

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 731 782**

51 Int. Cl.:

C12P 7/18 (2006.01)

C12P 7/42 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **19.11.2015** **E 15805601 (0)**

97 Fecha y número de publicación de la concesión europea: **03.04.2019** **EP 3221459**

54 Título: **Procedimiento de producción de al menos un metabolito de interés mediante la transformación de una pentosa en un microorganismo**

30 Prioridad:

19.11.2014 FR 1461183

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

19.11.2019

73 Titular/es:

**INSTITUT NATIONAL DE LA RECHERCHE
AGRONOMIQUE (INRA) (33.3%)
147, rue de l'Université
75007 Paris, FR;
INSTITUT NATIONAL DES SCIENCE
APPLIQUEES - INSA (33.3%) y
CENTRE NATIONAL DE LA RECHERCHE
SCIENTIFIQUE (33.3%)**

72 Inventor/es:

**WALTHER, THOMAS;
CAM, YVAN;
FRANCOIS, JEAN-MARIE y
ALKIM, CEREN**

74 Agente/Representante:

LINAGE GONZÁLEZ, Rafael

ES 2 731 782 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

Procedimiento de producción de al menos un metabolito de interés mediante la transformación de una pentosa en un microorganismo

Campo de la invención

- 5 La presente invención se refiere a un procedimiento de producción de al menos un metabolito de interés mediante transformación de una pentosa en un microorganismo. Más particularmente, proporciona un procedimiento para la biosíntesis de etilenglicol y/o ácido glicólico en un microorganismo que expresa una vía metabólica artificial/sintética de asimilación de pentosas procedentes, ventajosamente, de una fuente de carbono renovable.

Estado de la técnica

- 10 El etilenglicol (EG) y el ácido glicólico (AG) son compuestos utilizados en un amplio espectro de aplicaciones industriales en la industria petroquímica, como precursores de polímeros basados en EG, como el poli(tereftalato de etileno) (PET) o basados en AG como ciertas resinas termoplásticas, pero también como refrigerante en anticongelantes para automóviles para EG o para AG, en la industria textil, en las industrias de aceites y gases así como en un gran número de productos de belleza. Estos compuestos son todavía ampliamente producidos de forma
15 petroquímica. No obstante, la ingeniería metabólica representa un sector de gran desarrollo desde hace más de 15 años. Este procedimiento comprende la concepción de vías de biosíntesis artificiales y/o la optimización de vías naturales en un organismo hospedante e implica la transferencia, la expresión y el acoplamiento funcional de múltiples etapas enzimáticas en el seno de dicho organismo, para permitir la producción de moléculas de interés. La concepción de vías de biosíntesis totalmente artificiales en el microorganismo hospedante permite, más particularmente, constituir un conjunto metabólico autónomo de producción de moléculas de interés. Estos métodos de producción presentan la ventaja de utilizar, como sustrato, una fuente carbonada renovable y representan en lo sucesivo una alternativa real a las fuentes de energía fósiles. A este respecto, han sido descritas diferentes vías metabólicas para la bioproducción de EG. Implican, en particular, intermedios de glicólisis. La empresa Genomatica particularmente ha descrito en sus solicitudes de patentes WO2011130378 y WO2012177983, vías metabólicas, principalmente anaeróbicas de síntesis de EG a partir de serina, 3-fosfohidroxipiruvato, 3-fosfoglicerato o glioxilato. Han sido descritas también técnicas para la bioproducción de AG. Así, la empresa Mitsui Chemicals Inc. ha descrito un método que utiliza un microorganismo para producir ácidos hidrocarboxílicos a partir de un alcohol polihidroxilado alifático que tiene un grupo hidroxilo en su extremo (documentos EP 2.025.759 y EP 2.025.760). Se trata de un método de bioconversión, como se describe por Michihiko Kataoka en un artículo sobre la producción de AG utilizando microorganismos que oxidan etileno (Biosci Biotechnol Biochem, 2001). El AG puede ser producido también mediante bioconversión a partir de gliconitrilo, utilizando nitrilasas mutantes que tienen una actividad nitrilasa aumentada, como se describe por Dupont de Nemours en las solicitudes WO 2006/069110 y US 7.445.917.

- La biomasa lignocelulósica constituye una fuente de carbono duradera y una alternativa posible a las fuentes de carbono fósiles. Genera un interés creciente por la biosíntesis de metabolitos de interés industrial. Entre estos materiales carbonosos, la D-xilosa representa la segunda fuente más abundante después de la glucosa. No obstante, los microorganismos que no utilizan preferentemente pentosas y, cuando lo hacen, los rendimientos son extremadamente bajos. Por tanto, el desarrollo de vías metabólicas que asimilen este azúcar constituye un eje importante de investigación para la industria.

- En la actualidad, las principales fuentes de utilización de pentosas para la producción de EG o AG mediante microorganismos, se basan en la utilización de vías naturales, optimizadas para la producción de moléculas de interés y/o en la adición de enzimas para utilizar los productos de estas vías naturales.

- Por tanto, un método de producción AG orgánico optimizando las vías de asimilación de pentosas se describe en las solicitudes WO 2007140816 y WO 2007141336 presentadas por la entidad Metabolic Explorer. Estas solicitudes describen organismos genéticamente modificados a diferentes niveles para aumentar el flujo de la vía glioxilato, aumentar la conversión de glioxilato en glicolato y/o reducir el metabolismo de glicolato y de su intermedio, glioxilato. Una solicitud posterior de la entidad Metabolic Explorer (documento WO 2010108909) describe la posibilidad de tratar, además, para optimizar la producción de AG, mediante vías de producción de lactato (atenuando los genes que codifican metilglioxal-sintasa y D-lactato-deshidrogenasa), sobre la transición del metabolismo aeróbico/anaeróbico y atenuando el gen ArcA que regula el control de la vía respiratoria aeróbica y/o al atenuando

los genes que codifican proteínas importantes de glicolato.

Más recientemente, se han elaborado vías metabólicas de síntesis de EG basadas en vías metabólicas sintéticas de asimilación de pentosas. Así, Liu H *et al.* (Appl Microbiol Biotechnol, 2012, PMID: 23233208) describieron cepas de *E. coli* que expresan una vía metabólica de síntesis de etilenglicol a partir de D-xilosa y que implican enzimas (D)-xilosa deshidrogenasa, (D)-xilónico deshidratasa, 2-deshidro-3-desoxi-D-pentonato aldolasa y glicolaldehído reductasa. No obstante, los rendimientos de EG obtenidos mediante este método, de aproximadamente 275 mg por gramo de xilosa, pueden ser mejorados.

La institución Massachusetts Institute of Technology (MIT) ha descrito también en la solicitud WO 2013126721 una vía metabólica artificial de producción de EG a partir de pentosa. Esta vía implica la fosforilación en la posición 1 del ciclo de (D)-ribulosa o (L)-xilulosa. Esta vía no obstante necesita, para la asimilación de pentosas más abundantes como la (D)-xilosa y (L)-arabinosa, la expresión de varias isomerasas y epimerasas que permitan la conversión de D-xilosa en (D)-xilulosa y seguidamente en (D)-ribulosa y la conversión de (L)-arabinosa en (L)-ribulosa y seguidamente en (L)-xilulosa.

Por tanto, existe una necesidad en el estado de la técnica de nuevos procedimientos que efectúen la producción de metabolitos de interés mediante transformación de pentosas, particularmente de EG y/o de AG, ventajosamente a partir de fuentes carbonadas renovables y, más particularmente, basadas en la asimilación directa de pentosas más abundantes en estas fuentes naturales como (D)-xilosa y (L)-arabinosa.

Sumario de la invención

La invención descrita en la presente solicitud responde a los objetivos técnicos anteriormente mencionados. Se refiere a un procedimiento de transformación de al menos una pentosa en un microorganismo para la producción de al menos un metabolito de interés, mediante asimilación de dicha al menos una pentosa en un microorganismo. Más particularmente, suministra un procedimiento de biosíntesis simple y económico para producir de forma autónoma, en un microorganismo, un metabolito de interés y, particularmente etilenglicol y/o ácido glicólico, a partir de una fuente carbonada renovable como, por ejemplo, lignocelulosa y, en particular, hemicelulosa.

La utilización de este sustrato carbonado renovable suministra una alternativa duradera a la producción de metabolitos de interés (como etilenglicol y ácido glicólico), que son compuestos de alto valor añadido para la industria petroquímica y que en la actualidad continúan siendo producidos ampliamente por vía petroquímica.

La nueva vía de asimilación de pentosas descrita en la invención constituye una vía que es paralela a las vías naturales de asimilación de pentosas y que no existe en estado natural. Por tanto, es ampliamente independiente de las restricciones de regulación que rigen sobre las vías naturales de las células hospedantes. Por tanto, permite evitar vías naturales y sus regulaciones para producir metabolitos de interés, como etilenglicol y/o ácido glicólico.

Esta propiedad permite su portabilidad de forma simplificada a un amplio espectro de microorganismos hospedantes, ya que su metabolismo endógeno no interfiere con la vía sintética de la invención.

Además, los cálculos sobre los rendimientos teóricos de la vía sintética de asimilación de pentosas se descrita en la presente solicitud, para la producción de etilenglicol y ácido glicólico, permiten estimar una mejora considerable de los rendimientos con respecto a los procedimientos de biosíntesis basados en la optimización de vías naturales y/o una simplificación de la realización y de la producción.

El procedimiento de la invención se caracteriza por tanto porque comprende las etapas siguientes:

1. (i) una operación de cultivo de un microorganismo recombinante que expresa una vía sintética de asimilación de pentosas que comprende al menos una de las etapas de reacción siguientes:

1. a) fosforilación en la posición 1 de una pentosa escogida de (D)-xilulosa y/o (L)-ribulosa

2. b) escisión del pentosa-1-fosfato obtenido al final de la etapa a), para la obtención de glicolaldehído y de dihidroxiacetonafofato (DHAP), y

2. (ii) una operación de recuperación de al menos un metabolito de interés obtenido al final de la operación de cultivo (i).

5 Preferentemente, la etapa a) de la vía sintética de asimilación de pentosas puede ser catalizada por una enzima expresada de forma recombinante escogida entre el grupo constituido por cetohecho quinasa C (KHK-C), la ramnulosa quinasa (Rhab) y fuculosa quinasa (fucK).

De forma también preferente, la etapa b) es catalizada mediante una aldolasa, preferentemente de clase I. Una aldolasa de clase I según la invención se escoge normalmente entre el grupo constituido por aldolasa B (Aldo-B) y la fructosa-1,6-bis-fosfato-aldolasa (fructosa 1,6 bPaldolasa o FbaB).

Según un modo de realización, el procedimiento de producción se caracteriza porque:

- 10
- la etapa a) es catalizada por KhkC, y
 - la etapa b) es catalizada por aldolasa B (aldo-B) y/o la fructosa-1,6-bP aldolasa (FbaB)

Un modo particular de realización del procedimiento de la invención permite obtener etilenglicol (EG) y/o ácido glicólico (AG).

En este procedimiento, la vía sintética de asimilación de pentosas comprende además las etapas siguientes:

- 15
- c) reducción del glicolaldehído obtenido al final de la etapa b) en etilenglicol, y/o
 - c') oxidación del glicolaldehído obtenido al final de la etapa b) a ácido glicólico.

Los modos de realización descritos a continuación pueden ser combinados entre ellos salvo indicación en contra, en el procedimiento de la invención.

20 La etapa c) puede ser catalizada por un glicolaldehído reductasa, particularmente escogida entre el grupo constituido por: aldehído reductasa (YqhD), glicerol deshidrogenasa (GldA) y propano-1,2-dioloxidoreductasa (FucO).

La etapa c') puede ser catalizada por una glicolaldehído deshidrogenasa, particularmente lactaldehído deshidrogenasa (AldA).

En un procedimiento de la invención, el microorganismo preferentemente se pone en cultivo sobre un medio carbonado que contiene (D)-xilosa y/o (L)-arabinosa.

25 En un modo de realización preferido del procedimiento de la invención, el medio de cultivo comprende un hidrolizado de biomasa que comprende hemicelulosa. Normalmente, la vía sintética de asimilación de pentosas del procedimiento según la invención comprende, previamente a la etapa a), al menos una de las etapas siguientes:

- una etapa de transporte de (D)-xilosa y/o (L)-arabinosa en el microorganismo;
- una etapa de conversión de (D)-xilosa en (D)-xilulosa, y/o
- 30 • una etapa de conversión de (L)-arabinosa en (L)-ribulosa.

La conversión de (D)-xilosa en (D)-xilulosa puede ser catalizada por una (D)-xilosa isomerasa y/o por la acción de una (D)-xilosa reductasa y una xilitol deshidrogenasa.

La conversión de (L)-arabinosa en (L)-ribulosa puede ser catalizada por una (L)-arabinosa isomerasa y/o por la acción de una (L)-arabinosa reductasa y una arabitól deshidrogenasa.

35 Según el procedimiento de producción según la invención, el microorganismo recombinante puede ser una bacteria,

preferentemente seleccionada entre el grupo constituido por en *Enterobacteriaceae*, *Clostridiaceae*, *Bacillaceae*, *Streptomycetaceae*, *Streptococcaceae*, *Methylobacteriaceae* y *Corynebacteriaceae* preferentemente *Escherichia coli*, *Bacillus subtilis*, *Corynebacterium glutamicum*, *Clostridium acetobutylicum*, *Methylobacterium extorquens* o *Lactococcus lactis*. El microorganismo puede ser también una levadura, preferentemente escogida entre

5 *Saccharomycetaceae*, *Pichiaceae* y *Schizosaccharomycetaceae*, preferentemente *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Pichia jadinii*, *Scheffersomyces stipitis* o *Pichia pastoris*. El microorganismo de la invención puede ser también un hongo, preferentemente escogido entre el grupo cosntituido por *Penicillium*, *Aspergillus*, *Chrysosporium* o *Trichoderma*.

10 Preferentemente, las actividades de (D)-xilulosa-5-quinasa y/o (L)-ribulosa-5-quinasa del microorganismo utilizado en el procedimiento de la invención han sido suprimidas.

La invención se refiere también a un microorganismo recombinante que expresa una vía sintética de asimilación de pentosas que comprende al menos ácidos nucleicos que codifican las enzimas siguientes:

1. i) una enzima adecuada para fosforilar en posición 1 las pentosas escogidas entre (D)-xilulosa y/o (L)-ribulosa.

2. ii) una enzima adecuada para escindir dicha pentosa fosforilada en posición 1 en glicolaldehído y DHAP.

15 Ventajosamente, las vías naturales de asimilación de pentosas de este microorganismo han sido suprimidas.

Generalmente, un microorganismo que expresa la vía sintética de asimilación de la invención se caracteriza además porque las actividades de (D)-xilulosa-5-quinasa y/o (L)-ribulosa-5-quinasa han sido suprimidas y/o porque porta al menos uno de las modificaciones siguientes:

- sobreexpresión de un gen que codifica una glioxilato reductasa;

20 • sobreexpresión de un gen que codifica una isocitrato liasa;

- delección de los genes que codifican malato sintasa;

- delección de los genes que codifican glioxilato carboligasas;

- delección de los genes que codifican los genes que codifican glicolato oxidasas y/o glicolatos deshidrogenasas

25 • delección de los genes que codifican 2-ceto-4-hidroxi-glutarato aldolasas, particularmente la Entner-Doudouff aldolasa y/o fosfogluconato deshidratasa;

- delección de un gen represor de la respuesta aeróbica, particularmente el gen *arcA*;

- atenuación y, particularmente, delección de la expresión de isocitrato deshidrogenasa;

- delección de los genes codificadores para sistemas de internalización de ácido glicólico;

30 • atenuación de las vías metabólicas que conducen a la producción de productos secundarios como acetato, lactato o etanol;

- sobreexpresión de al menos un gen que codifica un transportador de azúcares.

