anisol, fenol, ácido antranílico, 3-hidroxiantranilato, ácido 2,3-dihidroxibenzoico, 2-aminofenol, 1,4-ciclohexanodiona, isopreno y estireno.

Un producto derivado de la xilosa puede convertirse en otro producto. Por ejemplo, la cetopentosa D-xilulosa, derivada de la aldopentosa D-xilosa, se puede convertir a través de la ruta de las pentosas fosfato en etanol.

Preferiblemente, dicho producto derivado de la xilosa es uno o más seleccionado del grupo que consiste en etanol, aminoácidos aromáticos, ácido láctico, ácido succínico, ácido acético, acetaldehído, furfural, ácido itacónico, ácido glutámico, ácido cítrico, cresol, lisina, ácido 3-hidroxi propiónico, poli-3-hidroxi alcanoatos, ácido protocatéutico, pirocatecol, guayacol, veratrol, resveratrol, vainillina, ácido vainílico, alcohol vainílico, ácido mucónico, ácido adípico, ácido 4-hidroxibenzoico, 4-hidroxibenzaldehído, ácido 4-metoxibenzoico, 4-aminobenzoato, 4-hidroxianilina, 4-metoxianilina, quinol, anisol, fenol, ácido antranílico, 3-hidroxiantranilato, ácido 2,3-dihidroxibenzoico, 2-aminofenol, 1,4-ciclohexandiona, isopreno y estireno.

Más preferiblemente, dicho producto derivado de la xilosa es etanol.

Medio de cultivo

En una realización, el medio de cultivo comprende xilosa y/o una fuente de xilosa.

Además, el medio de cultivo puede comprender al menos otra pentosa. En un aspecto preferido, el medio de cultivo comprende además al menos un aldopentosa.

Preferiblemente, los microorganismos transformados se cultivan en el medio de cultivo óptimo para dicho microorganismo. Se puede determinar el uso de técnicas de rutina, el medio de cultivo óptimo; además, se pueden determinar las condiciones de crecimiento óptimas.

En una realización, el medio de cultivo comprende uno o más sustratos capaces de activar la expresión de una secuencia de nucleótidos que codifica la xilosa isomerasa en un microorganismo de acuerdo con la presente invención, pero dicho medio de cultivo no comprende cualquier sustrato que es capaz de inhibir la expresión de una secuencia de nucleótidos que codifica la xilosa isomerasa en un microorganismo de acuerdo con la presente invención. Por ejemplo: si la secuencia de nucleótidos que codifica la xilosa isomerasa comprende un promotor inducible de xilosa, entonces el medio comprende xilosa.

En un aspecto dicho medio de cultivo comprende aproximadamente 1 %, aproximadamente 2 %, aproximadamente 4 %, aproximadamente 8 %, aproximadamente 15 % o aproximadamente 25 % de xilosa antes de la inoculación con el microorganismo (es decir, en el tiempo cero).

- 5 Preferiblemente, dicho medio de cultivo comprende las cantidades óptimas de sales, vitaminas y otros nutrientes necesarios para el microorganismo.

10 Los microorganismos se cultivan preferiblemente a su temperatura de crecimiento óptima. El experto hubiera podido determinar fácilmente la temperatura óptima a la que cultivar los microorganismos mencionados en la presente memoria.

En una realización, los microorganismos se cultivan a aproximadamente 20 °C, 25 °C, 30 °C, 35 °C o 37 °C.

- 15 En una realización, los microorganismos se cultivan a aproximadamente 35 °C a 39 °C, preferiblemente aproximadamente 36 °C a 38 °C, más preferiblemente a aproximadamente 35,5 °C a 37,5 °C.

20 En una realización, los microorganismos se cultivan durante aproximadamente 3 a 96 horas; preferiblemente aproximadamente 3 a 48 horas, aproximadamente 3 a 24 horas, aproximadamente 3 a 15 horas y aproximadamente 3 a 6 horas.

Preferiblemente, los microorganismos se cultivan durante aproximadamente 3 horas, aproximadamente 6 horas, aproximadamente 15 horas, aproximadamente 24 horas, aproximadamente 48 horas o aproximadamente 96 horas.

- 25 En un aspecto, el microorganismo, en particular, el microorganismo transformado, es tolerante al alcohol y/o tolerante al ácido.

La expresión “tolerante al alcohol” en relación con la presente invención se refiere a microorganismos que son capaces de crecer en un medio de cultivo que comprende al menos 2 %, 5 %, 10 % o 15 % de alcohol.

- 30 Como se ha mencionado en la presente memoria, la expresión “tolerante al ácido” se refiere a microorganismos que son capaces de crecer en un medio de cultivo que tiene un pH igual a o menor que 6,5, 6,0, 5,0, 4,0 o 3,0.

35 En un aspecto preferido, el medio de cultivo se inocula con al menos 5×10^7 a 5×10^{11} células por kg de medio de cultivo, preferiblemente 5×10^8 a 5×10^{10} células por kg de medio de cultivo, preferiblemente 1×10^9 a 1×10^{10} células por kg de medio de cultivo y más preferiblemente aproximadamente 5×10^9 células por kg de medio de cultivo.

El término “inóculo” y la expresión “cultivo iniciador” son intercambiables.

- 40 Las condiciones de cultivo permiten, por lo menos, el mantenimiento del microorganismo de acuerdo con la presente invención o el microorganismo preparado por un método de acuerdo con la presente invención. Las condiciones de cultivo pueden permitir opcionalmente el crecimiento del microorganismo de acuerdo con la presente invención o el microorganismo preparado por un método de acuerdo con la presente invención.

- 45 Fuentes de xilosa

50 La xilosa es una aldopentosa. La xilosa puede derivarse de: materiales vegetales habitualmente utilizados como fuentes de alimentación humana o animal (tales como: caña de azúcar, remolacha azucarera, sorgo, trigo y maíz, que son materiales vegetales en almidón y ricos en azúcar); plantas enteras (como las que se cultivan con fines energéticos, por ejemplo, pasto varilla) y, en particular, los materiales (vegetales) de desecho agrícola (por ejemplo: paja de cereales, por ejemplo, paja de trigo, pulpa de remolacha, bagazo, por ejemplo, bagazo de caña, rastrojos, por ejemplo, rastrojos de sorgo, soja o maíz y astillas de madera).

55 Fuentes de xilosa para el medio de cultivo descrito en la presente memoria incluyen materiales lignocelulósicos considerados normalmente como material de desecho agrícola. Los troncos, los tallos y las hojas contienen material lignocelulósico y son, por lo tanto, fuentes de material lignocelulósico. El bagazo de caña de azúcar, los rastrojos de maíz y las astillas de madera (hemicelulosa solamente) son tres fuentes de material lignocelulósico de fácil acceso, ya que éstas se recogen y almacenan en grandes cantidades, por diversas razones.

- 60 El material lignocelulósico consiste principalmente en cadenas de azúcar largas. Como término medio, dos tercios de estos azúcares son azúcares hexosa, los cuales están principalmente presentes en la celulosa y un tercio de los azúcares son azúcares pentosa presentes principalmente en polímeros de arabinoxilano.

Una cantidad significativa de azúcar pentosa derivado de la hemicelulosa es xilosa.

65

Los materiales lignocelulósicos se pueden hidrolizar con el fin de liberar los azúcares hexosa y/o pentosa en los azúcares de cadena larga de la celulosa, hemicelulosa y lignina.

5 La hidrólisis de los materiales lignocelulósicos se puede llevar a cabo por tratamiento ácido a temperatura elevada. Sin embargo, este tratamiento puede generar subproductos derivados del azúcar que son tóxicos para la mayoría de los microorganismos y que impiden la conversión de los azúcares en etanol. Estos subproductos tóxicos (si se generan) se pueden eliminar, pero generalmente es poco rentable.

10 Como alternativa, los materiales lignocelulósicos se pueden hidrolizar utilizando enzimas hidrolizantes de celulosa y hemicelulosa. Ventajosamente, este proceso evita la generación de subproductos tóxicos.

15 En un aspecto preferido, el medio de cultivo comprende material derivado de uno o más materiales lignocelulósicos que han sido tratado (ejemplos de tales técnicas de tratamiento incluyen: tratamiento con vapor, explosión de vapor, oxidación húmeda, hidrólisis ácida, oxidación húmeda alcalina y expansión de la fibra amoníaco) para liberar xilosa. Preferiblemente, el material lignocelulósico es tratado mediante un proceso de hidrólisis enzimática. Dicho material lignocelulósico hidrolizado se puede tratar adicionalmente con el fin de extraer los azúcares antes del uso de dicho extracto en un medio de cultivo.

20 *Hidrólisis del material lignocelulósico*

Tratamiento mecánico inicial:

25 El material lignocelulósico se corta en trozos más pequeños como y cuando se considere necesario. Por ejemplo, la paja de trigo se corta en trozos de aproximadamente 5 cm de longitud.

Pretratamiento hidrotérmico posterior:

30 El pretratamiento hidrotérmico del material lignocelulósico puede llevarse a cabo como un tratamiento previo de vapor seguido de una etapa de lavado, produciendo de este modo una fracción de fibra y una fracción líquida. La fracción de fibra contiene más de 90 % de la celulosa, la lignina originalmente presente en el material celulósico y algunas de las hemicelulosas. La fracción líquida contiene azúcares de las hemicelulosas (azúcares C5), más del 90 % de los cloruros alcalinos comprendidos en la biomasa lignocelulósica y la mayoría de los inhibidores de la fermentación derivados del pretratamiento de la materia prima lignocelulósica.

35 Generalmente, la paja de trigo se calienta mediante vapor de agua a una temperatura entre 180 y 200 °C con un tiempo de residencia de 5 a 15 min. La biomasa pretratada se descarga del reactor a presión, se lava y se prensa. El vapor liberado se recoge y se reutiliza para la evaporación de la fracción líquida para alimentar la melaza.

Hidrólisis enzimática:

40 La subsiguiente hidrólisis de los polímeros de azúcar se puede llevar a cabo mediante la adición de celulasas y hemicelulasas, ya sea antes de la fermentación o durante la fermentación o ambas, entre otros, mediante un proceso de sacarificación y fermentación simultáneo.

45 Hexosa

50 Los azúcares hexosa tienen 6 átomos de carbono. Las aldohexosas tienen un aldehído en la posición 1 y las cetohehexosas tienen una cetona en la posición 2. La glucosa es un ejemplo de una aldohexosa. La fructosa es un ejemplo de una cetohehexosa.

Pentosa

55 Los azúcares pentosa tienen 5 átomos de carbono. Las pentosas o bien tienen un grupo funcional aldehído en la posición 1 (aldopentosas) o un grupo funcional cetona en la posición 2 (cetopentosas). Ejemplos de pentosas son xilosa, arabinosa, ribosa, lixosa, xilulosa y ribulosa.

Aldopentosa

60 Ejemplos de aldopentosas son xilosa, arabinosa, ribosa y lixosa.

En un aspecto, dicha aldopentosa preferiblemente es xilosa.

La xilosa puede ser L-xilosa o D-xilosa.

65 Preferiblemente, dicha xilosa es D-xilosa.

Cetopentosa

Ejemplos de cetopentosas son xilulosa y ribulosa.

En un aspecto, dicha cetopentosa es xilulosa o ribulosa.

5

En un aspecto preferido dicha cetopentosa es xilulosa.

La xilulosa puede ser L-xilulosa o D-xilulosa.

10

Preferiblemente, dicha xilulosa es D-xilulosa.

En una realización alternativa, dicha cetopentosa es ribulosa. Más preferiblemente dicha ribulosa es D-ribulosa.

Enzimas

15

Los números de la nomenclatura enzimática (números EC) mencionados en la presente memoria se refieren a las recomendaciones del Comité de Nomenclatura de la Unión Internacional de Bioquímica y Biología Molecular sobre la nomenclatura y clasificación de las enzimas publicadas en 1992 (ISBN 0-12-227165-3).

20

Las enzimas mencionadas en la presente memoria pueden ser producidas por las secuencias de nucleótidos derivadas de una amplia variedad de fuentes. En un aspecto, las secuencias de nucleótidos que codifican las enzimas mencionadas en la presente memoria pueden ser derivadas o derivables de especies de *Lactococcus* (tales como *Lactococcus lactis* y *Lactococcus lactis* subsp. *lactis*), *Geobacillus stearothermophilus*, *Enterococcus faecalis*, *Piromyces* sp, especies de *Thermoanaerobacter* (tales como *Thermoanaerobacter thermohydrosulfuricus* y *Thermoanaerobacter thermosulphurigenes*), *Pichia stipitis* o *Saccharomyces cerevisiae*. Preferiblemente, las secuencias de nucleótidos que codifican los enzimas mencionados en la presente memoria se derivan o son derivables de especies de *Lactococcus* (tales como *Lactococcus lactis* y *Lactococcus lactis* subsp. *lactis*).

25

Xilosa isomerasa (EC 5.3.1.5)

30

La xilosa isomerasa tiene el número de nomenclatura EC EC 5.3.1.5. La xilosa isomerasa puede ser denominada como D-xilosa isomerasa, D-xilosa cetoisomerasa o D-xilosa cetol-isomerasa.

La expresión xilosa isomerasa se refiere a una enzima que es capaz de convertir D-xilosa en D-xilulosa y viceversa.

35

Una xilosa isomerasa mencionada en la presente memoria es capaz de actuar sobre la D-xilosa.

La xilosa isomerasa puede ser derivado de una especie bacteriana, tal como especies de *Lactococcus*. En particular, la xilosa isomerasa puede derivarse de *Lactococcus lactis* subsp. *lactis*; *Lactococcus lactis* subsp. *lactis* cepa NRRL B-449; *Lactococcus lactis* subsp. *lactis* cepa KF147; *Lactococcus lactis* subsp. *lactis* cepa DSM20175 y *Lactococcus lactis* subsp. *lactis* cepa IO-1.

40

Ejemplos de xilosa isomerasas adecuadas para su uso como se describe en la presente memoria incluyen xilosa isomerasas que comprenden: la secuencia de aminoácidos mostrada como SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o la SEQ ID No 20 o una variante, homóloga o derivada de la misma; la secuencia de aminoácidos mostrada como GenBank número de acceso ABX75758; o la secuencia de aminoácidos mostrada como GenBank número de acceso AAD20249.

45

Ejemplos de xilosa isomerasas adecuadas para su uso como se describe en la presente memoria incluyen xilosa isomerasas codificadas por: la secuencia de nucleótidos mostrada como SEC ID No 1, SEQ ID No 10, SEQ ID No 17, SEQ ID No 12, SEQ ID No 27 o la SEQ ID No 28 o una variante, homóloga o derivada de la misma; una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada como SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20, o una variante, homóloga o derivada de la misma.

50

Aldosa-1-epimerasa (EC 5.1.3.3)

Una aldosa-1-epimerasa mencionada en la presente memoria es capaz de actuar sobre la aldopentosa.

La aldosa-1-epimerasa tiene el número de nomenclatura EC 5.1.3.3. La aldosa-1-epimerasa puede ser denominada como un mutarrotasa o una aldosa mutarrotasa.

60

La expresión aldosa-1-epimerasa se refiere a una enzima que es capaz de convertir una α -aldopentosa en una β -aldopentosa y viceversa.

65

Preferiblemente, la aldosa-1-epimerasa es codificada por una secuencia de nucleótidos seleccionada del grupo que consiste en: AAD20257, ABX75760, AAK05605, AAD20245, AAD20251, ABJ73095, ABI49935 y AAO80762

(números de acceso de NCBI). Más preferiblemente, la aldosa-1-epimerasa se selecciona del grupo que consiste en: AAD20257, ABX75760, AAK05605, AAD20245 y AAD20251.

Ejemplos de aldosa-1-epimerasas adecuadas para su uso como se describe en la presente memoria incluyen aldosa-1-epimerasa codificada por: la secuencia de nucleótidos del gen de la aldosa-1-epimerasa de *Lactococcus lactis* (número de acceso del NCBI AAD20245); la secuencia de nucleótidos del gen GAL10 de *Saccharomyces cerevisiae* (en particular, la parte que codifica una secuencia de aminoácidos que tiene actividad mutarrotasa) y la secuencia de nucleótidos del gen GAL10 de la cepa D0002 de *Saccharomyces cerevisiae* (en particular, la parte que codifica una secuencia de aminoácidos que tiene actividad mutarrotasa).

Aldosa reductasa (EC 1.1.1.21)

La aldosa reductasa tiene el número de nomenclatura EC 1.1.1.21. La aldosa reductasa puede ser denominada como: poliol deshidrogenasa, aldehído reductasa, ALR2, NADPH-aldopentosa reductasa, NADPH-aldosa reductasa, alditol:NADP oxidorreductasa o alditol:NADP⁺1-oxidorreductasa.

La expresión aldosa reductasa se refiere a una enzima que es capaz de convertir un alditol en una aldosa y viceversa.

Una aldosa reductasa puede reducir más de un tipo de aldosa. Por ejemplo, la misma enzima puede ser capaz de reducir tanto la D-xilosa como la L-arabinosa, de modo que una enzima así puede ser denominada aldosa reductasa, o, puede ser denominada más específicamente según uno de los sustratos, por ejemplo, xilosa reductasa.

Xilosa reductasa (EC 1.1.1.21)

La aldosa reductasa puede ser una xilosa reductasa. La xilosa reductasa tiene el número de nomenclatura EC 1.1.1.21.

La expresión xilosa reductasa se refiere a una enzima que es capaz de convertir D-xilosa en xilitol y viceversa.

Una xilosa reductasa mencionada en la presente memoria es capaz de actuar sobre la D-xilosa.

Ejemplos de xilosa reductasas adecuadas para su uso como se describe en la presente memoria incluyen xilosa reductasa codificada por: la secuencia de nucleótidos del gen de la xilosa reductasa de *Pichia stipitis* (PsXR); la secuencia de nucleótidos del gen de la xilosa reductasa de *Pichia stipitis* cepa DSM3651 (PsXR) – número de acceso del NCBI X59465; la secuencia de nucleótidos de *Candida tenuis* (dicha secuencia de nucleótidos que codifica la xilosa reductasa se puede obtener como se describe por Kavanagh et al, 2003) y la secuencia de nucleótidos de *Neurospora crassa* (dicha secuencia de nucleótidos que codifica la xilosa reductasa se puede obtener como se describe por Woodyer et al, 2005).

Arabinosa reductasa (EC 1.1.1.21)

La aldosa reductasa puede ser una arabinosa reductasa. La arabinosa reductasa tiene el número de nomenclatura EC 1.1.1.21.

El término arabinosa reductasa se refiere a una enzima que es capaz de convertir L-arabinosa en L-arabitol y viceversa.

Una reductasa arabinosa mencionada en la presente memoria es capaz de actuar sobre la L-arabinosa.

Las D-xilosa reductasas actualmente conocidas en la técnica podrían actuar sobre la L-arabinosa como un sustrato con actividad similar. Por lo tanto, la expresión L-arabinosa reductasa puede referirse también a las enzimas que están clasificadas como D-xilosa reductasas y las xilosa reductasas mencionadas en la presente memoria como adecuadas para introducir el metabolismo de la xilosa son igualmente adecuadas para su uso en la introducción del metabolismo de la arabinosa en un microorganismo.

La aldosa reductasa puede ser capaz de convertir L-lixosa en L-arabitol y viceversa. En otra realización, la aldosa reductasa puede ser capaz de convertir D-lixosa en D-arabitol y viceversa.

La aldosa reductasa puede ser capaz de convertir la ribosa en ribitol (en particular, D-ribosa a D-ribitol) y viceversa.

Xilulosa reductasa (EC 1.1.1.9 y EC 1.1.1.10)

La expresión xilulosa reductasa abarca D-xilulosa reductasa y L-xilulosa reductasa.

D-xilulosa reductasa (EC 1.1.1.9)

La D-xilulosa reductasa tiene el número de nomenclatura EC 1.1.1.9. La D-xilulosa reductasa puede ser denominada como xilitol deshidrogenasa, xilitol-2-deshidrogenasa, 2,3-cis-poliol (DPN) deshidrogenasa (C3-5), xilitol deshidrogenasa dependiente de NAD, eritritol deshidrogenasa o pentitol-DPN deshidrogenasa.

La expresión D-xilulosa reductasa se refiere a una enzima que es capaz de convertir xilitol en D-xilulosa y viceversa.

Una D-xilulosa reductasa mencionada en la presente memoria es capaz de actuar sobre el xilitol.

Ejemplos de D-xilulosa reductasas adecuadas para su uso como se describe en la presente memoria incluyen la D-xilulosa reductasa codificada por: la secuencia de nucleótidos del gen de la D-xilulosa de *Pichia stipitis* (PsXDH); y la secuencia de nucleótidos del gen de la xilulosa reductasa de la cepa DSM3651 de *Pichia stipitis* (PsXDH) - Número de acceso del NCBI X55392.

L-xilulosa reductasa (EC 1.1.1.10)

La expresión L-xilulosa reductasa tiene el número de nomenclatura EC 1.1.1.10. La L-xilulosa reductasa puede ser denominada como L-xilitol deshidrogenasa.

La expresión L-xilulosa reductasa se refiere a una enzima que es capaz de convertir L-xilulosa en xilitol y viceversa.

Una L-xilulosa reductasa mencionada en la presente memoria es capaz de actuar sobre la L-xilulosa.

Una secuencia de nucleótidos que codifica la L-xilulosa reductasa se puede obtener de *Aspergillus niger* como se describe por Witteveen et al (1994) o de la levadura *Ambrosiozyma monospora* (Verho et al, 2004).

Xiluloquinasa (EC 2.7.1.17)

La xiluloquinasa tiene el número de nomenclatura EC 2.7.1.17. La xiluloquinasa puede ser denominada como D-xiluloquinasa.

La expresión xiluloquinasa se refiere a una enzima que es capaz de convertir D-xilulosa en D-xilulosa 5-fosfato y viceversa.

Una xiluloquinasa mencionada en la presente memoria es capaz de actuar sobre la D-xilulosa.

Ejemplos de xiluloquinasas adecuadas para su uso como se describe en la presente memoria incluyen la xiluloquinasa codificada por: la secuencia de nucleótidos del gen de la xiluloquinasa de *Pichia stipitis* (PsXKS); la secuencia de nucleótidos del gen de la xiluloquinasa de la cepa DSM3651 de *Pichia stipitis* (PsXKS) - Número de acceso del NCBI AF127802; la secuencia de nucleótidos del gen de la xiluloquinasa de *S. cerevisiae* (ScXKS); y la secuencia de nucleótidos del gen de la xiluloquinasa de la cepa D0002 de *S. cerevisiae* (ScXKS) - Número de acceso del NCBI X61377.

D-arabinitol 4-deshidrogenasa (EC 1.1.1.11)

La D-arabinitol 4-deshidrogenasa tiene el número de nomenclatura EC 1.1.1.11. La D-arabinitol 4-deshidrogenasa puede ser denominada como D-arabitol deshidrogenasa o arabitol deshidrogenasa.

La expresión D-arabinitol 4-deshidrogenasa se refiere a una enzima que es capaz de convertir D-arabinitol en D-xilulosa y viceversa.

Una D-arabinitol 4-deshidrogenasa mencionada en la presente memoria es capaz de actuar sobre el D-arabinitol.

Una D-arabinitol 4-deshidrogenasa adecuada y el gen correspondiente se describe por Cheng et al 2005.

L-arabinitol 4-deshidrogenasa (EC 1.1.1.12)

La L-arabinitol 4-deshidrogenasa tiene el número de nomenclatura EC 1.1.1.12. La L-arabinitol 4-deshidrogenasa puede ser denominada como L-arabitol 4-deshidrogenasa o pentitol-DPN deshidrogenasa.

La expresión L-arabinitol 4-deshidrogenasa se refiere a una enzima que es capaz de convertir L-arabinitol en L-xilulosa y viceversa.

Una L-arabinitol 4-deshidrogenasa mencionada en la presente memoria es capaz de actuar sobre el L-arabinitol.

Una L-arabinosa 4-deshidrogenasa adecuada y el gen correspondiente se describe en Richard et al (2001).

L-arabinosa isomerasa (EC 5.3.1.4)

5 La L-arabinosa isomerasa tiene el número de nomenclatura EC 5.3.1.4. La L-arabinosa isomerasa puede ser denominada como L-arabinosa cetol-isomerasa.

La expresión L-arabinosa isomerasa se refiere a una enzima que es capaz de convertir L-arabinosa en L-ribulosa y viceversa.

10

Una L-arabinosa isomerasa mencionada en la presente memoria es capaz de actuar sobre la L-arabinosa.

Un ejemplo de una secuencia de nucleótidos que codifica la L-arabinosa isomerasa es la secuencia de nucleótidos que se puede obtener a partir de la cepa NCIMB8826 de *Lactobacillus plantarum* (ATCC 14917) (gen descrito en el código de acceso del NCBI NC_004567).

15

Ribuloquinasa (EC 2.7.1.16)

La ribuloquinasa tiene el número de nomenclatura EC 2.7.1.16. La ribuloquinasa puede ser denominada como L-ribuloquinasa.

20

La expresión ribuloquinasa se refiere a una enzima que es capaz de convertir (i) L-ribulosa en L-ribulosa 5-fosfato y viceversa y/o (ii) D-ribulosa en D-ribulosa 5-fosfato y viceversa.

25 Una ribuloquinasa mencionada en la presente memoria es capaz de actuar sobre la L-ribulosa y/o la D-ribulosa.

Una secuencia de nucleótidos que codifica la ribuloquinasa adecuada se puede obtener de la cepa NCIMB8826 de *Lactobacillus plantarum* (ATCC 14917) (gen descrito en el código de acceso del NCBI NC_004567).

30 D-ribuloquinasa (EC 2.7.1.47)

La D-ribuloquinasa tiene el número de nomenclatura EC 2.7.1.47.

La expresión ribuloquinasa se refiere a una enzima que es capaz de convertir D-ribulosa en D-ribulosa 5-fosfato y viceversa.

35

Un ejemplo de una secuencia de nucleótidos que codifica la D-ribuloquinasa puede obtenerse a partir de *Botryotinia fuckeliana* (código de acceso de GenBank CH476984). El número de acceso de GenBank EDN20859 detalla la secuencia de aminoácidos de una D-ribuloquinasa.

40

L-ribulosa fosfato 4-epimerasa (EC 5.1.3.4)

La L-ribulosa fosfato 4-epimerasa tiene el número de nomenclatura EC 5.1.3.4. La L-ribulosa fosfato de 4-epimerasa puede ser denominada como ribulosa fosfato 4-epimerasa, fosforribulosa isomerasa, L-ribulosa 5-fosfato 4-epimerasa, AraD o L-Ru5P.

45

La expresión L-ribulosa fosfato 4-epimerasa se refiere a una enzima que es capaz de convertir L-ribulosa 5-fosfato en D-xilulosa 5-fosfato y viceversa.

50 Una L-ribulosa fosfato 4-epimerasa mencionada en la presente memoria es capaz de actuar sobre la L-ribulosa 5-fosfato.

Una secuencia de nucleótidos de la L-ribulosa fosfato 4-epimerasa adecuada puede obtenerse a partir la cepa NCIMB8826 de *Lactobacillus plantarum* (ATCC 14917) (gen descrito en el código de acceso del NCBI NC_004567).

55

D-ribitol 4-deshidrogenasa (EC 1.1.1.56)

La D-ribitol 4-deshidrogenasa tiene el número de nomenclatura EC 1.1.1.56. Esta enzima también puede ser denominada como ribitol 2-deshidrogenasa, adonitol deshidrogenasa o ribitol deshidrogenasa.

60

La expresión D-ribitol 4-deshidrogenasa se refiere a una enzima que es capaz de convertir ribitol en D-ribulosa y viceversa.

Una D-ribitol 4-deshidrogenasa mencionada en la presente memoria es capaz de actuar sobre el D-ribitol.

65

Una D-ribitol 4-deshidrogenasa adecuada y el gen correspondiente se describe por Dothie et al, 1985.