Preferentemente, el microorganismo recombinante comprende:

- al menos un ácido nucleico exógeno que codifica cetohecoquinasa C y

- al menos un ácido nucleico exógeno que codifica aldolasa B.

Ventajosamente, dicho microorganismo puede comprender las modificaciones siguientes:

- sobreexpresión del gen que codifica glicolaldehído reductasa principal;
- delección del gen que codifica glicolaldehído deshidrogenasa.

5 En una variante, el microorganismo puede comprender al menos una de las modificaciones siguientes para optimizar la producción de ácido glicólico:

- sobreexpresión del gen que codifica glicolaldehído deshidrogenasa;
- delección de al menos uno de los genes que codifican al menos una de las subunidades de glicolato oxidasa.

Normalmente, un microorganismo adecuado para optimizar la producción de ácido glicólico comprende las siguientes modificaciones:

- 10
- sobreexpresión de un gen que codifica glioxilato reductasa;
 - sobreexpresión de un gen que codifica una isocitrato liasa, acompañado o no de la delección de su represor transcripcional;
 - delección de los genes que codifican malato sintasas;
 - delección de los genes que codifican glioxilato carboligasas

- 15
- delección de los genes que codifican glicolato oxidasas y/o glicolato deshidrogenasas;
 - delección de los genes que codifican 2-ceto-4-hidroxi-glutaratoaldolasas, particularmente Entner-Doudouff aldolasa y/o fosfogluconato deshidratasa;
 - delección de un gen que codifica un represor de genes implicados en el metabolismo respiratorio, particularmente el gen *arcA*;

- 20
- atenuación y, particularmente, delección de la expresión de isocitrato deshidrogenasa;
 - ocasionalmente la sobreexpresión de un gen que codifica un transportador de azúcares.

Todavía, preferentemente, particularmente en el caso de que el microorganismo recombinante para la producción de ácido glicólico sea *Escherichia coli*, el microorganismo comprende las modificaciones siguientes:

- la sobreexpresión del gen *aldA* que codifica una glicolaldehído deshidrogenasa;
- 25
- la sobreexpresión del gen *ghrA*;
 - la sobreexpresión del gen *aceA* ocasionalmente acompañada de la delección del gen *icIR* que codifica un represor transcripcional de la vía del glioxilato;
 - la delección de los genes *aceB* y *glcB*;
 - la delección de al menos uno de los genes *glcDEFG*;
- 30
- la delección del gen *glc*;
 - la delección de los genes *edd-eda*;

- la delección del gen *arcA*;
- la delección del gen *icd*;
- la sobreexpresión del gen *galP*

Figuras

5 Figura 1. Vías naturales y vías sintéticas de asimilación de (D)-xilosa y (L)-arabinosa. Las reacciones catalizadas por las enzimas naturales están representadas en líneas discontinuas. Las reacciones catalizadas por enzimas sintéticas están representadas en trazos continuos. (1) (D)-xilosa isomerasa, (2) (D)-xilulosa-1-quinasa, (3) (D)-xilulosa-1-fosfato aldolasa, (4) glicolaldehído deshidrogenasa, (5) glicolaldehído reductasa, (6) (L)-arabinosa isomerasa, (7) (L)-ribulosa-1-quinasa, (8) (L)-ribulosa-1-fosfato aldolasa. El nombre de los genes bajo las reacciones
10 corresponde a los genes de *Escherichia coli* que codifican la enzima con la actividad correspondiente.

Figura 2. Síntesis *in vitro* de etilenglicol mediante la vía sintética de asimilación de (D)-xilosa. Cromatogramas HPLC de (A) una solución 10 mM de etilenglicol y (B) una mezcla de reacción compuesta por (D)-xilulosa-1-quinasa (Khk-C, 0,005 Unidades/ml), (D)-xilulosa-1-fosfato aldolasa (AldoB, 1 Unidad/ml (Sigma-Aldrich A6338)) y una glicolaldehído reductasa (GldA, 1 Unidad/ml (Sigma-A G3512-250U)) o (r) una (D)-xilulosa-1-quinasa (Khk-C, 0,005
15 Unidad/ml) y una glicolaldehído reductasa (GldA, 1 Unidad/ml (Sigma-Aldrich G3512-250U)). Las enzimas fueron incubadas durante 3 horas a 37 °C en un tampón de reacción Hepes que contenía Hepes 100 mM; KCl 85 mM; MgCl₂ 7,5 mM a pH = 7; ATP 2 mM; ZnCl₂ 5 mM; NADH 0,4 mM. Las reacciones fueron iniciadas mediante la adición de (D)-xilulosa 5 mM.

Leyenda: Intensidad (eje de ordenadas) en función del tiempo en minutos (eje de abscisas)

20 Figura 3. Síntesis *in vitro* de etilenglicol por vía sintética de asimilación de (L)-arabinosa. Cromatogramas HPLC de (A) una solución de etilenglicol 1 mM y (B) una mezcla de reacción compuesta por una (L)-ribulosa-1-quinasa, (L)-ribulosa-1-fosfato aldolasa y una glicolaldehído reductasa. Las enzimas fueron incubadas durante 3 horas a 37 °C en un tampón de reacción Hepes que contenía Hepes 55 mM; KCl 45 mM; MgCl₂ 4 mM a pH = 7; ATP 4 mM; NADH 0,4 mM. Las reacciones fueron iniciadas mediante la adición de (L)-ribulosa 20 mM. Leyenda: Intensidad (eje de
25 ordenadas) en función del tiempo en minutos (eje de abscisas).

Figura 4. Crecimiento de cepas de *E. coli* en medio mínimo que contenía (D)-xilosa Leyenda: DO a 600 nm (eje de ordenadas) en función del tiempo en minutos (eje de abscisas)

Figura 5. Seguimiento del crecimiento y producción de metabolitos sobre xilosa de una cepa de *E. coli* MG1655 $\Delta xylB$ *pEXT20 khk-C-aldB*.

30 Leyenda: DO a 600 nm (eje de ordenadas a la izquierda) y xilosa en mM, etilenglicol en mM y ácido glicólico en mM (eje de ordenadas a la derecha) en función del tiempo en horas (eje de abscisas)

Figura 6. Producción de etilenglicol (en mM sobre el eje de ordenadas) por mutantes de glicolaldehído reductasa candidatos después de la incubación durante 12 h en presencia de glicolaldehído 10 mM.

35 Figura 7. Optimización de la producción de etilenglicol en diferentes mutantes según la invención expresada en rendimiento en mol/mol de xilosa (en eje de ordenadas) a través de la vía metabólica sintética de asimilación de (D)-xilosa.

Figura 8. Optimización de la producción de ácido glicólico en diferentes mutantes según la invención expresada en rendimiento en mol/mol de xilosa (en eje de ordenadas) a través de la vía metabólica sintética de asimilación de (D)-xilosa.

40 Figura 9. Crecimiento de la cepa 905 en medio mineral en presencia de glucosa y (D)-xilosa. Leyenda: DO a 600 nm (eje de ordenadas a la izquierda) y glucosa, xilosa, ácido glicólico y acetato en mM y mM (eje de ordenadas a la derecha) en función del tiempo en horas (eje de abscisas).

Figura 10. Crecimiento de la cepa 979 en medio mineral en presencia de glucosa y (D)-xilosa. Leyenda: DO a 600 nm (eje de ordenadas de la izquierda) y glucosa, xilosa, ácido glicólico y acetato en mM y mM (eje de ordenadas a la derecha) en función del tiempo en horas (eje de abscisas).

5 Figura 11. Crecimiento en medio mineral en presencia de xilosa como única fuente de carbono de cepas CEN.PK-2-1. Leyenda: DO a 600 nm (eje de ordenadas a la derecha) y xilosa en mM (eje de ordenadas a la izquierda) en función del tiempo en horas (eje de abscisas).

Figura 12. Crecimiento en medio mineral y xilosa como única fuente de carbono de cepas TMB3001. Leyenda: DO a 600 nm en función del tiempo en horas.

10 Figura 13. Crecimiento en medio mineral M9 y (L)-arabinosa de cepas Δ araB con o sin vía sintética. Leyenda: DO a 600 nm en función del tiempo en horas.

Definiciones

Salvo indicación particular, los términos técnicos y científicos empleados en la presente solicitud tienen el significado habitual, entendido por un experto en la técnica adecuado para realizar la invención.

15 Salvo que se indique igualmente en contra, los diferentes modos de realización descritos a continuación pueden ser combinados entre ellos en la realización de la invención.

Mediante "vía de asimilación de pentosas" se entiende, según la invención, una vía metabólica, es decir, un conjunto de reacciones químicas que se desarrollan en el microorganismo, catalizadas por una serie de enzimas que intervienen de forma secuencial, que utilizan las pentosas como sustrato inicial y que dan lugar a su transformación para la formación de metabolitos de interés.

20 Mediante vía natural de asimilación de pentosas se entiende la asimilación de pentosas que implica su fosforilación en la posición 5, seguido de su utilización en la vía metabólica denominada pentosas fosfatos, que existe de forma natural en la mayoría de las células eucariotas y procariotas. Normalmente, la (L)-arabinosa es isomerizada en (L)-ribulosa, que seguidamente es fosforilada en la posición 5. El (L)-ribulosa-5-fosfato obtenido es epimerizado en el carbono C3 para producir (D)-xilulosa-5-fosfato, sustrato de la vía de pentosas fosfatos.

25 La expresión "ruta sintética" o "vía metabólica sintética" significa, según la invención, que dicha vía metabólica no se realiza de forma natural por el microorganismo. Esta condición es normalmente respetada cuando al menos una de las enzimas de dicho microorganismo que cataliza al menos una de las etapas a) o b) de la vía metabólica de la invención no es expresada de forma natural o cuando dicha al menos una enzima, cuando es expresada, no cataliza dicha al menos una etapa a) o b).

30 Como ejemplos, (i) la vía metabólica que implica la fosforilación en posición 1 de una pentosa escogida entre (D)-xilulosa y (L)-ribulosa, por la Kxk-C de origen humano en una célula no humana y normalmente en un microorganismo, (ii) la vía metabólica que implica la fosforilación en posición 1 de una pentosa escogida entre (D)-xilulosa y (L)-ribulosa, por ramnulosa quinasa RhaB en un microorganismo y (iii) la vía metabólica que implica la fosforilación en la posición 1 de una pentosa escogida entre (D)-xilulosa y (L)-ribulosa, mediante fuculosa quinasa FucK en un microorganismo, son vías sintéticas según la invención.

35 La expresión de las enzimas RhaB y FucK en el seno de *E. coli* depende de la presencia de su sustrato natural, a saber, respectivamente (L)-ramnulosa y (L)-fucosa. En ausencia de su sustrato natural (o cuando este está presente en una concentración demasiado baja), la expresión de estas enzimas en un microorganismo y particularmente en el seno de *E. coli* puede ser obtenida de forma recombinante bajo el control de un promotor de forma inducible o constitutiva.

40 El término "transformación" o "transfección" se refiere a la adquisición de nuevos genes funcionales en una célula, después de la incorporación de ácidos nucleicos exógenos.

El término "modificación" o "modificar" relacionado con el nivel de proteína o de actividad enzimática producida por

una célula hospedante se refiere al control de los niveles de proteína o de actividad enzimática producida durante el cultivo, de forma que estos niveles sean aumentados o disminuidos como se desee.

5 El término “modificado”, referido a un ácido nucleico o un polinucleótido, significa que el ácido nucleico se ha modificado con respecto a la versión salvaje contenida en la célula hospedante, particularmente mediante una mutación de tipo sustitución, inserción, delección de una parte o la totalidad de dicho ácido nucleico o que dicho ácido nucleico ha sido unido de forma funcional a una zona de control de la transcripción.

10 Mediante “gen” se entiende, según la invención, un segmento de ADN implicado en la codificación de ARN ribosómicos, ARN reguladores, ARN de transferencia, secuencias reguladoras (que comprenden normalmente una región promotora) unidos de una manera funcional a la expresión de un péptido o un polipéptido o de una proteína, que incluyen regiones codificadoras (transcritas en ARN mensajero) y no codificadoras que preceden o que terminan la región codificadora, así como los intrones (regiones no codificadoras que separan las regiones codificadoras o exones).

15 La expresión “unido de forma funcional” se refiere a una yuxtaposición de elementos de forma que su disposición les permite estar funcionalmente unidos. Una secuencia reguladora, que contiene normalmente una región promotora, está funcionalmente unida a una región codificadora cuando controla la transcripción de la región codificadora, y un sitio de unión de un ribosoma está funcionalmente unido a una región codificadora cuando está colocado de forma que permita la traslación del ARNm.

20 El término “inactivación” o “supresión” o “atenuación” se refiere a la expresión disminuida o reducida o significativamente reducida de un gen o la actividad disminuida o reducida de una proteína o del producto de un gen, normalmente una enzima. Diferentes métodos conocidos por el experto en la técnica pueden ser empleados con este fin, como:

- la introducción de una mutación en el gen que da lugar a una expresión de una proteína cuya actividad es reducida,
- la sustitución del promotor natural por un promotor poco activo que da lugar a una baja expresión del gen,
- 25 • la utilización del elemento desestabilizante de ARNm correspondiente a la proteína, o
- la delección del gen.

Normalmente, la atenuación de un gen o de una proteína se define mediante una actividad de la proteína expresada por dicho gen disminuida en al menos 50%, preferentemente en al menos 60%.

30 La inactivación o la delección o la supresión de un gen o de una proteína se define mediante una actividad residual de la proteína producto de dicho gen inferior a 20%, particularmente inferior a 10%, particularmente inferior a 5%.

El término “expresión” corresponde a la transcripción y a la traslación de un gen hacia una proteína, producto de dicho gen.

35 El término “sobrexpresión” corresponde a una expresión aumentada con respecto a la expresión natural de dicho gen en la misma célula hospedante. Normalmente, la sobreexpresión de una proteína se define mediante una actividad de al menos 200% de dicha proteína con respecto a su expresión natural en la célula hospedante.

La sobreexpresión de una proteína puede ser obtenida según diferentes técnicas conocidas por un experto en la técnica, como:

- la mutación de una proteína, con el fin de obtener una forma más activa o resistente a la inhibición,
- 40 • el aumento de la expresión del gen codificador para dicha proteína (por ejemplo, mediante la introducción de un promotor específico que regula la expresión del gen),

- la adición de múltiples copias de genes en la célula hospedante, etc.

Las células hospedadoras compatibles con la invención pueden expresar una copia endógena de uno o varios genes que codifican una proteína de interés según la invención, así como ocasionalmente una copia recombinante de dicho gen.

- 5 Un ácido nucleico que codifica una enzima asociada a la presente invención puede ser introducido en una célula hospedante según cualquier técnica estándar conocida por un experto en la técnica. Como ejemplo, los ácidos nucleicos pueden ser introducidos mediante transformación (química o electroporación), transfección, infección, transducción, etc.

- 10 Los genes que codifican proteínas asociadas a la invención pueden ser expresados de forma extracromosómica en un vector de expresión recombinante o pueden ser integrados en cromosomas.

- 15 Mediante “vector”, según la invención, se entiende un ácido nucleico en el que es insertada una secuencia de interés, mediante restricción y ligadura de forma que esté asociado de forma operativa a secuencias de regulación, para la expresión como transcrito de ARNm en una célula hospedante. Los vectores son compuestos por ARN o , preferentemente, ADN e incluyen de forma no limitativa plásmidos, fagémidos, genomas virales, bacteriófagos y cromosomas bacterianos.