D-lixosa isomerasa (EC 5.3.1.15)

La D-lixosa isomerasa tiene el número de nomenclatura EC 5.3.1.15. Esta enzima también puede ser denominada como D-lixosa cetol-isomerasa.

5 La expresión D-lixosa isomerasa se refiere a una enzima que es capaz de convertir D-lixosa en D-xilulosa y viceversa.

La D-lixosa isomerasa mencionada en la presente memoria es capaz de actuar sobre la D-lixosa.

10 Una secuencia de nucleótidos que codifica una D-lixosa/-L-ribosa isomerasa se puede clonar a partir del organismo *Acinetobacter* sp. cepa DL-28 (Shimonishi y Izumori, 1996) o de *Aerobacter aerogenes* (Anderson y Allison, 1965).

Ribosa isomerasa (EC 5.3.1.20)

15 La ribosa isomerasa tiene el número de nomenclatura EC 5.3.1.20. Esta enzima también puede ser denominada como D-ribosa isomerasa o D-ribosa cetol-isomerasa.

20 La expresión ribosa isomerasa se refiere a una enzima que es capaz de convertir la D-ribosa en D-ribulosa y viceversa.

Una ribosa isomerasa mencionada en la presente memoria es capaz de actuar sobre la D-ribosa.

25 La D-ribosa isomerasa se ha encontrado en el organismo *Mycobacterium smegmatis* (Izumori et al, 1975), desde donde puede ser clonado.

Ribulosa-5-fosfato de 3-epimerasa (EC 5.1.3.1)

30 La ribulosa-5-fosfato 3-epimerasa tiene el número de nomenclatura EC 5.1.3.1. Esta enzima también puede ser denominada: pentosa-5-fosfato 3-epimerasa, fosfocetopentosa 3-epimerasa, fosfocetopentosa epimerasa, fosforribulosa epimerasa, ribulosa-fosfato 3-epimerasa; D-ribulosa 5-fosfato epimerasa; D-ribulosa fosfato-3-epimerasa; D-ribulosa-5-P 3-epimerasa; D-xilulosa-5-fosfato 3-epimerasa; eritrosa-4-fosfato isomerasa; ribulosa 5-fosfato 3-epimerasa y xilulosa fosfato 3-epimerasa.

35 La expresión ribulosa-5-fosfato 3-epimerasa se refiere a una enzima que es capaz de convertir D-ribulosa 5-fosfato en D-xilulosa 5-fosfato y viceversa.

40 Ejemplos de ribulosa-5-fosfato 3-epimerasa adecuados para su uso como se describe en la presente memoria incluyen ribulosa-5-fosfato 3-epimerasa codificada por: la secuencia de nucleótidos del gen *RPE1* de *S. cerevisiae*; la secuencia de nucleótidos del gen *RPE1* de la cepa D0002 de *S. cerevisiae*; la secuencia de nucleótidos del código de acceso del NCBI NP_012414 y la ribulosa-5-fosfato isomerasa de *P. stipitis* que se pueden encontrar en número de acceso del NCBI NP_012414.

Ribosa-5-fosfato isomerasa (EC 5.3.1.6)

45 La ribosa-5-fosfato isomerasa tiene el número de nomenclatura EC 5.3.1.6. Esta enzima también puede ser denominada como fosfopentoisomerasa, fosfopentosisomerasa, fosfopentosisomerasa, fosforiboisomerasa, ribosa fosfato isomerasa, 5-fosforribosa isomerasa o D-ribosa-5-fosfato cetol-isomerasa.

50 La expresión ribosa-5-fosfato isomerasa se refiere a una enzima que es capaz de convertir D-ribosa 5-fosfato en D-ribulosa 5-fosfato y viceversa.

55 Ejemplos de ribosa-5-fosfato isomerasa adecuados para su uso como se describe en la presente memoria incluyen ribosa-5-fosfato cetol-isomerasa codificada por: la secuencia de nucleótidos del gen *RK11* de *S. cerevisiae*; la secuencia de nucleótidos del gen *RK11* de *S. cerevisiae* cepa D0002; la secuencia de nucleótidos del código de acceso del NCBI X94335; la secuencia de nucleótidos del código de acceso del NCBI NP_014738 y la ribosa-5-fosfato isomerasa de *P. stipitis* que se puede encontrar en el número de acceso NC_009043.

Transcetolasa (EC 2.2.1.1)

60 La transcetolasa tiene el número de nomenclatura EC 2.2.1.1. Esta enzima también puede ser denominada como glicolaldehído transferasa.

65 El término transcetolasa se refiere a una enzima que es capaz de convertir sedoheptulosa 7-fosfato + D-gliceraldehído 3-fosfato en D-ribosa 5-fosfato + D-xilulosa 5-fosfato y viceversa.

Ejemplos de transcetolasas adecuadas para su uso como se describe en la presente memoria incluyen transcetolasa codificada por: la secuencia de nucleótidos del gen *TKL1* de *Saccharomyces cerevisiae*; la secuencia de nucleótidos del gen *TKL1* de *Saccharomyces cerevisiae* cepa D0002; la secuencia de nucleótidos del código de acceso del NCBI X73224; la secuencia de nucleótidos del código de acceso del NCBI NP_015399 y la transcetolasa de *P. stipitis* que se puede encontrar en el número de acceso CP000496.

Transaldolasa (EC 2.2.1.2)

La transaldolasa tiene el número de nomenclatura EC 2.2.1.2. Esta enzima también puede ser denominada como dihidroxiacetona transferasa, dihidroxiacetona sintasa, formaldehído transcetolasa o glicerona transferasa.

El término transaldolasa se refiere a una enzima que es capaz de convertir sedoheptulosa 7-fosfato + D-gliceraldehído 3-fosfato en D-eritrosa 4-fosfato + D-fructosa 6-fosfato y viceversa.

Ejemplos de transaldolasas adecuadas para su uso como se describe en la presente memoria incluyen transaldolasa codificada por: la secuencia de nucleótidos del gen *TAL1* de *Saccharomyces cerevisiae*; la secuencia de nucleótidos del gen *TAL1* de *Saccharomyces cerevisiae* cepa D0002; la secuencia de nucleótidos del código de acceso del NCBI X15953; la secuencia de nucleótidos del código de acceso del NCBI NP_013458 y la transaldolasa de *P. stipitis* que se puede encontrar en el número de acceso CP000502.

Xilosa isomerasa exógena

En un aspecto, la xilosa isomerasa mencionada en la presente memoria es una xilosa isomerasa exógena.

En un aspecto, la secuencia de nucleótidos que codifica la xilosa isomerasa mencionada en la presente memoria codifica una xilosa isomerasa exógena.

Preferiblemente, la xilosa isomerasa exógena se deriva de un microorganismo mesófilo. Más preferiblemente, la xilosa isomerasa exógena se deriva de una bacteria mesófila.

Tal como se utiliza en la presente memoria, la expresión "microorganismo mesófilo" se refiere a un microorganismo que crece y/o se desarrolla mejor a una temperatura entre 15 °C y 40 °C. Ejemplos de microorganismos mesófilos incluyen especies de *Lactococcus*, especies de *Saccharomyces* (tales como *S. cerevisiae*), especies de *Escherichia* (tales como *E. coli*) y especies de *Bacillus* (tales como *B. subtilis*).

En una realización, la xilosa isomerasa exógena se deriva de una especie de *Lactococcus*. Preferiblemente, la xilosa isomerasa exógena se deriva de una especie de *Lactococcus* capaz de crecer en xilosa. Preferiblemente, la xilosa isomerasa exógena se deriva de un *Lactococcus lactis*. Más preferiblemente, la xilosa isomerasa exógena se deriva de un *Lactococcus lactis* subsp. *lactis*. En una realización muy preferida, la xilosa isomerasa exógena se deriva de *Lactococcus lactis* cepa NRRL B-4449, la cepa DSM 20175, la cepa KF147 o la cepa IO-1.

El microorganismo *Lactococcus lactis* subsp. *lactis* puede denominarse *Lactobacillus xylosus*.

Ensayo de aldopentosa

La cantidad de una aldopentosa (como xilosa) en una solución (tal como un medio de cultivo) puede determinarse colorimétricamente por el método del floroglucinol, como se describe por Ebert et al. (A Simplified, Colorimetric Micromethod for Xylose in Serum or Urine, with Phloroglucinol, 1979, Clin. Chem. 25, no.8, pp. 1440-1443).

El reactivo de color consiste en 0,5 g de floroglucinol (1,3,5 trihidroxibenceno), 100 ml de ácido acético glacial y 10 ml de HCl conc. Se añaden 50 µl de muestra y 950 µl de reactivo de color. La mezcla se calienta a 100 °C durante 4 minutos y la absorbancia de la mezcla se lee a 554 nm. La cantidad de una aldopentosa (como xilosa) en la muestra se determina de acuerdo con una curva patrón hecha con la misma aldopentosa como patrón. Este método también se puede utilizar para determinar la cantidad de arabinosa, lixosa y ribosa en un medio de cultivo.

Ensayo de cetopentosa

La cantidad de una cetopentosa (como xilulosa) en una solución (tal como un medio de cultivo) puede determinarse colorimétricamente por el método de la cisteína-carbazol como se describe por Zacarías Dische and Ellen Borenfreund (1951; J. Biol.Chem.192 (2): 583).

Ensayo de xilosa isomerasa

La cantidad de actividad de xilosa isomerasa en una solución puede ser determinada como sigue.

La conversión de xilosa en xilulosa se determina en una solución que contiene una cantidad de la xilosa isomerasa, xilosa 2 % y $MgCl_2$ 2 mM en tampón TRIS 20 mM, pH 7,4 y se incubaron a la temperatura de elección. La reacción se inicia por la adición de la xilosa y se detiene mediante la colocación de la reacción o de la muestra de la reacción en hielo. El tiempo de reacción es normalmente de 30 o 60 minutos, o se pueden tomar muestras para seguir la reacción en diferentes puntos temporales. La xilulosa generada se cuantificó utilizando el método de cisteína/carbazol (Dische y Borenfreund, 1951). Una unidad de actividad de xilosa isomerasa se define como la cantidad de actividad que cataliza la conversión de 1 nanomol de xilosa en xilulosa por minuto.

En un aspecto muy preferido, el microorganismo transformado tiene una actividad xilosa isomerasa de al menos 0,2 unidades de xilosa isomerasa por mg de proteína de microorganismo. La cantidad de actividad xilosa isomerasa se ensaya como se ha mencionado en la presente memoria y el contenido de proteína se determina como se describe a continuación. Los microorganismos han sido cultivados previamente durante 96 horas o menos en un medio de cultivo que comprende xilosa.

15 Contenido de proteínas

La cantidad de proteína (es decir, el contenido de proteína) en una solución se determinó mediante el uso del kit de ensayo BCA comercializado (Pierce Biotechnology Inc., EE.UU.), basado en el ensayo de ácido bicinconínico (Smith et al, 1985), usando seroalbúmina bovina como patrón.

20 Ensayo de etanol

La cantidad del biocombustible como etanol, en una solución (tal como un medio de cultivo) se puede determinar mediante el uso de un ensayo de etanol comercializado, el Kit de K-ETOH, fabricado y vendido por Megazyme International, Bray Business Park, Bray, Co. Wicklow, Irlanda o se puede determinar mediante, por ejemplo, el uso de cromatografía de gases.

Ruta de las pentosas fosfato

30 Las cetopentosas (tales como xilulosa) se convierten en etanol a través de la ruta de las pentosas fosfato. Un ejemplo de esto se muestra en la Figura 4.

Ejemplos de enzimas implicadas en la ruta de las pentosas fosfato incluyen: transcetolasa, transaldolasa, ribosa-5-fosfato cetol-isomerasa y ribulosa-5-fosfato 3-epimerasa.

35 En una realización, el microorganismo de acuerdo con la presente invención también ha sido transformado para expresar o sobreexpresar una o más enzimas implicadas en la ruta de las pentosas fosfato.

40 En una realización, el microorganismo de acuerdo con la presente invención también ha sido transformado con una o más secuencias de nucleótidos que hacen que el microorganismo sobreexpresa una o más enzimas implicadas en la ruta de las pentosas fosfato. Por ejemplo, se inserta un promotor en el genoma de un microorganismo que permite que el microorganismo sobreexpresa una secuencia de nucleótidos endógena que codifica una enzima implicada en la ruta de las pentosas fosfato.

45 En otra realización, el microorganismo ha sido transformado con una o más secuencias de nucleótidos que codifican una o más enzimas implicadas en la ruta de las pentosas fosfato. Por ejemplo, el microorganismo se transforma con un vector de expresión que comprende una secuencia de nucleótidos que codifica una o más enzimas implicadas en la ruta de las pentosas fosfato.

50 Preferiblemente, el vector de expresión mencionado en la presente memoria comprende uno o más promotores capaces de sobreexpresar una o más secuencias de nucleótidos que codifican una o más enzimas implicadas en la ruta de las pentosas fosfato. Ejemplos de tales promotores incluyen el promotor GPD, el promotor TEF y el promotor ADP. Promotores preferidos que pueden ser usados para sobreexpresar una o más secuencias de nucleótidos que codifican una o más enzimas implicadas en la ruta de las pentosas fosfato puede ser cualquiera de los elementos reguladores que controlan la expresión de secuencias de nucleótidos que codifican proteínas implicadas en la glicolisis y la fermentación de la glucosa.

60 Tal como se utiliza en la presente memoria, el término "sobreexpresa" en la frase "una o más secuencias de nucleótidos que hacen que el microorganismo sobreexpresa una o más enzimas implicadas en la ruta de las pentosas fosfato" y "uno o más promotores capaces de sobreexpresar una o más secuencias de nucleótidos que codifican una o más enzimas implicadas en la ruta de las pentosas fosfato" se refiere a un aumento en la expresión desde cero hasta un nivel de expresión o que va desde un nivel más bajo de expresión hasta un nivel más alto de expresión (por ejemplo, regulación positiva) cuando el microorganismo transformado se compara con el microorganismo equivalente antes de la transformación. Los microorganismos que sobreexpresan una o más enzimas implicadas en la ruta de las pentosas fosfato tienen una mayor capacidad para catalizar la conversión de una cetopentosa (como xilulosa 5-fosfato) en un biocombustible (tal como etanol).

Preferiblemente, dicho microorganismo transformado que sobreexpresa una o más enzimas implicadas en la ruta de las pentosas fosfato es capaz de catalizar la conversión de una cetopentosa (como xilulosa 5-fosfato) en un biocombustible (tal como etanol) a una tasa que es al menos 10 %, 15 %, 20 % o 25 % mayor que un microorganismo no transformado.

5 Ejemplos de microorganismos que sobreexpresan una o más enzimas implicadas en la ruta de las pentosas fosfato incluyen: (i) microorganismos transformados con uno o más vectores de expresión que codifican una o más de transcetolasa, transaldolasa, ribosa-5-fosfato cetol-isomerasa y ribulosa-5-fosfato 3-epimerasa y (ii) microorganismos transformados para regular positivamente la expresión de una o más secuencias de nucleótidos endógenas que
10 codifican una o más de transcetolasa, transaldolasa, ribosa-5-fosfato cetol-isomerasa y ribulosa-5-fosfato 3-epimerasa (antes de la transformación dicho microorganismo era capaz de expresar una o más de estas enzimas para un determinado conjunto de condiciones de cultivo durante el crecimiento exponencial, pero después de la transformación dicho microorganismo es capaz de la expresión de una o más de estas enzimas en un nivel superior, en las mismas condiciones de cultivo, durante el crecimiento exponencial).

15 SECUENCIAS VARIANTES/HOMÓLOGAS/DERIVADAS

La presente invención abarca el uso de secuencias variantes, homólogas y derivadas de cualquier secuencia de aminoácidos de una proteína o de cualquier secuencia de nucleótidos que codifica una proteína tal.

20 En particular, la presente invención abarca el uso de secuencias variantes, homólogas y derivadas de la secuencia de aminoácidos codificada por la SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o la SEQ ID No 20.

25 Además, la presente invención abarca el uso de secuencias variantes, homólogas y derivadas de la secuencia de nucleótidos que codifican la secuencia de aminoácidos mostrada como SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20 y la secuencia de nucleótidos mostrada como SEC ID No 1, SEQ ID No 10, SEQ ID No 17, SEQ ID No 12, SEQ ID No 27 o SEQ ID No 28.

30 Por ejemplo, una secuencia variante, homóloga o derivada de una secuencia de nucleótidos (tal como, SEQ ID No 1, SEQ ID No 10, SEQ ID No 17, SEQ ID No 12, SEQ ID No 27 o SEQ ID No 28, o la secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada como SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20) puede comprender hasta 30, 21, 15, 9, 6 o 3 sustituciones de ácidos nucleicos.

35 Por ejemplo, una secuencia variante, homóloga o derivada de una secuencia de aminoácidos (tal como, SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20) pueden comprender hasta a 10, 7, 5, 3, 2 o 1 sustituciones de aminoácidos.

40 Sin desear estar limitado por la teoría, dos residuos de aminoácidos (concretamente, D296 y D339) de la xilosa isomerasa pueden estar implicados en la unión de iones metálicos divalentes en el sitio activo (Meng, M. et al, 1993). En una realización, D296 y D339 son importantes para la unión de iones metálicos divalentes en el sitio activo. Por lo tanto, en una realización, la variante, homóloga o derivada de una secuencia de aminoácidos (por ejemplo, SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20) comprende el aminoácido ácido aspártico (D) en el residuo 296 y/o el aminoácido ácido aspártico d(D) en el residuo 339 de SEQ
45 ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 y SEQ ID No 20 o sustitución(es) equivalentes en otras secuencias de aminoácidos adecuadas.

Además, sin desear estar ligado por la teoría, el residuo de aminoácido (histidina 101) en la xilosa isomerasa puede estar implicado en el sitio activo (Lee et al1989). En una realización, H101 es importante para el sitio activo. Por lo tanto, en una realización, la variante, homóloga o derivada de una secuencia de aminoácidos (tal como, SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20) comprende el aminoácido histidina (H) en el residuo 101 de SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 y
50 SEQ ID No 20 o una sustitución equivalente en otras secuencias de aminoácidos adecuadas .

55 En una realización, la variante, homóloga o derivada de una secuencia de aminoácidos (tal como, SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20) comprende el aminoácido arginina (R) en el residuo 391 y/o el aminoácido lisina (K) en el residuo 407 de la SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 y SEQ ID No 20 o sustituciones equivalentes en otras secuencias de aminoácidos adecuadas.

60 En una realización, una variante, homóloga o derivada de una secuencia de nucleótidos (tal como, SEQ ID No 1, SEQ ID No 10, SEQ ID No 17, SEQ ID No 12, SEQ ID No 27 o SEQ ID No 28, o la secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada como SEC ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20) comprende cambios en la secuencia de codones silenciosos con el fin de optimizar
65 la secuencia para las preferencias de codones de una célula hospedadora particular en la que se expresa la

secuencia. Por ejemplo, el uso de codones en una secuencia puede ser optimizada basándose en la tabla de uso de codones de levadura de la base de datos de uso de codones de Kazusa (Nakamura et al., 2000).

5 Aquí, el término "homóloga" significa una entidad que tiene una cierta homología con las secuencias de aminoácidos objeto y las secuencias de nucleótidos objeto. Aquí, el término "homología" se puede equiparar con la "identidad".

10 En el presente contexto, una secuencia homóloga se considera que incluye una secuencia de aminoácidos que puede ser al menos 75, 80, 85 o 90 % idéntica, preferiblemente al menos 95, 96, 97, 98 o 99 % idéntica a la secuencia objeto; y en una realización altamente preferida, al menos 98, 99 o 99,5 % idéntica a la secuencia objeto. Generalmente, los homólogos comprenderán los mismos sitios activos, etc. como la secuencia de aminoácidos objeto. Aunque la homología también puede considerarse en términos de similitud (es decir, residuos de aminoácidos que tienen propiedades/funciones químicas similares), en el contexto de la presente invención se prefiere expresar la homología en términos de identidad de secuencia.

15 En el presente contexto, una secuencia homóloga se considera que incluye una secuencia de nucleótidos que puede ser al menos 75, 80, 85 o 90 % idéntica; preferiblemente al menos 95, 96, 97, 98 o 99 % idéntica; y en una realización altamente preferida, al menos 98, 99 o 99,5 % idéntica con una secuencia de nucleótidos que codifica una enzima de la presente invención (la secuencia objeto). Generalmente, los homólogos comprenderán las mismas secuencias que codifican para los sitios activos, etc. que la secuencia objeto. Aunque la homología también puede considerarse en términos de similitud (es decir, residuos de aminoácidos que tienen propiedades/funciones químicas similares), en el contexto de la presente invención se prefiere expresar la homología en términos de identidad de secuencia.

20 Las comparaciones de homología pueden realizarse a ojo, o más normalmente, con la ayuda de programas de comparación de secuencias fácilmente disponibles. Estos programas informáticos comercializados pueden calcular el % de homología entre dos o más secuencias.

25 El % de homología puede calcularse sobre secuencias contiguas, es decir, se alinea una secuencia con la otra secuencia y cada aminoácido en una secuencia se compara directamente con el aminoácido correspondiente en la otra secuencia, un residuo cada vez. Esto se denomina alineación "sin huecos". Generalmente, tales alineaciones sin huecos se realizan sólo sobre un número relativamente corto de residuos.

30 Aunque este es un método muy simple y sistemático, no tiene en cuenta que, por ejemplo, en un par de secuencias por lo demás idénticas, una inserción o delección provocará que los siguientes residuos de aminoácidos salgan fuera de la alineación, provocando probablemente una gran reducción en el % de homología cuando se realiza una alineación global. En consecuencia, la mayoría de los métodos de comparación de la secuencia se diseñan para producir alineaciones óptimas que tengan en cuenta las posibles inserciones y delecciones sin penalizar indebidamente la puntuación global de homología. Esto se logra insertando "huecos" en la alineación de la secuencia para intentar maximizar la homología local.

35 Sin embargo, estos métodos más complejos asignan "penalizaciones por hueco" a cada hueco que se produce en la alineación de modo que, para el mismo número de aminoácidos idénticos, una alineación de secuencias con tan pocos huecos como sea posible, lo que refleja una mayor relación entre las dos secuencias comparadas, logrará una puntuación más alta que una con muchos huecos. Se utilizan normalmente "costes por huecos afines" para cargar un coste relativamente alto por la existencia de un hueco y una penalización más pequeña para cada residuo posterior en el hueco. Este es el sistema de puntuación de huecos más comúnmente utilizado. Las penalizaciones por un alto número de huecos, por supuesto, producirán alineaciones optimizadas con menos huecos. La mayoría de los programas de alineación permiten modificar las penalizaciones por hueco. Sin embargo, cuando se utiliza un software de este tipo para las comparaciones de secuencias se prefiere utilizar los valores por defecto. Por ejemplo, cuando se utiliza el paquete GCG Wisconsin Bestfit la penalización por hueco por defecto para secuencias de aminoácidos es -12 para un hueco y -4 para cada extensión.

40 El cálculo del % de homología máxima, por tanto, requiere en primer lugar la producción de una alineación óptima, teniendo en consideración las penalizaciones por hueco. Un programa informático adecuado para llevar a cabo dicha alineación es el paquete GCG Wisconsin Bestfit (Devereux et al 1984 Nuc. Acids Research 12 P387). Ejemplos de otros software que pueden realizar comparaciones de secuencias incluyen, pero no se limitan a, el paquete BLAST (ver Ausubel et al, 1999 Short Protocols in Molecular Biology, 4ª Ed - Capítulo 18), FASTA (Altschul et al., 1990 J. Mol.Biol.403-410) y la suite GENWORKS de herramientas de comparación. Tanto BLAST como FASTA están disponibles para la búsqueda online y offline (ver Ausubel et al., 1999, Short Protocols in Molecular Biology, páginas 7-58 a 7-60). Sin embargo, para algunas aplicaciones, se prefiere utilizar el programa GCG Bestfit. Una nueva herramienta, denominada BLAST 2 Sequences está también disponible para comparar secuencias de proteínas y de nucleótidos (ver FEMS Microbiol Lett 1999 174(2): 247-50; FEMS Microbiol Lett 1999 177(1):187-8 y tatiana@ncbi.nlm.nih.gov).

65 Aunque el % de homología final puede medirse en términos de identidad, el proceso de alineación en sí generalmente no se basa en una comparación de pares de todo o nada. En su lugar, se utiliza generalmente una

matriz de puntuación de similitud a escala que asigna puntuaciones a cada comparación por parejas basándose en la similitud química o la distancia evolutiva. Un ejemplo de tal matriz comúnmente utilizada es la matriz BLOSUM62 - la matriz por defecto para la suite de programas BLAST. Los programas GCG Wisconsin generalmente utilizan cualquiera de los valores por defecto públicos o una tabla de comparación de símbolos personalizada si se suministra (véase el manual del usuario para más detalles). Para algunas aplicaciones, se prefiere utilizar los valores por defecto públicos para el paquete GCG, o en el caso de otro software, la matriz por defecto, tal como BLOSUM62.

Como alternativa, las homologías porcentuales pueden calcularse utilizando la característica de alineación múltiple en DNASIS™ (Hitachi Software), basado en un algoritmo, análogo a CLUSTAL (Higgins DG y Sharp PM (1988), Gene 73(1), 237-244).

Una vez que el software ha producido una alineación óptima, es posible calcular el % de homología, preferiblemente el % de identidad de secuencia. El software hace normalmente esto como parte de la comparación de secuencias y genera un resultado numérico.

Las secuencias también pueden tener deleciones, inserciones o sustituciones de residuos de aminoácidos que producen un cambio silencioso y tienen como resultado una sustancia funcionalmente equivalente. Pueden realizarse sustituciones de aminoácidos deliberadas en base a la similitud de polaridad, carga, solubilidad, hidrofobicidad, hidrofiliidad y/o la naturaleza anfipática de los residuos siempre que se retenga la actividad de unión secundaria de la sustancia. Por ejemplo, los aminoácidos con carga negativa incluyen ácido aspártico y ácido glutámico; los aminoácidos con carga positiva incluyen lisina y arginina y los aminoácidos con grupos de cabeza polares no cargados que tienen valores de hidrofiliidad similares incluyen leucina, isoleucina, valina, glicina, alanina, asparagina, glutamina, serina, treonina, fenilalanina y tirosina.

Se pueden hacer sustituciones conservadoras, por ejemplo, según la siguiente Tabla. Los aminoácidos del mismo bloque de la segunda columna y preferiblemente en la misma línea en la tercera columna pueden sustituirse entre sí:

ALIFÁTICO	No polar	G A P
		I L V
	Polar - no cargado	C S T M
		N Q
	Polar - cargado	D E
		K R
AROMÁTICO		H F W Y

La presente invención también abarca la sustitución homóloga (sustitución y reemplazamiento se utilizan ambos en la presente memoria para indicar el intercambio de un residuo de aminoácido existente, con un residuo alternativo) que puede ocurrir, es decir, sustitución de igual a igual, tal como básico por básico, ácido por ácido, polar por polar, etc. La sustitución no homóloga puede producirse también, es decir, de una clase de residuo a otro, o que implica alternativamente la inclusión de aminoácidos no naturales tales como ornitina (en lo sucesivo denominado Z), ácido diaminobutírico ornitina (en lo sucesivo denominado B), norleucina ornitina (en lo sucesivo denominado O), pirilalanina, tienilalanina, naftilalanina y fenilglicina.

Los reemplazos pueden hacerse también por aminoácidos no naturales e incluyen: aminoácidos alfa* y alfa-disustituidos*, N-alkil aminoácidos*, ácido láctico*, derivados haluro de aminoácidos naturales tales como trifluorotirosina*, p-Cl-fenilalanina*, p-Br-fenilalanina*, p-I-fenilalanina* , L-alil-glicina*, β-alanina*, ácido L-α-amino butírico*, ácido L-γ-amino butírico*, ácido L-α-amino isobutírico*, ácido L-ε-amino caproico#, ácido 7-amino ácido heptanoico*, L-metionina sulfona#, L-norleucina*, L-norvalina*, p-nitro-L-fenilalanina*, L-hidroxi prolina#, L-tioprolina*, derivados metílicos de fenilalanina (Phe) tales como 4-metil-Phe*, pentametil-Phe*, L-Phe (4-amino)#, L-Tyr (metil)*, L-Phe (4-isopropil)*, L-Tic (ácido 1,2,3,4-tetrahidroisoquinolin-3-carboxílico)*, ácido L-diaminopropiónico# y L-Phe (4-bencil)*. La notación* se ha utilizado con el propósito de la discusión anterior (en relación con la sustitución homóloga o no homóloga), para indicar la naturaleza hidrófoba del derivado mientras que # se ha utilizado para indicar la naturaleza hidrófila del derivado, #* indica características anfipáticas.

Las secuencias de aminoácidos variantes pueden incluir grupos espaciadores adecuados que pueden insertarse entre cualesquiera dos residuos de aminoácidos de la secuencia incluyendo grupos alquilo, tales como grupos metilo, etilo o propilo, además de espaciadores de aminoácidos tales como residuos de glicina o β-alanina. Una forma adicional de variación, que implica la presencia de uno o más residuos de aminoácidos en forma peptoide, será bien conocida por los expertos en la técnica. Para evitar dudas, "la forma peptoide" se utiliza para referirse a residuos de aminoácidos variantes en los que el grupo sustituyente del carbono α está en el átomo de nitrógeno del residuo en lugar de en el carbono α. Los procesos para preparar péptidos en forma peptoide son conocidos en la

técnica, por ejemplo, Simon RJ et al., PNAS (1992) 89 (20), 9367-9371 y Horwell DC, Trends Biotechnol. (1995) 13(4), 132-134.

Las secuencias de nucleótidos para su uso en la presente invención pueden incluir dentro de ellas nucleótidos sintéticos o modificados. En la técnica se conocen varios tipos diferentes de modificación de oligonucleótidos. Estos incluyen cadenas principales de metilfosfonato y fosforotioato y/o la adición de cadenas de acridina o polilisina en los extremos 3' y/o 5' de la molécula. Para los fines de la presente invención, se entiende que las secuencias de nucleótidos descritas en la presente memoria pueden ser modificadas por cualquier método disponible en la técnica. Tales modificaciones pueden ser llevadas a cabo con el fin de mejorar la actividad *in vivo* o la duración de las secuencias de nucleótidos de la presente invención.

La presente invención también abarca el uso de secuencias de nucleótidos que son complementarias a las secuencias presentadas en la presente memoria, o cualquier derivado, fragmento o derivado de las mismas. Si la secuencia es complementaria a un fragmento de la misma entonces la secuencia puede utilizarse como una sonda para identificar secuencias de codificación similares en otros organismos, etc.

Los polinucleótidos que no son 100 % homólogos con las secuencias mencionadas en la presente memoria se pueden obtener de varias maneras. Otras variantes de las secuencias descritas en la presente memoria pueden obtenerse, por ejemplo, buscando mediante sondas en bibliotecas de ADN. Además, otros homólogos pueden obtenerse y dichos homólogos y fragmentos de los mismos en general serán capaces de hibridar selectivamente con las secuencias mostradas en la lista de secuencias en la presente memoria. Tales secuencias pueden obtenerse buscando en bibliotecas de ADNc con sondas que comprenden toda o parte de una cualquiera de las secuencias de la lista de secuencias adjunta en condiciones de media a alta rigurosidad.

También pueden obtenerse variantes y homólogos de cepa/especies usando PCR degenerada que usará cebadores diseñados para dirigirse a secuencias diana dentro de las variantes y homólogos que codifican secuencias de aminoácidos conservados dentro de las secuencias mencionadas en la presente memoria. Las secuencias conservadas pueden predecirse, por ejemplo, mediante la alineación de las secuencias de aminoácidos de varias variantes/homólogos. Los alineamientos de secuencias pueden realizarse usando el software informático conocido en la técnica. Por ejemplo el programa GCG Wisconsin PileUp es ampliamente usado.

Los cebadores utilizados en la PCR degenerada contendrán una o más posiciones degeneradas y se utilizarán en condiciones de rigurosidad inferiores a las utilizadas para la clonación de secuencias con cebadores de secuencia única frente a las secuencias conocidas.

Como alternativa, tales polinucleótidos pueden obtenerse por mutagénesis dirigida al sitio de secuencias caracterizadas. Esto puede ser útil cuando por ejemplo se requieren cambios en la secuencia de codones silenciosos para optimizar las preferencias de codón para una célula hospedadora particular en la que se expresan las secuencias de polinucleótidos. Otros cambios de secuencia pueden desearse con el fin de introducir sitios de reconocimiento de enzimas de restricción, o para alterar la propiedad o función de los polipéptidos codificados por los polinucleótidos.

Los polinucleótidos (secuencias de nucleótidos) mencionados en la presente memoria pueden usarse para producir un cebador, por ejemplo, un cebador de PCR, un cebador para una reacción de amplificación alternativa, una sonda, por ejemplo, etiquetada con un marcador de revelado por medios convencionales usando marcadores radiactivos o no radiactivos, o los polinucleótidos pueden clonarse en vectores. Tales cebadores, sondas y otros fragmentos serán al menos de 15, preferiblemente al menos 20, por ejemplo al menos 25, 30 o 40 nucleótidos de longitud y están también comprendidos por la expresión polinucleótidos tal como se utiliza en la presente memoria.

Los polinucleótidos tales como los polinucleótidos de ADN y las sondas mencionadas en la presente memoria pueden producirse de forma recombinante, de forma sintética, o por cualquier medio disponible para los expertos en la técnica. También pueden ser clonados por técnicas estándar.

En general, los cebadores se producirán por medios sintéticos, lo que implica una fabricación por etapas de la secuencia de ácido nucleico deseada, un nucleótido cada vez. Las técnicas para conseguir esto usando técnicas automatizadas están fácilmente disponibles en la técnica.

Los polinucleótidos más largos se producirán generalmente usando medios recombinantes, por ejemplo, usando técnicas de clonación PCR (reacción en cadena de la polimerasa). Los cebadores pueden ser diseñados para contener sitios de reconocimiento de enzima de restricción adecuados de modo que el ADN amplificado se puede clonar en un vector de clonación adecuado.

La invención se describirá ahora adicionalmente por medio de ejemplos, que están destinados a servir de ayudar a una persona con experiencia ordinaria en la técnica en la realización de la invención y no se pretende de ninguna manera limitar el alcance de la invención.

Ejemplos

Técnicas de la metodología de ADN recombinante general

5 La presente invención emplea, a menos que se indique lo contrario, técnicas convencionales de química, biología molecular, microbiología, ADN recombinante e inmunología, que están dentro de las capacidades de una persona de experiencia ordinaria en la técnica. Tales técnicas se explican en la literatura. Véase, por ejemplo, J. B. Roe, J. Crabtree, and A. Kahn, 1996, DNA Isolation and Sequencing: Essential Techniques, John Wiley & Sons; M. J. Gait (Editor), 1984, Oligonucleotide Synthesis: A Practical Approach, Irl Press; D. M. J. Lilley and J. E. Dahlberg, 1992, Methods of Enzymology: DNA Structure Part A: Synthesis and Physical Analysis of DNA Methods in Enzymology, Academic Press; Sambrook J., et al (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., Vol. 1-3. Cold Spring Harbor Laboratory Press y Ausubel F. M., et al (1995 (y suplementos periódicos)) Current Protocols in Molecular Biology, Vol. ch 9, 13 y 16. John Wiley & Sons, Nueva York, N.Y.

Organismos:

15 *Lactococcus* es una familia de bacterias grampositivas, que se encuentra comúnmente en muchas partes del mundo. Las cepas de esta familia tienen una larga historia de uso seguro para la fermentación de productos lácteos para la producción de, por ejemplo, queso. Una cepa de este tipo se usa en estos ejemplos, *Lactococcus lactis*, cepa DSM 20175, que se puede obtener del depósito del Instituto DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Colección Alemana de Microorganismos y Cultivos Celulares).

25 *Saccharomyces* es una familia de cepas de levadura que se encuentra comúnmente en muchas partes del mundo. Tienen una larga historia de uso seguro para fines de cocción, elaboración de la cerveza y vino, y son en la actualidad el organismo eucariota mejor caracterizado. Por lo tanto, es un organismo muy popular para su uso experimental en los laboratorios de biología y microbiología molecular. La cepa *S. cerevisiae* BY4741 es la que se utiliza aquí en los ejemplos. La cepa se puede obtener de Euroscarf (**EURO**pean **Saccharomyces Cerevisiae AR**chive for Functional Analysis - Alemania).

30 Las cepas *S. cerevisiae* cepa T0040, *S. cerevisiae* cepa T0028, *S. cerevisiae* cepa T0049 y *S. cerevisiae* cepa T0004 utilizadas en los siguientes ejemplos son todas variantes isogénicas, construidas a partir de BY4741 obtenida de Euroscarf.

35 *Escherichia coli* es una bacteria gramnegativa que se encuentra comúnmente en el intestino grueso de los animales superiores y los seres humanos. Muchas cepas del organismo tienen una larga historia de uso seguro en los laboratorios de biología molecular y microbiología. La cepa utilizada en los siguientes ejemplos es la cepa DH5-alfa de *E. coli*, que se utiliza comúnmente en los laboratorios de biología molecular para la clonación y la propagación de material genético. La cepa se puede obtener en la empresa "Invitrogen" (Invitrogen Corporation, California, EE.UU.).

40 *Thermoanaerobacter* es una familia de cepas bacterias grampositivas termófilas y de crecimiento anaerobio. Las dos cepas, cuyas enzimas se utilizan en los siguientes ejemplos, *Thermoanaerobacter thermohydrosulfuricus* (DSM N.º 15750) y *Thermoanaerobacterium thermosulfurigenes* (ATCC 33743) se aíslan originalmente de baños termales, DSM N.º 15750 en Ayas, Turquía y ATCC 33743 en el Parque Nacional de Yellowstone, EE.UU. Ambas pueden obtenerse en la DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Colección Alemana de Microorganismos y Cultivos Celulares). Sin embargo, el material genético que codifica las enzimas de estas cepas y que se utiliza en los ejemplos, se produjo sintéticamente, en base a una retrotraducción de las secuencias de aminoácidos como se especifica en la lista de secuencias como SEC ID No 15 y SEQ ID No 16.

50 *Pichia stipitis* es una especie de levadura que se ha aislado del intestino de larvas de insectos y termitas. La cepa de *Pichia stipitis* cuyos genes se han utilizado en los siguientes ejemplos es la cepa *Pichia stipitis* DSM N.º 3651, que se puede obtener del DSMZ - Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (Colección Alemana de Microorganismos y Cultivos Celulares).

55 *Pseudomonas* es una familia diversa de bacterias gramnegativas aerobias en forma de bacilo que se encuentran comúnmente en el agua y en semillas de plantas y tejidos de plantas. La cepa utilizada en estos ejemplos es la *Pseudomonas syringae* cepa DSM 50315, aislada originalmente de tejido vegetal herido. La cepa puede obtenerse del depósito del Instituto DSMZ - Deutsche Sammlung von Mikroorganismen und Cultivos Celulares GmbH (Colección Alemana de Microorganismos y Cultivos Celulares).

60 Ejemplo 1 - Construcción de una cepa de *S. cerevisiae* que expresa la XI de *Lactobacillus xylosus* y la xilulosa quinasa de *Pichia stipitis*.

Ejemplo 1a – Clonación TOPO del gen de la D-xilosa isomerasa a partir de *Lactobacillus xylosus* (*Lactococcus lactis*, cepa DSM 20175), utilizando secuencias de cebadores que se basan en el código de acceso del NCBI AF092042 (SEQ ID No 1).

65

La totalidad del gen de la D-xilosa isomerasa (Lx XI) de *Lactobacillus xylosus* se amplificó por PCR a partir del ADN obtenido de la cepa DSM 20175 usando los cebadores identificados por la SEQ ID No 2 y SEQ ID No 3. La región de codificación del gen de la xilosa isomerasa clonado por PCR de *Lactococcus lactis*, cepa DSM 20175, usando cebadores que se muestran en la SEQ ID No 2 y SEQ ID No 3, se muestra en la SEQ ID No 17. Se introdujo un sitio de restricción de NheI, proximal al codón de inicio ATG y un sitio de restricción de XhoI, distal al codón de parada flanqueando al gen Lx XI. Como molde, se usó el ADN de la cepa de *L. xylosus* en una concentración de 0,2 ng/μl para la reacción de PCR. La PCR se realizó en 30 ciclos de 30 segundos a 96 °C, 30 segundos a 50 °C y 150 segundos a 72 °C, seguido de una incubación final de 10 minutos a 72 °C utilizando ADN polimerasa de alta fidelidad Phusion (Finnzymes Oy , Finlandia). El producto de PCR se separó por electroforesis en un gel de agarosa de baja fusión al 1 % y se aisló un fragmento de 1339 pb. El fragmento de ADN se clonó con TOPO en el vector pCR-Blunt II-TOPO (Invitrogen, EE.UU.) según las instrucciones del fabricante y el plásmido resultante se utilizó para la transformación de *E. coli* TOP10. El plásmido se denominó pCR-Blunt 2 LxXI.

Ejemplo 1b - Construcción del plásmido LxXI-2a que contiene el gen de la xilosa isomerasa (LxXI) de *L. xylosus* bajo el control del promotor GPD y el terminador CYC1 de *S. cerevisiae*.

El vector lanzadera de alto número de copias de *E. coli*/*S. cerevisiae* P426-GPD (Mumberg et al.,1995) se digirió con SpeI y XhoI y los extremos resultantes se desfosforilaron con fosfatasa alcalina. De un modo similar, el fragmento de ADN que codifica el gen de la xilosa isomerasa de *L. xylosus* fue liberado a partir del vector pCR-Blunt 2 LxXI (que se describe en el Ejemplo 1a) mediante digestión con NheI y XhoI. El plásmido linealizado P426-GPD resultante y el fragmento de ADN que codifica ThXI fueron separados por electroforesis en un gel de agarosa de baja fusión al 1 % y aislados. Los dos fragmentos de ADN se ligaron juntos dando como resultado el plásmido denominado LxXI-2a.

Ejemplo 1c – Clonación TOPO del gen de la D-xiluloquinasa de *Pichia stipitis* cepa DSM3651 basado en el código de acceso del NCBI AF127802.

La totalidad del gen de la D-xiluloquinasa (PsXKS) de *P. stipitis* se amplificó por PCR a partir del ADN obtenido de la cepa DSM3651 usando los cebadores identificados por la SEQ ID No 4 y SEQ ID No 5. Se introdujo un sitio de restricción de NheI, proximal al codón de inicio ATG y un sitio de restricción de XhoI, distal al codón de parada flanqueando al gen PsXKS. Como molde, se usó el ADN de la cepa de *P. stipitis* en una concentración de 0,2 ng/μl para la reacción de PCR. La PCR se realizó en 30 ciclos de 30 segundos a 96 °C, 30 segundos a 50 °C y 150 segundos a 72 °C, seguido de una incubación final de 10 minutos a 72 °C utilizando ADN polimerasa de alta fidelidad Phusion (Finnzymes Oy , Finlandia). El producto de PCR se separó por electroforesis en un gel de agarosa de baja fusión al 1 % y se aisló un fragmento de 1891 pb. El fragmento de ADN se clonó con TOPO en el vector pCR-Blunt II-TOPO (Invitrogen, EE.UU.) según las instrucciones del fabricante y el plásmido resultante se utilizó para la transformación de *E. coli* TOP10. El plásmido se denominó pCR-Blunt 2 P.stip XKS.

Ejemplo 1d - Construcción del plásmido PsXKS-14a que contiene el gen de la D-xiluloquinasa (PsXKS) de *P. stipitis* bajo el control del promotor GPD y el terminador de CYC1 de *S. cerevisiae*.

El vector lanzadera de alto número de copias de *E. coli*/*S. cerevisiae* P425-GPD (Mumberg et al.,1995) se digirió con SpeI y XhoI y los extremos resultantes se desfosforilaron con fosfatasa alcalina. De un modo similar, el fragmento de ADN que codifica el gen de la xilulosa quinasa de *P. stipitis* fue liberado a partir del vector pCR-Blunt 2 P.stip XKS (descrito en el Ejemplo 1c) mediante digestión con NheI y XhoI. El plásmido linealizado P425-GPD resultante y el fragmento de ADN que codifica PsXKS fueron separados por electroforesis en un gel de agarosa de baja fusión al 1 % y aislados. Los dos fragmentos de ADN se ligaron juntos dando como resultado el plásmido denominado PsXKS-14a.

Ejemplo 1e - Construcción de cepas de *S. cerevisiae* que contienen LxXI-2a y PsXKS-14a.

Se combinaron 200 ng de cada uno de los plásmidos y se utilizaron para la transformación de la cepa BY4741 de la levadura *S. cerevisiae* (Euroscarf, Alemania) por medio de electroporación utilizando el sistema Biorad Gene Pulser II (BioRad, EE.UU.) de acuerdo con las instrucciones del fabricante. Las células de levadura se hicieron competentes de acuerdo con un protocolo estándar (Becker and Guarente, 1991). La selección de clones transformados con ambos plásmidos se llevó a cabo en medios excluyentes completos sintéticos sólidos sin uracilo y sin leucina y complementado con D-glucosa 2 % (SC-Ura, Leu) (Rose et al.,1990). Los principales clones de tamaño medio se volvieron a sembrar en SC-Ura, Leu y se obtuvo una colonia de la cepa T0004 transformada con los plásmidos LxXI-2a y PsXKS-14a.

Ejemplo 2 - Construcción de una cepa de *S. cerevisiae* que expresa la XI de *Thermoanaerobacter thermohydrosulfuricus* y la xilulosa quinasa de *Pichia stipitis*.

Ejemplo 2a - Construcción del gen sintético de la xilosa isomerasa que codifica la xilosa isomerasa de *Thermoanaerobacter thermohydrosulfuricus* basado en el código de acceso del NCBI D00756.

La totalidad del gen de la xilosa isomerasa de *T. thermohydrosulfuricus* (ThXI) se sintetizó y fue ensamblado por Genearth AG (Regensburg, Alemania). El uso de codones en la secuencia fue optimizado basándose en la tabla de uso de codones de levadura de la base de datos de uso de codones de Kazusa (Nakamura et al., 2000). Flanqueando al marco de lectura abierto, se incluyeron en la construcción sintética un sitio de restricción de NheI, proximal al codón de inicio ATG y un sitio de restricción de XhoI, distal al codón de parada. La integridad del gen sintético ThXI se determinó por secuenciación de ambas cadenas. La secuencia de nucleótidos de ThXI incluyendo los sitios de restricción flanqueantes se identifica como SEC ID NO 6, que muestra la secuencia de ADN sintetizada con la traducción de aminoácidos de la región de codificación mostrada anteriormente de la secuencia de nucleótidos. El plásmido que lo alberga fue denominado 0717046pGA14.

Ejemplo 2b - Construcción del plásmido ThXI-5a que contiene el gen de la xilosa isomerasa (ThXI) de *T. thermohydrosulfuricus* bajo el control del promotor GPD y el terminador CYC1 de *S. cerevisiae*.

El vector lanzadera de alto número de copias de *E. coli/S. cerevisiae* P426-GPD (Mumberg et al., 1995) se digirió con SpeI y XhoI y los extremos resultantes se desfosforilaron con fosfatasa alcalina. De un modo similar, el fragmento de ADN que codifica el gen de la xilosa isomerasa de *T. thermohydrosulfuricus* fue liberado a partir del vector 0717046pGA14 (descrito en el Ejemplo 2a) mediante digestión con NheI y XhoI. El plásmido linealizado P426-GPD resultante y el fragmento de ADN que codifica ThXI fueron separados por electroforesis en un gel de agarosa de baja fusión al 1 % y aislados. Los dos fragmentos de ADN se ligaron juntos dando como resultado el plásmido denominado ThXI-5a.

Ejemplo 2c - Construcción de cepas de *S. cerevisiae* que contienen ThXI-5a y PsXKS-14a.

Se combinaron 200 ng de cada uno de los plásmidos y se utilizaron para la transformación de la cepa BY4741 de la levadura *S. cerevisiae* (Euroscarf, Alemania) por medio de electroporación utilizando el sistema Biorad Gene Pulser II (BioRad, EE.UU.) de acuerdo con las instrucciones del fabricante. Las células de levadura se hicieron competentes de acuerdo con un protocolo estándar (Becker and Guarente, 1991). La selección de clones transformados con ambos plásmidos se llevó a cabo en medios excluyentes completos sintéticos sólidos sin uracilo y sin leucina y complementado con D-glucosa 2 % (SC-Ura, Leu) (Rose et al., 1990). Los principales clones de tamaño medio se volvieron a sembrar en SC-Ura, Leu y se obtuvo una colonia de la cepa T0049 transformada con los plásmidos ThXI-5a y PsXKS-14a.

Ejemplo 3 - Construcción de una cepa de *S. cerevisiae* que expresa la xilosa isomerasa de *Thermoanaerobacter thermosulphurigenes* y la xilulosa quinasa de *Pichia stipitis*.

Ejemplo 3a - Construcción de un gen sintético de la xilosa isomerasa que codifica la xilosa isomerasa de *Thermoanaerobacter thermosulphurigenes* basado en el código de acceso del NCBI J05650.

La totalidad del gen de la xilosa isomerasa de *T. thermosulphurigenes* (TsXI) se sintetizó y fue ensamblado por Genearth AG (Regensburg, Alemania). El uso de codones en la secuencia fue optimizado basándose en la tabla de uso de codones de levadura de la base de datos de uso de codones de Kazusa (Nakamura et al., 2000). Flanqueando al marco de lectura abierto, se incluyeron en la construcción sintética un sitio de restricción de NheI, proximal al codón de inicio ATG y un sitio de restricción de XhoI, distal al codón de parada. La integridad del gen sintético TsXI se determinó por secuenciación de ambas cadenas. La secuencia de nucleótidos de TsXI incluyendo los sitios de restricción flanqueantes se identifica como SEC ID NO 7, que muestra la secuencia de ADN sintetizada con la traducción de aminoácidos de la región de codificación mostrada anteriormente de la secuencia de nucleótidos. El plásmido que lo alberga fue denominado 0717047pGA 18.

Ejemplo 3b - Construcción del plásmido TsXI-28a que contiene el gen de la xilosa isomerasa (TsXI) de *T. thermosulphurigenes* bajo el control del promotor GPD y el terminador CYC1 de *S. cerevisiae*.

El vector lanzadera de alto número de copias de *E. coli/S. cerevisiae* P426-GPD (Mumberg et al., 1995) se digirió con SpeI y XhoI y los extremos resultantes se desfosforilaron con fosfatasa alcalina. De un modo similar, el fragmento de ADN que codifica el gen de la xilosa isomerasa de *T. thermosulphurigenes* fue liberado a partir del vector 0717047pGA18 (descrito en el Ejemplo 3a) mediante digestión con NheI y XhoI. El plásmido linealizado P426-GPD resultante y el fragmento de ADN que codifica TsXI fueron separados por electroforesis en un gel de agarosa de baja fusión al 1 % y aislados. Los dos fragmentos de ADN se ligaron juntos dando como resultado el plásmido denominado TsXI-28a.

Ejemplo 3c - Construcción de cepas de *S. cerevisiae* que contienen TsXI-28a y PsXKS-14a.

Se combinaron 200 ng de cada uno de los plásmidos y se utilizaron para la transformación de la cepa BY4741 de la levadura *S. cerevisiae* (Euroscarf, Alemania) por medio de electroporación utilizando el sistema Biorad Gene Pulser II (BioRad, EE.UU.) de acuerdo con las instrucciones del fabricante. Las células de levadura se hicieron competentes de acuerdo con un protocolo estándar (Becker and Guarente, 1991). La selección de clones transformados con ambos plásmidos se llevó a cabo en medios excluyentes completos sintéticos sólidos sin uracilo

y sin leucina y complementado con D-glucosa 2 % (SC-Ura, Leu) (Rose et al.,1990). Los principales clones de tamaño medio se volvieron a sembrar en SC-Ura, Leu y se obtuvo una colonia de la cepa T0028 transformada con los plásmidos TsXI-28a y PsXKS-14a.

- 5 Ejemplo 4 - Construcción de una cepa de *S. cerevisiae* que expresa la XylA (PsXI) de *Pseudomonas syringae* y la xilulosa quinasa de *Pichia stipitis*.

Ejemplo 4a - Clonación TOPO del gen de la D-xilosa isomerasa (PsxI) a partir de *Pseudomonas syringae* pv *Tomato*

- 10 La totalidad del gen de la xilosa isomerasa (Ps XI) de *P. syringae* se amplificó por PCR a partir del ADN obtenido de la cepa DSM 50315 usando los cebadores identificados por la SEQ ID No 8 y SEQ ID No 9. Se introdujo un sitio de restricción de NheI, proximal al codón de inicio ATG y un sitio de restricción de XhoI, distal al codón de parada flanqueando al gen PsXI. Como molde, se usó el ADN de la cepa de *P. syringae* en una concentración de 0,2 ng/μl para la reacción de PCR. La PCR se realizó en 30 ciclos de 30 segundos a 96 °C, 30 segundos a 50 °C y 150 segundos a 72 °C, seguido de una incubación final de 10 minutos a 72 °C utilizando ADN polimerasa de alta fidelidad Phusion (Finnzymes Oy, Finlandia). El producto de PCR se separó por electroforesis en un gel de agarosa de baja fusión al 1 % y se aisló un fragmento de 1323 pb. El fragmento de ADN se clonó con TOPO en el vector pCR-Blunt II-TOPO (Invitrogen, EE.UU.) según las instrucciones del fabricante y el plásmido resultante se utilizó para la transformación de *E. coli* TOP10. El plásmido se denominó pCR-Blunt 2 PsXI.

- 20 Ejemplo 4b - Construcción del plásmido PsXI-34a que contiene el gen de la xilosa isomerasa (PsXI) de *P. syringae* bajo el control del promotor GPD y el terminador de CYC1 de *S. cerevisiae*.

- 25 El vector lanzadera de alto número de copias de *E. coli*/*S. cerevisiae* P426-GPD (Mumberg et al.,1995) se digirió con SpeI y XhoI y los extremos resultantes se desfosforilaron con fosfatasa alcalina. De un modo similar, el fragmento de ADN que codifica el gen de la xilosa isomerasa de *P. syringae* fue liberado a partir del vector pCR-Blunt 2 PsXI (descrito en el Ejemplo 4a) mediante digestión con NheI y XhoI. El plásmido linealizado P426-GPD resultante y el fragmento de ADN que codifica PsXI fueron separados por electroforesis en un gel de agarosa de baja fusión al 1 % y aislados. Los dos fragmentos de ADN se ligaron juntos dando como resultado el plásmido denominado PsXI-34a.

- 30 Ejemplo 4c - Construcción de cepas de *S. cerevisiae* que contienen PsXI-34a y PsXKS-14a.

- 35 Se combinaron 200 ng de cada uno de los plásmidos y se utilizaron para la transformación de la cepa BY4741 de la levadura *S. cerevisiae* (Euroscarf, Alemania) por medio de electroporación utilizando el sistema Biorad Gene Pulser II (BioRad, EE.UU.) de acuerdo con las instrucciones del fabricante. Las células de levadura se hicieron competentes de acuerdo con un protocolo estándar (Becker and Guarente, 1991). La selección de clones transformados con ambos plásmidos se llevó a cabo en medios excluyentes completos sintéticos sólidos sin uracilo y sin leucina y complementado con D-glucosa 2 % (SC-Ura, Leu) (Rose et al.,1990). Los principales clones de tamaño medio se volvieron a sembrar en SC-Ura, Leu y se obtuvo una colonia de la cepa T0040 transformada con los plásmidos PsXI-34a y PsXKS-14a.

Ejemplo 5 - Determinación y comparación de las actividades xilosa isomerasa específicas obtenidas por la expresión de cuatro xilosa isomerasas bacterianas diferentes en *Saccharomyces cerevisiae*.