- 20 Un vector según la invención puede comprender una o varias secuencias marcadoras para la identificación de células transformadas con dicho vector. Estos marcadores incluyen, por ejemplo, genes que codifican proteínas que aumentan o disminuyen la resistencia o la sensibilidad a compuestos antibióticos, genes que codifican enzimas cuya actividad es detectable mediante ensayos estándar (por ejemplo, luciferasa, galactosidasa, fosfatasa alcalina, etc.) o genes que modifican el fenotipo de las células transformadas (por ejemplo, que codifican GFP, es decir, “green fluorescent protein”).

Las secuencias reguladoras (promotoras) utilizadas en la expresión de proteínas recombinantes de la invención pueden ser secuencias reguladoras endógenas (es decir, el promotor nativo del gen al que está asociado) o exógenas. El promotor puede ser inducible o constituido.

- 25 Mediante “célula hospedante” o “microorganismo hospedante” se entiende cualquier tipo de célula susceptible de someterse a una transformación, transfección, transducción, etc., con una construcción de ácido nucleico o un vector de expresión que comprende uno o varios polinucleótidos, en particular uno o varios polinucleótidos que codifican enzimas descritas en la solicitud.

- 30 Las quinasas son enzimas del grupo de las transferasas que catalizan reacciones de fosforilación mediante la adición de un ión fosfato a una molécula diana.

Las oxidorreductasas son enzimas que catalizan reacciones de oxidorreducción y que transfieren H^+ y electrones. Están asociadas a coenzimas de oxidorreducción (NAD, FAD, FMN, etc.)

La deshidrogenasa (o deshidrogenasa) es una enzima que oxida un sustrato mediante transferencia de uno o varios iones (H^+) a un aceptor, generalmente una coenzima de tipo $NAD^+/NADP^+$ o flavina como FAD o FMN.

- 35 Una aldehído deshidrogenasa es una enzima de tipo deshidrogenasa que cataliza la oxidación de aldehídos.

- 40 Cuando las enzimas mencionadas en la presente solicitud son identificadas por su actividad específica, esta definición incluye todos los polipéptidos que tienen la misma actividad específica y que están presentes en diferentes células y, particularmente, diferentes microorganismos. La invención por tanto se refiere también a las proteínas homólogas de las proteínas de referencia mencionadas en la presente solicitud que tienen la misma actividad que las proteínas de referencia, así como a los genes que codifican dichas proteínas homólogas.

En ausencia de precisión, los genes y las proteínas mencionadas en la presente solicitud son identificadas en referencia a *E. coli* (particularmente a la cepa MG1655). La Khk-C y la aldolasa B son identificadas en referencia a *H. sapiens*. No obstante, las proteínas y, por tanto, los genes homólogos a las proteínas (y a los genes que las

codifican) identificadas en la presente solicitud pueden encontrarse en diferentes microorganismos.

Una proteína homóloga a una proteína de referencia según la invención posee la misma función, es decir, en su caso para una enzima, cataliza la misma reacción que la enzima de referencia. Un gen homólogo de un gen que codifica una proteína de referencia según la presente invención codifica una proteína homóloga como se definió con anterioridad.

5

Normalmente, a partir del nombre de la proteína y de su secuencia, el experto en la técnica es capaz de identificar en otros microorganismos equivalentes de las proteínas mencionadas en la presente solicitud. Este trabajo rutinario se efectúa habitualmente utilizando secuencias de consenso identificadas por alineaciones de secuencias con otras proteínas procedentes de diferentes organismos.

10 De forma igualmente preferente, una proteína homóloga de una proteína de referencia corresponde a una enzima que tiene al menos 30% de identidad de secuencia, preferentemente 40%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% o 95% de identidad con la secuencia de la proteína de referencia.

15 Con el fin de determinar el porcentaje de identidad de dos secuencias de aminoácidos para las necesidades de la invención, las secuencias serán alineadas para permitir una comparación óptima. Pueden ser introducidos espacios (huevos) en una u otra de las secuencias que van a ser alineadas con el fin de permitir una alineación óptima y las secuencias no homólogas pueden ser ignoradas para la comparación.

20 El porcentaje de identidad de dos secuencias de aminoácidos comparadas puede ser obtenido como se describe en el libro de D. Voet y J. G. Voet, Biochemistry (2ª Edición, De Boeck & Larcier, 2005, sección 7.4, párrafo B). Las alineaciones se realizan con el programa CLUSTAL W (versión 1.82) con los parámetros siguientes: (1) CPU MODE = ClustalW mp ; (2) ALIGNMENT = "full"; (3) OUTPUT FORMAT = "aln w/numbers"; (4) OUTPUT ORDER = "aligned"; (5) COLOR ALIGNMENT = "no"; (6) KTUP (word size) = "default"; (7) WINDOW LENGTH = "default"; (8) SCORE TYPE = "percent"; (9) TOPDIAG = "default"; (10) PAIRGAP = "default"; (11) PHYLOGENETIC TREE/TREE TYPE = "none"; (12) MATRIX = "default"; (13) GAP OPEN = "default"; (14) END GAPS = "default"; (15) GAP EXTENSION = "default"; (16) GAP DISTANCES = "default"; (17) TREE TYPE = "cladogram" y (18) TREE GRAP DISTANCES = "hide".

25

La lignocelulosa está compuesta por ligninas, hemicelulosas y celulosa en proporciones variables. Las hemicelulosas son uno de los tres componentes principales de la biomasa lignocelulósica y representan aproximadamente 20-40% en peso de dicha biomasa.

30 Mediante "hemicelulosa" se entiende según la invención, un grupo de polisacáridos complejos que se caracterizan por su solubilidad en soluciones alcalinas (por ejemplo, KOH 1 M) y su insolubilidad en agua. Las hemicelulosas se definen estructuralmente como polisacáridos cuya cadena principal está compuesta por residuos β -(1,4)-D-piranosas, en la que el O4 está en la posición ecuatorial. Están fijadas cadenas laterales cortas sobre la cadena principal. Las hemicelulosas comprenden xilanos, arabinoxilanos, xiloglucanos, glucoronoxilanos y glucomananos. Su hidrólisis realizada, por ejemplo, mediante contacto de un material lignocelulósico con ácido sulfúrico diluido a presiones y temperaturas elevadas, conduce a la liberación de azúcares monómeros. Según la naturaleza de la materia prima y las condiciones de hidrólisis, los porcentajes de xilosa, glucosa y arabinosa varían respectivamente de 60 a 80%, de 10 a 30% y de 10 a 30% en peso con respecto al peso total del hidrolizado de lignocelulosa.

35

40 Las lignocelulosas procedentes de maderas duras (normalmente árboles caducifolios), mazorcas de maíz, hierbas, hojas y periódicos tienen particularmente un contenido elevado de azúcares hemicelulósicos (Jorgensen H *et al.*, Enzymatic conversion of lignocellulose into fermentable sugars: Challenges and opportunities, Biofuels, Bioprod Bioref, 2007, 1, 119-134). Representan fuentes de materias primas preferidas para la utilización de microorganismos modificados según la invención. Por "iniciador" o "cebador" se entiende una secuencia corta de ADN, complementaria del comienzo de una matriz, que sirve de punto de partida para la síntesis de la cadena complementaria de dicha matriz por una ADN polimerasa.

40

45 Descripción detallada

Vía sintética de asimilación de pentosas

La presente invención se refiere a un procedimiento de transformación de una pentosa en un microorganismo recombinante que expresa una vía sintética de asimilación de pentosas para la producción de al menos un metabolito de interés.

Este procedimiento según la invención comprende:

5 1. (i) una operación de cultivo un microorganismo recombinante que expresa una vía sintética de asimilación de pentosas, ilustrada de forma general en la figura 1, que comprende al menos las siguientes etapas:

1. a) la fosforilación en posición 1 de una pentosa escogida entre (D)-xilulosa y/o (L)-ribulosa, para la obtención respectivamente de (D)-xilulosa-1P y/o (L)-ribulosa-1P,

10 2. b) la escisión del pentosa-1-fosfato obtenido después de la etapa a) ((D)-xilulosa-1P y/o (L)-ribulosa-1P), para la obtención de glicolaldehído y de dihidroxiacetona fosfato (DHAP),

permitiendo dicha vía obtener al menos un metabolito de interés, y

2. (ii) una operación de recuperación de dicho al menos un metabolito de interés obtenido al final de la operación de cultivo (i).

15 Mediante "fosforilación" se entiende, ventajosamente, la adición de un grupo fosfato, en la forma de un fosforilo PO_3^{2-} .

Mediante "metabolito de interés" se entiende, particularmente, el glicolaldehído y el DHAP, pero también sus derivados susceptibles de ser obtenidos mediante reacciones de oxidación o reducción de estos compuestos, en particular etilenglicol (EG), ácido glicólico (AG) y sus derivados.

Mediante "derivados del ácido glicólico" se entiende en particular:

20 • ésteres de glicolato, como el éster glicolato de etilo o el éster glicolato de metilo, así como:

• polímeros que contienen glicolato como poli(ácido glicólico)

• así como el ácido glioxílico procedente de una oxidación de ácido glicólico.

Debe apreciarse que en la presente solicitud, las expresiones "ácido glicólico" y "glicolato", así como las expresiones "ácido glioxílico" y "glioxilato" son utilizadas como sinónimos.

25 Preferentemente, la vía sintética de asimilación de pentosas expresada por el microorganismo comprende por lo tanto, además, de forma ventajosa, las etapas siguientes:

c) reducción del glicolaldehído obtenido al final de la etapa b) en etilenglicol, o

c') oxidación del glicolaldehído obtenido al final de la etapa b) a ácido glicólico.

30 En estos modos de realización, los metabolitos de interés obtenidos al final de la vía sintética de asimilación de pentosas según la invención son etilenglicol y/o ácido glicólico, y sus derivados.

Enzimas de la invención

La vía sintética de asimilación de pentosas, representada en la figura 1, es catalizada por un conjunto de enzimas.

La enzima recombinante que cataliza la etapa de fosforilación a) de la vía sintética de asimilación de pentosas de la invención es una quinasa que fosforila la (D)-xilulosa o la (L)-arabinosa en la posición 1.

Esta enzima se escoge, por ejemplo, entre el grupo que constituido por:

- cetohecoquinasa, preferentemente la isoforma C de la cetohecoquinasa (KhK-C),
- ramnulosa quinasa (RhaB) y
- fuculosa quinasa (FucK).

5 La cetohecoquinasa C es codificada por el gen *khk*, que se encuentra normalmente en el *H. sapiens*.

En un modo de realización preferido, se utiliza el gen de *H. sapiens* que codifica *Khk-C*, de secuencia SEQ ID N° 1.

10 La ramnulosa quinasa así como la fuculosa quinasa de la invención son codificadas, respectivamente, por los genes *rhaB* y *fucK*, que se encuentran normalmente en *E. coli*. Por tanto, en ciertos modos de realización, se utiliza el gen *rhaB* de *E. coli* que codifica la ramnulosa quinasa B (RhaB) de secuencia SEQ ID N° 6 o el gen *fucK* de *E. coli* que codifica la fuculosa quinasa (FucK) de la secuencia SEQ ID. N° 5.

La enzima que cataliza la etapa de escisión b) es una aldolasa que escinde (D)-xilulosa-1P o (L)-ribulosa-1P en glicolaldehído y DHAP.

15 Una aldolasa según la invención puede ser escogida entre aldolasa B, codificada por el gen *aldoB*, que se encuentra normalmente en *Homo sapiens* y fructosa-1,6 bisfosfato aldolasa B de *E. coli*, codificada por el gen *fbaB* que se encuentra normalmente en *E. coli*.

Por tanto, en ciertos modos de realización particulares, se usa el gen *H. sapiens aldoB* que codifica la aldolasa B, de secuencia SEQ ID N° 2 o el gen *E. coli fbaB* de que codifica la fructosa-1,6-bifosfato aldolasa B, de secuencia SEQ ID N° 9.

La enzima que cataliza la etapa de reducción c) es una glicolaldehído reductasa.

20 Una glicolaldehído (o aldehído) reductasa conveniente para la invención puede ser escogida, por ejemplo, entre la aldehído reductasa codificada por:

- el gen *yqhD*, que se encuentra normalmente en *E. coli*, que codifica la aldehído reductasa YqhD de la secuencia SEQ ID N° 4,
- 25 • la glicerol deshidrogenasa codificada por el gen *gldA*, que se encuentra normalmente en *E. coli*, que codifica la glicerol deshidrogenasa GldA y de secuencia SEQ ID N° 51 y
- la L-1,2-propanodiol oxidorreductasa codificada por el gen *FucO*, que se encuentra normalmente en *E. coli*, que codifica la L-1,2-propanodiol oxidorreductasa FucO y de secuencia SEQ ID N° 52.

La enzima que cataliza la etapa de oxidación c') es una glicolaldehído deshidrogenasa.

30 Una glicolaldehído deshidrogenasa que conviene para la invención es, por ejemplo, la glicolaldehído deshidrogenasa codificada por el gen *aldA*, que se encuentra normalmente en *E. coli*, que codifica la lactaldehído deshidrogenasa AldA, de secuencia SEC ID N° 3.

Las enzimas catalizan, respectivamente, las etapas de reducción y oxidación ventajosamente en presencia de la forma reducida u oxidada, respectivamente, de nicotinamida adenina dinucleótido (fosfato) (NAD(P)H), coenzima de oxidorreducción.

35 En ciertos modos de realización, el microorganismo utilizado en la presente invención expresa, ventajosamente, de forma nativa o recombinante, al menos una de las siguientes enzimas:

- una xilosa isomerasa, que convierte la (D)-xilosa en (D)-xilulosa como, por ejemplo, la enzima codificada por el

gen *xylA* de *E. coli*,

- una xilosa reductasa y una xilitol deshidrogenasa como, por ejemplo, las enzimas codificadas por los genes *XYL1* y *XYL2* de la levadura *Scheffersomyces stipitis*.

5 • una L-arabinosa isomerasa, que convierte la (L)-arabinosa en (L)-ribulosa como, por ejemplo, la enzima codificada por el gen *E. coli araA*, y/o

- una arabinosa reductasa y una arabitól deshidrogenasa.

Todavía de forma preferida, el microorganismo de la invención expresa, de forma nativa o recombinante, al menos una proteína que transporta las pentosas al interior de la célula y, particularmente:

10 • proteínas que transportan (L)-arabinosa como, por ejemplo, las enzimas codificadas por los genes *araE*, *araF*, *araG* o *araH* de *E. coli*; y /o

- proteínas que transportan (D)-xilosa como, por ejemplo, proteínas codificadas por los genes *xylE*, *xylF*, *xylG* o *xylH* de *E. coli* o incluso el gen *galP* que codifica una permeasa de azúcares, o el gen *gal-2a* de *S cerevisiae*.

Microorganismo recombinante

15 Mediante “microorganismo” se entiende, según la invención, una célula hospedante escogida entre las células procariotas, particularmente las arqueobacterias, las bacterias o microalgas procariotas y células eucariotas, particularmente los hongos, las levaduras y las células vegetales y las microalgas eucariotas.

20 Mediante “microorganismo recombinante” o “microorganismo genéticamente modificado” o “microorganismo modificado” se entiende, según la invención, una célula hospedante que ha sido sometida a una modificación de su genoma, por ejemplo, mediante la adición de un ácido nucleico exógeno (o recombinante) o mediante la modificación de un ácido nucleico endógeno.

Las bacterias adecuadas para la invención pueden ser escogidas, por ejemplo, entre las familias de las *Enterobacteriaceae*, *Clostridiaceae*, *Bacillaceae*, *Streptomycetaceae*, *Streptococcaceae*, *Methylobacteriaceae* y *Corynebacteriaceae*.

25 Las bacterias particularmente adecuadas para la invención pueden ser escogidas normalmente entre el grupo constituido por *Escherichia coli*, *Bacillus subtilis*, *Corynebacterium glutamicum*, *Clostridium acetobutylicum*, *Methylobacterium extorquens* y *Lactococcus lactis*.

Las levaduras adecuadas para la invención se pueden escogidas, por ejemplo, entre las familias de las *Saccharomycetaceae*, *Pichiaceae* y *Schizosaccharomycetaceae*.

30 Las levaduras particularmente adecuadas para la invención pueden ser escogidas normalmente entre el grupo constituido por *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Pichia jadinii*, *Scheffersomyces stipitis* y *Pichia pastoris*.

Los géneros de hongos adecuados para la invención pueden ser escogidos normalmente entre el grupo constituido por *Penicillium*, *Aspergillus*, *Chrysosporium* y *Trichoderma*.

35 En un modo de realización preferido de la invención, se utiliza *Escherichia coli*, *Scheffersomyces stipitis* o *Saccharomyces cerevisiae* como microorganismos hospedante.

Ventajosamente, se utiliza un microorganismo capaz por naturaleza de asimilar (D)-xilosa y/o (L)-arabinosa.

Preferentemente, la vía sintética de asimilación de pentosas según la invención implica que al menos una enzima que cataliza una de las etapas de fosforilación a), de escisión b), de reducción c) o de oxidación c') es expresada de forma recombinante por el microorganismo.

Normalmente, al menos la enzima que cataliza la etapa de fosforilación a) y/o al menos la enzima que cataliza la etapa de escisión b) del procedimiento de la invención es expresada de forma recombinante.

En un modo de realización particular, las enzimas que catalizan las etapas de fosforilación a), de escisión b) y de reducción c) y/o de oxidación c') son enzimas recombinantes.

- 5 En ciertos modos de realización, al menos una de dichas enzimas recombinantes es una enzima codificada por un gen heterólogo (es decir, no expresado de forma natural en el organismo hospedante de referencia), en particular al menos una de las enzimas que catalizan las etapas a) y b) es codificada por un gen heterólogo.

Preferentemente, el microorganismo expresa al menos la *KhkC*.

Incluso preferentemente, el microorganismo expresa de forma recombinante:

- 10 • la *KhkC* (particularmente *KhkC*, *H. sapiens* codificada por el gen *khkC* de la secuencia SEQ ID N° 1) que cataliza la etapa a) de la vía de asimilación de pentosas, y
- la aldolasa B, codificada por el gen *aldo-B* (particularmente por el gen *aldoB* de la secuencia SEQ ID N° 2) o la fructosa-1,6, bis-fosfato aldolasa, codificada por el gen *fbpA* (particularmente el gen *fbpA* de la secuencia SEQ ID N° 9).

- 15 En un modo de realización particular, la glicolaldehído reductasa y/o la glicolaldehído deshidrogenasa catalizan, respectivamente, las etapas c) y c') de la vía sintética de asimilación de pentosas de la invención son enzimas endógenas, que se expresan de forma natural por el microorganismo.

- 20 En ciertos modos de realización, las enzimas endógenas de la vía sintética de asimilación de la invención pueden ser sobreexpresadas, particularmente las enzimas que codifican las etapas de reducción c) o de oxidación c'). En particular, la glicolaldehído deshidrogenasa puede ser sobreexpresada para estimular la etapa c') de oxidación. Por ejemplo, en *E. coli*, se puede sobreexpresar el gen *aldA* que codifica una glicolaldehído deshidrogenasa.

- 25 En ciertos modos de realización de la invención, las enzimas que convierten la (D)-xilosa o la (L)-arabinosa en (D)-xilulosa o (L)-ribulosa, respectivamente, a saber, las isomerasas o epimerasas como se describen con anterioridad, o incluso las proteínas que importan la (D)-xilosa o la (L)-arabinosa en la célula, son sobreexpresadas en el microorganismo.

En un modo de realización de la invención, los ácidos nucleicos que codifican las enzimas que catalizan las etapas a) y b) son clonados en operón en un vector de expresión bajo el control de un mismo promotor. En ciertos modos de realización, los ácidos nucleicos que codifican las enzimas que catalizan las etapas a), b) y c) y/o c') son clonados con operón.

- 30 La expresión de las proteínas recombinantes es controlada por un promotor inducible o, preferentemente, constitutivo.

Optimización de la vía sintética de asimilación de pentosas

En ciertos modos de realización, la actividad de una o varias enzimas endógenas de la célula hospedante puede ser modificada también de forma que se optimice la producción de etilenglicol y/o ácido glicólico.

- 35 Se describen a continuación ciertas modificaciones que pueden ser aportadas a un microorganismo de la invención.

1. A) Preferentemente, el microorganismo utilizado es genéticamente modificado de forma que se atenúe o suprima la actividad de las enzimas endógenas implicadas en las vías naturales de asimilación de pentosas de fosfatos y, particularmente, la o las enzima(s) que cataliza(n) la fosforilación en posición 5 del ciclo de pentosas y, más particularmente, la (L)-ribulosa-5-quinasa y/o la (D)-xilulosa-5-quinasa.

- 40 Como ejemplo, los genes *araB* y/o *xyIB*, que codifican la ribulosa-5-quinasa y la xilulosa-5-quinasa, respectivamente,

que se encuentran normalmente en *E. coli*, se pueden ser atenuados o, preferentemente, inactivados.

Esta modificación permite orientar el flujo de carbono preferentemente hacia vía sintética de asimilación de pentosas de la invención y optimizar la producción de etilenglicol y/o ácido glicólico mediante dicha vía sintética.

5 En un modo de realización de la invención, se utiliza un microorganismo en el que ha sido suprimido el gen *xyIB* y, particularmente, el gen *xyIB* que codifica la secuencia de xilosa quinasa de secuencia SEQ ID NO: 53.

2. B) La actividad de las enzimas de tipo glicolaldehído reductasa y/o glicolaldehído deshidrogenasa puede ser modificada también con el fin de orientar la vía sintética de asimilación de la invención hacia la producción de ácido glicólico o etilenglicol.

10 Como ejemplo, las enzimas codificadas por el gen *aldA* que codifica una glicolaldehído deshidrogenasa, así como los genes *gldA*, *fucO* y/o *yqhD* que codifican una glicolaldehído reductasa, pueden ser particularmente sobreexpresados para favorecer la producción de etilenglicol o atenuados o inactivados para favorecer la producción de ácido glicólico.

15 Ventajosamente, la producción de etilenglicol es optimizada utilizando un microorganismo en el que se aportan además al menos una (y preferentemente las dos) de las modificaciones siguientes en cuanto a la expresión de enzimas endógenas:

- sobreexpresión del gen que codifica al menos una glicolaldehído reductasa, preferentemente la glicolaldehído reductasa principal, expresada por el microorganismo;

- inactivación o delección del gen que codifica una glicolaldehído deshidrogenasa catalizando la etapa c', por ejemplo, el gen *aldA*.

20 Ventajosamente, la producción de ácido glicólico puede ser optimizada utilizando un microorganismo en el que se aportan al menos una de las modificaciones siguientes (y preferentemente al menos las dos primeras) en cuanto a la expresión de enzimas endógenas:

- sobreexpresión del gen que codifica la glicolaldehído deshidrogenasa, por ejemplo, la glicolaldehído deshidrogenasa codificada por el gen *aldA*;

25 • reducción de la degradación de ácido glicólico, particularmente mediante la atenuación o inactivación de la glicolato oxidasa, por ejemplo, mediante inactivación de al menos uno de los genes *glcDEFG* que codifica al menos una de las subunidades de la glicolato oxidasa;

- ocasionalmente, una inactivación del o de los genes que codifican una glicolaldehído reductasa.

30 Preferentemente, un microorganismo según la invención comprende, además de las modificaciones propias para la expresión de la vía sintética de asimilación de pentosas de la invención y, particularmente, apropiadas para catalizar las etapas de fosforilación a) y de escisión b) de esta vía, al menos una de las modificaciones complementarias descritas en los apartados A) o B) anteriores.

35 Incluso preferentemente, el microorganismo comprende al menos una de las modificaciones descritas en el apartado A). Según los modos de realización, esta modificación puede ser combinada con las modificaciones del apartado B) para optimizar la producción de etilenglicol o de ácido glicólico.

Las modificaciones anteriormente mencionadas pueden ser combinadas.

En particular:

40 • para optimizar la producción de etilenglicol, la inactivación de las enzimas que catalizan la fosforilación en posición 5 de pentosas, particularmente la (L)-ribulosa-5-quinasa y/o (D)-xilulosa-5-quinasa (como los genes *xyIB* o *araB*) puede ser combinada con la inactivación de la vía de síntesis del ácido glicólico (particularmente, mediante la

inactivación del o de los genes que codifican una glicolaldehído deshidrogenasa, por ejemplo el gen *aldA*). Estas inactivaciones pueden estar combinadas a una sobreexpresión del gen que codifica la glicolaldehído reductasa que cataliza etapa c) del procedimiento de la invención.

5 • para optimizar la producción de ácido glicólico, la inactivación de las enzimas que catalizan la fosforilación en posición 5 de pentosas, particularmente la (L)-ribulosa-5-quinasa y/o (D)-xilulosa-5-quinasa (particularmente codificadas por los genes *xyiB* o *araB*) pueden ser combinada con la inactivación de la glicolato oxidasa mediante inactivación de al menos uno de los genes *glcDEFG* que codifican sus subunidades y/o a la sobreexpresión de la glicolaldehído deshidrogenasa codificada por el gen *aldA*. Ocasionalmente, el gen que codifica la glicolaldehído reductasa que cataliza la etapa c) del procedimiento de la invención es también inactivado.

10 Por ejemplo, un microorganismo modificado para optimizar la producción de etilenglicol expresa al menos las actividades de fosforilación, de escisión y de reducción correspondientes a las etapas a), b) y c) anteriormente descritas, preferentemente las enzimas KhkC, aldolasa B (particularmente codificada por el gen *aldoB*) así como una glicolaldehído reductasa (como la aldehído reductasa YqhD o la glicerol deshidrogenasa GldA o incluso la enzima codificada por el gen *fucO*) y que comprende las modificaciones siguientes:

15 • la delección del gen que codifica la aldehído deshidrogenasa, particularmente del gen *aldA*;

• la delección de los genes que codifican la o las enzima(s) que cataliza(n) la fosforilación en posición 5 de la (D)-xilulosa y/o la (L)-ribulosa y, más particularmente, la (D)-xilulosa-5-quinasa y/o la (L)-ribulosa-5-quinasa, particularmente los genes *araB* y/o *xyiB*.

20 C) Es también posible utilizar un microorganismo genéticamente modificado con el fin de favorecer, además de la vía sintética de la invención, la producción de ácido glicólico por vías naturales.

En efecto, los inventores han descubierto además que la producción de ácido glicólico puede ser también incluso aumentada combinando el funcionamiento de la vía sintética de asimilación según la invención con modificaciones genéticas que conduzcan en paralelo a la producción de ácido glicólico por la vía denominada del glioxilato.

25 Un microorganismo según la invención, por tanto, puede presentar modificaciones que favorezcan la producción de ácido glicólico a partir del glioxilato, como se describe en la solicitud WO 2010/108909.

30 Por tanto, la dihidroxiacetona fosfato (DHAP o glicerona fosfato), particularmente obtenida al final de la etapa de escisión b), podría integrar las vías naturales de la glicólisis, del ciclo de los ácidos tricarbónicos (CAT) y de la vía del glioxilato, un shunt del CAT (véase para una revisión Neidhardt, F. C. (Ed. in Chief), R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter, and H. E. Umberger (eds). 1996. *Escherichia coli and Salmonella: Cellular and Molecular Biology*. American Society for Microbiology).

La optimización de la vía del glioxilato puede ser obtenida por tanto mediante al menos una y, preferentemente, mediante una combinación, de las modificaciones siguientes:

1. i) sobreexpresión de un gen que codifica una glioxilato reductasa;
- 35 2. ii) sobreexpresión de un gen que codifica una isocitrato liasa, ocasionalmente acompañada de la delección del gen que codifica su represor transcripcional;
3. iii) delección de los genes que codifican malato sintasas;
4. iv) delección de los genes que codifican glioxilato carboligasas;
5. v) delección de los genes que codifican genes que codifican glicolato oxidasas o glicolato deshidrogenasas;
- 40 6. vi) delección de los genes que codifican 2-ceto-4-hidroxi-glutarato aldolasas, particularmente Entner-Doudouroff aldolasa y/o para fosfogluconato deshidratasas;

7. vii) delección de un gen represor de respuesta aeróbica (es decir, un gen que codifica un represor de genes implicados en el metabolismo respiratorio), particularmente el gen *arcA*;

8. viii) atenuación y, particularmente, delección de la expresión de isocitrato deshidrogenasa;

9. ix) ocasionalmente, delección de los genes que codifican sistemas de internalización del ácido glicólico; y

5 10. x) ocasionalmente atenuación de las vías metabólicas que conducen a la preparación de productos secundarios como acetato, lactato o etanol.

En el caso de que el microorganismo recombinante, para la producción del ácido glicólico, sea *Escherichia coli*, las modificaciones anteriormente descritas corresponden a

1. i) la sobreexpresión del gen *ghrA* y/o el gen *ycdW*;

10 2. ii) la sobreexpresión del gen *aceA* acompañada o no por la delección de su represor transcripcional, *icIR*;

3. iii) la delección de los genes *aceB* y *glcB*;

4. iv) la delección del gen *gcl*;

5. v) la delección de al menos un gen escogido entre *glcD*, *glcE*, *glcF* o *glcG*;

6. vi) la delección de los genes *edd-eda*;

15 7. vii) la delección del gen *arcA*;

8. viii) la atenuación y, preferentemente, la delección de la expresión del gen *icd*;

9. ix) la delección de los genes *glcA*, *lldP* y/o *yjcG*;

10. x) la delección de los genes *ackA-pta*, *poxB*, *ldha* y/o *adhE*.

20 Ventajosamente, las modificaciones siguientes son aportadas para optimizar la producción de ácido glicólico en un microorganismo según la presente invención, por ejemplo, en *E. coli*:

1. i) la sobreexpresión del gen *ghrA*

2. ii) la sobreexpresión de los genes *aceA* ocasionalmente acompañada por la delección del gen *icIR*;

3. iii) la delección de los genes *aceB* y *glcB*

4. iv) la delección del gen *gcl*;

25 V) la delección de al menos un gen escogido entre *glcD*, *glcE*, *glcF* o *glcG*;

6. vi) la delección de los genes *edd-eda*;

7. vii) la delección del gen *arcA*

8. viii) la delección del gen *icd*.

30 En un modo de realización, el microorganismo portador de las modificaciones que anteceden no expresa una enzima que cataliza la fosforilación en posición 5 de (L)-ribulosa y/o (D)-xilulosa y, más particularmente, (L)-ribulosa-5-quinasa y/o (D)-xilulosa-5-quinasa. Por tanto, en ciertos modos de realización se utiliza un microorganismo en el

que los genes *xyIB* y/o *araB* son suprimido.

En una variante, el microorganismo (ventajosamente *E. coli*) expresa una o varias enzimas que catalizan la fosforilación en posición 5 de (L)-ribulosa y/o (D)-xilulosa y, más particularmente, (L)-ribulosa-5-quinasa y/o (D)-xilulosa-5-quinasa. Normalmente, el microorganismo expresa el gen *xyIB* y/o el gen *araB*. Este microorganismo ofrece un excelente rendimiento de ácido glicólico cuando es cultivado en un medio que contiene glucosa, particularmente que contiene al menos glucosa y xilosa o xilulosa. Preferentemente, el medio contiene mayoritariamente glucosa.

D) Finalmente es posible, para aumentar la producción de etilenglicol o de ácido glicólico, utilizar un microorganismo en el que al menos un gen que codifica un transportador de azúcares (por ejemplo, el gen *galP* que codifica una permeasa de azúcares y/o el gen *gal-2a* de *S cerevisiae*) es sobreexpresado. Ventajosamente, dicho gen es expresado de forma constitutiva.