- 45 Las cuatro cepas diferentes:

- 50 *S. cerevisiae* cepa T0040 que contiene el PsXI-34a y el PsXKS-14a;
S. cerevisiae cepa T0028 contiene el TsXI-28a y el PsXKS-14a;
S. cerevisiae cepa T0049 contiene el ThXI-5a y el PsXKS-14a y
S. cerevisiae cepa T0004 que contiene el LxXI-2a y el PsXKS-14a;

fueron sembradas en placas y se cultivaron durante la noche en un medio excluyente completo sintético sólido sin uracilo y sin leucina y complementado con un D-glucosa 2 % (SC-Ura, Leu) (Rose et al.,1990).

- 55 A continuación, los tubos de 50 ml que contenían 10 ml de medios excluyentes completos sintéticos sin uracilo y sin leucina y complementado con un D-glucosa 2 % (SC-Ura, Leu) (Rose et al,1990), se inocularon con colonias de las placas sembradas hasta una DO600 ajustada desde 0,1 y se cultivaron durante 18 horas con agitación, 200 rpm, a 30 °C. Los cultivos se recogieron por centrifugación, 4.000 rpm durante 15 minutos, se lavaron mediante resuspensión en 2 ml de agua y se centrifugaron de nuevo. Las células se lisaron por adición de 300 μl de CelLytic Y Plus Kit (Sigma-Aldrich) seguido de incubación a temperatura ambiente durante 30 minutos con agitación suave y centrifugación a 16.000 x g durante 10 minutos para eliminar los restos celulares. El sobrenadante se aisló y se utilizó para la determinación de la actividad de la xilosa isomerasa y el contenido de proteína. Se determinó la actividad xilosa isomerasa usando el método de la cisteína-carbazol como se describe por Dische y Borenfreund (Dische y Borenfreund, 1951). El contenido de proteína se determinó mediante el uso del ensayo comercializado del ácido bicinonínico a base de proteínas (kit de ensayo de BCA, Pierce Biotechnology Inc., EE.UU.).

- 65

Los resultados se muestran en la Figura 5. La actividad xilosa isomerasa total más elevada por mg de proteína de levadura total, obtenida mediante la expresión de la xilosa isomerasa de *L. xylosus*, es aproximadamente 10 veces mayor que el segundo mejor resultado, obtenido mediante la expresión de la xilosa isomerasa de *T. thermohydrosulphuricus*. Está claro que la xilosa isomerasa de *L. xylosus* es muy superior a las otras isomerasas bacterianas. Los resultados obtenidos, unidades de xilosa isomerasa/mg de proteína de levadura para las dos bacterias termófilas, están en línea con otros resultados obtenidos utilizando xilosa isomerasas de bacterias termófilas (Walfriedsson et al, 1996; Lönn et al, 2003) y la actividad específica isomerasa de *L. xylosus* es igual o supera a la actividad específica obtenida por el uso de la xilosa isomerasa del hongo *Piromyces* sp. Ell (Kuyper et al. 2003).

Ejemplo 6 - Crecimiento en xilosa como la fuente de carbono de las cepas de *S. cerevisiae* que expresan LxXI-2a y PsXKS-14a (cepa T0004).

La cepa T0004 de *S. cerevisiae*, descrita en el Ejemplo 1e, comprende un casete de expresión que contiene la SEQ ID No 17 (que codifica la xilosa isomerasa) y un casete de expresión que contiene la secuencia de nucleótidos que codifica la xiluloquinasa de *P. stipitis*.

La cepa T0004, descrita en el Ejemplo 1e, se adaptó gradualmente al crecimiento en xilosa mediante dilución en serie del contenido de glucosa, seguido por el crecimiento en xilosa como única fuente de carbono. Esto se llevó a cabo usando el siguiente procedimiento:

Receta de *Synthetic Complete growth medium without uracil and leucine* (medio SC-Ura y Leu):

Bacto - base de nitrógeno de levadura sin aminoácidos: 6,7 g/l;
suministrado con 40 mg/l de todos los aminoácidos, excepto leucina;
y suministrado con ácido *p*-aminobenzoico, 40 mg/l.

Inoculación de un primer tubo de cultivo que contiene 10 ml de medio SC-Ura y Leu que contiene xilosa 2 % + glucosa 0,5 % con la cepa T0004. Incubación del tubo a 30 °C y agitación rotatoria a 150 rpm hasta que la DO600 alcanzó 1,5. Este cultivo se denomina aquí como "Cultivo 1".

Inoculación de un segundo tubo de cultivo que contiene 9 ml de medio SC-Ura y Leu que contiene xilosa 2 % con 1 ml del "Cultivo 1" anteriormente descrito. Incubación del tubo de cultivo a 30 °C y agitación rotatoria a 150 rpm hasta que la DO600 alcanzó aproximadamente 1,0. Este cultivo se denomina en la presente memoria como "Cultivo 2".

Inoculación de un tercer tubo de cultivo que contiene 9 ml de medio SC-Ura y Leu que contiene xilosa 2 % con 1 ml del "Cultivo 2" anteriormente descrito. Incubación del tubo de cultivo a 30 °C y agitación rotatoria a 150 rpm hasta que la DO600 estaba entre 0,5 y 1,0. Este cultivo se denomina en la presente memoria como "Cultivo 3".

Inoculación de un matraz de agitación de 250 ml con deflectores, que contiene 50 ml de medio SC-Ura y Leu que contiene xilosa 2 % con una cantidad del "Cultivo 3" hasta alcanzar un valor de la DO600 de 0,05. Incubación del tubo de cultivo a 30 °C y agitación rotatoria a 150 rpm y seguimiento del crecimiento por mediciones de la DO600 durante un período de cultivo de 240 horas.

El crecimiento de la cepa T0004 en xilosa como la fuente de carbono se puede ver en la Figura 6. Después de una fase de latencia, la máxima tasa de crecimiento alcanzó 0,05/h sin otras modificaciones de la cepa.

Ejemplo 7 - Demostración de la elevada actividad de la xilosa isomerasa en células de *S. cerevisiae* por la expresión intracelular de tres variantes de origen natural y dos variantes artificiales de una isomerasa de *Lactococcus*.

Las xilosa isomerasas naturales con secuencias de aminoácidos como se muestra en SEQ ID No 13, SEQ ID No 14 y SEQ ID No 18, donde se expresan intracelularmente en *Saccharomyces cerevisiae*, cepa BY4741 (Euroscarf, Alemania). Además, se generaron dos variantes relacionadas, que no se encuentran en ningún organismo (como se detalla más adelante) y también se expresaron. Las secuencias de aminoácidos de estas dos variantes se muestran en la SEQ ID No 19 y SEQ ID No 20. Excepto por la variación en la posición del aminoácido 391 (arginina o lisina), la posición 407 (ácido glutámico o lisina) o 416 (tirosina o histidina), todas las variantes tienen secuencias de aminoácidos idénticas. Esto significa que cuando se compara con la SEQ ID No 18, la SEQ ID No 13 tiene el aminoácido K en la posición 391, el aminoácido K en la posición 407 y el aminoácido H en la posición 416. Estos aminoácidos de la SEQ ID No 13 están subrayados en la lista de secuencias. En comparación con la SEQ ID No 18, la SEQ ID No 14 tiene el aminoácido K en la posición 391. Este aminoácido en la SEQ ID No 14 está subrayado en la lista de secuencias. En comparación con la SEQ ID No 18, la SEQ ID No 19 tiene el aminoácido K en la posición 407 y el aminoácido H en la posición 416. Estos aminoácidos en la SEQ ID No 19 están subrayados en la lista de secuencias. En comparación con la SEQ ID No 18, la SEQ ID No 20 tiene el aminoácido K en la posición 407. Este aminoácido en la SEQ ID No 20 está subrayado en la lista de secuencias.

Las secuencias de ADN que codifican las secuencias de aminoácidos SEQ ID No 13, SEQ ID No 14, SEQ ID No 19 y SEQ ID No 20 utilizadas en este ejemplo se generaron todas por mutagénesis dirigida al sitio de la secuencia de ADN mostrada como SEQ ID No 17, utilizando la construcción obtenida como se describe en el Ejemplo 1a como plantilla inicial. (La secuencia de aminoácidos SEQ ID No 18 es codificada por la SEQ ID No 17.) La mutagénesis dirigida al sitio se realizó utilizando el kit "Quick Change" (Stratagene, La Jolla, EE.UU.) para cambiar los codones que codifican los residuos de aminoácidos 391, 407 y 416, respectivamente. El kit se utilizó de acuerdo con el protocolo del fabricante. Las secuencias de los cebadores de ADN mutagénicos utilizados se muestran en la SEQ ID No 21 y SEQ ID No 22 (sustitución R391K y eliminación de un sitio de restricción Psp14061), SEQ ID No 23 y SEQ ID No 24 (sustitución E407K y adición de un sitio de restricción *Hind*III) y SEQ ID No 25 y SEQ ID No 26 (sustitución combinada de E407K y Y416H y adición de un sitio de restricción *Hind*III). La adición o eliminación de los sitios de restricción permite una fácil identificación de los vectores mutados. Las secuencias de ADN generales correctas de las construcciones de ADN generadas se confirmaron posteriormente mediante secuenciación de ADN.

Se utilizó una versión modificada del vector lanzada de alto número de copias de *E. coli/S. cerevisiae* pRS426-gpd (Mumberg et al., 1995) para la expresión intracelular de las variantes de la xilosa isomerasa. El vector se modificó mediante el intercambio del marcador URA3 en pRS426-gpd con el marcador de resistencia a nourseotricina del vector pAG35 (Euroscarf Alemania), utilizando técnicas de subclonación de biología molecular convencionales. El vector lanzadera resultante confiere resistencia a la nourseotricina a *S. cerevisiae*.

Los diferentes cinco genes de xilosa isomerasa se aislaron y se insertaron en el vector de expresión modificada de acuerdo con la descripción del Ejemplo 1b. Las construcciones de plásmidos resultantes se denominaron pRS426gpd_nat::X113, pRS426gpd_nat::X114, pRS426gpd_nat::X118, pRS426gpd_nat::X119 y pRS426gpd_nat::X120, respectivamente.

Cada construcción de plásmido variante se transformó en una cepa BY4741 de *S. cerevisiae* (Euroscarf, Alemania) por medio de electroporación utilizando el sistema Biorad Gene Pulser II (BioRad, EE.UU.) de acuerdo con las instrucciones del fabricante. Las células de levadura se hicieron competentes de acuerdo con un protocolo estándar (Becker y Guarente, 1991). La selección de clones transformados se realizó en placas YPD (20 g/l de triptona, 10 g/l de extracto de levadura y 20 g/l de glucosa con 15 g/l de bactoagar añadido para la solidificación) que contienen 100 mg/l de nourseotricina.

Se obtuvieron cinco líneas celulares, cada una expresando una de las xilosa isomerasas de *Lactococcus* variantes.

La actividad xilosa isomerasa específica de cada línea celular se determinó por duplicado y se comparó tal como se describe en el Ejemplo 5, excepto que el medio utilizado para el crecimiento de la levadura era YPD (20 g/l de triptona, 10 g/l de extracto de levadura y 20 g/l de glucosa) suministrado con 100 mg/l de nourseotricina.

La actividad xilosa isomerasa específica promedio obtenida de las determinaciones por duplicado, calculada como unidades de xilosa isomerasa/mg de proteína total de levadura extraída, de las cinco cepas se muestra en la Figura 7.

Todas las xilosa isomerasas de *Lactococcus* variantes son muy activas cuando se expresan en *S. cerevisiae*. Por lo tanto, esto puede ser considerado como general para los genes xilosa isomerasa de *Lactococcus* que codifican xilosa isomerasas activas en *Lactococcus*. Parece que, comparando la actividad de la SEQ ID 19 y SEQ ID 20, la sustitución Y416H representa muy poca diferencia en términos de actividad. Por el contrario, la sustitución K391R y la sustitución E407K aumentan la actividad específica en la levadura en aproximadamente un 15 a 20 %, un efecto que parece ser aditivo si esas dos sustituciones se combinan como en la SEQ ID No 19 y SEQ ID No 20.

REFERENCIAS:

- Anderson R. L. and Allison D. P. (1965) Purification and Characterization of D-Lyxose Isomerase. *J. Biol. Chem.* 240, 2367-2372.
- Ausubel F. M., Brent R., Kingston R. E., Moore D., Seidman J. G., Smith J. A., and Struhl K., eds (1995 (and periodic supplements)) *Current Protocols in Molecular Biology*, Vol. ch. 9, 13 and 16. John Wiley & Sons, Nueva York, N.Y.
- Bao X., Gao D., and Wang Z. (1999) Expression of xylose isomerase gene (*xyIA*) in *Saccharomyces cerevisiae* from *Clostridium thermohydrosulphuricum*. *Wei Sheng Wu Xue Bao* 39, 49-54.
- Becker D. M. and Guarente L. (1991) High-efficiency transformation of yeast by electroporation. *Methods Enzymol.* 194, 182-187.
- Cheng H., Jiang N., Shen A., and Feng Y. (2005) Molecular cloning and functional expression of D-arabitol dehydrogenase gene from *Gluconobacter oxydans* in *Escherichia coli*. *FEMS Microbiol. Lett.* 252, 35-42.
- Dische Z. and Borenfreund E. (1951) A new spectrophotometric method for the detection and determination of keto sugars and trioses. *J. Biol. Chem.* 192, 583.
- Dothie J. M., Giglio J. R., Moore C. B., Taylor S. S., and Hartley B. S. (1985) Ribitol dehydrogenase of *Klebsiella aerogenes*. Sequence and properties of wild-type and mutant strains. *Biochem. J.* 230, 569-578.
- Gait M. J., ed (1984) *Oligonucleotide Synthesis: A Practical Approach*. Oxford University Press.

- Izumori K., Rees A. W., and Elbein A. D. (1975) Purification, Crystallization, and Properties of D-Ribose Isomerase from *Mycobacterium smegmatis*. *J. Biol. Chem.* 250, 8085-8087.
- Kavanagh K., Klimcek M., Nidetsky B., and Wilson D. K. (2003) Structure of xylose reductase bound to NAD⁺ and the basis for single and dual co-substrate specificity in family 2 aldo-keto reductases. *Biochem. J.* 373, 319-326.
- 5 Kuyper M., Harhangi H. R., Stave A. K., Winkler A. A., Jetten M. S., de Laat W. T., den Ridder J. J., Op den Camp H. J., van Dijken J. P., and Pronk J. T. (2003) High-level functional expression of a fungal xylose isomerase: the key to efficient ethanol fermentation of xylose by *Saccharomyces cerevisiae*. *FEMS Yeast Research* 4, 69-78.
- Lee C., Bagdasarian M., Meng M., and Zeikus J. G. (1989) Catalytic Mechanism of Xylose (Glucose) Isomerase from *Clostridium thermosulfurogenes*. Characterization of the structural gene and function of active site histidine. *J. Biol. Chem.* 265, 19082-19090.
- 10 Lee Y. E., Jan M. K., Lee C., Lowe S. E., and Zeikus J. G. (1993) Taxonomic distinction of saccharolytic thermophilic anaerobes: description of *Thermoanaerobacterium xylanolyticum* gen. nov., sp. nov., and *Thermoanaerobacterium saccharolyticum* gen. nov., sp. nov.; reclassification of *Thermoanaerobacterium brockii*, *Clostridium thermosulfurogenes*, and *Clostridium thermohydrosulfuricum* E100-69 as *Thermoanaerobacter brockii* comb. nov., *Thermoanaerobacterium thermosulfurigenes* comb. nov., and *Thermoanaerobacter thermohydrosulfuricus* comb. nov., respectively; and transfer of *Clostridium thermohydrosulfuricum* 39E to *Thermoanaerobacter ethanolicus*. *Int. J. Syst. Bacteriol.* 43, 41-51.
- Lilley D. M. J. and Dahlberg J. E., eds (1992) *Methods in Enzymology: DNA Structures Part A: Synthesis and Physical Analysis of DNA*, Vol 211. Academic Press.
- Lönn A., Gardonyi M., van Zyl W., Hahn-Hägerdahl B., and Otero R. C. (2002) Cold adaptation of xylose isomerase from *Thermus thermophilus* through random PCR mutagenesis. *Eur. J. Biochem.* 269, 157-163.
- 20 Meng M., Bagdasarian M., and Zeikus J. G. (1993) The role of active-site aromatic and polar residues in catalysis and substrate discrimination by xylose isomerase. *Proc. Natl. Acad. Sci. USA* 90, 8459-8463.
- Mumberg D., Mailer R., and Funk M. (1995) Yeast vectors for the controlled expression of heterologous proteins in different genetic backgrounds. *Gene* 156, 119-122.
- 25 Nakamura Y., Gojohori T., and Ikemura T. (2000) Codon Usage tabulated from international DNA sequence databases: status for the year 2000. *Nucleic Acids Res.* 28, 292.
- Richard P., Londesborough J., Putkonen M., Kalkkinen N., and Penttilä M. (2001) Cloning and Expression of a Fungal L-Arabinitol 4-Dehydrogenase Gene. *J. Biol. Chem.* 276, 40631-40637.
- Roe B., Crabtree J., and Kahn A., eds (1996) *DNA Isolation and Sequencing: Essential Techniques*. John Wiley & Sons.
- 30 Rose M. D., Winston F., and Hieter P., eds (1990) *Methods in Yeast Genetics: A Laboratory Course Manual*. Cold Spring Harbor Laboratory Press.
- Sambrook J., Fritsch E. F., and Maniatis T., eds (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Vol. 1-3. Cold Spring Harbor Laboratory Press.
- 35 Shimonishi T. and Izumori K. (1996) A new enzyme, L-ribose isomerase from *Acinetobacter* sp. strain DL-28. *J. Ferment. Bioeng.* 81, 493-497.
- Smith P. K., Krohn R. I., Hermanson G. T., Mallia A. K., Gartner F. H., Provenzano M. D., Fujimoto E. K., Goeke N. M., Olson B. J., and Klenk D. C. (1985) Measurement of protein using bicinchoninic acid. *Anal. Biochem.* 150, 76-85.
- 40 Verho R., Putkonen M., Londesborough J., Penttilä M., and Richard P. (2004) A Novel NADH-linked L-Xylulose Reductase in the L-Arabinose Catabolic Pathway of Yeast. *J. Biol. Chem.* 279, 14746-14751.
- Walfridsson M., Bao X., Anderlund M., Lilius G., and Hahn-Hägerdahl B. (1996) Ethanol fermentation of Xylose with *Saccharomyces cerevisiae* Harboring the *Thermus thermophilus* xylA Gene, Which Expresses an Active Xylose (Glucose) Isomerase. *Appl. Environ. Microbiol.* 62, 4648-4651.
- 45 Wang P. Y. and Schneider H. (1980) Growth of yeasts on D-xylulose. *Can. J. Microbiol.* 26, 1165-1168.
- Witteveen C. F. B., Weber F., Busink R., and Visser J. (1994) Isolation and characterisation of two xylitol dehydrogenases from *Aspergillus niger*. *Microbiol.* 140, 1679-1685.
- Woodyer R., Simurdiak M., van der Donk W. A., and Zhao H. (2005) Heterologous Expression, Purification, and Characterization of a Highly Active Xylose Reductase from *Neurospora crassa*. *Appl. Environ. Microbiol.* 71, 1642-1647.

50 Aunque la invención se ha descrito en conexión con realizaciones preferidas específicas, debe entenderse que la invención según se reivindica no debe limitarse indebidamente a dichas realizaciones específicas. De hecho, se pretende que diversas modificaciones de los modos descritos para llevar a cabo la invención, que son obvias para los expertos en los campos de la bioquímica y la biología molecular, o relacionados, estén dentro del alcance de las siguientes reivindicaciones.

LISTADO DE SECUENCIAS

60 SEQ ID No 1: Región codificante del gen de la xilosa isomerasa de *Lactococcus lactis* subsp. *lactis* (*Lactobacillus xylosus*) cepa NRRL-4449 como se describe en el número de acceso AF092042. La secuencia de aminoácidos codificada por esta secuencia de nucleótidos se muestra como SEQ ID No 14.

ES 2 582 782 T3

M A Y F N D I A P I K Y E G T K
 1 ATG GCT TAC TTT AAC GAC ATC GCA CCT ATC AAA TAC GAA GGT ACA AAA
 T K N M F A F R H Y N P E E V V
 49 ACT AAA AAT ATG TTT GCC TTT CGC CAT TAT AAT CCA GAA GAA GTA GTT
 A G K T M E E Q L H F A L A F W
 97 GCT GGT AAA ACA ATG GAA GAA CAA CTT CAT TTT GCC CTT GCA TTT TGG
 H T I T M D G S D P F G G A T M
 145 CAT ACA ATT ACA ATG GAT GGG TCA GAT CCC TTT GGG GGA GCA ACA ATG
 E R P W D L E G G S E L D R A H
 193 GAA CGC CCT TGG GAT TTG GAA GGT GGT TCT GAA CTT GAC CGT GCT CAC
 R R V D A F F E I A E K L G V K
 241 CGT CGA GTA GAT GCT TTC TTT GAA ATT GCT GAA AAA TTA GGT GTT AAA
 Y Y C F H D I D I A P T G N S L
 289 TAT TAT TGT TTC CAT GAT ATT GAT ATT GCA CCT ACT GGA AAT TCT TTG
 K E F Y A N L D E I T D H L L E
 337 AAA GAA TTT TAT GCT AAT TTG GAC GAA ATT ACT GAC CAC CTT CTT GAA
 K Q K A T G I K L L W N T A N M
 385 AAA CAA AAA GCA ACA GGG ATT AAA TTA CTT TGG AAT ACA GCA AAC ATG
 F S N P R Y M N G V S T S N R A
 433 TTT TCA AAT CCC CGC TAT ATG AAT GGT GTT TCA ACT TCT AAT CGT GCT
 E V F A Y G A A Q V K K G L E L
 481 GAA GTC TTT GCT TAT GGT GCT GCA CAA GTT AAA AAA GGT CTT GAA CTT
 S K K L G G E N Y V F W G G R E
 529 TCT AAA AAA CTC GGT GGT GAA AAT TAC GTC TTC TGG GGT GGT CGT GAA
 G Y E S L L N T D M G L E M D H
 577 GGT TAT GAA TCA CTT TTG AAT ACA GAT ATG GGT CTT GAA ATG GAT CAT
 M A K F F H L A I D Y A K S I N
 625 ATG GCA AAA TTC TTC CAT TTG GCA ATT GAT TAT GCA AAA TCA ATC AAC
 H L P I F L I E P K P K E P M T
 673 CAC TTG CCC ATT TTC TTG ATT GAA CCA AAA CCA AAA GAA CCA ATG ACT
 H Q Y D F D S A T A L A F L Q K

ES 2 582 782 T3

```

721  CAC CAA TAT GAT TTT GAC TCA GCA ACA GCT CTT GCT TTC TTG CAA AAA
      Y  D  L  D  K  Y  F  K  L  N  L  E  T  N  I  A
769  TAT GAT TTG GAC AAA TAT TTC AAA CTC AAT CTT GAA ACA AAT CAT GCT
      W  L  A  G  H  T  F  E  H  E  L  N  T  A  R  T
817  TGG TTG GCT GGA CAC ACT TTT GAA CAC GAA TTA AAT ACT GCT CGT ACT
      F  N  A  L  G  S  I  D  A  N  Q  G  N  Y  L  L
865  TTC AAT GCT TTG GGT TCT ATT GAT GCC AAT CAA GGA AAT TAC TTG CTT
      G  W  D  T  D  E  F  P  T  L  V  I  D  I  T  L
913  GGT TGG GAT ACA GAT GAA TTC CCA ACA CTT GTT ATT GAT ATC ACA CTT
      A  M  H  Q  I  L  L  N  G  G  L  G  K  G  G  I
961  GCG ATG CAC CAA ATT CTT TTG AAC GGT GGA CTT GGC AAA GGT GGA ATT
      N  F  D  A  K  V  R  R  T  S  F  K  A  E  D  L
1009 AAC TTT GAT GCG AAA GTA CGT CGT ACA AGT TTC AAA GCA GAA GAT TTA
      I  L  A  H  I  A  G  M  D  T  Y  A  R  A  L  K
1057 ATT CTT GCT CAT ATT GCA GGG ATG GAT ACT TAT GCG CGT GCT TTG AAA
      G  A  A  A  I  I  E  D  K  F  L  S  D  I  V  D
1105 GGT GCA GCA GCA ATC ATT GAA GAT AAA TTC TTG TCT GAT ATT GTT GAC
      E  R  Y  S  S  Y  K  N  T  E  V  G  Q  S  I  E
1153 GAA CGT TAT AGT TCA TAC AAA AAT ACA GAA GTT GGT CAA TCC ATT GAA
      N  G  T  A  T  F  E  S  L  A  A  F  A  L  E  Y
1201 AAT GGA ACA GCA ACT TTT GAA AGT CTT GCC GCA TTT GCA CTT GAA TAT
      G  D  D  I  E  L  D  S  N  H  L  E  Y  I  K  S
1249 GGT GAT GAT ATT GAA CTT GAT TCT AAT CAC TTG GAA TAC ATC AAA TCA
      V  L  N  D  Y  L  V  *
1297 GTA TTG AAT GAC TAT CTT GTT TAA

```

SEQ ID No. 2:

GCTAGCCATGGCTTACTTTAACGACATCGCACCTATC

SEQ ID No. 3:

CTCGAGCCTAGGCTAAACAAGATAGTCATTCAATACTGATTTG

SEQ ID No. 4:

GCTAGCCATGGCCACTACCCCATTTGATGCTCCAGATAAG

SEQ ID No. 5:

CTCGAGCCTAGGCTAGTGTTCATTTCAATTCATTTCCATCTTGGCC

SEQ ID No 6:

5 Gen de la xilosa isomerasa artificial de *Thermoanaerobacter thermohydrosulfuricus*, con regiones flanqueantes N-terminal y C-terminal. La región codificante se basa en la secuencia de aminoácidos de código de acceso de GenBank D00756. La secuencia de aminoácidos codificada por la secuencia de nucleótidos codificante se muestra como SEQ ID No 15.

ES 2 582 782 T3

1 GAATTCCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGGAGCTAGCC
 M E Y F K N V P Q I K Y E G P K
 49 ATG GAA TAT TTC AAA AAC GTG CCA CAG ATC AAG TAT GAA GGT CCT AAA
 S N N P Y A F K F Y N P D E I I
 97 AGC AAT AAC CCT TAT GCA TTT AAG TTC TAT AAC CCA GAT GAA ATT ATA
 D G K P L K E H L R F S V A Y W
 145 GAT GGA AAA CCA TTA AAA GAA CAC TTA AGA TTT AGC GTA GCC TAC TGG
 H T F T A N G T D P F G A P T M
 193 CAT ACA TTT ACC GCT AAC GGA ACG GAT CCA TTT GGT GCA CCG ACT ATG
 Q R P W D H F T D P M D I A K A
 241 CAG CGT CCT TGG GAT CAT TTT ACC GAC CCT ATG GAC ATA GCA AAA GCA
 R V E A A F E L F E K L D V P F
 289 CGT GTG GAA GCC GCA TTC GAG CTT TTT GAA AAA TTG GAT GTT CCA TTC
 F C F H D R D I A P E G E T L R
 337 TTC TGT TTT CAT GAC AGA GAT ATA GCT CCG GAA GGT GAA ACA TTG AGA
 E T N K N L D T I V A M I K D Y
 385 GAA ACC AAC AAA AAC TTA GAT ACT ATC GTT GCT ATG ATT AAA GAC TAC
 L K T S K T K V L W G T A N L F
 433 TTA AAA ACG TCA AAG ACT AAA GTT CTT TGG GGC ACT GCT AAT TTG TTT
 S N P R F V H G A A T S C N A D
 481 TCT AAT CCA CGT TTC GTG CAT GGC GCT GCC ACA TCA TGT AAT GCA GAC
 V F A Y A A A Q V K K A L E I T
 529 GTA TTT GCT TAT GCA GCC GCT CAA GTT AAA AAG GCC TTA GAG ATT ACC
 K E L G G Q N Y V F W G G R E G
 577 AAA GAG TTA GGA GGC CAG AAT TAT GTT TTC TGG GGT GGT CGT GAG GGA
 Y E T L L N T D M E L E L D N L
 625 TAT GAG ACA CTT TTA AAT ACT GAT ATG GAG TTG GAA TTA GAT AAT TTA
 A R F L H M A V E Y A Q E I G F
 673 GCA AGA TTC TTA CAC ATG GCA GTA GAA TAT GCT CAG GAA ATT GGT TTT
 E G Q F L I E P K P K E P T K H

```

721  GAA GGA CAG TTC TTG ATC GAG CCT AAA CCA AAG GAA CCA ACA AAG CAT
      Q  Y  D  F  D  A  A  S  V  H  A  F  L  K  K  Y
769  CAG TAT GAT TTC GAC GCT GCT TCT GTA CAC GCC TTT TTG AAG AAG TAT
      D  L  D  K  Y  F  K  L  N  I  E  A  N  H  A  T
817  GAT TTG GAT AAA TAC TTT AAG TTG AAC ATA GAG GCT AAT CAC GCA ACG
      L  A  G  H  D  F  Q  H  E  L  R  Y  A  R  I  N
865  TTG GCA GGT CAC GAT TTT CAA CAC GAA TTG AGA TAC GCC CGT ATT AAT
      N  M  L  G  S  I  D  A  N  M  G  D  M  L  L  G
913  AAC ATG TTA GGT TCC ATA GAT GCC AAC ATG GGT GAC ATG TTG CTG GGT
      W  D  T  D  Q  Y  P  T  D  I  R  M  T  T  L  A
961  TGG GAT ACT GAT CAA TAC CCA ACG GAT ATT AGA ATG ACA ACT TTA GCA
      M  Y  E  V  I  K  M  G  G  F  N  K  G  G  L  N
1009 ATG TAC GAG GTC ATT AAA ATG GGA GGT TTT AAC AAA GGA GGT TTG AAT
      F  D  A  K  V  R  R  A  S  F  E  P  E  D  L  F
1057 TTC GAT GCT AAA GTG CGT CGT GCC TCT TTT GAA CCT GAA GAC CTT TTT
      L  G  H  I  A  G  M  D  A  F  A  K  G  F  K  V
1105 CTT GGA CAT ATT GCC GGA ATG GAT GCA TTT GCA AAA GGT TTC AAG GTC
      A  Y  K  L  V  K  D  G  V  F  D  R  F  I  E  E
1153 GCT TAT AAG CTT GTT AAG GAT GGT GTA TTT GAT AGA TTC ATT GAA GAG
      R  Y  K  S  Y  R  E  G  I  G  A  E  I  V  S  G
1201 AGA TAC AAA TCC TAT CGT GAA GGT ATA GGT GCT GAA ATC GTT TCA GGT
      K  A  N  F  K  T  L  E  E  Y  A  L  N  N  P  K
1249 AAG GCC AAT TTT AAG ACT TTA GAG GAA TAT GCA TTG AAT AAC CCA AAA
      I  E  N  K  S  G  K  Q  E  L  L  E  S  I  L  N
1297 ATC GAA AAC AAA AGC GGT AAA CAG GAA CTG CTG GAA TCT ATT TTG AAT
      Q  Y  L  F  S  E  *
1345 CAA TAT TTG TTC TCT GAA TAG CCTAGGCTCGAGGAATTC
    
```

SEQ ID No 7:

Gen de la xilosa isomerasa artificial de *Thermoanaerobacter thermosulphurigenes*, con regiones flanqueantes N-terminal y C-terminal. La región codificante se basa en la secuencia de aminoácidos de código de acceso de GenBank J05650, a excepción del segundo aminoácido que se ha cambiado de N a A, con el fin de permitir el uso de la "secuencia de consenso de Kozak" para la expresión eucariota. *C. thermosulfurogenes* es un nombre alternativo a *Thermoanaerobacter thermosulphurigenes* (Lee *et al* 1993). La secuencia de aminoácidos codificada por la región codificante de esta secuencia de nucleótidos se muestra como SEQ ID No 16.

10

ES 2 582 782 T3

1 GAATTCCTAGAAATAATTTTGTTTAACTTTAAGAAGGAGGAGCTAGCC
 M A K Y F E N V S K I K Y E G P
 49 ATG GCC AAG TAT TTT GAG AAT GTT TCC AAG ATT AAG TAT GAA GGT CCT
 K S N N P Y S F K F Y N P E E V
 97 AAG TCA AAC AAC CCT TAC TCC TTC AAA TTC TAT AAT CCA GAA GAA GTT
 I D G K T M E E H L R F S I A Y
 145 ATA GAC GGT AAA ACG ATG GAG GAA CAC TTA AGA TTT TCT ATT GCT TAC
 W H T F T A D G T D Q F G K A T
 193 TGG CAC ACA TTT ACT GCC GAC GGC ACA GAC CAA TTC GGA AAG GCT ACT
 M Q R P W N H Y T D P M D I A K
 241 ATG CAA AGA CCG TGG AAC CAT TAC ACT GAT CCA ATG GAC ATA GCC AAA
 A R V E A A F E F F D K I N A P
 289 GCC AGA GTG GAG GCT GCA TTC GAG TTC TTC GAT AAG ATA AAC GCC CCT
 Y F C F H D R D I A P E G D T L
 337 TAC TTC TGC TTC CAT GAT CGT GAC ATT GCT CCG GAA GGC GAT ACC TTG
 R E T N K N L D T I V A M I K D
 385 AGA GAA ACC AAC AAA AAT CTT GAC ACC ATT GTC GCA ATG ATA AAA GAT
 Y L K T S K T K V L W G T A N L
 433 TAT TTG AAG ACG TCT AAA ACC AAA GTT TTG TGG GGT ACG GCA AAC TTA
 F S N P R F V H G A S T S C N A
 481 TTT TCT AAT CCG AGA TTT GTT CAT GGT GCT TCC ACA TCC TGC AAC GCA
 D V F A Y S A A Q V K K A L E I
 529 GAT GTC TTT GCT TAT TCC GCT GCT CAA GTC AAA AAG GCT CTT GAG ATA
 T K E L G G E N Y V F W G G R E
 577 ACC AAA GAA TTA GGT GGT GAA AAC TAC GTT TTC TGG GGT GGC AGA GAA
 G Y E T L L N T D M E F E L D N
 625 GGT TAT GAG ACT CTT TTA AAT ACA GAT ATG GAA TTT GAA CTT GAC AAT
 F A R F L H M A V D Y A K E I G
 673 TTC GCA AGA TTC TTG CAC ATG GCT GTC GAT TAT GCC AAG GAG ATA GGT
 F E G Q F L I E P K P K E P T K
 721 TTT GAA GGC CAG TTC TTG ATT GAA CCG AAA CCA AAG GAA CCT ACC AAA
 H Q Y D F D V A N V L A F L R K
 769 CAT CAG TAT GAC TTT GAT GTC GCA AAT GTT TTG GCT TTC TTG AGA AAA
 Y D L D K Y F K V N I E A N H A
 817 TAC GAT TTA GAT AAG TAT TTC AAA GTA AAT ATT GAA GCT AAT CAT GCA
 T L A F H D F Q H E L R Y A R I
 865 ACG TTG GCC TTC CAT GAT TTC CAG CAC GAG TTG AGA TAC GCT AGA ATC
 N G V L G S I D A N T G D M L L
 913 AAT GGA GTC TTA GGA TCT ATC GAT GCT AAT ACC GGT GAT ATG CTG CTG

ES 2 582 782 T3

G W D T D Q F P T D I R M T T L
 961 GGA TGG GAT ACT GAT CAA TTT CCG ACC GAT ATT CGT ATG ACT ACC CTG
 A M Y E V I K M G G F D K G G L
 1009 GCA ATG TAT GAA GTT ATT AAG ATG GGT GGA TTT GAT AAA GGC GGT CTG
 N F D A K V R R A S F E P E D L
 1057 AAC TTC GAT GCA AAA GTA AGA AGA GCT TCT TTT GAA CCT GAA GAT TTG
 F L G H I A G M D A F A K G F K
 1105 TTT TTA GGA CAC ATC GCT GGC ATG GAC GCA TTT GCT AAA GGT TTT AAG
 V A Y K L V K D R V F D K F I E
 1153 GTA GCC TAT AAG CTT GTA AAA GAT AGA GTG TTC GAC AAA TTC ATC GAA
 E R Y A S Y K D G I G A D I V S
 1201 GAG AGA TAT GCT TCT TAT AAG GAC GGA ATA GGA GCC GAT ATA GTT TCC
 G K A D F R S L E K Y A L E R S
 1249 GGT AAG GCC GAT TTC AGA TCT CTT GAG AAA TAC GCC TTG GAA AGA TCA
 Q I V N K S G R Q E L L E S I L
 1297 CAA ATC GTG AAC AAA TCA GGC CGT CAA GAA TTG TTA GAA TCA ATT CTT
 N Q Y L F A E *
 1345 AAT CAA TAC CTG TTC GCT GAA TAG CCTAGGCTCGAGGAATTC

SEQ ID No. 8:

TAGCTAGCATGTCGTACTIONTCCCCACTGTCGAC

SEQ ID No. 9:

ATAAGCTTTCAGGTGTAGATAAAGCGGTTGACC

SEQ ID No 10:

Gen de la xilosa isomerasa de *Lactococcus lactis* subsp. *lactis*, cepa KF147. Basado en el número de acceso de GenBank EU255918. La secuencia de aminoácidos codificada por esta secuencia de nucleótidos se muestra como SEQ ID No 11.

5

1 atg gct tac ttt aac gac atc gca cct atc aaa tac gaa ggt aca aaa
 49 act aaa aat atg ttt gcc ttt cgt cat tat aat cca gaa gaa gta gtt
 97 gct ggt aaa aca atg gaa gaa caa ctt cat ttt gcc ctt gca ttt tgg
 145 cat aca att acg atg gat ggg tca gat ccc ttt ggg gga gca aca atg
 193 gaa cgt cct tgg gat ttg gaa ggt ggt tct gaa ctt gac cgt gct cac
 241 cgt cga gta gat gct ttc ttt gaa att gct gaa aaa tta ggt gtt aaa
 289 tat tat tgt ttc cat gat att gat att gca cct act gga aat tct ttg

ES 2 582 782 T3

337 aaa gaa ttt tat gct aat ttg gac gaa att act gac cac ctt ctt gaa
 385 aaa caa aaa gca aca ggc alt aaa tta ctt tgg aat aca gca aac atg
 433 ttt tca aat ccc cgc tat atg aat ggt gtt tca act tct aat cgt gct
 481 gaa gtc ttt gct tat ggt gct gca caa gtt aaa aaa ggt ctt gaa ctt
 529 tct aaa aaa ctc ggt ggt gaa aat tat gtc ttc tgg ggt ggt cgt gaa
 577 ggt tat gaa tca ctt ttg aat aca gat atg ggt ctt gaa atg gat cat
 625 atg gca aaa ttc ttc cat ttg gca att gat tat gca aaa tca atc aac
 673 cac ttg cct att ttc ttg att gaa cca aaa cca aaa gaa cca atg act
 721 cac caa tat gat ttt gac tca gca aca gct ctt gct ttc ttg caa aaa
 769 tat gac ttg gac aaa tac ttc aaa ctc aat ctt gaa aca aat cat gct
 817 tgg ttg gct ggg cac act ttt gaa cac gaa tta aat act gca cgt act
 865 ttc aat gct ttg ggt tct att gat gcc aat caa gga aat tac ttg ctt
 913 ggt tgg gat aca gat gaa ttc cca aca ctt gtt att gat atc aca ctt
 961 gcg atg cac caa att ctt ttg aac ggt gga ctt ggc aaa ggt gga att
 1009 aac ttt gat gcg aaa gta cgt cgt aca agt ttc aaa gca gaa gat tta
 1057 att ctt gct cat att gca ggg atg gat act tat gcg cgt gct ttg aaa
 1105 ggt gca gca gca atc att gaa gat aaa ttc ttg tct gat att gtt gac
 1153 gaa cgt tat agt tca tac aaa aat aca gaa gtt ggt caa tcc att gaa
 1201 aat gga aca gca act ttt gaa agt ctt gct gca ttt gca ctt gaa cat
 1249 ggt gac gat att gaa ctt gat tct aat cac ttg gaa tac atc aaa tca
 1297 gta ttg aat gac tat ctt gtt taa

SEQ ID No 11:

5 Gen de la xilosa isomerasa de *Lactococcus lactis* subsp. *lactis*, cepa KF147. Basado en el número de acceso de GenBank ABX75758. Secuencia de aminoácidos codificada por la secuencia de nucleótidos de SEQ ID No 10.

MAYFNDJAPIKYEGTKTKNMFAFRHYNPPEVVAGKTMEEQLHFALAFWHTITMDGSDPFG
 GATMERPWDLEGGSELDRARRVDAFFEIAEKLGVKYYCFHDIDIAPTGNSLKEFYANLD
 EITDHLLEKQKATGIKLLWNTANMFSNPRYMNGVSTSNRAEVFAYGAAQVKKGLELSKKL
 GGENYVFWGGREGYESLLNTDMGLEMDHMAKFFHLAIDYAKSINHLPFLIEPKPKPMT
 HQYDFDSATALAFLQKYDLKYFKLNLETNHAWLAGHTFEHELNTARTFNALGSIDANQG
 NYLLGWDTDEFPTLVIDITLAMHQILLNGGLGKGGINFDAKVRRTSFKAEDLILAHIAGM
 DTYARALKGAAAI IEDKFLSDJVDERYSSYKNTVEVGQSIENGTATFESLAAFALEHGDDI
 ELDSNHLEYIKSVLNDYLV

SEQ ID No 12:

10 Gen de la xilosa isomerasa de *Lactococcus lactis* subsp. *lactis*, cepa IO-1. Basado en el número de acceso de GenBank AF092041. La secuencia de aminoácidos codificada por esta secuencia de nucleótidos se muestra como SEQ ID No 13.

ES 2 582 782 T3

1 atg gct tac ttt aac gac atc gca cct atc aaa tac gaa ggt aca aaa
 49 act aaa aat atg ttt gcc ttt cgc cat tat aat cca gaa gaa gta gtt
 97 gct ggt aaa aca atg gaa gaa caa ctt cat ttt gcc ctt gca ttt tgg
 145 cat aca att aca atg gat ggg tca gac ccc ttt ggg gga gca aca atg
 193 gaa cgt cct tgg gat ttg gaa ggt ggt tct gaa ctt gac cgt gct cac
 241 cgt cga gta gat gct ttc ttt gaa att gct gaa aaa tta ggt gtt aaa
 289 tat tat tgt ttc cat gat att gat att gca cct act gga aat tct ttg
 337 aaa gaa ttt tat gct aat ttg gac gaa att act gac cac ctt ctt gaa
 385 aaa caa aaa gca aca ggg att aaa tta ctt tgg aat aca gca aac atg
 433 ttt tca aat ccc cgc tat atg aat ggt gtt tca act tct aac cgt gct
 481 gaa gtc ttt gct tat ggt gct gca caa gtt aaa aaa ggt ctt gaa ctt
 529 tct aaa aaa ctc ggt ggt gaa aat tac gtc ttc tgg ggt ggt cgt gaa
 577 ggt tat gaa tca ctt ttg aat aca gat atg ggt ctt gaa atg gat cat
 625 atg gca aaa ttc ttc cat ttg gca att gat tat gca aaa tca atc aac
 673 cac ttg ccc att ttc ttg atc gaa cca aaa cca aaa gaa cca atg act
 721 cac caa tat gat ttt gac tca gca aca gct ctt gct ttc ttg caa aaa
 769 tat gat ttg gac aaa tat ttc aaa ctc aat ctt gaa aca aat cat gct
 817 tgg ttg gct gga cac act ttt gaa cac gaa tta aat act gct cgt act
 865 ttc aat gct ttg ggt tct att gat gcc aat caa gga aat tac ttg ctt
 913 ggt tgg gat aca gat gaa ttc cca aca ctt gtt att gat atc aca ctt
 961 gcg atg cac caa att ctt ttg aac ggt gga ctt ggc aaa ggt gga att
 1009 aac ttt gat gcg aaa gta cgt cgt aca agt ttc aaa gca gaa gat tta
 1057 att ctt gct cat att gca ggg atg gat act tat gcg cgt gct ttg aaa
 1105 ggt gca gca gca atc att gaa gat aaa ttc ttg tct gat att gtt gac
 1153 gaa cgt tat agt tca tac aaa aat aca gaa gtt gga caa tcc att gaa
 1201 aat gga aca gca act ttt aaa agt ctt gcc gca ttt gca ctt gaa cat
 1249 ggt gac gat att gaa ctt gat tct aat cac ttg gaa tac atc aaa tca
 1297 gta ttg aat gac tat ctt gtt taa

SEQ ID No 13:

Gen de la xilosa isomerasa de *Lactococcus lactis* subsp. *lactis*, cepa IO-1. Basado en el número de acceso de GenBank AAD20249. La secuencia de aminoácidos codificada por la secuencia de nucleótidos de SEQ ID No 12. Los aminoácidos que difieren respecto a la SEQ ID No 18 están subrayados.

5

MAYFN⁵DIAP¹IKYEGTKTKNMFAFRHYNPEEVVAGKTMEEQLHFALAFWHTITMDGSDPFG
 GATMERPWDLEGGSELDRAHRRVDAFFEIAEKLGVKY¹YCFHDIDIAPTGNSLKEFYANLD
 EITDHLLEKQKATGIKLLWNTANMFSNPRYMNGVSTSNRAEVFAYGAAQVKKGLELSKKL
 GGENYVFWGGREGYESLLNTDMGLEMDHMAKFFHLAIDYAKSINHLPIFLIEPKPKPMT
 HQYDFDSATALAFLOKYDLDKYFKLNLETNHAWLAGHTFEHELNTARTFNALGSIDANQG
 NYLLGWD¹TDEFPTLVIDITFLAMHQILLNGGLGKGGINFD¹AKVRRTSFKAEDLILAHIAGM
 DTYARALKGAAAI¹IEDKFLSDIVDERYSSYK¹NTEVQGSIENGTATFKSLAAFALEHGDDI
 ELDSNHLEYIKSVLNDYLV

10 SEQ ID No 14:

La secuencia de aminoácidos codificada por la secuencia de nucleótidos de SEQ ID No 1. Los aminoácidos que difieren respecto a la SEQ ID No 18 están subrayados.

M A Y F N D I A P I K Y E G T K
T K N M F A F R H Y N P E E V V
A G K T M E E Q L H F A L A F W
H T I T M D G S D P F G G A T M
E R P W D L E G G S E L D R A H
R R V D A F F E I A E K L G V K
Y Y C F H D I D I A P T G N S L
K E F Y A N L D E I T D H L L E
K Q K A T G I K L L W N T A N M
F S N P R Y M N G V S T S N R A
E V F A Y G A A Q V K K G L E L
S K K L G G E N Y V F W G G R E
G Y E S L L N T D M G L E M D H
M A K F F H L A I D Y A K S I N
H L P I F L I E P K P K E P M T
H Q Y D F D S A T A L A F L Q K
Y D L D K Y F K L N L E T N H A
W L A G H T F E H E L N T A R T
F N A L G S I D A N Q G N Y L L
G W D T D E F P T L V I D I T L
A M H Q I L L N G G L G K G G I
N F D A K V R R T S F K A E D L

I L A H I A G M D T Y A R A L K
G A A A I I E D K F L S D I V D
E R Y S S Y K N T E V G Q S I E
N G T A T F E S L A A F A L E Y
G D D I E L D S N H L E Y I K S
V L N D Y L V

SEQ ID No 15:

- 5 Número de acceso de GenBank BAA00652. La secuencia de aminoácidos codificada por la región codificante de la secuencia de nucleótidos de SEQ ID No 6.

ES 2 582 782 T3

M E Y F K N V P Q I K Y E G P K
 S N N P Y A F K F Y N P D E I T
 D G K P L K E H L R F S V A Y W
 H T F T A N G T D P F G A P T M
 Q R P W D H F T D P M D I A K A
 R V E A A F E L F E K L D V P F
 F C F H D R D I A P E G E T L R
 E T N K N L D T I V A M I K D Y
 L K T S K T K V L W G T A N L F
 S N P R F V H G A A T S C N A D
 V F A Y A A A Q V K K A L E I T
 K E L G G Q N Y V F W G G R E G
 Y E T L L N T D M E L E L D N L
 A R F L H M A V E Y A Q E I G F
 E G Q F L I E P K P K E P T K H
 Q Y D F D A A S V H A F L K K Y
 D L D K Y F K L N T E A N H A T
 L A G H D F Q H E L R Y A R I N
 N M L G S I D A N M G D M L L G
 W D T D Q Y P T D I R M T T L A
 M Y E V I K M G G F N K G G L N
 F D A K V R R A S F E P E D L F
 L G H I A G M D A F A K G F K V
 A Y K L V K D G V F D R F I E E
 R Y K S Y R E G I G A E I V S G
 K A N F K T L E E Y A L N N P K
 I E N K S G K Q E L L E S I L N

Q Y L F S E

SEQ ID No 16:

Número de acceso de GenBank AAA23285. La secuencia de aminoácidos codificada por la región codificante de la secuencia de nucleótidos de SEQ ID No 7.

5

ES 2 582 782 T3

M A K Y F E N V S K I K Y E G P
K S N N P Y S F K F Y N P E E V
I D G K T M E E H L R F S I A Y
W H T F T A D G T D Q F G K A T
M Q R P W N H Y T D P M D I A K
A R V E A A F E F F D K I N A P
Y F C F H D R D I A P E G D T L
R E T N K N L D T I V A M I K D
Y L K T S K T K V L W G T A N L
F S N P R F V H G A S T S C N A
D V F A Y S A A Q V K K A L E I
T K E L G G E N Y V F W G G R E
G Y E T L L N T D M E F E L D N
F A R F L H M A V D Y A K E I G
F E G Q F L I E P K P K E P T K
H Q Y D F D V A N V L A F I R K
Y D L D K Y F K V N I E A N H A
T L A F H D F Q H E L R Y A R I
N G V L G S I D A N T G D M L L
G W D T D Q F P T D I R M T T L
A M Y E V I K M G G F D K G G L
N F D A K V R R A S F E P E D L
F L G H I A G M D A F A K G F K
V A Y K L V K D R V F D K F I E
E R Y A S Y K D G I G A D I V S
G K A D F R S L E K Y A L E R S
Q I V N K S G R Q E L L E S I L
N Q Y L F A E

SEQ ID No 17:

Región codificante del gen de la xilosa isomerasa clonado por PCR de la cepa designada como DSM 20175 de *Lactococcus lactis*, usando los cebadores marcados en la SEQ ID No 2 y SEQ ID No 3.

5

1 atg gct tac ttt aac gac atc gca cct atc aaa tac gaa ggt aca aaa
 49 act aaa aat atg ttt gcc ttt cgc cat tat aat cca gaa gaa gta gtt
 97 gct ggt aaa aca atg gaa gaa caa ctt cat ttt gcc ctt gca ttt tgg
 145 cat aca att aca atg gat ggg tca gat ccc ttt ggg gga gca aca atg
 193 gaa cgc cct tgg gat ttg gaa ggt ggt tct gaa ctt gac cgt gct cac
 241 cgt cga gta gat gct ttc ttt gaa att gct gaa aaa tta ggt gtt aaa
 289 tat tat tgt ttc cat gat att gat att gca cct act gga aat tct ttg
 337 aaa gaa ttt tat gct aat ttg gac gaa att act gac cac ctt ctt gaa
 385 aaa caa aaa gca aca ggg att aaa tta ctt tgg aat aca gca aac atg
 433 ttt tca aat ccc cgc tat atg aat ggt gtt tca act tct aat cgt gct
 481 gaa gtc ttt gct tat ggt gct gca caa gtt aaa aaa ggt ctt gaa ctt
 529 tct aaa aaa ctc ggt ggt gaa aat tac gtc ttc tgg ggt ggt cgt gaa
 577 ggt tat gaa tca ctt ttg aat aca gat atg ggt ctt gaa atg gat cat
 625 atg gca aaa ttc ttc cat ttg gca att gat tat gca aaa tca atc aac
 673 cac ttg ccc att ttc ttg att gaa cca aaa cca aaa gaa cca atg act
 721 cac caa tat gat ttt gac tca gca aca gct ctt gct ttc ttg caa aaa
 769 tat gat ttg gac aaa tat ttc aaa ctc aat ctt gaa aca aat cat gct
 817 tgg ttg gct gga cac act ttt gaa cac gaa tta aat act gct cgt act
 865 ttc aat gct ttg ggt tct att gat gcc aat caa gga aat tac ttg ctt
 913 ggt tgg gat aca gat gaa ttc cca aca ctt gtt att gat atc aca ctt
 961 gcg atg cac caa att ctt ttg aac ggt gga ctt gcc aaa ggt gga att
 1009 aac ttt gat gcg aaa gta cgt cgt aca agt ttc aaa gca gaa gat tta
 1057 att ctt gct cat att gca ggg atg gat act tat gcg cgt gct ttg aaa
 1105 ggt gca gca gca atc att gaa gat aaa ttc ttg tct gat att gtt gac
 1153 gaa cgt tat agt tca tac aga aat aca gaa gtt ggt caa tcc att gaa
 1201 aat gga aca gca act ttt gaa agt ctt gcc gca ttt gca ctt gaa tat
 1249 ggt gat gat att gaa ctt gat tct aat cac ttg gaa tac atc aaa tca
 1297 gta ttg aat gac tat ctt gtt taa

SEQ ID No 18:
 Secuencia de aminoácidos codificada por la SEQ ID No 17.