En un modo de realización particularmente preferido de la invención, se utiliza para la producción de ácido glicólico un microorganismo que combina las modificaciones descritas en los párrafos A) a D) anteriores. En particular, se utiliza un microorganismo que combina las modificaciones recogidas en los párrafos B), C) y D) anteriores para la producción de ácido glicólico.

Ventajosamente, este microorganismo expresa la KhkC y la aldolasa B y comprende las modificaciones siguientes:

- la sobreexpresión del gen que codifica glicolaldehído deshidrogenasa, por ejemplo, glicolaldehído deshidrogenasa codificada por el gen *aldA*;
- la sobreexpresión del gen *ghrA*;
- la sobreexpresión del gen *aceA* ocasionalmente acompañada de la delección del gen *icIR*;
- la delección de los genes *aceB* y *glcB*;
- la delección del gen *glc*;
- la delección de al menos un gen escogido entre *glcD*, *glcE*, *glcF* o *glcG*;
- la delección de los genes *edd-eda*;
- la delección del gen *arcA*;
- la delección del gen *icd*;
- la sobreexpresión del gen *galP*.

En función del sustrato de cultivo escogido, este microorganismo puede portar también o no una delección del gen *xyIB* y/o del gen *araB*. Preferentemente, la expresión del gen *xyIB* y/o del gen *araB* es conservada cuando el microorganismo es cultivado en un sustrato que comprende glucosa.

Igualmente en función del sustrato de cultivo, se utiliza un microorganismo que expresa, de forma recombinante o no, enzimas que convierten la (D)-xilosa o la (L)-arabinosa en (D)-xilulosa o (L)-ribulosa, respectivamente, a saber, (D)-xilulosa isomerasas o (L)-arabinosa isomerasas o (D)-xilosa reductasas/(D)-xilitol deshidrogenasas, o (L)-arabinosa reductasas/(L)-arabitol deshidrogenasas, como se describen con anterioridad.

De forma general, sin ser limitativo, el rendimiento teórico del procedimiento de la invención para etilenglicol es de aproximadamente 1 mol de etilenglicol por mol de xilosa o de arabinosa.

El rendimiento teórico del procedimiento de la invención para el ácido glicólico es de aproximadamente 1 mol de ácido glicólico por mol de xilosa o arabinosa, sin activación del ciclo de glioxilato. Durante un funcionamiento

paralelo de la vía sintética y del ciclo glioxilato, el rendimiento teórico es de aproximadamente 2 moles de ácido glicólico por mol de xilosa o arabinosa.

Cultivo del microorganismo

5 Las condiciones de cultivo del microorganismo según la invención pueden ser adaptadas según técnicas habituales, conocidas por un experto en la técnica.

Normalmente, las bacterias utilizadas como células hospedantes en la presente invención pueden ser cultivadas en medios de todos los tipos y de cualquier composición.

10 Los medios de cultivo son normalmente medios carbonados que comprenden o están complementados con diversos compuestos, que incluyen, particularmente, diferentes fuentes de carbono y, particularmente, pentosas como (D)-glucosa, (D)-xilosa, (L)-arabinosa y/o hidrolizados de biomasa lignocelulósica, particularmente hemicelulosa, almidón y sus derivados.

En ciertos modos de realización, el medio de cultivo comprende menos de 5%, particularmente menos de 4%, menos de 3%, menos de 2% o incluso menos de 1% de ramnulosa.

15 Mediante "hidrolizado de biomasa" se entiende, en particular, hidrolizados lignocelulósicos, en particular hidrolizados que contienen al menos 20% de xilosa y/o al menos 5%, particularmente al menos 10% de arabinosa, en peso con respecto al peso total del hidrolizado. En un modo de realización preferido, se utilizan así hidrolizados lignocelulósicos de maderas duras, de mazorcas de maíz o de papel.

Otros parámetros relativos a las condiciones de cultivo pueden ser optimizados mediante experimentación rutinaria, como el pH o la temperatura.

20 En ciertos modos de realización, la temperatura del cultivo varía de 25 a 43 °C y depende esencialmente del tipo de célula hospedante y del medio de cultivo. Como ejemplo, cuando la célula hospedante es *E. coli*, la temperatura de cultivo óptima varía generalmente entre 30 a 38 °C.

La duración del cultivo depende también de los parámetros de cultivo anteriormente mencionados. Normalmente, los cultivos pueden ser cultivados entre 6 y 300 horas.

25 Preferentemente, el o los metabolitos de interés obtenidos al final de la operación de cultivo del microorganismo según la invención son recuperados en el medio de cultivo.

La presente solicitud se refiere también a un microorganismo recombinante como se describe en la presente solicitud.

30 En particular, la presente invención se refiere a un microorganismo que expresa una vía sintética de asimilación de pentosas según la invención.

Según diferentes modos de realización, las vías naturales de asimilación de pentosas son conservadas o son inactivadas (por ejemplo, mediante la delección de los genes que codifican xilulosa-5-quinasa y/o ribulosa-5-quinasa).

Un microorganismo según la invención expresa al menos:

35 ● un ácido nucleico que codifica una enzima adecuada para fosforilar, en posición 1, una pentosa escogida entre (D)-xilulosa y/o (L)-ribulosa, y

● un ácido nucleico que codifica una enzima de tipo aldolasa, adecuada para escindir el (D)-xilulosa-1-fosfato y/o el (L)-ribulosa-1-fosfato en glicolaldehído y DHAP, como se describió con anterioridad.

Preferentemente, al menos una de estas enzimas es expresada de forma recombinante.

En ciertos modos de realización, al menos una de estas enzimas es codificada por un ácido nucleico exógeno y, preferentemente, el microorganismo expresa al menos:

- un ácido nucleico que codifica la isoforma C de cetohecoquinasa (que se encuentra normalmente en *H. sapiens*) y
- al menos un ácido nucleico que codifica la aldolasa B.

5 En ciertos modos de realización, el microorganismo es modificado además como se describe en los apartados A) a E) anteriores.

Por ejemplo, un microorganismo adecuado para la producción de ácido glicólico puede comprender ventajosamente las modificaciones siguientes:

- 10
- la sobreexpresión del gen que codifica la glicolaldehído deshidrogenasa, por ejemplo, la glicolaldehído deshidrogenasa codificada por el gen *aldA*;
 - la sobreexpresión del gen *ghrA*;
 - la sobreexpresión del gen *aceA*;
 - ocasionalmente, la eliminación del gen *iclR*;
 - la deleción del gen *glc*;
- 15
- la deleción de al menos un gen escogido entre los genes *glcD*, *glcE*, *glcF* o *glcG*;
 - la deleción de los genes *aceB* y *glcB*.
 - la deleción de los genes *edd-eda*;
 - la deleción del gen *ArcA*;
 - la deleción del gen *icd*;
- 20
- ocasionalmente la deleción de los genes *xyIB* y/o *araB*.

Ejemplos

Composición de los medios

Medio Luria-Bertani (LB)

25 Para un litro de medio: 10 g de triptona, 5 g de extracto de levaduras, 5 g de cloruro de sodio en un litro de agua purificada. El medio es tratado en autoclave antes de su utilización. Para ser utilizado en medio sólido, se añade 2% de agar al medio antes del tratamiento en autoclave.

Medio mínimo M9

30 Para un litro de medio: 18 g de $\text{Na}_2\text{HPO}_4 \cdot 12\text{H}_2\text{O}$; 3 g de KH_2PO_4 ; 0,5 g de NaCl ; 2 g de NH_4Cl ; 0,5 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$; 0,015 g de $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$; 1 ml de una solución de elementos residuales (que contiene por litro 0,04 g de $\text{NaEDTA} \cdot 2\text{H}_2\text{O}$, 0,18 g de $\text{CoCl}_2 \cdot 6\text{H}_2\text{O}$; $\text{ZnCl}_2\text{SO}_4 \cdot 7\text{H}_2\text{O}$; 0,04 g de $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 0,01 g de H_3BO_3 , 0,12 g de $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0,12 g de $\text{CuCl}_2 \cdot \text{H}_2\text{O}$); cantidades de (D)-glucosa, (D)-xilosa y (L)-arabinosa indicadas en el texto. El medio se ajusta a pH 7 y se filtra.

YPD

El medio YPD es utilizado con medio rico para el crecimiento de *S. cerevisiae*. Para un litro, 10 g de extractos de levaduras, 20 g de bacto-peptona. El medio es tratado en autoclave antes de su utilización en 20 g de glucosa filtrada sin adiciones. Para ser utilizado en medio sólido, se agrega 2% de agar al medio antes del tratamiento en autoclave.

5 Medio mínimo de SCD para *Saccharomyces cerevisiae*

Para un litro, 1,7 g de base de nitrógeno de levaduras sin aminoácidos, 5 g de sulfato de amonio sin aminoácidos, se hacen gotear 0,940 g de aminoácidos esenciales excepto los utilizados para poner de manifiesto una auxotrofia, 900 ml de agua y seguidamente la totalidad se somete a tratamiento en autoclave. En el caso de una utilización en medio sólido, se añaden 20 g de bacto-agar. Se añaden 100 ml de una solución de azúcares al 20%.

10 Ensayo de crecimiento en medio M9 + xilosa

Todos los cultivos se realizaron en matraces Erlenmeyer de 250 ml que contenían 50 ml de medio de cultivo y agitando los cultivos a 200 rpm.

Las células que iban a ser ensayadas se pusieron en cultivo durante la noche a 37 °C en medio LB. Este precultivo se utiliza seguidamente para inocular a medio DO_{600nm} ~ 0,2 del medio M9 + 10 g/l de glucosa. En fase exponencial de crecimiento (OD entre 0,6 y 1), se añade el IPTG 1 mM y los cultivos se incuban así durante 16 a 18 horas. Después de este período de incubación, las células se lavan dos veces con agua esterilizada y se inoculan a DO_{600nm} ~ 0,2 en medio M9 + IPTG 1 mM + glucosa y/o xilosa y/o arabinosa en las cantidades indicadas en el texto. La DO_{600nm} es sometida a seguimiento y se tomaron partes alícuotas, se centrifugaron y se inyectaron en HPLC para un análisis de los metabolitos.

20 Métodos de construcción de cepas

Transformación bacteriana

Las transformaciones bacterianas se realizan sobre células quimiocompetentes comerciales o preparadas en laboratorio. Las células hechas quimicaocompetentes se preparan según el protocolo de cloruro de calcio (Dagert and Ehrlich, 1979). La transformación se efectúa seguidamente dejando transformar durante 20 minutos ADN plasmídico en contacto bacterias competentes sobre hielo y seguidamente se realiza un choque térmico a 42 °C durante 45 segundos. Las células se dejan 5 minutos sobre hielo y seguidamente se añade 1 ml de medio LB antes de incubar durante 1 hora a 37 °C. Seguidamente, las células se extienden sobre una cámara LB sólida complementada con el marcador de selección correspondiente.

De forma general, además de los plásmidos desarrollados en el contexto de la presente invención, se utilizaron los siguientes plásmidos: pACT3 (Dykxhoorn *et al.*, 1996), pEXT20 (Dykxhoorn *et al.*, 1997), pGEM-T (Promega), pET28a (Novagen), pCP20 (Cherepanov & Wackernagel, 1995), peX-A-aldoB (Eurofins) y pET11-KHK-C (Asipu *et al.*, 2003).

Delección de los genes mediante transducción de una cassette de kanamicina de una cepa KEIO

Para transferir una delección de gen portada por una cepa KEIO de *E. coli* hacia una cepa receptora dada de MG1655 de *E. coli*, se realiza una transducción.

A partir de una cepa Keio dispuesta en cultivo en LB + kanamicina 50 µM a 37°C durante una noche, se genera un lisado de fagos. Sobre un precultivo de 10 ml de LB inoculado en la mañana a partir de 200 µl del cultivo durante la noche en presencia de 2 g/l de glucosa y CaCl₂ 5 mM se añaden 200 µl de fago P1. El cultivo se deja 2 h que es el tiempo de lisis celular debida al fago. La reacción se detiene con 200 µl de cloroformo. El conjunto se centrifuga 10 min a 4500 x g y se recuperan 9 ml del condensado sobrenadante de los fagos y se almacena con 200 µl de cloroformo a 4 °C. La cepa receptora se pone en precultivo durante la noche. De este cultivo se recupera 1,5 ml y se centrifuga. El sedimento se recoge en 600 µl de MgSO₄ 10 mM + CaCl₂ 5 mM. La transducción se realiza disponiendo 100 µl de células en presencia de 100 µl del lisado de fagos. El conjunto se incuba 30 minutos a 30 °C sin agitación. Seguidamente, se añaden 100 µl de citrato de sodio 1 M así como 1 ml de LB. La expresión fenotípica de las cepas que han integrado la cassette de kanamicina se hace dejando reposar las células durante 1 hora a 37

°C bajo agitación. Las células seguidamente se extienden sobre la cámara de medio LB que contiene el marcador de selección y se dejan reposar durante una noche. A la mañana siguiente, las colonias formadas son ensayadas mediante PCR en cuanto a la presencia de la cassette de selección y en cuanto a la ausencia del gen suprimido.

Protocolo de escisión de la cassette de selección flanqueada por secuencia FRT

- 5 La cassette es escindida del cromosoma utilizando la recombinación FLP portada por el plásmido pCP20 (Cherepanov y Wackernagel, 1995) que deja una región de cicatriz que contiene el sitio de FRT. El pCP20 es un plásmido que porta resistencia a la ampicilina y al cloranfenicol que presenta una replicación termosensible y una expresión de la recombinación FLP termoinducida. Los mutantes marcadores resistentes que contienen así la cassette son transformados con pCP20 y los transformantes resistentes a ampicilina que portan la resistencia del plásmido se seleccionan a 30 °C. Seguidamente se ponen en cultivo a 37 °C sobre LB sólido y seguidamente se ensayan en cuanto a la pérdida de resistencia a la ampicilina. La escisión de la cassette de selección se verifica seguidamente mediante PCR con los cebadores que han servido para amplificar con Taq polimerasa (NEB). Las delecciones múltiples se obtienen repitiendo la operación.

Clonación de genes sobre plásmido en *S. cerevisiae*

- 15 La clonación de un gen en *S. cerevisiae* utiliza las capacidades de recombinación genética de la levadura. El gen que va a ser clonado es asociado a una secuencia promotora y una secuencia terminadora que proporcionan tres fragmentos que van a ser ligados en un plásmido previamente linealizado. Para esto, se diseñan regiones homólogas de 40 nucleótidos sobre los cebadores. Estos 40 nucleótidos de homología permiten a los sistemas de recombinación de la levadura ligar el conjunto de los fragmentos después de la transformación. Cada fragmento es amplificado mediante PCR utilizando la Phusion™ polimerasa. El conjunto de los fragmentos y el vector receptor linealizado es transformado en una cepa competente de *S. cerevisiae* según el método descrito por Gietz y Woods (2002). Después del crecimiento de los transformantes, los plásmidos son extraídos según el método descrito por Zeugin y Hartley (1985). Los plásmidos seguidamente se utilizan para transformar una cepa DH10B de *E. coli*. Los plásmidos son extraídos de *E. coli*, verificados mediante secuenciado y utilizados para transformar la cepa receptora de *S. cerevisiae*.