5

MAYFNDIAPIKYEGTKTKNMFAFRHYNPEEVVAGKTMEEQLHFALAFWHTITM
 DGSDPFGGATMERPWDLEGGSELDRAHRRVDAFFEIAEKLGVKYYCFHDDIAP
 TGNSLKEFYANLDEITDHILLEKQKATGIKLLWNTANMFSNPRYMNGVSTSNRAE
 VFAYGAAQVKKGLELSKKGGENYVFWGGREGYESLLNTDMGLEMDHMAKFF

HLAIDYAKSINHLPIFLIEPKPKPEMTHQYDFDSATALAFLQKYDLDKYFKLNLET
 NHAWLAGHTFEHELNTARTFNALGSIDANQGNYLLGWDTDEFPTLVIDITLAMH
 QILLNGGLGKGGINFDAKVRRTSFKAEIDLILAHAGMDTYARALKGAAAIIEDKF
 LSDIVDERYSSYRNTEVVGQSIENGTATFESLAFALEYGDDIELDSNHLEYIKSVL
 NDYLV

SEQ ID No 19:

Variante artificial de la xilosa isomerasa de *Lactococcus*, secuencia de aminoácidos (SEQ ID No 18). Los aminoácidos que difieren respecto a la SEQ ID No 18 están subrayados. La secuencia de nucleótidos que codifica esta secuencia de aminoácidos se muestra como SEQ ID No 27

5

MAYFN⁵NDI¹⁰APIKYEGTKTKNMFAFRHYNPEEVVAGKTMEEQLHFALAFWHTITMDGSDPFG
 GATMERPW¹⁵DLEGGSELDRARRVD²⁰AFFEIAEKLGVKYYCFHDI²⁵DIAPTGN³⁰SLKEFYANLD
 EITDHLLEKQKATGIKLLWNTANMFSNPRYMNGVSTSNRAEV³⁵FAYGAAQVKKGLELSKKL
 GGENYVFWGGREGYESLLNTDMGLEMDHMAKFFHLAIDYAKSINHLPIFLIEPKPKPMT
 HQYDFDSATALAFLQKYDL⁴⁰DKYFKLNLETNHAWLAGHTFEHELNTARTFNALGSIDANQG
 NYLLGWD⁴⁵TDEFPTLV⁵⁰IDI⁵⁵TLAMHQILLNGGLGKGGINFDAKVRRTSFKAE⁶⁰D⁶⁵LILAHIAGM
 DTYARAL⁷⁰KGAAAI⁷⁵IEDKFLSDIVDERYSSYRNTEV⁸⁰QSIENGTATFKSLAAFALEHGDDI
 ELDSNHLEYIKSVLNDYLV

SEQ ID No 20:

Variante artificial de la xilosa isomerasa de *Lactococcus*, secuencia de aminoácidos (SEQ ID No 18). Los aminoácidos que difieren respecto a la SEQ ID No 18 están subrayados. La secuencia de nucleótidos que codifica esta secuencia de aminoácidos se muestra como SEQ ID No 28

10

MAYFN⁵NDI¹⁰APIKYEGTKTKNMFAFRHYNPEEVVAGKTMEEQLHFALAFWHTITMDGSDPFG
 GATMERPW¹⁵DLEGGSELDRARRVD²⁰AFFEIAEKLGVKYYCFHDI²⁵DIAPTGN³⁰SLKEFYANLD
 EITDHLLEKQKATGIKLLWNTANMFSNPRYMNGVSTSNRAEV³⁵FAYGAAQVKKGLELSKKL
 GGENYVFWGGREGYESLLNTDMGLEMDHMAKFFHLAIDYAKSINHLPIFLIEPKPKPMT
 HQYDFDSATALAFLQKYDL⁴⁰DKYFKLNLETNHAWLAGHTFEHELNTARTFNALGSIDANQG
 NYLLGWD⁴⁵TDEFPTLV⁵⁰IDI⁵⁵TLAMHQILLNGGLGKGGINFDAKVRRTSFKAE⁶⁰D⁶⁵LILAHIAGM
 DTYARAL⁷⁰KGAAAI⁷⁵IEDKFLSDIVDERYSSYRNTEV⁸⁰QSIENGTATFKSLAAFALEYGDDI
 ELDSNHLEYIKSVLNDYLV

15 SEQ ID No 21:

Cebador sentido R391K:

GTTGACGAACGATATAGTTCATACAAAAATACAGAAGTTGG

20 SEQ ID No 22:

Cebador antisentido R391K:

CCA⁵ACTTCTGTATTTTTGTATGAACTATATCGTTCGTC¹⁰CAAC

SEQ ID No 23:

Cebador sentido E407K:

25

ACAGCAACTTTTAAAAGCTTAGCCGCATTTGCACTTGAATATGGTGATGATATTG

SEQ ID No 24:

Cebador antisentido E407K:

30

CAATATCATCACCATATTCAAGTGCAAATGCGGCTAAGCTTTTAAAAGTTGCTGT

SEQ ID No 25:

Cebador sentido E407K-Y416H:

ACAGCAACTTTTAAAAGCTTAGCCGCATTTGCACTTGAACATGGTGATGATATTG

35 SEQ ID No 26:

Cebador antisentido E407K-Y416H:

CAATATCATCACCATGTTCAAGTGCAAATGCGGCTAAGCTTTTAAAAGTTGCTGT

ES 2 582 782 T3

SEQ ID No 27:

La secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada como SEQ ID No 19

```
1  atg gct tac ttt aac gac atc gca cct atc aaa tac gaa ggt aca aaa
49  act aaa aat atg ttt gcc ttt cgc cat tat aat cca gaa gaa gta gtt
97  gct ggt aaa aca atg gaa gaa caa ctt cat ttt gcc ctt gca ttt tgg
145 cat aca att aca atg gat ggg tca gat ccc ttt ggg gga gca aca atg
193 gaa cgc cct tgg gat ttg gaa ggt ggt tct gaa ctt gac cgt gct cac
241 cgt cga gta gat gct ttc ttt gaa att gct gaa aaa tta ggt gtt aaa
289 tat tat tgt ttc cat gat att gat att gca cct act gga aat tct ttg
337 aaa gaa ttt tat gct aat ttg gac gaa att act gac cac ctt ctt gaa
385 aaa caa aaa gca aca ggg att aaa tta ctt tgg aat aca gca aac atg
433 ttt tca aat ccc cgc tat atg aat ggt gtt tca act tct aat cgt gct

481 gaa gtc ttt gct tat ggt gct gca caa gtt aaa aaa ggt ctt gaa ctt
529 tct aaa aaa ctc ggt ggt gaa aat tac gtc ttc tgg ggt ggt cgt gaa
577 ggt tat gaa tca ctt ttg aat aca gat atg ggt ctt gaa atg gat cat
625 atg gca aaa ttc ttc cat ttg gca att gat tat gca aaa tca atc aac
673 cac ttg ccc att ttc ttg att gaa cca aaa cca aaa gaa cca atg act
721 cac caa tat gat ttt gac tca gca aca gct ctt gct ttc ttg caa aaa
769 tat gat ttg gac aaa tal ttc aaa ctc aat ctt gaa aca aat cat gct
817 tgg ttg gct gga cac act ttt gaa cac gaa tta aat act gct cgt act
865 ttc aat gct ttg ggt tct att gat gcc aat caa gga aat tac ttg ctt
913 ggt tgg gat aca gat gaa ttc cca aca ctt gtt att gat atc aca ctt
961 gcg atg cac caa att ctt ttg aac ggt gga ctt ggc aaa ggt gga att
1009 aac ttt gat gcg aaa gta cgt cgt aca agt ttc aaa gca gaa gat tta
1057 att ctt gct cat att gca ggg atg gat act tat gcg cgt gct ttg aaa
1105 ggt gca gca gca atc att gaa gat aaa ttc ttg tct gat att gtt gac
1153 gaa cgt tat agt tca tac aga aat aca gaa gtt ggt caa tcc att gaa
1201 aat gga aca gca act ttt aaa agc tta gcc gca ttt gca ctt gaa cat
1249 ggt gat gat att gaa ctt gat tct aat cac ttg gaa tac atc aaa tca
1297 gta ttg aat gac tat ctt gtt taa
```

5

SEQ ID No 28:

La secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada como SEQ ID No 20

ES 2 582 782 T3

```

1  atg gct tac ttt aac gac atc gca cct atc aaa tac gaa ggt aca aaa
49 act aaa aat atg ttt gcc ttt cgc cat tat aat cca gaa gaa gta gtt.
97 gct ggt aaa aca atg gaa gaa caa ctt cat ttt gcc ctt gca ttt tgg
145 cat aca att aca atg gat ggg tca gat ccc ttt ggg gga gca aca atg
193 gaa cgc cct tgg gat ttg gaa ggt ggt tct gaa ctt gac cgt gct cac
241 cgt cga gta gat gct ttc ttt gaa att gct gaa aaa tta ggt gtt aaa
289 tat tat tgt ttc cat gat att gat att gca cct act gga aat tct ttg
337 aaa gaa ttt tat gct aat ttg gac gaa att act gac cac ctt ctt gaa
385 aaa caa aaa gca aca ggg att aaa tta ctt tgg aat aca gca aac atg
433 ttt tca aat ccc cgc tat atg aat ggt gtt tca act tct aat cgt gct
481 gaa gtc ttt gct tat ggt gct gca caa gtt aaa aaa ggt ctt gaa ctt
529 tct aaa aaa ctc ggt ggt gaa aat tac gtc ttc tgg ggt ggt cgt gaa
577 ggt tat gaa tca ctt ttg aat aca gat atg ggt ctt gaa atg gat cat
625 atg gca aaa ttc ttc cat ttg gca att gat tat gca aaa tca atc aac
673 cac ttg ccc att ttc ttg att gaa cca aaa cca aaa gaa cca atg act
721 cac caa tat gat ttt gac tca gca aca gct ctt gct ttc ttg caa aaa

769 tat gat ttg gac aaa tat ttc aaa ctc aat ctt gaa aca aat cat gct
817 tgg ttg gct gga cac act ttt gaa cac gaa tta aat act gct cgt act
865 ttc aat gct ttg ggt tct att gat gcc aat caa gga aat tac ttg ctt
913 ggt tgg gat aca gat gaa ttc cca aca ctt gtt att gat atc aca ctt
961 gcg atg cac caa att ctt ttg aac ggt gga ctt ggc aaa ggt gga att
1009 aac ttt gat gcg aaa gta cgt cgt aca agt ttc aaa gca gaa gat tta
1057 att ctt gct cat att gca ggg atg gat act tat gcg cgt gct ttg aaa
1105 ggt gca gca gca atc att gaa gat aaa ttc ttg tct gat att gtt gac
1153 gaa cgt tat agt tca tac aga aat aca gaa gtt ggt caa tcc att gaa
1201 aat gga aca gca act ttt aaa agc tta gcc gca ttt gca ctt gaa tat
1249 ggt gat gat att gaa ctt gat tct aat cac ttg gaa tac atc aaa tca
1297 gta ttg aat gac tat ctt gtt taa

```

LISTADO DE SECUENCIAS

- 5 <110> Terranol A/S
- <120> Microorganismo
- <130> P035262WO
- 10 <150> GB0822937.9
- <151> 16-12-2008
- <160> 28
- 15 <170> PatentIn versión 3.5
- <210> 1
- <211> 1320
- 20 <212> ADN
- <213> *Lactococcus lactis* subsp. *lactis*
- <220>
- <221> CDS
- 25 <222> (1)..(1320)

ES 2 582 782 T3

<400> 1

atg gct tac ttt aac gac atc gca cct atc aaa tac gaa ggt aca aaa	48
Met Ala Tyr Phe Asn Asp Ile Ala Pro Ile Lys Tyr Glu Gly Thr Lys	
1 5 10 15	
act aaa aat atg ttt gcc ttt cgc cat tat aat cca gaa gaa gta gtt	96
Thr Lys Asn Met Phe Ala Phe Arg His Tyr Asn Pro Glu Glu Val Val	
20 25 30	
gct ggt aaa aca atg gaa gaa caa ctt cat ttt gcc ctt gca ttt tgg	144
Ala Gly Lys Thr Met Glu Glu Gln Leu His Phe Ala Leu Ala Phe Trp	
35 40 45	
cat aca att aca atg gat ggg tca gat ccc ttt ggg gga gca aca atg	192
His Thr Ile Thr Met Asp Gly Ser Asp Pro Phe Gly Gly Ala Thr Met	
50 55 60	
gaa cgc cct tgg gat ttg gaa ggt ggt tct gaa ctt gac cgt gct cac	240
Glu Arg Pro Trp Asp Leu Glu Gly Gly Ser Glu Leu Asp Arg Ala His	
65 70 75 80	
cgt cga gta gat gct ttc ttt gaa att gct gaa aaa tta ggt gtt aaa	288
Arg Arg Val Asp Ala Phe Phe Glu Ile Ala Glu Lys Leu Gly Val Lys	
85 90 95	
tat tat tgt ttc cat gat att gat att gca cct act gga aat tct ttg	336
Tyr Tyr Cys Phe His Asp Ile Asp Ile Ala Pro Thr Gly Asn Ser Leu	
100 105 110	
aaa gaa ttt tat gct aat ttg gac gaa att act gac cac ctt ctt gaa	384
Lys Glu Phe Tyr Ala Asn Leu Asp Glu Ile Thr Asp His Leu Leu Glu	
115 120 125	
aaa caa aaa gca aca ggg att aaa tta ctt tgg aat aca gca aac atg	432
Lys Gln Lys Ala Thr Gly Ile Lys Leu Leu Trp Asn Thr Ala Asn Met	
130 135 140	
ttt tca aat ccc cgc tat atg aat ggt gtt tca act tct aat cgt gct	480

ES 2 582 782 T3

Phe 145	Ser	Asn	Pro	Arg	Tyr 150	Met	Asn	Gly	Val	Ser 155	Thr	Ser	Asn	Arg	Ala 160	
gaa	gtc	ttt	gct	tat	ggt	gct	gca	caa	gtt	aaa	aaa	ggt	ctt	gaa	ctt	528
Glu	Val	Phe	Ala	Tyr 165	Gly	Ala	Ala	Gln	Val 170	Lys	Lys	Gly	Leu	Glu	Leu 175	
tct	aaa	aaa	ctc	ggt	ggt	gaa	aat	tac	gtc	ttc	tgg	ggt	ggt	cgt	gaa	576
Ser	Lys	Lys	Leu 180	Gly	Gly	Glu	Asn	Tyr 185	Val	Phe	Trp	Gly	Gly	Arg	Glu 190	
ggt	tat	gaa	tca	ctt	ttg	aat	aca	gat	atg	ggt	ctt	gaa	atg	gat	cat	624
Gly	Tyr	Glu 195	Ser	Leu	Leu	Asn	Thr 200	Asp	Met	Gly	Leu	Glu	Met	Asp	His	
atg	gca	aaa	ttc	ttc	cat	ttg	gca	att	gat	tat	gca	aaa	tca	atc	aac	672
Met	Ala	Lys 210	Phe	Phe	His 215	Leu	Ala	Ile	Asp	Tyr 220	Ala	Lys	Ser	Ile	Asn	
cac	ttg	ccc	att	ttc	ttg	att	gaa	cca	aaa	cca	aaa	gaa	cca	atg	act	720
His	Leu	Pro	Ile	Phe	Leu 230	Ile	Glu	Pro	Lys	Pro 235	Lys	Glu	Pro	Met	Thr 240	
cac	caa	tat	gat	ttt	gac	tca	gca	aca	gct	ctt	gct	ttc	ttg	caa	aaa	768
His	Gln	Tyr	Asp 245	Phe	Asp	Ser	Ala	Thr	Ala 250	Leu	Ala	Phe	Leu	Gln	Lys 255	
tat	gat	ttg	gac	aaa	tat	ttc	aaa	ctc	aat	ctt	gaa	aca	aat	cat	gct	816
Tyr	Asp	Leu 260	Asp	Lys	Tyr	Phe	Lys	Leu 265	Asn	Leu	Glu	Thr	Asn	His	Ala	
ttg	ttg	gct	gga	cac	act	ttt	gaa	cac	gaa	tta	aat	act	gct	cgt	act	864
Trp	Leu	Ala 275	Gly	His	Thr	Phe	Glu	His 280	Glu	Leu	Asn	Thr	Ala	Arg	Thr	
ttc	aat	gct	ttg	ggt	tct	att	gat	gcc	aat	caa	gga	aat	tac	ttg	ctt	912
Phe	Asn	Ala 290	Leu	Gly	Ser	Ile 295	Asp	Ala	Asn	Gln	Gly 300	Asn	Tyr	Leu	Leu	
ggt	tgg	gat	aca	gat	gaa	ttc	cca	aca	ctt	gtt	att	gat	atc	aca	ctt	960
Gly	Trp	Asp 305	Thr	Asp	Glu	Phe 310	Pro	Thr	Leu	Val 315	Ile	Asp	Ile	Thr	Leu 320	
gcg	atg	cac	caa	att	ctt	ttg	aac	ggt	gga	ctt	ggc	aaa	ggt	gga	att	1008
Ala	Met	His	Gln 325	Ile	Leu	Leu	Asn	Gly 330	Gly	Leu	Gly	Lys	Gly	Gly	Ile 335	
aac	ttt	gat	gcg	aaa	gta	cgt	cgt	aca	agt	ttc	aaa	gca	gaa	gat	tta	1056
Asn	Phe	Asp 340	Ala	Lys	Val	Arg	Arg	Thr 345	Ser	Phe	Lys	Ala	Glu	Asp	Leu 350	
att	ctt	gct	cat	att	gca	ggg	atg	gat	act	tat	gcg	cgt	gct	ttg	aaa	1104
Ile	Leu	Ala 355	His	Ile	Ala	Gly	Met	Asp 360	Thr	Tyr	Ala	Arg	Ala	Leu	Lys	
ggt	gca	gca	gca	atc	att	gaa	gat	aaa	ttc	ttg	tct	gat	att	gtt	gac	1152
Gly	Ala	Ala 370	Ala	Ile	Ile	Glu	Asp 375	Lys	Phe	Leu	Ser 380	Asp	Ile	Val	Asp	
gaa	cgt	tat	agt	tca	tac	aaa	aat	aca	gaa	ggt	ggt	caa	tcc	att	gaa	1200
Glu	Arg	Tyr 385	Ser	Ser	Tyr 390	Lys	Asn	Thr	Glu	Val 395	Gly	Gln	Ser	Ile	Glu 400	

ES 2 582 782 T3

aat gga aca gca act ttt gaa agt ctt gcc gca ttt gca ctt gaa tat	1248
Asn Gly Thr Ala Thr Phe Glu Ser Leu Ala Ala Phe Ala Leu Glu Tyr	
405 410 415	
ggt gat gat att gaa ctt gat tct aat cac ttg gaa tac atc aaa tca	1296
Gly Asp Asp Ile Glu Leu Asp Ser Asn His Leu Glu Tyr Ile Lys Ser	
420 425 430	
gta ttg aat gac tat ctt gtt taa	1320
Val Leu Asn Asp Tyr Leu Val	
435	

5 <210> 2
 <211> 37
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Cebador

<400> 2
 gctagccatg gcttacttta acgacatcgc acctatc 37

15 <210> 3
 <211> 43
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> Cebador

<400> 3
 ctcgagccta ggctaaacaa gatagtcatt caatactgat ttg 43

25 <210> 4
 <211> 40
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223>

35 <400> 4
 gctagccatg gccactaccc catttgatgc tccagataag 40

40 <210> 5
 <211> 43
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> Cebador

<400> 5
 ctcgagccta ggctagtggt tcaattcaact ttccatcttg gcc 43

50 <210> 6
 <211> 1383
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Gen de la xilosa isomerasa artificial de *Thermoanaerobacter thermohydrosulfuricus*

55 <220>
 <221> CDS

ES 2 582 782 T3

<222> (49)..(1365)

<400>6

gaattcctag aaataat	ttt gttaacttt	aagaaggagg	agctagcc	atg gaa	tat	57
				Met Glu Tyr		
				1		
ttc aaa aac gtg cca cag atc aag tat gaa ggt cct aaa agc aat aac						105
Phe Lys Asn Val Pro Gln Ile Lys Tyr Glu Gly Pro Lys Ser Asn Asn						
5		10		15		
cct tat gca ttt aag ttc tat aac cca gat gaa att ata gat gga aaa						153
Pro Tyr Ala Phe Lys Phe Tyr Asn Pro Asp Glu Ile Ile Asp Gly Lys						
20		25		30		35
cca tta aaa gaa cac tta aga ttt agc gta gcc tac tgg cat aca ttt						201
Pro Leu Lys Glu His Leu Arg Phe Ser Val Ala Tyr Trp His Thr Phe						
	40		45			50
acc gct aac gga acg gat cca ttt ggt gca cgg act atg cag cgt cct						249
Thr Ala Asn Gly Thr Asp Pro Phe Gly Ala Pro Thr Met Gln Arg Pro						
	55		60			65
tgg gat cat ttt acc gac cct atg gac ata gca aaa gca cgt gtg gaa						297
Trp Asp His Phe Thr Asp Pro Met Asp Ile Ala Lys Ala Arg Val Glu						
	70		75			80
gcc gca ttc gag ctt ttt gaa aaa ttg gat gtt cca ttc ttc tgt ttt						345
Ala Ala Phe Glu Leu Phe Glu Lys Leu Asp Val Pro Phe Phe Cys Phe						
	85		90			95
cat gac aga gat ata gct cgg gaa ggt gaa aca ttg aga gaa acc aac						393
His Asp Arg Asp Ile Ala Pro Glu Gly Glu Thr Leu Arg Glu Thr Asn						
100		105		110		115
aaa aac tta gat act atc gtt gct atg att aaa gac tac tta aaa acg						441
Lys Asn Leu Asp Thr Ile Val Ala Met Ile Lys Asp Tyr Leu Lys Thr						
	120			125		130
tca aag act aaa gtt ctt tgg ggc act gct aat ttg ttt tct aat cca						489
Ser Lys Thr Lys Val Leu Trp Gly Thr Ala Asn Leu Phe Ser Asn Pro						
	135			140		145
cgt ttc gtg cat ggc gct gcc aca tca tgt aat gca gac gta ttt gct						537
Arg Phe Val His Gly Ala Ala Thr Ser Cys Asn Ala Asp Val Phe Ala						
	150			155		160
tat gca gcc gct caa gtt aaa aag gcc tta gag att acc aaa gag tta						585
Tyr Ala Ala Ala Gln Val Lys Lys Ala Leu Glu Ile Thr Lys Glu Leu						
	165			170		175
gga gcc cag aat tat gtt ttc tgg ggt ggt cgt gag gga tat gag aca						633
Gly Gly Gln Asn Tyr Val Phe Trp Gly Gly Arg Glu Gly Tyr Glu Thr						

ES 2 582 782 T3

180		185		190		195	
ctt tta aat act gat atg gag ttg gaa tta gat aat tta gca aga ttc							681
Leu Leu Asn Thr Asp Met Glu Leu Glu Leu Asp Asn Leu Ala Arg Phe							
		200		205		210	
tta cac atg gca gta gaa tat gct cag gaa att ggt ttt gaa gga cag							729
Leu His Met Ala Val Glu Tyr Ala Gln Glu Ile Gly Phe Glu Gly Gln							
		215		220		225	
ttc ttg atc gag cct aaa cca aag gaa cca aca aag cat cag tat gat							777
Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys His Gln Tyr Asp							
		230		235		240	
ttc gac gct gct tct gta cac gcc ttt ttg aag aag tat gat ttg gat							825
Phe Asp Ala Ala Ser Val His Ala Phe Leu Lys Lys Tyr Asp Leu Asp							
		245		250		255	
aaa tac ttt aag ttg aac ata gag gct aat cac gca acg ttg gca ggt							873
Lys Tyr Phe Lys Leu Asn Ile Glu Ala Asn His Ala Thr Leu Ala Gly							
		260		265		270	
275							
cac gat ttt caa cac gaa ttg aga tac gcc cgt att aat aac atg tta							921
His Asp Phe Gln His Glu Leu Arg Tyr Ala Arg Ile Asn Asn Met Leu							
		280		285		290	
ggg tcc ata gat gcc aac atg ggt gac atg ttg ctg ggt tgg gat act							969
Gly Ser Ile Asp Ala Asn Met Gly Asp Met Leu Leu Gly Trp Asp Thr							
		295		300		305	
gat caa tac cca acg gat att aga atg aca act tta gca atg tac gag							1017
Asp Gln Tyr Pro Thr Asp Ile Arg Met Thr Thr Leu Ala Met Tyr Glu							
		310		315		320	
gtc att aaa atg gga ggt ttt aac aaa gga ggt ttg aat ttc gat gct							1065
Val Ile Lys Met Gly Gly Phe Asn Lys Gly Gly Leu Asn Phe Asp Ala							
		325		330		335	
aaa gtg cgt cgt gcc tct ttt gaa cct gaa gac ctt ttt ctt gga cat							1113
Lys Val Arg Arg Ala Ser Phe Glu Pro Glu Asp Leu Phe Leu Gly His							
		340		345		350	
355							
att gcc gga atg gat gca ttt gca aaa ggt ttc aag gtc gct tat aag							1161
Ile Ala Gly Met Asp Ala Phe Ala Lys Gly Phe Lys Val Ala Tyr Lys							
		360		365		370	
ctt gtt aag gat ggt gta ttt gat aga ttc att gaa gag aga tac aaa							1209
Leu Val Lys Asp Gly Val Phe Asp Arg Phe Ile Glu Glu Arg Tyr Lys							
		375		380		385	
tcc tat cgt gaa ggt ata ggt gct gaa atc gtt tca ggt aag gcc aat							1257
Ser Tyr Arg Glu Gly Ile Gly Ala Glu Ile Val Ser Gly Lys Ala Asn							
		390		395		400	
ttt aag act tta gag gaa tat gca ttg aat aac cca aaa atc gaa aac							1305
Phe Lys Thr Leu Glu Glu Tyr Ala Leu Asn Asn Pro Lys Ile Glu Asn							
		405		410		415	
aaa agc ggt aaa cag gaa ctg ctg gaa tct att ttg aat caa tat ttg							1353
Lys Ser Gly Lys Gln Glu Leu Leu Glu Ser Ile Leu Asn Gln Tyr Leu							
		420		425		430	
435							
ttc tct gaa tag cctaggctcg aggaattc							1383

Phe Ser Glu

- <210> 7
- <211> 1386
- 5 <212> ADN
- <213> Secuencia artificial
- <220>
- <223> Gen de la xilosa isomerasa artificial de *Thermoanaerobacter thermosulphurigenes*
- 10 <220>
- <221> CDS

ES 2 582 782 T3

<222> (49)..(1368)

<400>7

gaattcctag aaataatttt gtttaacttt aagaaggagg agctagcc atg gcc aag	57
Met Ala Lys	
1	
tat ttt gag aat gtt tcc aag att aag tat gaa ggt cct aag tca aac	105
Tyr Phe Glu Asn Val Ser Lys Ile Lys Tyr Glu Gly Pro Lys Ser Asn	
5 10 15	
aac cct tac tcc ttc aaa ttc tat aat cca gaa gaa gtt ata gac ggt	153
Asn Pro Tyr Ser Phe Lys Phe Tyr Asn Pro Glu Glu Val Ile Asp Gly	
20 25 30 35	
aaa acg atg gag gaa cac tta aga ttt tct att gct tac tgg cac aca	201
Lys Thr Met Glu Glu His Leu Arg Phe Ser Ile Ala Tyr Trp His Thr	
40 45 50	
ttt act gcc gac ggc aca gac caa ttc gga aag gct act atg caa aga	249
Phe Thr Ala Asp Gly Thr Asp Gln Phe Gly Lys Ala Thr Met Gln Arg	
55 60 65	
ccg tgg aac cat tac act gat cca atg gac ata gcc aaa gcc aga gtg	297
Pro Trp Asn His Tyr Thr Asp Pro Met Asp Ile Ala Lys Ala Arg Val	
70 75 80	
gag gct gca ttc gag ttc ttc gat aag ata aac gcc cct tac ttc tgc	345
Glu Ala Ala Phe Glu Phe Phe Asp Lys Ile Asn Ala Pro Tyr Phe Cys	
85 90 95	
ttc cat gat cgt gac att gct ccg gaa ggc gat acc ttg aga gaa acc	393
Phe His Asp Arg Asp Ile Ala Pro Glu Gly Asp Thr Leu Arg Glu Thr	
100 105 110 115	
aac aaa aat ctt gac acc att gtc gca atg ata aaa gat tat ttg aag	441
Asn Lys Asn Leu Asp Thr Ile Val Ala Met Ile Lys Asp Tyr Leu Lys	
120 125 130	
acg tct aaa acc aaa gtt ttg tgg ggt acg gca aac tta ttt tct aat	489
Thr Ser Lys Thr Lys Val Leu Trp Gly Thr Ala Asn Leu Phe Ser Asn	
135 140 145	
ccg aga ttt gtt cat ggt gct tcc aca tcc tgc aac gca gat gtc ttt	537
Pro Arg Phe Val His Gly Ala Ser Thr Ser Cys Asn Ala Asp Val Phe	
150 155 160	

ES 2 582 782 T3

gct tat tcc gct gct caa gtc aaa aag gct ctt gag ata acc aaa gaa 585
Ala Tyr Ser Ala Ala Gln Val Lys Lys Ala Leu Glu Ile Thr Lys Glu
165 170 175

tta ggt ggt gaa aac tac gtt ttc tgg ggt ggc aga gaa ggt tat gag 633
Leu Gly Gly Glu Asn Tyr Val Phe Trp Gly Arg Glu Gly Tyr Glu
180 185 190 195

act ctt tta aat aca gat atg gaa ttt gaa ctt gac aat ttc gca aga 681
Thr Leu Leu Asn Thr Asp Met Glu Phe Glu Leu Asp Asn Phe Ala Arg
200 205 210

ttc ttg cac atg gct gtc gat tat gcc aag gag ata ggt ttt gaa ggc 729
Phe Leu His Met Ala Val Asp Tyr Ala Lys Glu Ile Gly Phe Glu Gly
215 220 225

cag ttc ttg att gaa ccg aaa cca aag gaa cct acc aaa cat cag tat 777
Gln Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Thr Lys His Gln Tyr
230 235 240

gac ttt gat gtc gca aat gtt ttg gct ttc ttg aga aaa tac gat tta 825
Asp Phe Asp Val Ala Asn Val Leu Ala Phe Leu Arg Lys Tyr Asp Leu
245 250 255

gat aag tat ttc aaa gta aat att gaa gct aat cat gca acg ttg gcc 873
Asp Lys Tyr Phe Lys Val Asn Ile Glu Ala Asn His Ala Thr Leu Ala
260 265 270 275

ttc cat gat ttc cag cac gag ttg aga tac gct aga atc aat gga gtc 921
Phe His Asp Phe Gln His Glu Leu Arg Tyr Ala Arg Ile Asn Gly Val
280 285 290

tta gga tct atc gat gct aat acc ggt gat atg ctg ctg gga tgg gat 969
Leu Gly Ser Ile Asp Ala Asn Thr Gly Asp Met Leu Leu Gly Trp Asp
295 300 305

act gat caa ttt ccg acc gat att cgt atg act acc ctg gca atg tat 1017
Thr Asp Gln Phe Pro Thr Asp Ile Arg Met Thr Thr Leu Ala Met Tyr
310 315 320

gaa gtt att aag atg ggt gga ttt gat aaa ggc ggt ctg aac ttc gat 1065
Glu Val Ile Lys Met Gly Gly Phe Asp Lys Gly Glu Leu Asn Phe Asp
325 330 335

gca aaa gta aga aga gct tct ttt gaa cct gaa gat ttg ttt tta gga 1113
Ala Lys Val Arg Arg Ala Ser Phe Glu Pro Glu Asp Leu Phe Leu Gly
340 345 350 355

cac atc gct ggc atg gac gca ttt gct aaa ggt ttt aag gta gcc tat 1161
His Ile Ala Gly Met Asp Ala Phe Ala Lys Gly Phe Lys Val Ala Tyr
360 365 370

aag ctt gta aaa gat aga gtg ttc gac aaa ttc atc gaa gag aga tat 1209
Lys Leu Val Lys Asp Arg Val Phe Asp Lys Phe Ile Glu Glu Arg Tyr
375 380 385

gct tct tat aag gac gga ata gga gcc gat ata gtt tcc ggt aag gcc 1257
Ala Ser Tyr Lys Asp Gly Ile Gly Ala Asp Ile Val Ser Gly Lys Ala
390 395 400

gat ttc aga tct ctt gag aaa tac gcc ttg gaa aga tca caa atc gtg 1305
Asp Phe Arg Ser Leu Glu Lys Tyr Ala Leu Glu Arg Ser Gln Ile Val
405 410 415

aac aaa tca ggc cgt caa gaa ttg tta gaa tca att ctt aat caa tac 1353
Asn Lys Ser Gly Arg Gln Glu Leu Leu Glu Ser Ile Leu Asn Gln Tyr
420 425 430 435

ctg ttc gct gaa tag cctaggctcg aggaattc 1386
Leu Phe Ala Glu

ES 2 582 782 T3

<211> 32
 <212> ADN
 <213> Secuencia artificial

5 <220>
 <223> Cebador

<400>8
 tagctagcat gtcgtacttc cccactgtcg ac 32

10 <210> 9
 <211> 33
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <223> Cebador

20 <400> 9
 ataagctttc aggtgtagat aaagcgggtg acc 33

25 <210> 10
 <211> 1320
 <212> ADN
 <213> *Lactococcus lactis* subsp. *lactis*

<400> 10

atggcttact ttaacgacat cgcacctatc aaatacgaag gtacaaaaac taaaaatatg 60
 ttgctcttc gtcattataa tccagaagaa gtagttgctg gtaaaacaat ggaagaacaa 120
 cttcattttg cccttgcat tggcataca attacgatgg atgggtcaga tccctttggg 180
 ggagcaacaa tggaaactcc ttgggatttg gaaggtggt ctgaacttga ccgtgctcac 240
 cgtcgagtag atgctttctt tgaaattgct gaaaaattag gtgttaata ttattgtttc 300
 catgatattg atattgcacc tactggaaat tctttgaaag aattttatgc taatttgac 360
 gaaattactg accaccttct tgaaaaacaa aaagcaacag gcattaaatt actttggaat 420
 acagcaacaa tgttttcaaa tccccgctat atgaatggtg tttcaacttc taatcgtgct 480
 gaagtctttg cttatggtgc tgcacaagtt aaaaaaggtc ttgaactttc taaaaaactc 540
 ggtggtgaaa attatgtctt ctggggtggt cgtgaaggt atgaatcact ttggaatata 600
 gatatgggtc ttgaaatgga tcatatggca aaattcttcc atttggaat tgattatgca 660
 aaatcaatca accacttgcc tattttcttg attgaaccaa aaccaaaaga accaatgact 720
 caccaatag attttgactc agcaacagct ctgctttct tgcaaaaata tgacttggac 780
 aaatacttca aactcaatct tgaacaaat catgcttggg ttgctgggca cacttttgaa 840
 cacgaattaa atactgcacg tactttcaat gctttgggtt ctattgatgc caatcaagga 900
 aattacttgc ttggttggga tacagatgaa tcccaacac ttgttattga taccacactt 960
 gcgatgcacc aaattctttt gaacggtgga cttggcaaaag gtggaattaa ctttgatgcg 1020
 aaagtacgtc gtacaagttt caaagcagaa gatttaattc ttgctcatat tgcagggatg 1080
 gatacttatg cgcgtgcttt gaaaggtgca gcagcaatca ttgaagataa attcttgtct 1140
 gatattgttg acgaacgtta tagttcatac aaaaatacag aagttggtca atccattgaa 1200
 aatggaacag caacttttga aagtcttgc gcatttgcac ttgaacatgg tgacgatatt 1260
 gaacttgatt ctaatcactt ggaatacatc aaatcagtat tgaatgacta tcttgtttaa 1320

30

ES 2 582 782 T3

<210> 11
 <211> 439
 <212> PRT
 <213> *Lactococcus lactis* subsp. *lactis*

5

<400> 11

```

Met Ala Tyr Phe Asn Asp Ile Ala Pro Ile Lys Tyr Glu Gly Thr Lys
1           5           10           15

Thr Lys Asn Met Phe Ala Phe Arg His Tyr Asn Pro Glu Glu Val Val
          20           25           30

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

His Thr Ile Thr Met Asp Gly Ser Asp Pro Phe Gly Gly Ala Thr Met
          50           55           60

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

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

Tyr Tyr Cys Phe His Asp Ile Asp Ile Ala Pro Thr Gly Asn Ser Leu
          100          105          110

Lys Glu Phe Tyr Ala Asn Leu Asp Glu Ile Thr Asp His Leu Leu Glu
          115          120          125

Lys Gln Lys Ala Thr Gly Ile Lys Leu Leu Trp Asn Thr Ala Asn Met
130          135          140
    
```

ES 2 582 782 T3

Phe Ser Asn Pro Arg Tyr Met Asn Gly Val Ser Thr Ser Asn Arg Ala
 145 150 155 160
 Glu Val Phe Ala Tyr Gly Ala Ala Gln Val Lys Lys Gly Leu Glu Leu
 165 170 175
 Ser Lys Lys Leu Gly Gly Glu Asn Tyr Val Phe Trp Gly Gly Arg Glu
 180 185 190
 Gly Tyr Glu Ser Leu Leu Asn Thr Asp Met Gly Leu Glu Met Asp His
 195 200 205
 Met Ala Lys Phe Phe His Leu Ala Ile Asp Tyr Ala Lys Ser Ile Asn
 210 215 220
 His Leu Pro Ile Phe Leu Ile Glu Pro Lys Pro Lys Glu Pro Met Thr
 225 230 235 240
 His Gln Tyr Asp Phe Asp Ser Ala Thr Ala Leu Ala Phe Leu Gln Lys
 245 250 255
 Tyr Asp Leu Asp Lys Tyr Phe Lys Leu Asn Leu Glu Thr Asn His Ala
 260 265 270
 Trp Leu Ala Gly His Thr Phe Glu His Glu Leu Asn Thr Ala Arg Thr
 275 280 285
 Phe Asn Ala Leu Gly Ser Ile Asp Ala Asn Gln Gly Asn Tyr Leu Leu
 290 295 300
 Gly Trp Asp Thr Asp Glu Phe Pro Thr Leu Val Ile Asp Ile Thr Leu
 305 310 315 320
 Ala Met His Gln Ile Leu Leu Asn Gly Gly Leu Gly Lys Gly Gly Ile
 325 330 335
 Asn Phe Asp Ala Lys Val Arg Arg Thr Ser Phe Lys Ala Glu Asp Leu
 340 345 350
 Ile Leu Ala His Ile Ala Gly Met Asp Thr Tyr Ala Arg Ala Leu Lys
 355 360 365
 Gly Ala Ala Ala Ile Ile Glu Asp Lys Phe Leu Ser Asp Ile Val Asp
 370 375 380
 Glu Arg Tyr Ser Ser Tyr Lys Asn Thr Glu Val Gly Gln Ser Ile Glu
 385 390 395 400
 Asn Gly Thr Ala Thr Phe Glu Ser Leu Ala Ala Phe Ala Leu Glu His
 405 410 415
 Gly Asp Asp Ile Glu Leu Asp Ser Asn His Leu Glu Tyr Ile Lys Ser
 420 425 430
 Val Leu Asn Asp Tyr Leu Val
 435

ES 2 582 782 T3

<210> 12
 <211> 1320
 <212> ADN
 <213> *Lactococcus lactis* subsp. *lactis*

5

<400> 12

```

atggcttact ttaacgacat cgcaacctatc aaatacgaag gtacaaaaac taaaaaatatg      60
tttgcccttc gccattataa tccagaagaa gtagtgtctg gtaaaacaat ggaagaacaa      120
cttcattttg ccottgcatt ttggcataca attacaatgg atgggtcaga cccctttggg      180
ggagcaacaa tggaacgtcc ttgggatttg gaaggtggtt ctgaacttga ccgtgctcac      240
cgtcagatag atgctttctt tgaaattgct gaaaaattag gtgttaataa ttattgtttc      300
catgatattg atattgcacc tactggaaat tctttgaaag aattttatgc taatttggac      360
gaaattactg accaccttct tgaaaaacaa aaagcaacag ggattaaatt actttggaat      420
acagcaacaa tgttttcaaa tccccgctat atgaatggtg tttcaacttc taaccgtgct      480
gaagtctttg cttatggtgc tgcacaagtt aaaaaaggtc ttgaactttc taaaaaactc      540
ggtggtgaaa attacgtctt ctgggggtgt cgtgaagggt atgaatcact tttgaatata      600
gatattgggtc ttgaaatgga tcatatggca aaattcttcc atttggcaat tgattatgca      660
aatcaatca accacttgcc cattttcttg atcgaaccaa aaccaaaga accaatgact      720
caccaaatatg attttgactc agcaacagct cttgctttct tgcaaaaata tgatttggac      780
aatatttca aactcaatct tgaacaacaa catgcttggg ttgctggaca cacttttgaa      840
cacgaattaa atactgctcg tactttcaat gctttgggtt ctattgatgc caatcaagga      900
aattacttgc ttggttggga tacagatgaa ttcccaacac ttgttattga tatcacactt      960
gcgatgcacc aaattctttt gaacggtgga cttggcaaag gtggaattaa ctttgatgcg     1020
aaagtacgtc gtacaagttt caaagcagaa gatttaattc ttgctcatat tgcagggatg     1080
gatacttatg cgcgtgcttt gaaaggtgca gcagcaatca ttgaagataa attcttgtct     1140
gatattgttg acgaacgtta tagttcatac aaaaatacag aagttggaca atccattgaa     1200
aatggaacag caacttttaa aagtcttgcc gcatttgcac ttgaacatgg tgacgatatt     1260

gaacttgatt ctaatcactt ggaatacatc aaatcagtat tgaatgacta tcttgtttaa     1320
    
```

10 <210> 13
 <211> 439
 <212> PRT
 <213> *Lactococcus lactis* subsp. *lactis*

15 <400> 13

ES 2 582 782 T3

Met Ala Tyr Phe Asn Asp Ile Ala Pro Ile Lys Tyr Glu Gly Thr Lys
 1 5 10 15

Thr Lys Asn Met Phe Ala Phe Arg His Tyr Asn Pro Glu Glu Val Val
 20 25 30

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

His Thr Ile Thr Met Asp Gly Ser Asp Pro Phe Gly Gly Ala Thr Met
 50 55 60

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

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

Tyr Tyr Cys Phe His Asp Ile Asp Ile Ala Pro Thr Gly Asn Ser Leu
 100 105 110

Lys Glu Phe Tyr Ala Asn Leu Asp Glu Ile Thr Asp His Leu Leu Glu
 115 120 125

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

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

Glu Val Phe Ala Tyr Gly Ala Ala Gln Val Lys Lys Gly Leu Glu Leu
 165 170 175

Ser Lys Lys Leu Gly Gly Glu Asn Tyr Val Phe Trp Gly Gly Arg Glu
 180 185 190

Gly Tyr Glu Ser Leu Leu Asn Thr Asp Met Gly Leu Glu Met Asp His
 195 200 205

Met Ala Lys Phe Phe His Leu Ala Ile Asp Tyr Ala Lys Ser Ile Asn

ES 2 582 782 T3

Met Ala Tyr Phe Asn Asp Ile Ala Pro Ile Lys Tyr Glu Gly Thr Lys
1 5 10 15

Thr Lys Asn Met Phe Ala Phe Arg His Tyr Asn Pro Glu Glu Val Val
20 25 30

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

His Thr Ile Thr Met Asp Gly Ser Asp Pro Phe Gly Gly Ala Thr Met
50 55 60

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

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

Tyr Tyr Cys Phe His Asp Ile Asp Ile Ala Pro Thr Gly Asn Ser Leu
100 105 110

Lys Glu Phe Tyr Ala Asn Leu Asp Glu Ile Thr Asp His Leu Leu Glu
115 120 125

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

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

Glu Val Phe Ala Tyr Gly Ala Ala Gln Val Lys Lys Gly Leu Glu Leu
165 170 175

Ser Lys Lys Leu Gly Gly Glu Asn Tyr Val Phe Trp Gly Gly Arg Glu
180 185 190

Gly Tyr Glu Ser Leu Leu Asn Thr Asp Met Gly Leu Glu Met Asp His
195 200 205

Met Ala Lys Phe Phe His Leu Ala Ile Asp Tyr Ala Lys Ser Ile Asn
210 215 220

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

ES 2 582 782 T3

His Gln Tyr Asp Phe Asp Ser Ala Thr Ala Leu Ala Phe Leu Gln Lys
 245 250 255

Tyr Asp Leu Asp Lys Tyr Phe Lys Leu Asn Leu Glu Thr Asn His Ala
 260 265 270

Trp Leu Ala Gly His Thr Phe Glu His Glu Leu Asn Thr Ala Arg Thr
 275 280 285

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

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

Ala Met His Gln Ile Leu Leu Asn Gly Gly Leu Gly Lys Gly Gly Ile
 325 330 335

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

Ile Leu Ala His Ile Ala Gly Met Asp Thr Tyr Ala Arg Ala Leu Lys
 355 360 365

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

Glu Arg Tyr Ser Ser Tyr Lys Asn Thr Glu Val Gly Gln Ser Ile Glu
 385 390 395 400

Asn Gly Thr Ala Thr Phe Glu Ser Leu Ala Ala Phe Ala Leu Glu Tyr
 405 410 415

Gly Asp Asp Ile Glu Leu Asp Ser Asn His Leu Glu Tyr Ile Lys Ser
 420 425 430

Val Leu Asn Asp Tyr Leu Val
 435

<210> 15
 <211> 438
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Construcción Sintética

<400> 15

Met Glu Tyr Phe Lys Asn Val Pro Gln Ile Lys Tyr Glu Gly Pro Lys

ES 2 582 782 T3

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

ES 2 582 782 T3

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

Leu Ala Gly His Asp Phe Gln His Glu Leu Arg Tyr Ala Arg Ile Asn
275 280 285

Asn Met Leu Gly Ser Ile Asp Ala Asn Met Gly Asp Met Leu Leu Gly
290 295 300

Trp Asp Thr Asp Gln Tyr Pro Thr Asp Ile Arg Met Thr Thr Leu Ala
305 310 315 320

Met Tyr Glu Val Ile Lys Met Gly Gly Phe Asn Lys Gly Gly Leu Asn
325 330 335

Phe Asp Ala Lys Val Arg Arg Ala Ser Phe Glu Pro Glu Asp Leu Phe
340 345 350

Leu Gly His Ile Ala Gly Met Asp Ala Phe Ala Lys Gly Phe Lys Val
355 360 365

Ala Tyr Lys Leu Val Lys Asp Gly Val Phe Asp Arg Phe Ile Glu Glu
370 375 380

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

Lys Ala Asn Phe Lys Thr Leu Glu Glu Tyr Ala Leu Asn Asn Pro Lys
405 410 415

Ile Glu Asn Lys Ser Gly Lys Gln Glu Leu Leu Glu Ser Ile Leu Asn
420 425 430

Gln Tyr Leu Phe Ser Glu
435

<210> 16
<211> 439
<212> PRT
<213> Secuencia artificial

<220>
<223> Construcción Sintética

<400> 16

Met Ala Lys Tyr Phe Glu Asn Val Ser Lys Ile Lys Tyr Glu Gly Pro
1 5 10 15

Lys Ser Asn Asn Pro Tyr Ser Phe Lys Phe Tyr Asn Pro Glu Glu Val

5

10

ES 2 582 782 T3

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

ES 2 582 782 T3

Thr Leu Ala Phe His Asp Phe Gln His Glu Leu Arg Tyr Ala Arg Ile
 275 280 285

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

Gly Trp Asp Thr Asp Gln Phe Pro Thr Asp Ile Arg Met Thr Thr Leu
 305 310 315 320

Ala Met Tyr Glu Val Ile Lys Met Gly Gly Phe Asp Lys Gly Gly Leu
 325 330 335

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

Phe Leu Gly His Ile Ala Gly Met Asp Ala Phe Ala Lys Gly Phe Lys
 355 360 365

Val Ala Tyr Lys Leu Val Lys Asp Arg Val Phe Asp Lys Phe Ile Glu
 370 375 380

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

Gly Lys Ala Asp Phe Arg Ser Leu Glu Lys Tyr Ala Leu Glu Arg Ser
 405 410 415

Gln Ile Val Asn Lys Ser Gly Arg Gln Glu Leu Leu Glu Ser Ile Leu
 420 425 430

Asn Gln Tyr Leu Phe Ala Glu
 435

<210> 17
 <211> 1320
 <212> ADN
 <213> *Lactococcus lactis*
 <400> 17

5

ES 2 582 782 T3

atggcttact ttaacgacat cgcacctatc aaatacgaag gtacaaaaac taaaaatg	60
tttgcccttc gccattataa tccagaagaa gtagtgtctg gtaaaacaat ggaagaacaa	120
cttcattttg cccttgcatt ttggcataca attacaatgg atgggtcaga tccctttggg	180
ggagcaacaa tggaacgccc ttgggatttg gaaggtggtt ctgaacttga ccgtgctcac	240
cgtcgagtag atgctttctt tgaattgct gaaaaattag gtgttaata ttattgtttc	300
catgatattg atattgcacc tactggaaat tctttgaaag aattttatgc taatttggac	360
gaaattactg accaccttct tgaaaaacaa aaagcaacag ggattaaatt actttggaat	420
acagcaaaaa tgttttcaaa tccccgctat atgaatgggt tttcaacttc taatcgtgct	480
gaagtcttg cttatggtgc tgcacaagtt aaaaaaggtc ttgaactttc taaaaaactc	540
ggtggtgaaa attacgtctt ctgggggtgt cgtgaagggt atgaatcact tttgaataca	600
gatatgggtc ttgaaatgga tcatatggca aaattcttcc atttggcaat tgattatgca	660
aatcaatca accacttgcc cattttcttg attgaaccaa aacaaaaaga accaatgact	720
caccaatatg attttgactc agcaacagct cttgctttct tgcaaaaata tgatttggac	780
aatatttca aactcaatct tgaacaaat catgcttgggt tggctggaca cacttttgaa	840
cacgaattaa atactgctcg tactttcaat gctttgggtt ctattgatgc caatcaagga	900
aattactgct ttggttggga tacagatgaa ttcccaacac ttgttattga taccacactt	960
gcgatgcacc aaattctttt gaacggtgga cttggcaaag gtggaattaa ctttgatgag	1020
aaagtacgct gtacaagttt caaagcagaa gatttaattc ttgctcatat tgcagggatg	1080
gatacttatg cgcgtgcttt gaaaggtgca gcagcaatca ttgaagataa attcttgtct	1140
gatattgttg acgaacgtta tagttcatac agaaatacag aagtgggtca atccattgaa	1200
aatggaacag caacttttga aagtcttgcc gcatttgcac ttgaatatgg tgatgatatt	1260
gaacttgatt ctaatcactt ggaatacatc aaatcagtat tgaatgacta tcttgtttaa	1320

<210> 18
 <211> 439
 <212> PRT
 <213> *Lactococcus lactis*

5

<400> 18

ES 2 582 782 T3

Tyr Tyr Cys Phe His Asp Ile Asp Ile Ala Pro Thr Gly Asn Ser Leu
 100 105 110

Lys Glu Phe Tyr Ala Asn Leu Asp Glu Ile Thr Asp His Leu Leu Glu
 115 120 125

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

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

Glu Val Phe Ala Tyr Gly Ala Ala Gln Val Lys Lys Gly Leu Glu Leu
 165 170 175

Ser Lys Lys Leu Gly Gly Glu Asn Tyr Val Phe Trp Gly Gly Arg Glu
 180 185 190

Gly Tyr Glu Ser Leu Leu Asn Thr Asp Met Gly Leu Glu Met Asp His
 195 200 205

Met Ala Lys Phe Phe His Leu Ala Ile Asp Tyr Ala Lys Ser Ile Asn
 210 215 220

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

His Gln Tyr Asp Phe Asp Ser Ala Thr Ala Leu Ala Phe Leu Gln Lys
 245 250 255

Tyr Asp Leu Asp Lys Tyr Phe Lys Leu Asn Leu Glu Thr Asn His Ala
 260 265 270

Trp Leu Ala Gly His Thr Phe Glu His Glu Leu Asn Thr Ala Arg Thr
 275 280 285

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

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

Ala Met His Gln Ile Leu Leu Asn Gly Gly Leu Gly Lys Gly Gly Ile
 325 330 335

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

ES 2 582 782 T3

Ile Leu Ala His Ile Ala Gly Met Asp Thr Tyr Ala Arg Ala Leu Lys
355 360 365

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

Glu Arg Tyr Ser Ser Tyr Arg Asn Thr Glu Val Gly Gln Ser Ile Glu
385 390 395 400

Asn Gly Thr Ala Thr Phe Glu Ser Leu Ala Ala Phe Ala Leu Glu Tyr
405 410 415

Gly Asp Asp Ile Glu Leu Asp Ser Asn His Leu Glu Tyr Ile Lys Ser
420 425 430

Val Leu Asn Asp Tyr Leu Val
435

<210> 19

<211> 439

<212> PRT

<213> Secuencia artificial

<220>

<223> Variante de xilosa isomerasa de *Lactococcus*

<400> 19

Met Ala Tyr Phe Asn Asp Ile Ala Pro Ile Lys Tyr Glu Gly Thr Lys
1 5 10 15

Thr Lys Asn Met Phe Ala Phe Arg His Tyr Asn Pro Glu Glu Val Val
20 25 30

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

His Thr Ile Thr Met Asp Gly Ser Asp Pro Phe Gly Gly Ala Thr Met
50 55 60

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

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

Tyr Tyr Cys Phe His Asp Ile Asp Ile Ala Pro Thr Gly Asn Ser Leu
100 105 110

5

10

ES 2 582 782 T3

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

ES 2 582 782 T3

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

Glu Arg Tyr Ser Ser Tyr Arg Asn Thr Glu Val Gly Gln Ser Ile Glu
 385 390 395 400

Asn Gly Thr Ala Thr Phe Lys Ser Leu Ala Ala Phe Ala Leu Glu His
 405 410 415

Gly Asp Asp Ile Glu Leu Asp Ser Asn His Leu Glu Tyr Ile Lys Ser
 420 425 430

Val Leu Asn Asp Tyr Leu Val
 435

<210> 20
 <211> 439
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Variante de xilosa isomerasa de *Lactococcus*

<400> 20

Met Ala Tyr Phe Asn Asp Ile Ala Pro Ile Lys Tyr Glu Gly Thr Lys
 1 5 10 15

Thr Lys Asn Met Phe Ala Phe Arg His Tyr Asn Pro Glu Glu Val Val
 20 25 30

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

His Thr Ile Thr Met Asp Gly Ser Asp Pro Phe Gly Gly Ala Thr Met
 50 55 60

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

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

Tyr Tyr Cys Phe His Asp Ile Asp Ile Ala Pro Thr Gly Asn Ser Leu
 100 105 110

Lys Glu Phe Tyr Ala Asn Leu Asp Glu Ile Thr Asp His Leu Leu Glu
 115 120 125

ES 2 582 782 T3

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

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

Glu Val Phe Ala Tyr Gly Ala Ala Gln Val Lys Lys Gly Leu Glu Leu
 165 170 175

Ser Lys Lys Leu Gly Gly Glu Asn Tyr Val Phe Trp Gly Gly Arg Glu
 180 185 190

Gly Tyr Glu Ser Leu Leu Asn Thr Asp Met Gly Leu Glu Met Asp His
 195 200 205

Met Ala Lys Phe Phe His Leu Ala Ile Asp Tyr Ala Lys Ser Ile Asn
 210 215 220

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

His Gln Tyr Asp Phe Asp Ser Ala Thr Ala Leu Ala Phe Leu Gln Lys
 245 250 255

Tyr Asp Leu Asp Lys Tyr Phe Lys Leu Asn Leu Glu Thr Asn His Ala
 260 265 270

Trp Leu Ala Gly His Thr Phe Glu His Glu Leu Asn Thr Ala Arg Thr
 275 280 285

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

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

Ala Met His Gln Ile Leu Leu Asn Gly Gly Leu Gly Lys Gly Gly Ile
 325 330 335

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

Ile Leu Ala His Ile Ala Gly Met Asp Thr Tyr Ala Arg Ala Leu Lys
 355 360 365

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

ES 2 582 782 T3

Glu Arg Tyr Ser Ser Tyr Arg Asn Thr Glu Val Gly Gln Ser Ile Glu
385 390 395 400

Asn Gly Thr Ala Thr Phe Lys Ser Leu Ala Ala Phe Ala Leu Glu Tyr
405 410 415

Gly Asp Asp Ile Glu Leu Asp Ser Asn His Leu Glu Tyr Ile Lys Ser
420 425 430

Val Leu Asn Asp Tyr Leu Val
435

5 <210> 21
<211> 41
<212> ADN
<213> Secuencia artificial

10 <220>
<223> Cebador

<400> 21
gttgacgaac gatatagttc atacaaaaat acagaagttg g 41

15 <210> 22
<211>41
<212> ADN
<213> Secuencia artificial

20 <220>
<223> Cebador

<400> 22
ccaacttctg tatttttga tgaactatat cgttcgtcaa c 41

25 <210> 23
<211> 55
<212> ADN
<213> Secuencia artificial

30 <220>
<223> Cebador

35 <400> 23
acagcaactt ttaaagctt agccgcattt gcacttgaat atgggatga tattg 55

40 <210> 24
<211> 55
<212> ADN
<213> Secuencia artificial

45 <220>
<223> Cebador

<400> 24
caatatcatc accatattca agtgcaaatg cggctaagct tttaaaagtt gctgt 55

50 <210> 25
<211> 55
<212> ADN
<213> Secuencia artificial

ES 2 582 782 T3

<220>
<223> Cebador

5 <400> 25
acagcaactt taaaagctt agccgcattt gcactgaac atggtgatga tattg 55

10 <210> 26
<211> 55
<212> ADN
<213> Secuencia artificial

15 <220>
<223> Cebador

<400> 26
caatatcatc accatgttca agtgcaaatg cggctaagct tttaaaagtt gctgt 55

20 <210> 27
<211> 1320
<212> ADN
<213> Secuencia artificial

25 <220>
<223> Variante de xilosa isomerasa de *Lactococcus*

<400> 27

ES 2 582 782 T3

atggcttact ttaacgacat cgcacctatc aaatacgaag gtacaaaaac taaaaatatg 60
 ttgctcttcc gccattataa tccagaagaa gtagttgctg gtaaaaaat ggaagaacaa 120
 cttcattttg cccttgcatc ttggcataca attacaatgg atgggtcaga tccctttggg 180
 ggagcaacaa tgaacgccc ttgggatttg gaaggtggtt ctgaacttga ccggtgctcac 240
 cgtcgagtag atgctttctt tgaattgct gaaaattag gtgttaata ttattgtttc 300
 catgatattg atattgcacc tactggaaat tctttgaaag aattttatgc taatttggac 360
 gaaattactg accaccttct tgaaaaacaa aaagcaacag ggattaaatt actttggaat 420
 acagcaacaa tgttttcaaa tccccgctat atgaatggtg tttcaacttc taatcgtgct 480
 gaagtctttg cttatggtgc tgcacaagtt aaaaaaggtc ttgaactttc taaaaaactc 540
 ggtggtgaaa attacgtctt ctggggtggt cgtgaagggt atgaatcact tttgaataca 600
 gatatgggtc ttgaaatgga tcatatggca aaattcttcc atttggcaat tgattatgca 660
 aatcaatca accacttgcc cttttctttg attgaaccaa aacccaaaga accaatgact 720
 caccaatag attttgactc agcaacagct cttgctttct tgcaaaaata tgatttggac 780
 aaatatttca aactcaatct tgaacaacat catgcttggg ttgctggaca cacttttgaa 840

 cacgaattaa atactgctcg tactttcaat gctttgggtt ctattgatgc caatcaagga 900
 aattacttgc ttggttggga tacagatgaa ttccaacac ttgttattga tatcacactt 960
 gcgatgcacc aaattctttt gaacggtgga cttggcaaag gtggaattaa ctttgatgcg 1020
 aaagtacgtc gtacaagttt caaagcagaa gatttaattc ttgctcatat tgcagggatg 1080
 gatacttatg cgcgtgcttt gaaagtgca gcagcaatca ttgaagataa attcttgtct 1140
 gatattgttg acgaacgta tagttcatac agaaatacag aagttggtca atccattgaa 1200
 aatggaacag caacttttaa aagcttagcc gcatttgcac ttgaacatgg tgatgatatt 1260
 gaacttgatt ctaatcactt ggaatacatc aatcagtat tgaatgacta tcttgtttaa 1320