Extracción de plásmidos en *S. cerevisiae*

- Después de la transformación, las colonias obtenidas se vuelven a poner en suspensión en agua y seguidamente se centrifugan con el fin de una extracción de plásmido. El sedimento de células se vuelve a poner en suspensión en 400 µl de un tampón a 4 °C que contiene glucosa 50 mM, EDTA 10 mM, Tris-HCl 25 mM, pH 8 y se complementa con 0,1 mg/ml de RnaseA. Se utilizan 400 µl de una solución que contiene NaOH 0,2 M y 1% de SDS para hacer lisar las células. Seguidamente se añaden bolas de vidrio hasta una altura de un tercio del volumen total y las células se centrifugan a 4 °C durante 10 minutos. Esta etapa está seguida de una centrifugación de 60 segundos a 13.000 rpm. Se retiran 700 µl de la materia sobrenadante y se disponen en un nuevo tubo de 2 ml. Se añaden 325 µl de una solución a 4 °C de KAC 3 M a pH 5,5. La mezcla se deja incubar durante 10 minutos sobre hielo antes de ser centrifugada durante 10 minutos a 13.000 rpm a 4 °C. Se retiran 700 µl de materia sobrenadante y se depositan en un nuevo tubo. Se añaden 700 µl de isopropanol y el conjunto se agita intensamente antes de dejarlo incubar 10 minutos a temperatura ambiente. Seguidamente se realiza una centrifugación a temperatura ambiente durante 30 minutos a 13.000 rpm. La materia sobrenadante seguidamente se elimina y el sedimento se vuelve a poner en suspensión en 500 µl de etanol al 70% a -20 °C y seguidamente se centrifuga 5 minutos. La materia sobrenadante se elimina. Esta etapa se repite una vez más y seguidamente el sedimento se seca hasta la desaparición del etanol. Este sedimento seguidamente se recoge en 30 µl de H₂O.

De forma general, los iniciadores utilizados para la expresión de la vía en *E. coli* se recogen en la tabla 3 y los iniciadores utilizados para la expresión de la vía en *S. cerevisiae* se recogen en la tabla 4.

Expresión y purificación de proteínas a partir de un plásmido pET28a

- 45 La cepa de *E. coli* BL21 (DE3) que contiene el pET28a que porta el gen de interés se pone en cultivo durante una noche a 37°C en 100 ml de medio LB en un Erlenmeyer de 500 ml bajo agitación a 200 rpm.

En la mañana siguiente, 10 a 50 ml de este precultivo de *E. coli* BL21 (DE3) que contiene el pET28a que porta el

gen de interés se ponen en cultivo en medio LB complementado con 50 µg/ml de kanamicina a 37°C.

La expresión proteica es provocada mediante la adición de IPTG 1 mM a los cultivos que alcanzan una $DO_{600nm} \sim 0,7$. Después de 3 horas a 37°C, las células se centrifugan y los sedimentos se congelan a -80°C. Para la purificación de las proteínas así expresadas, las células se recogen en 1 ml de tampón de lisis (50 mmol/l de Hepes [pH 7,5], 3 mol/l de NaCl, 0,25 mmol/l) y se guardan durante una hora en hielo.

Las células se someten a ultrasonidos y los residuos se separan mediante centrifugación durante 10 minutos a 13.000 RPM a 4 °C. Seguidamente se prepara una resina Talon © con 0,3 ml de resina en 3 ml de tampón de lisis. Toda la suspensión se deposita seguidamente sobre la resina y se incuba 20 minutos a temperatura ambiente antes de ser centrifugada 5 minutos a 2500 rpm a 4 °C. El sedimento se lava con 10 veces el volumen de la resina con un tampón de lisis y se incuba durante 10 minutos a temperatura ambiente. La operación se repite con el tampón de lisis que contiene imidazol 15 mM. Seguidamente el depósito se lava con 500 µl de imidazol 200 mM. El conjunto se vuelve a centrifugar durante 5 minutos a 2500 rpm a 4°C. La materia sobrenadante se recupera proporcionando la elución 1. La operación se repite proporcionando una elución 2. Las proteínas purificadas se encuentran en las diferentes eluciones que serán ensayadas.

Métodos analíticos

Para dosificar los productos de la vía sintética a partir de las materias sobrenadantes de los cultivos, se utilizó una HPLC (Ultimate 3000, Dionex) equipada con un dispositivo automático de toma de muestras y un horno (Shimadzu CTO-20A) y acoplado al detector RID-10A (Shimadzu) y UV SPD-20A (Shimadzu). Los compuestos se separaron en una columna Aminex HPX-87H (300 mm x 7,8 mm) equipada con una precolumna (Aminex). Los análisis se realizaron a una temperatura de 35 °C con un caudal de 0,5 ml/min de H₂SO₄ a 1,25 mM. Se inyectaron muestras de 20 µl en el aparato.

Ejemplo 1: Demostración de las actividades de (D)-xilulosa-1-quinasa y (L)-ribulosa-1-quinasa

Clonación de xiluloquinasas candidatas: *khkC*, *rhaB* y *fucK* en pET28a

La clonación de los genes *khkC* (SEQ ID N° 1), *rhaB* (SEQ ID N° 6) y *fucK* (SEQ ID N° 5) en pET28a se realiza como se indica con posterioridad. El gen *khk-C* es digerido a partir de pET11a-hk-C (Asipu *et al.*, 2003) mediante NdeI/EcoRI. Se inserta mediante ligadura a través de la ligasa T4 (Biolabs) a continuación de tag-histidina del pET28a previamente digerido con estas mismas enzimas. La clonación de *rhaB* y *fucK* se realizó mediante amplificación por PCR *rhaB* y *fucK* a partir de ADN genómico de *E. coli* con los cebadores siguientes, respectivamente, P1/P2; P3/P4 citados en la tabla 3. Los fragmentos seguidamente se clonaron en el vector pGEM-T (Invitrogen). Seguidamente se digieren mediante NcoI y BamHI y seguidamente se ligan en el plásmido pET28a a nivel de MCS previamente digerido por estas mismas enzimas. El producto de ligadura se transforma en una cepa BL21 (DE3) de *E. coli*. Los vectores pET28-*khk-C*, pET28-*rhaB* y pET28-*fucK* así obtenidos son verificados en cuanto al secuenciado que contiene los genes con las secuencias correctas.

Determinación de los parámetros cinéticos de las (D)-xiluloquinasas y (L)-ribuloquinasas candidatas

Las proteínas se expresan y purifican como se describió anteriormente y los parámetros cinéticos de las enzimas se determinan mediante la reacción de acoplamiento de piruvato quinasa/lactato deshidrogenasa que se basa en el principio siguiente (el ejemplo proporcionado corresponde a la dosificación de una actividad de (D)-xilulosa-1P aldolasa):

Xilulosa quinasa:	(D)-xilulosa + ATP → (D)-xilulosa-1P + ADP
Piruvato quinasa:	fosfoenolpiruvato + ADP → piruvato + ATP
Lactato deshidrogenasa:	piruvato + NADH → lactato + NAD

5 La reacción se realiza en el medio siguiente: NADH 0,4 mM (Sigma), PEP 2 mM (Sigma), ATP 4 mM (Sigma) en un tampón Hepes (Hepes 90 mM, KCl 77 mM, MgCl₂ 12 mM, ajustado a pH 7 con una solución de KOH). Se añade un volumen de 1,25 µl de una mezcla de enzimas piruvato quinasa/lactato deshidrogenasa (Sigma) a una mezcla de reacción total de 250 µl. La reacción arranca con la adición de 100 µl de (D)-xilulosa (Carbosynth) o (L)-ribulosa (Sigma) 10 mM. El consumo de NADH es seguido mediante espectrofluorometría a 340 nm.

Tabla 1A. Parámetros cinéticos de las quinasas sobre su sustrato natural y sobre (D)-xilulosa. Los sustratos naturales de Khk-C, RhaB y FucK son, respectivamente, fructosa, ramnulosa y fuculosa.

Enzima Candidata	Sustrato Natural		(D) -xilulosa	
	Vmax [U/mg]	Km [mM]	Vmax [U/mg]	Km [mM]
Cetohexokinasa Khk-C, <i>H. sapiens</i>	6	0,72	3,1	0,6
(L)-ramnulosa quinasa RhaB, <i>E. coli</i>	n/a	n/a	18,7	ns
L-fuculoquinasa FucK, <i>E. coli</i>	20	0,06	0,1	ns

10 Estos resultados muestran que pueden ser identificadas las quinasas tienen la capacidad de fosforilar (D)-xilulosa en la posición 1.

Se utilizó la misma propuesta para caracterizar estos parámetros sobre (L)-ribulosa:

Tabla 1B. Parámetros cinéticos de las quinasas sobre su sustrato natural y sobre (L)-ribulosa.

Enzima Candidata	sustrato natural		(L) -ribulosa	
	Vmax [U/mg]	Km [mM]	Vmax [U/mg]	Km [mM]
Cetohexokinasa, Khk-C, <i>H. sapiens</i>	6,59 ± 1,4	0,31 ± 0,1	3,5	0,55

15 La Khk-C es funcionalizada a la vez sobre (D)-xilulosa como en (L)-ribulosa y presenta características apropiadas para su utilización en la vía sintética.

Ejemplo 2: Demostración de las actividades de (D)-xilulosa-1P-aldolasa y (L)-ribulosa-1P-aldolasa.

Clonación de genes que codifican las aldolasas candidatas AldoB, FbaB y AgaY

20 La clonación de las aldolasas candidatas se hace amplificando mediante PCR *aldoB*, *fbaB* y *agaY* (utilizando, respectivamente, los pares de cebadores P5/P6, P7/P8 y P11/P12 recogidos en la tabla 3). Esta amplificación se realizó a partir del plásmido peX-A-aldob que porta *aldoB* con optimización de codones (Eurofins) o ADN genómico de *E. coli* para *fbaB* y *agaY*, respectivamente. Los fragmentos se clonan seguidamente en pGEM (Invitrogen). Seguidamente son digeridos mediante BamHI y HindIII (para *aldoB*) o NdeI y BamHI (para *fbaB* y *agaY*) y seguidamente son ligados en plásmido pET28a a nivel de MCS previamente digerido mediante BamHI y HindIII o NdeI y BamHI para clonar respectivamente *aldoB* (SEQ ID N° 2) o *fbaB* (SEQ ID N° 9) y *agaY* (SEQ ID N° 8). El producto de ligadura se transforma en cepa BL21 (DE3) de *E. coli*. Los vectores pET28-aldob, pET28-fbaB y pET28-agaY así obtenidos son verificados mediante secuenciado en cuanto a si contienen los genes con las secuencias correctas.

Determinación de los parámetros cinéticos de (D)-xilulosa-1-fosfato aldolasas y (L)-ribulosa-1-fosfato aldolasa

candidatas

Las proteínas se expresan y purifican como se describe con anterioridad y los parámetros cinéticos de las enzimas se determinan sobre (D)-fructosa-1,6bP, (D)-xilulosa-1P y (L)-ribulosa-1P basándose en el principio siguiente (el ejemplo proporcionado corresponde a la dosificación de una actividad de (L)-ribulosa-1P aldolasa).

(L)-ribulosa-1-quinasa:	(L)-ribulosa + ATP → (L)-ribulosa-1P + ADP
(L)-ribulosa-1P aldolasa:	(L)-ribulosa-1P → DHAP + glicolaldehído
Glicerol-3P deshidrogenasa:	DHAP + NADH → glicerol-3P + NAD

5

La reacción se realiza en la mezcla siguiente: NADH 0,4 mM, PEP 2 mM, ATP 4 mM (todos de Sigma) en tampón de Hepes (Hepes 90 mM, KCl 77 mM, MgCl₂ 6,8 mM, ajustado a pH 7 con una solución de KOH). Las enzimas Khk-C purificadas y GldA (glicerol deshidrogenasa de *Cellulomonas sp.* de la empresa Sigma) se añaden a una cantidad de 15 µl de Khk-C (0,005 U) y 4 µl de una solución concentrada a 84 U/mg de GldA para una mezcla de reacción total de 250 µl. La reacción arranca con la adición de 100 µl de D-xilulosa o (L)-ribulosa 20 mM (Sigma). El consumo de NADH fue sometido a seguimiento mediante espectrofluorometría a 340 nm.

10

Tabla 2A. Parámetros cinéticos de aldolasas sobre (D)-fructosa-1,6-bifosfonato y sobre (D)-xilulosa-1-fosfato

Enzima candidata	(D)-fructosa-1,6bP		(D)-xilulosa-1P	
	Vmax [U/mg]	Km [mM]	Vmax [U/mg]	Km [mM]
Fructosa-16bP aldolasa, Aldo-B, <i>H. sapiens</i>	0,46 ± 0,05	0,03 ± 0,01	0,81 ± 0,2	nd
Fructosa--16bP aldolasa, FbaB, <i>E. coli</i>	0,52 ± 0,1	0,33 ± 0,07	0,18 ± 0,04	nd

Tabla 2B. Parámetros cinéticos de aldolasas sobre (D)-fructosa-1,6-bifosfonato y sobre (L)-ribulosa-1-fosfato

Enzima candidata	(D)-Fructose-1,6bP		(L)-Ribulosa-1P	
	Vmax [U/mg]	Km [mM]	Vmax [U/mg]	Km [mM]
Fructosa-16bP aldolasa, Aldo-B, <i>H. sapiens</i>	0,46 ± 0,05	0,03 ± 0,01	0,00	nd
Fructosa-16bP aldolasa, FbaB, <i>E. coli</i>	0,52 ± 0,1	0,33 ± 0,07	0,06 ± 0,01	nd

15

La AldoB posee una actividad sobre (D)-xilulosa-1-P y, por tanto, puede ser utilizada para la vía sintética de asimilación de xilosa. La FbaB posee una actividad sobre (D)-xilulosa-1P y (L)-ribulosa-1P y, por tanto, puede ser utilizada para la construcción de una vía sintética para la asimilación de (D)-xilosa o (L)-arabinosa.

Ejemplo 3: Funcionamiento *in vitro* de la vía metabólica sintética de asimilación de pentosas

20 Funcionamiento *in vitro* de la vía metabólica sintética de asimilación de (D)-xilosa

La vía metabólica de asimilación de D-xilosa ha sido reconstituida *in vitro* a partir de enzimas purificadas

(comerciales y expresadas y seguidamente purificadas a partir de *E. coli*) a partir de (D)-xilulosa para demostrar su funcionamiento mediante la producción de etilenglicol (figura 1).

Las enzimas utilizadas para la llevar a cabo la vía metabólica sintética son las siguientes:

- Khk-C (cetohecoquinasa/*H.sapiens*), codificada por el gen *khkC* de secuencia SEQ ID N° 1 o KHK-A (Prospecbio);
- 5 • Aldolasa B (AldoB/*H.sapiens*), codificada por el gen *aldoB* de secuencia SEQ ID N° 2 o aldolasa de conejo (Sigma-Aldrich-A2714);
- Glicerol deshidrogenasa *Cellulomonas sp.* (Sigma-Aldrich/G3512-250U).

10 El medio de reacción comprendía el tampón Hepes (Hepes 90 mM, KCl 77 mM, MgCl₂ 6,8 mM) a pH = 7; ATP 4 mM; NADH 0,4 mM; 0,005 Unidades/ml de Khk-A (Prospecbio) o Khk-C (purificada a partir de pET28a); 1 Unidad/ml de aldolasa (AldoB, Sigma A6338) y 1 Unidad/ml de GldA (Sigma-G3512-250U). La reacción se inicia mediante la adición de (D)-xilulosa 5 mM (Cabosynth). Después de un tiempo de incubación de 3 h, el etilenglicol producido en el transcurso de la reacción se cuantifica mediante HPLC (figura 2).

La aparición de etilenglicol en la reacción que contiene (D)-xilulosa-1-quinasa, (D)-xilulosa-1P aldolasa y glicolaldehído reductasa demuestra el funcionamiento de la vía sintética.

15 Funcionamiento *in vitro* de la vía metabólica de asimilación de (L)-arabinosa

La vía metabólica de asimilación de (L)-arabinosa se reconstituye *in vitro* a partir de enzimas purificadas (comerciales y expresadas y seguidamente purificadas a partir de *E. coli*) a partir de (L)-ribulosa.

El funcionamiento de la vía se comprobó mediante la medición por HPLC de etilenglicol producido.

Las enzimas utilizadas para la realización de la vía metabólica sintética son las siguientes:

- 20 • Khk-C (cetohecoquinasa/*H.sapiens*), codificada por el gen *khkC* de secuencia SEQ ID N° 1;
- FbaB (fructosa-1,6-bifosfato aldolasa, *E. coli*) codificada por el gen *fbaB*, de secuencia SEC ID N° 9;
- GldA (glicerol deshidrogenasa *Cellulomonas sp.* (Sigma-Aldrich/G3512-250U).

25 Las enzimas son incubadas en medio de reacción que contiene: NADH 0,4 mM, ATP 4 mM, tampón Hepes pH 7 (Hepes 55 mM, KCl 45 mM, MgCl₂ 4 mM ajustado a pH 7 con KOH) para un volumen final de 500 µl. Se añade Khk-C a 0,005 Unidades/ml mientras que FbaB a 100 µg/ml. Se añade GldA a 1 U/ml. La reacción se inicia mediante la adición de L-ribulosa 20 mM (Sigma-Aldrich). Después de un tiempo de incubación de 3 h, el etilenglicol producido en el transcurso de la reacción se cuantifica mediante HPLC. Los resultados se presentan en la figura 3.

La aparición de etilenglicol en la reacción que contiene (L)-ribulosa-1-quinasa, (L)-ribulosa-1P aldolasa y glicolaldehído reductasa demuestra el funcionamiento de la vía sintética.

30 Ejemplo 4: Funcionamiento *in vivo* de la vía metabólica sintética de asimilación de (D)-xilosa

Clonación de genes de la vía sintética en operón

35 Los genes *H. sapiens khk-C*, que codifica la isoforma C de la enzima cetohecoquinasa (Khk), de secuencia SEQ ID N° 1 y *aldoB* que codifica la isoforma B de fructosa-1,6 aldolasa de secuencia SEQ ID N° 2 son clonados en operón sobre un plásmido pEXT20 (Dykxhoorn, (1996)) bajo el control de un promotor inducible de IPTG construido como sigue. El gen humano *khk-C* fue suministrado por el Dr. Asipu (Asipu *et al.*, 2003) y amplificado con iniciadores P13 y P14 (tabla 3). La aldolasa es sintetizada con optimización de codones para *E. coli* mediante Eurofins® y es amplificada mediante PCR con los iniciadores P15 y P16 (tabla 3). Los iniciadores para la amplificación de los dos genes son conocidos para proporcionar fragmentos de PCR que pueden ser utilizados con el estuche de ensayo In-

Fusion de Clonotech para la adición de una cola de 17 nt de homología con el fragmento contiguo. Se añade un RBS canónico (AGGAGG) a las secuencias de *khk-C* y *aldoB*. El plásmido pEXT20 es digerido por las enzimas de restricción BamHI y SacI. Se utiliza el estuche de ensayo In-Fusion de Clonotech para ligar mediante recombinación los dos fragmentos de PCR y el pEXT20 linealizado proporcionando el plásmido *pEXT20-khk-C-aldob*. El vector *pEXT20-khk-C-aldob* así obtenido es verificado mediante secuenciado que contiene los genes con las secuencias correctas. Este plásmido es transformado en una cepa $\Delta xyIB$ de *E. coli* MG1655 en la que *xyIB*, de secuencia SEQ ID N° 53, que codifica la xilulosa-5-quinasa, es suprimida. Los dos genes *khk-c* y *aldob* son igualmente clonados individualmente sobre el pEXT20 y han sido amplificados en primer lugar por P60 y P61 y P62 y P63, respectivamente, y seguidamente ligados en pEXT20 mediante restricción con las enzimas BamHI y Sall proporcionando los plásmidos pEXT20-*khk-Cet* y pEXT20-*aldob*.

Ensayo de la vía sintética mediante seguimiento del crecimiento de una cepa $\Delta xyIB$ sobre (D)-xilosa

El crecimiento bacteriano, así como la producción de etilenglicol, se ensaya en medio líquido M9 que comprende (D)-xilosa 120 mM como la única fuente carbonada, en presencia de IPTG.

Para controlar la capacidad de las cepas que no poseen la vía de asimilación natural de (D)-xilosa para ser impulsada en presencia de (D)-xilosa, se ensayan las cepas MG1655, MG1655 $\Delta xyIB$ y MG1655 $\Delta XyIB$ que portan el pEXT20-*khk-C*, pEXT20-*aldob* o pEXT20-*khk-C-aldob*.

Sin la vía sintética, solamente la cepa salvaje es impulsada en estas condiciones. La parte de *XyIB* no permite el crecimiento sobre xilosa. Además, ni la presencia de *Khk-C*, ni la presencia de *AldoB* permite restaurar un crecimiento mediante la utilización de la vía natural de asimilación de xilosa (figura 4). Por el contrario, la cepa MG1655 $\Delta xyIB$, que porta el plásmido pEXT20-*khkC-aldob*, es capaz de ser impulsada sobre xilosa.

Mediante un seguimiento de la producción de metabolitos durante el crecimiento mediante HPLC, el etilenglicol es identificado como producto principal de la fermentación de xilosa pasando por la vía sintética producida con un rendimiento de 0,45 moles por mol de xilosa (0,19 g de EG por g de xilosa) (figura 5).

La vía sintética de asimilación de (D)-xilosa por tanto es funcional *in vivo* y restaura el crecimiento de un mutante $\Delta xyIB$ sobre este mismo azúcar.

Ejemplo 5: Identificación de la glicolaldehído reductasa principal

La optimización de la producción de etilenglicol depende de la identificación y la sobreexpresión de glicolaldehído reductasa responsable de la conversión de glicolaldehído en etilenglicol. Varias oxidorreductasas que tienen una actividad sobre glicolaldehído presentes de forma natural en *E. coli*, *GldA*, *YqhD*, *FucO*, *DkgA*, *DkgB*, *YghZ*, *YeaE*, *YajO* han sido identificadas (Lee *et al.*, 2013). Se recuperaron mutantes de estos genes de la colección KEIO para determinar si uno de estos genes codifica la glicolaldehído reductasa principal. Para esto, se ensaya la capacidad de estos mutantes para generar etilenglicol a partir de glicolaldehído. Estas cepas son cultivadas en medio M9 en presencia de xilosa 133 mM y, cuando la DO_{600nm} alcanza el valor de 1, las células son expuestas a glicolaldehído 10 mM. La cantidad de etilenglicol se mide seguidamente mediante HPLC después de 12 h de cultivo.

La ausencia de *YqhD* disminuye enormemente la producción de etilenglicol a partir de glicolaldehído (figura 6), sugiriendo que es la glicolaldehído reductasa principal en *E. coli* en las presentes condiciones de cultivo.

Ejemplo 6: Optimización de la cepa para la producción de etilenglicol por la vía sintética de la invención

Para mejorar la producción de etilenglicol, la glicolaldehído reductasa principal es sobreexpresada sobre un plásmido pACT3. Para esto, se amplifica *yqhD* (SEQ ID N° 4) con P11 y P12 y seguidamente es clonado en pACT3 mediante In-Fusion, previamente digerido con PstI y HindIII. La inserción de *yqhD* en el pACT3 linealizado se efectúa mediante recombinación por medio del estuche de ensayo In-Fusion (Clonotech), proporcionando el plásmido pACT3-*yqhD*. El vector pACT3-*yqhD* así obtenido se verifica mediante secuenciado. El producto de ligadura es seguidamente transformado en las cepas MG1655de interés. De la misma forma, se clonan *gldA* (SEQ ID N° 51) y *fucO* (SEQ ID N° 52) (respectivamente amplificadas por P24 y P25; y P26 y P27) mediante In-Fusion en el pACT3 previamente digerido con PstI y HindIII con el fin de ensayar su efecto.

En las presentes condiciones de cultivo, la reductasa principal de glicolaldehído es YqhD, pero ni su sobreexpresión ni la de otras reductasas GldA y FucO aumentan la producción de etilenglicol. El rendimiento, en efecto, es de solo 0,45 mol/mol de xilosa (0,19 g de EG por g de xilosa), un rendimiento comparable al de la cepa MG1655 $\Delta xyIB$ que expresa el plásmido pEXT20-*khkC-aldB* (figura 7).

- 5 Para aumentar la producción de etilenglicol, la vía de la oxidación del glicolaldehído a ácido glicólico debe ser bloqueada. Para esto, se ensaya el impacto de la delección del gen aldehído deshidrogenasa *AldA* (SEQ ID N° 3). Para cuantificar la producción residual de ácido glicólico, se bloquea el reconsumo de este ácido mediante la inactivación de la glicolato deshidrogenasa (véase ejemplo 7) mediante la delección de su subunidad GlcD (codificada por la secuencia del gen de la SEC ID N° 7). Las delecciones de *aldA* y/o *glcD* en una cepa $\Delta xyIB$ por tanto son emprendidas para la transducción a partir de una cepa KEIO. Estas construcciones proporcionan las cepas MG1655 $\Delta xyIB \Delta aldA$ y MG1655 $\Delta xyIB \Delta aldA \Delta glcD$ que son seguidamente transformadas por pEXT20-*khkC-aldB*.

15 Gracias a la delección de *aldA*, aumenta enormemente el rendimiento de EG. En efecto, la cepa MG1655 $\Delta xyIB \Delta aldA$ que portan el plásmido pEXT20-*khkC-aldB* produce etilenglicol hasta un rendimiento de 0,88 mol por mol de xilosa (0,36 g EG por g de xilosa) (figura 7). La sobreexpresión de YqhD y FucO en estas condiciones hace posible alcanzar un rendimiento de 0,9 y 0,94, respectivamente, mol/mol (0,38 y 0,39 g EG por g de xilosa, respectivamente). Esto está muy cercano al rendimiento teórico máximo alcanzado de 1 mol/mol.

Ejemplo 7: Optimización de la cepa para la producción de ácido glicólico por la vía sintética de asimilación de D-xilosa

- 20 Para aumentar la producción de ácido glicólico a través de la vía sintética de la asimilación de pentosas, se sobreexpresa la glicolaldehído deshidrogenasa de *E. coli*. Para esto, se amplifica *aldA* a partir del ADN genómico de *E. coli* en MG1655 utilizando el par de cebadores (P17 y P18, tabla 3) y el fragmento obtenido es ligado en pGEM-T (Promega) según las instrucciones del fabricante. El fragmento es seguidamente digerido por las enzimas KpnI y HindIII y seguidamente es ligado en el pACT3 en sí mismo linealizado por las dos mismas enzimas. El vector pACT3-*aldA* así obtenido es verificado por secuenciado en cuanto al contenido del gen con la secuencia correcta. Seguidamente se transforma pACT3-*aldA* en la cepa MG1655 $\Delta xyIB$ pEXT20-*khkC-aldB* que proporciona la cepa $\Delta xyIB$ pEXT20-*khkC-aldB* pACT3-*aldA*. Durante la puesta en cultivo de esta cepa sobre un medio M9 + 10 g/l de xilosa, la producción de etilenglicol disminuye significativamente (rendimiento de 0.2 mol/mol) pero la producción de ácido glicólico aumenta solo transitoriamente, indicando el reconsumo del ácido glicólico producido (figura 8).

- 30 Para evitar el reconsumo de ácido glicólico, la glicolato oxidasa, codificada por *glcDEF*, es atenuada. Para esto, la cepa MG1655 $\Delta xyIB \Delta glcD$ es construida mediante delección de *glcD* (SEC ID N° 7) a través de transducción de la mutación a partir de una cepa de la colección KEIO. La cepa MG1655 $\Delta xyIB \Delta glcD$ es transformada con plásmidos pEXT20-*khkC-aldB* o pEXT20-*aldA-khkC-aldB*. Este plásmido es construido a partir de pEXT20-*khkC-aldB* escindido por las enzimas de restricción EcoRI y SmaI en las que es clonado mediante el método In-Fusion el gen *aldA* con un RBS en sentido ascendente. Este gen en sí mismo es amplificado mediante PCR a través de los cebadores P24 y P25. Durante la realización de esta cepa que contiene pEXT20-*khkC-aldB* sobre un medio M9-xilosa (10 g/l), la producción de ácido glicólico aumenta significativamente y alcanza un rendimiento de 0,35 mol/mol (0,19 g de AG por g de xilosa) (figura 8). Cuando *aldA* es sobreexpresado sobre pEXT20, el rendimiento alcanza 0,92 moles/mol de xilosa (0,47 g de AG por g de xilosa).

- 40 La delección de *glcD* permite la acumulación de ácido glicólico debido a la sobreexpresión de *aldA* utilizando 92% del flujo de carbono procedente de la parte C2 de la xilosa.

Ejemplo 8: Optimización de la cepa para la producción de ácido glicólico a través del ciclo de glioxilato

- 45 La producción de ácido glicólico puede ser aumentada todavía más combinando el funcionamiento de la vía sintética optimizada anteriormente descrita con intervenciones genéticas que conducen a la producción de ácido glicólico a través de la vía de glioxilato como se describe en las patentes US 20090155867 (Soucaille, 2009) y US 20120315682 (Dischert *et al.*, 2012). Apoyándose en estos datos publicados, los genes *aceB* y *glcB*, que codifican malato sintasas, el gen *glc* que codifica glioxilato carboligasa, el gen *arcA* que codifica un represor de la respuesta aeróbica y el gen *icd* que codifica una isocitrato deshidrogenasa son suprimidos mediante el protocolo de transducción de fago P1 en la cepa MG1655. El operón *glcDEFG* que codifica una glicolato oxidasa, *edd-eda* que codifica, respectivamente, una fosfonogluconato deshidratasa y Entner-Doudoroff aldolasa así como el *iclR*, que

codifica el represor transcripcional de la vía del glioxilato, son suprimidos a través del procedimiento de delección Datsenko. (Datsenko *et al.*, 2000) utilizando los cebadores P52 y P53 y P54 y P55 y 64 y 65, respectivamente. Los plásmidos para la sobreexpresión paralela de isocitrato liasa, codificada por *aceA* y glioxilato reductasa, codificada por *ghrA* (o *ycdW*) son construidos para mejorar la producción de ácido glicólico a través de los ciclos de Krebs y del glioxilato. Para esto, un plásmido pACT3 es digerido por las enzimas BamHI y HindIII. El gen *ghrA* es amplificado mediante PCR como se describe con anterioridad usando el par de cebadores P40 y P41, mientras que el gen *aceA* es amplificado mediante el par de iniciadores P21 y P22. Los dos fragmentos amplificados y el plásmido linealizado son conjuntamente ligados mediante la utilización del estuche de ensayo In-Fusion (Clontech). Esta construcción proporciona el plásmido pACT3-*ghrA-aceA*.

Este plásmido es seguidamente transformado en la cepa que porta las delecciones $\Delta aceB \Delta glcDEFGB \Delta gcl \Delta edd-eda \Delta iclR \Delta arcA \Delta icd$. La cepa resultante es la cepa 1054. Cuando esta cepa se pone en cultivo sobre M9 + glucosa, se producen 1,17 mol/mol de ácido glicólico (0,49 g AG por g de glucosa) sin producción de acetato (tabla 6).