<210> 28
 <211> 1320
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Variante de xilosa isomerasa de *Lactococcus*

10

<400> 28

ES 2 582 782 T3

atggcttact ttaacgacat cgcacctatc aaatacgaag gtacaaaaac taaaaatatg	60
tttgcctttc gccattataa tccagaagaa gtagttgctg gtaaaacaat ggaagaacaa	120
cttcattttg cccttgcatt ttggcataca attacaatgg atgggtcaga tccctttggg	180
ggagcaacaa tggaacgcc ttgggatttg gaaggtggtt ctgaacttga ccgtgctcac	240
cgtcgagtag atgctttctt tgaaattgct gaaaaattag gtgttaataa ttattgtttc	300
catgatattg atattgcacc tactggaaat tctttgaaag aattttatgc taatttggac	360
gaaattactg accaccttct tgaaaaacaa aaagcaacag ggattaaatt actttggaat	420
acagcaacaa tgttttcaaa tccccgctat atgaatggtg tttcaacttc taatcgtgct	480
gaagtctttg cttatggtgc tgcacaagtt aaaaaaggtc ttgaactttc taaaaaactc	540
ggtggtgaaa attacgtctt ctgggggtgt cgtgaaggtt atgaatcact tttgaataca	600
gatatgggtc ttgaaatgga tcatatggca aaattcttcc atttggcaat tgattatgca	660
aatcaatca accacttgcc cattttcttg attgaaccaa aaccaaaga accaatgact	720
caccaatag attttgactc agcaacagct cttgctttct tgcaaaaata tgatttggac	780
aatatttca aactcaatct tgaaacaaat catgcttggg ttgctggaca cacttttgaa	840
cacgaattaa atactgctcg tactttcaat gctttgggtt ctattgatgc caatcaagga	900
aattacttgc ttggttggga tacagatgaa ttcccaacac ttgttattga tatcacactt	960
gcgatgcacc aaattctttt gaacggtgga cttggcaaag gtggaattaa ctttgatgcg	1020
aaagtacgtc gtacaagttt caaagcagaa gatttaattc ttgctcatat tgcagggatg	1080
gatacttatg cgcgtgcttt gaaaggtgca gcagcaatca ttgaagataa attcttgtct	1140
gatattggtg acgaacgtta tagttcatac agaaatacag aagttggtca atccattgaa	1200
aatggaacag caacttttaa aagcttagcc gcatttgcac ttgaatatgg tgatgatatt	1260
gaacttgatt ctaatcactt ggaatacatc aaatcagtat tgaatgacta tcttgtttaa	1320

REIVINDICACIONES

1. Un microorganismo transformado capaz de:

- 5 (a) una actividad xilosa isomerasa superior a la del microorganismo equivalente antes de la transformación y/o
 (b) una tasa de crecimiento superior, en o sobre un medio de crecimiento que comprende xilosa, que la del
 microorganismo equivalente antes de la transformación y/o
 (c) un metabolismo de la xilosa más rápido que el del microorganismo equivalente antes de la transformación y/o
 10 (d) una producción de etanol superior, cuando se cultiva en condiciones anaerobias en xilosa como la fuente de
 carbono, que la del microorganismo equivalente antes de la transformación;

en el que dicho microorganismo ha sido transformado con una secuencia de nucleótidos que codifica una xilosa
 isomerasa en el que dicha xilosa isomerasa es una xilosa isomerasa exógena derivada de una bacteria mesófila y
 en el que dicha xilosa isomerasa es una xilosa isomerasa exógena derivada de una especie de *Lactococcus*;
 15 y en el que dicho microorganismo transformado es una levadura transformada en el que dicha levadura
 transformada es *Saccharomyces*.

2. El microorganismo de acuerdo con la reivindicación 1, en el que dicho microorganismo ha sido transformado con
 una secuencia de nucleótidos que codifica la secuencia de aminoácidos mostrada como SEQ ID No 14, SEQ ID No
 20 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20, o una variante, homóloga o derivada de la misma
 que codifica una secuencia de aminoácidos que tiene al menos 75 % de identidad con la secuencia de aminoácidos
 mostrada como SEQ ID No 14, SEQ ID No 11, SEQ ID No 18, SEQ ID No 13, SEQ ID No 19 o SEQ ID No 20.

3. El microorganismo de acuerdo con la reivindicación 1 o 2, en el que dicho microorganismo ha sido transformado
 con la secuencia de nucleótidos mostrada como SEC ID No 1, SEQ ID No 10, SEQ ID No 17, SEQ ID No 12, SEQ
 25 ID No 27 o SEQ ID No 28, o una variante, homóloga o derivada de la misma que tiene al menos 75 % de identidad
 con la secuencia de nucleótidos mostrada como SEC ID No 1, SEQ ID No 10, SEQ ID No 17, SEQ ID No 12, SEQ
 ID No 27 o SEQ ID No 28.

- 30 4. Un inóculo que comprende un microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 3.

5. Un medio de cultivo que comprende un microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a
 3.

- 35 6. Un método para preparar un microorganismo transformado, comprendiendo dicho método la etapa de transformar
 un microorganismo de tal modo que dicho microorganismo transformado es capaz de:

- (a) una actividad xilosa isomerasa superior a la del microorganismo antes de la transformación y/o
 (b) una tasa de crecimiento superior, en o sobre un medio de crecimiento que comprende xilosa, que la del
 40 microorganismo antes de la transformación y/o
 (c) un metabolismo de la xilosa más rápido que el del microorganismo antes de la transformación y/o
 (d) una producción de etanol superior, cuando se cultiva en condiciones anaerobias en xilosa como la fuente de
 carbono, que la del microorganismo antes de la transformación;

- 45 en el que dicho microorganismo es transformado con una secuencia de nucleótidos que codifica una xilosa
 isomerasa en el que dicha xilosa isomerasa es una xilosa isomerasa exógena derivada de una bacteria mesófila y
 en el que dicha xilosa isomerasa es una xilosa isomerasa exógena derivada de una especie de *Lactococcus*;
 y en el que dicho microorganismo transformado es una levadura transformada en el que dicha levadura
 transformada es *Saccharomyces*.

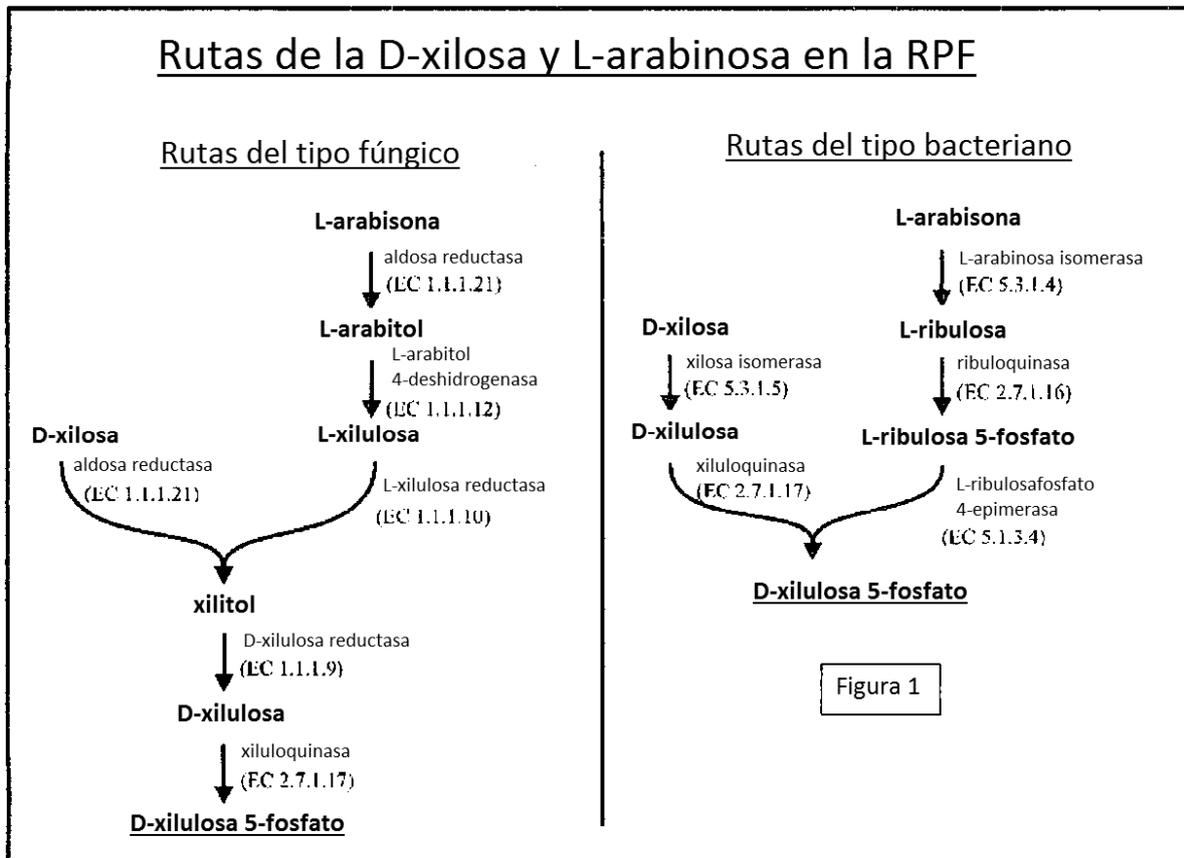
- 50 7. El método de acuerdo con la reivindicación 6, en el que dicha secuencia de nucleótidos que codifica una xilosa
 isomerasa comprende la secuencia de nucleótidos mostrada como SEC ID No 1, SEQ ID No 10, SEQ ID No 17,
 SEQ ID No 12, SEQ ID No 27 o SEQ ID No 28, o una variante, homóloga o derivada de la misma que tiene al menos
 75 % de identidad con la secuencia de nucleótidos mostrada como SEC ID No 1, SEQ ID No 10, SEQ ID No 17,
 55 SEQ ID No 12, SEQ ID No 27 o SEQ ID No 28.

8. El método de acuerdo con la reivindicación 6 o 7, en el que dicha secuencia de nucleótidos que codifica una xilosa
 isomerasa está en un vector de expresión que codifica la misma.

- 60 9. Un método de fermentación que comprende cultivar, en un medio de cultivo, un microorganismo de acuerdo con
 una cualquiera de las reivindicaciones 1 a 3, o un microorganismo preparado por el método de una cualquiera de las
 reivindicaciones 6 a 8.

- 65 10. Un método para producir un producto derivado de la xilosa que comprende cultivar, en un medio de cultivo, un
 microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 3, o un microorganismo preparado
 mediante el método de una cualquiera de las reivindicaciones 6 a 8.

11. Un método para producir un biocombustible, en el que dicho método comprende la etapa de cultivar, en un medio de cultivo, un microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 3, o un microorganismo preparado por el método de una cualquiera de las reivindicaciones 6 a 8.
- 5 12. El método de acuerdo con la reivindicación 11, en el que dicho método comprende además la etapa de obtener el biocombustible del medio de cultivo.
13. Uso de un microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 3 o un microorganismo preparado mediante el método de una cualquiera de las reivindicaciones 6 a 8 para la producción de un producto derivado de la xilosa.
- 10 14. Uso de un microorganismo de acuerdo con una cualquiera de las reivindicaciones 1 a 3 o un microorganismo preparado mediante el método de una cualquiera de las reivindicaciones 6 a 8 para la producción de un biocombustible.
- 15



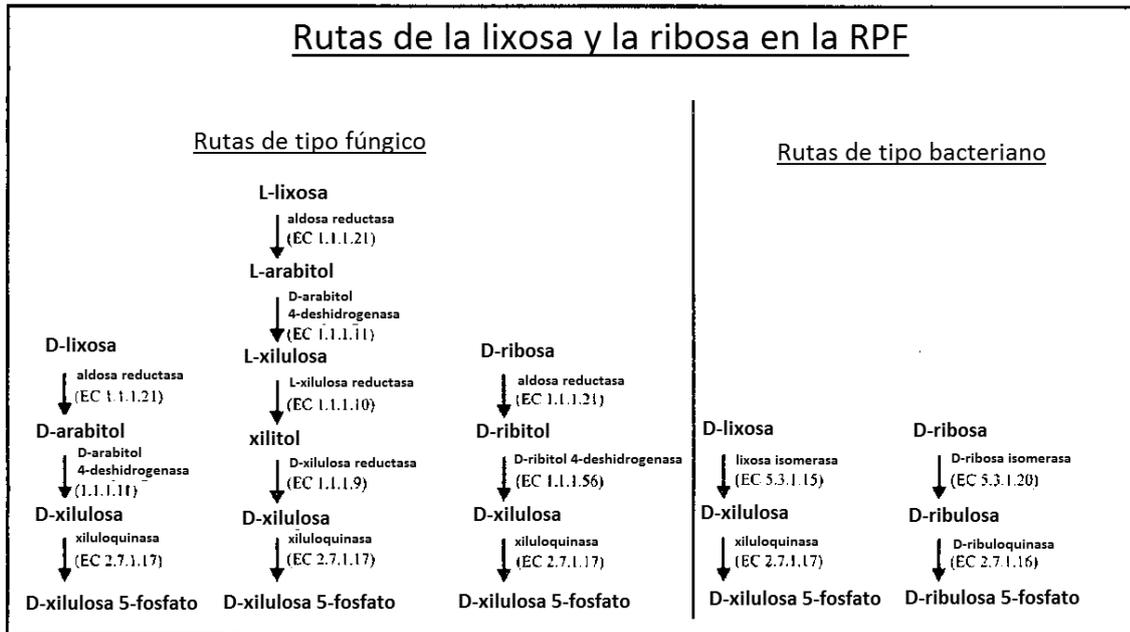


Figura 2

Figura 3

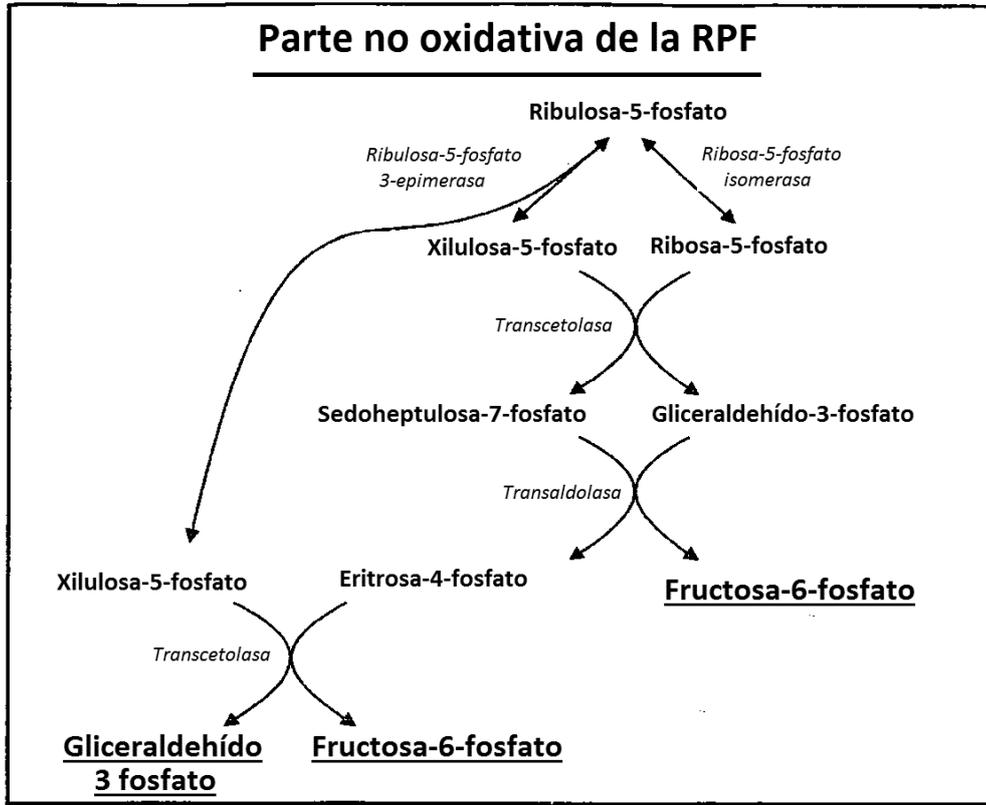


Figura 4

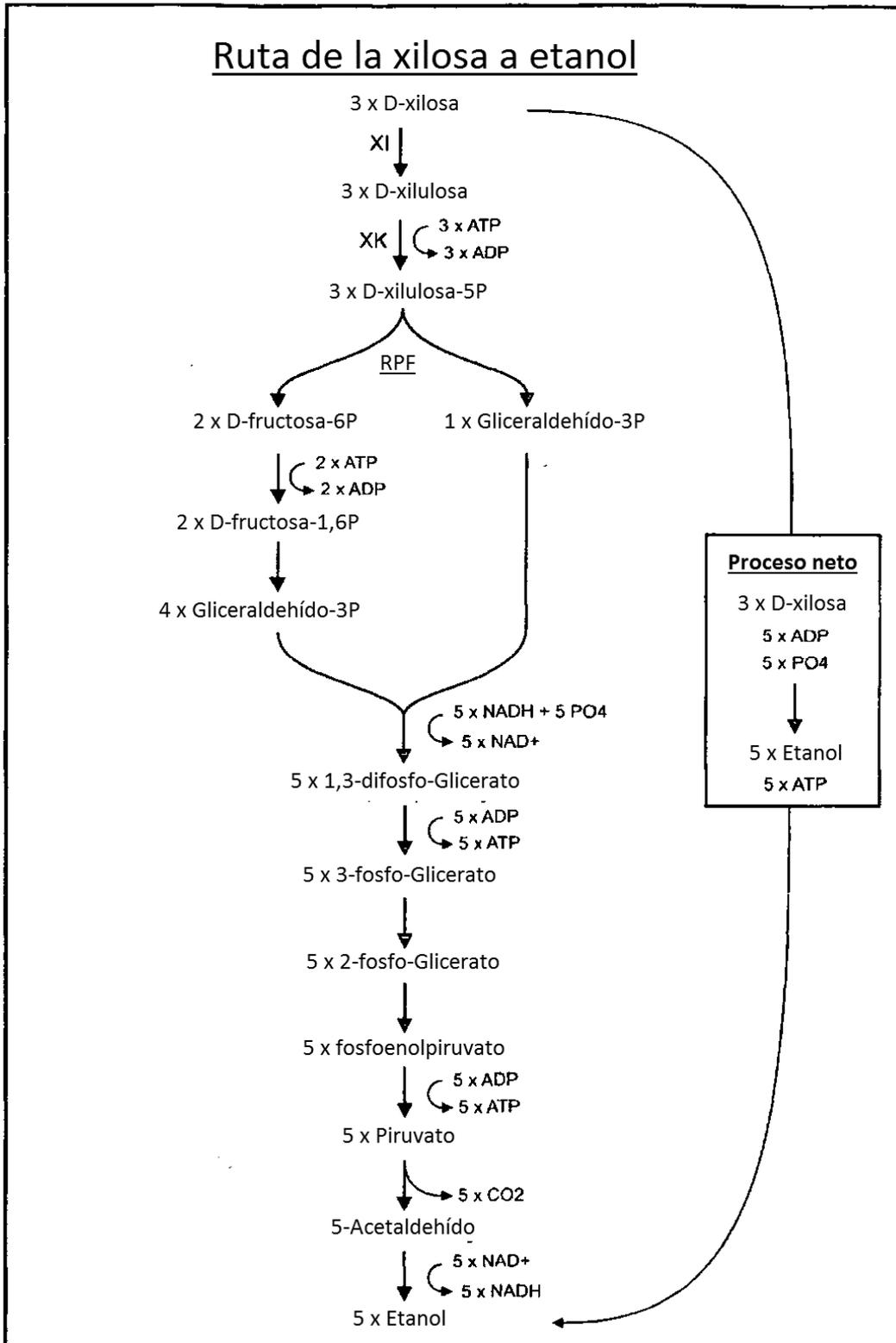


Figura 5

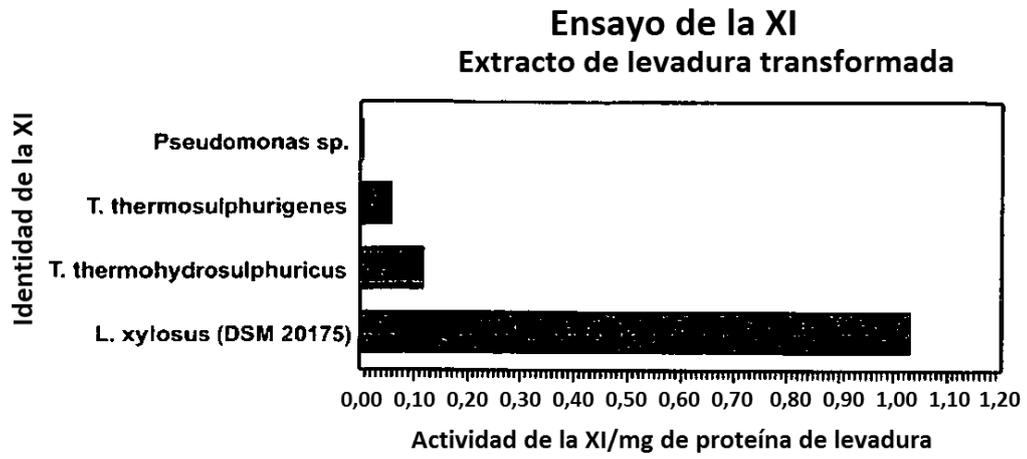
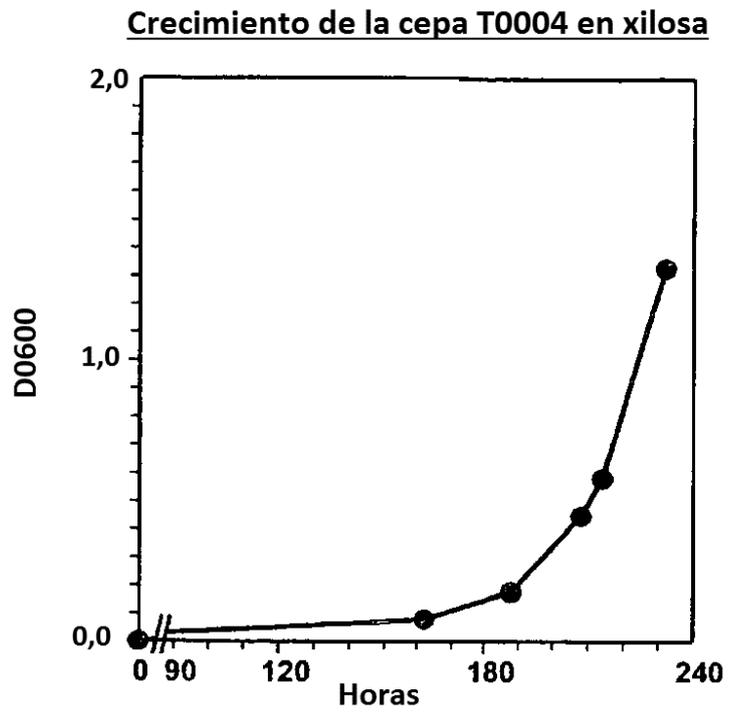


Figura 6.



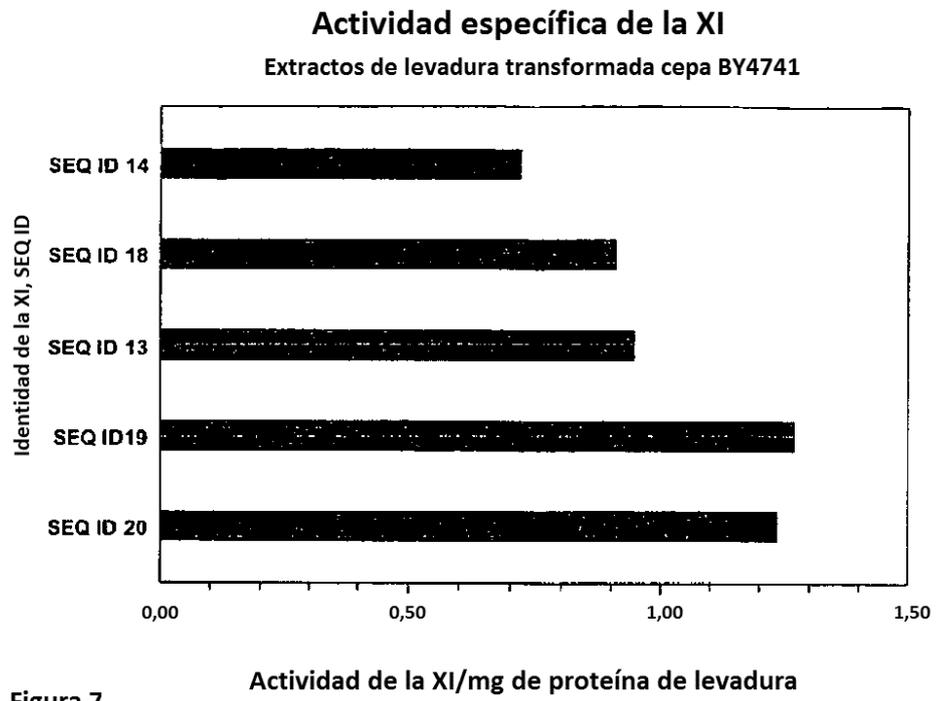


Figura 7