Tabla 6. Producción de ácido glicólico y de acetato de las cepas de *E. coli* $\Delta aceB \Delta glcDEFGB \Delta gcl \Delta edd-eda \Delta iclR$ con mutaciones adicionales en medio M9 + 1% de glucosa

Cepa	Mutaciones adicionales	Plásmidos	AG [mol/mol]	AG [g/g]	Acetato mol/mol
1052	-	<i>pACT3-ghrA aceA</i>	0,15 ± 0,03	0,06 ± 0,01	0,02
1053	$\Delta arcA$	<i>pACT3-ghrA aceA</i>	0,15 ± 0,02	0,06 ± 0,01	0,00
1054	$\Delta arcA \Delta icd$	<i>pACT3-ghrA aceA</i>	1,17	0,49	0,00

Ejemplo 9: Optimización de la cepa para la producción de ácido glicólico a través de la utilización del ciclo de glioxilato y de la vía sintética de asimilación de (D)-xilosa sobre glucosa y D-xilosa

Para la aplicación de la vía sintética de asimilación de xilosa descrita en este documento, es preferido realizar la producción de ácido glicólico sobre hidrolizados celulósicos o hemicelulósicos que contienen normalmente glucosa y xilosa en porcentajes diferentes. Para demostrar que el rendimiento de ácido glicólico sobre un sustrato que contiene a la vez los azúcares glucosa y xilosa puede ser aumentado a través de la producción simultánea de ácido glicólico por medio de la vía sintética y el ciclo del glioxilato, se construye una cepa que coexpresa las dos vías de forma simultánea. La cepa *E. coli* $\Delta xylB \Delta aceB \Delta glcDEFGB \Delta gcl \Delta edd-eda \Delta iclR \Delta arcA \Delta icd$ es cotransformada por los plásmidos *pACT3-ghrA-aceA* y *pEXT20-khk-c-aldob-aldA*. La cepa así obtenida es la cepa 905. El seguimiento de la producción de ácido glicólico mediante HPLC se realiza durante un crecimiento de esta cepa sobre un medio mineral M9 + 0,1% de extractos de levaduras + 0,2% de triptona y en presencia de 0,25% de glucosa y 0,5% de xilosa (figura 9)

Tabla 7. Producción de ácido glicólico a partir de cepas de *E. coli* en medio M9 + 2,5 g/l de glucosa + 5 g/l (D)-xilosa + 1 g/l de extracto de levaduras + 2 g/l de triptona. proD: *galP-galP* sobreexpresada con el promotor proD constitutivo (Davis *et al.*, 2011) (**) Rendimiento calculado sobre la base de azúcar total consumido

Cepa	Genotipo	Plásmidos	AG** [g/g]
1054	$\Delta aceB \Delta glcDEFGB \Delta gcl \Delta edd-eda \Delta iclR \Delta arcA \Delta icd$	<i>pACT3- aceA ghrA</i>	0,4
1044	$\Delta aceB \Delta glcDEFGB \Delta gcl \Delta edd-eda \Delta iclR \Delta arcA \Delta icd$	<i>pACT3- aceA ghrA pEXT20-khkC-aldob- aldA</i>	0,43
905	$\Delta xylB \Delta aceB \Delta glcDEFGB \Delta gcl \Delta edd-eda \Delta iclR \Delta arcA \Delta icd$	<i>pACT3- aceA ghrA pEXT20-khkC-aldob- aldA</i>	0,51

979	$\Delta xyIB \Delta aceB \Delta glcDEFG \Delta gcl \Delta edd-eda \Delta iclR \Delta arcA \Delta icd proD :galP$	$pACT3- aceA ghrA pEXT20-khkC-aldB- aldA$	0,66
-----	--	---	------

5 La cepa consume primeramente la glucosa y seguidamente la xilosa a pesar de la ausencia de *xilB*, mostrando así que la vía sintética es activa incluso en estas condiciones. Después de 100 h de cultivo, se producen 2,35 g/l de ácido glicólico por la cepa con un rendimiento sobre el azúcar utilizado total de 0,51 g/g (tabla 7). Este rendimiento es superior al obtenido con una cepa casi isogénica que no posee la vía sintética de asimilación de (D)-xilosa (tabla 7, cepa 1054), o que no porta la delección del gen *xyIB* que codifica la enzima que cataliza la entrada en la vía natural de asimilación de (D)-xilosa (tabla 7, cepa 1044).

10 La tasa de asimilación de xilosa después del consumo total de la glucosa se continúa siendo relativamente baja con un valor de aproximadamente 0,19 mmol/(l h). Para acelerar la asimilación de xilosa, la expresión de la permeasa de azúcares *galP* se hizo constitutiva utilizando el método siguiente: un fragmento de ADN que codifica el promotor constitutivo *proD* descrito por Davis (Davis *et al.*, 2011) y sintetizado mediante Eurofins (SEC ID N° 90) es amplificado con iniciadores P56 y P57. Una cassette de expresión del gen *kan* es amplificado utilizando los iniciadores P58 y P59 y el plásmido pKD4 (SEQ ID N° 95) como matriz. Los dos fragmentos de PCR se fusionan mediante una extensión solapada PCR utilizando los iniciadores P59 y P57. El producto de PCR así obtenido es transformado en la cepa 905 utilizando el método de Datsenko y Wanner (Datsenko *et al.*, 2000). Se recuperan clones resistentes a la kanamicina y se verifica si contienen el promotor sintético y constitutivo ante *galP*.

20 La nueva cepa así obtenida es cotransformada por $pACT3 aceA-ghrA$ y $pEXT20-khk-C-aldB-aldA$ proporcionando la cepa 979. Su crecimiento y el seguimiento de la producción de ácido glicólico mediante HPLC se realizan sobre un medio mineral M9 + 0,1% de extracto de levaduras + 0,2% de triptona en presencia de 0,25% de glucosa y 0,5% de xilosa (figura 10).

La cepa acumula 3,84 g/l de ácido glicólico con un rendimiento de 0,66 g/g sobre azúcar total. La tasa de asimilación de xilosa después del consumo total de la glucosa es aumentado gracias a la sobreexpresión de *Galp* y alcanza un valor de 0,32 mmol/(l h).

Ejemplo 10: Expresión de la vía sintética de asimilación de xilosa en *Saccharomyces cerevisiae*

25 Para ensayar la portabilidad de la vía sintética de asimilación de xilosa, se ensaya su expresión en otro microorganismo de interés, la levadura *Saccharomyces cerevisiae*.

30 La *S. cerevisiae* no posee un sistema enzimático natural para convertir (D)-xilosa en (D)-xilulosa y, por tanto, no es capaz de impulsarse sobre este azúcar. Habitualmente son expresadas dos vías metabólicas de forma heteróloga en esta levadura para realizar la conversión de (D)-xilosa en (D)-xilulosa y para facilitar su crecimiento sobre xilosa. La xilosa isomerasa (XI) cataliza la conversión de xilosa en xilulosa directa de forma redox neutra. Como alternativa, la acción secuencial de la xilosa reductasa (XR) y la xilitol deshidrogenasa (XDH) permite también una conversión de xilosa en xilulosa, utilizando los cofactores NADPH y NAD, respectivamente. Para mostrar el funcionamiento de la vía sintética de asimilación de xilosa en la levadura, esta vía sintética se complementa mediante una isomerasa de xilosa o bien mediante el sistema XR/XDH.

35 Para completar la vía sintética de asimilación de xilosa con una XI, se utiliza la XI de *Clostridium phytofermentans* codón optimizado para *S. cerevisiae* conocida por el equipo de Eckhard Boles (Brat *et al.*, 2009). Se expresa *khk-C* bajo el control del promotor de la triosa fosfato isomerasa *Tpi* (*Ptpi*: SEC ID N° 15) y se usa el terminador *Trk1* (SEC ID N° 12).

40 El gen de aldolasa *AldoB* se coloca bajo el control del promotor de la aldolasa natural de *S. cerevisiae* (*pFab*: SEC ID N° 14) y se utiliza también un terminador *Tik1* (SEC ID N° 13).

Estos genes son clonados mediante recombinación en levadura en el plásmido p425 linealizado por *KpnI* y *HindIII* proporcionando *p425-khk-aldB* utilizando los cebadores siguientes (véase también la tabla 4): P1' y P2' para *Ptpi*, P3' y P4' para *khk-C*, P5' y P6' para *Trk1*, P7' y P8' para *pFab*, P9' y P10' para *aldB* y P11' y P12' para *tik1*.

El plásmido p425-khk-aldob se usa seguidamente en una cepa competente incapaz de impulsarse sobre xilosa, ya que está mutada para su xilulosa quinasa natural, Xks1 (SEQ ID N° 89).

Se realiza así cotransformación con un plásmido construido por el equipo de Eckhard Boles que contiene a la vez la xilulosa isomerasa y un transportador de xilosa Gal- en la cepa CEN.PK-2 *xks1*-.

- 5 Se ensaya la capacidad de la cepa para impulsarse en un medio mínimo que contiene únicamente xilosa (figura 11).

En este caso, aunque el mutante *xks1* no es capaz de impulsarse sobre D-xilosa, el mutante *xks1* que porta la vía sintética restablece su crecimiento, indicando que la vía es funcional en la levadura.

La utilización de una XI no es el único medio posible para asimilar la xilosa convirtiéndola en xilulosa. En efecto, a continuación de una reacción de reducción catalizada por una xilosa reductasa (XR) y una deshidrogenación del xilitol así obtenido catalizada por una xilitol deshidrogenasa (XDH), se obtiene xilulosa. Por tanto, se utiliza la cepa TMB3001 (Eliasson *et al.*, 2000) que expresa el sistema XR/XDH para ensayar si la vía sintética es capaz de ser utilizada en estas condiciones para asimilar la xilosa. Se construye un mutante *xks1*⁻ en la cepa TMB3001 transformando un fragmento de PCR que contiene los bordes flanqueantes homólogos a las extremidades no codificadores de *xks1* para generar una delección mediante recombinación homóloga. El fragmento de PCR es amplificado a partir de la cepa BY de entidad Yeast collection knockout (YKO) (Winzeler *et al.*, 1999) con los cebadores P26' y P27' y que contiene una cassette *kanMX* para la selección de los recombinantes. La vía sintética es expresada en el plásmido pYCP-khk-C-aldob construido de la forma siguiente. El plásmido pYCP-TPS1 (SEC ID N°: 80) es digerido por *AgeI* y *XbaI* para extraer la cassette TPS1. Se diseñan los cebadores Fw pTpi y rev Tlk1 para tener una cola flotante de 40 nucleótidos homóloga al plásmido pYCP linealizado por *AgeI* y *XbaI* con el fin de que se puede recombinar en el plásmido. Seguidamente se expresa *khk-C* codón optimizado para la levadura (SEQ ID N° 83) amplificado por P15' y P16' bajo el control del promotor de la triosa fosfato isomerasa Tpi (amplificada por P13' y P14') y del terminador Trk1 (amplificado por P17' y P18') y se expresa la aldolasa *aldob* codón optimizada para la levadura (SEQ ID N°: 84) amplificado por P21' y P22' bajo el control del promotor de la aldolasa natural de *S. cerevisiae* pFbaB1 (amplificado por P19' y P20') y del terminador Tlk1 (amplificado por P23' y P24'). Estas construcciones son clonadas mediante recombinación en levadura en el plásmido pYCP proporcionando pYCP-khk-C-aldob. Este plásmido se utiliza seguidamente en la cepa TMB3001 *xks1*⁻ incapaz de impulsarse sobre xilosa ya que es mutada para xilulosa quinasa, Xks1.

Aunque que un mutante TMB3001 *xks1*⁻ ha perdido su capacidad de impulsarse sobre xilosa, contrariamente a la cepa TMB3001, la cepa TMB3001 *xks1*⁻ que lleva *pYCP-khk-C-aldob* recupera un crecimiento sobre xilosa (figura 12). Esto sugiere que la vía sintética también es funcional en una cepa que asimila la xilosa a través del sistema XR/XDH.

Ejemplo 11: Funcionamiento *in vivo* de la vía metabólica sintética de la asimilación de (L)-arabinosa

Clonación de los genes de la vía sintética en el operón

Los genes *H. sapiens khk-C* que codifican isoforma C de la enzima cetohexoquinasa (Khk) de secuencia SEQ ID N° 1 y *fbaB* que codifica una fructosa-1,6-bifosfonato aldolasa isoforma B de *E. coli* de secuencia SEQ ID N°: 9 son clonados en operón sobre un plásmido pEXT20 (Dykhooorn *et al.*, 1996) bajo el control de un promotor inducible por IPTG construido como sigue. El gen *khk-C* humano es proporcionado por el Dr. Asipu (Asipu *et al.*, 2003) y amplificado con los iniciadores P13 y P14 (tabla 3). La aldolasa es amplificada a partir de ADN genómico de *E. coli* de la cepa MG1655 con los cebadores P28 y P29 (tabla 3).

Los iniciadores para la amplificación de estos dos genes están diseñados para proporcionar fragmentos PCR que pueden ser utilizados con el estuche de ensayo In-Fusion de Clontech mediante la adición de una cola de 17 nucleótidos de homología con el fragmento contiguo. Se escogen RBS a través del programa informático RBS Calculator (Salis *et al.*, 2011). El gen *fbaB*, por lo tanto, precedido de un RBS de secuencia (TTAGGAGGTATACT) previsto para proporcionar una expresión máxima y *khk-C* está precedido por un RBS de secuencia (ACAGCTTTTAATTATACACTTTAAGGAGGACAGAC) previsto para minimizar la expresión. Los iniciadores para la amplificación de los dos genes con los nuevos RBS (P30 y P22 para *khk-C* y P31 y P32 para *fbaB*) están diseñados para proporcionar fragmentos PCR que pueden ser utilizados con el estuche de ensayo de Fusion mediante la adición de una cola de 15 nucleótidos de homología con el fragmento contiguo. El plásmido pEXT20 es digerido por

5 las enzimas de restricción BamHI y Sall. El estuche de ensayo In-Fusion de Clontech es utilizado para ligar mediante recombinación los dos fragmentos de PCR y el pEXT20 linealizado, proporcionando el plásmido pEXT20-*khk-C-RBSmax-fbaB*. El plásmido pEXT20-*khk-C-RBSmax-fbaB* es transformado en MG1655 1655 Δ *araB*. El vector pEXT20-*khk-C-RBSmax-fbaB* así obtenido es verificado mediante secuenciado. De forma análoga, los plásmidos pEXT20-*khk-C* y pEXT20-*fbaB* son construidos, utilizando los iniciadores 83/84 y 33/34 (tabla 3) mediante In-Fusion. Estos plásmidos son transformados en una cepa Δ *araB* de *E. coli* MG1655 en la que *araB*, de secuencia SEQ ID N° 88, que codifica la ribulo-5-quinasa, es suprimido.

Ensayo de la vía sintética mediante seguimiento del crecimiento de una cepa Δ *araB* sobre (L) -arabinosa

10 El crecimiento bacteriano es ensayado en medio líquido M9 que comprende (L)-arabinosa 60 mM como única fuente carbonada, en presencia de IPTG. Para controlar la capacidad de las cepas que no poseen la vía natural de asimilación de arabinosa para ser impulsada en presencia de (L)-arabinosa en medio M9, se ensayan las cepas MG1655 Δ *araB* y MG1655 Δ *araB* que portan los plásmidos pEXT20-*khk-C*, pEXT2-*fbaB* o pEXT20-*khk-C-RBSmax-fbaB*. Se efectúa un precultivo durante una noche en LB + IPTG 100 μ M + 2% de L-arabinosa y seguidamente las células se transfirieren a medio M9 + glucosa 2% + arabinosa 2% + IPTG 100 μ M a DO_{600nm} = 0,2 y seguidamente a 15 DO_{600nm} = 1, las células son transferidas a un medio M9 + L-arabinosa 2% + IPTG 100 μ M a DO_{600nm} = 0,2. Solo la cepa que expresa simultáneamente las enzimas KhkC y FbaB, que tienen actividades de (L)-ribulosa-1 quinasa y (L)-ribulosa-1P aldolasa, respectivamente, es impulsada en estas condiciones (figura 13). Estos resultados ponen de manifiesto el funcionamiento de la vía sintética de la asimilación de la (L)-arabinosa.

Tabla 3. Iniciadores utilizados para la expresión de la vía en *E. coli*

AA	Iniciadores	SEQ ID
P1	YC-49 rhaB fw-ndel	16
P2	YC-50 rhaB rev-BamHI	17
P3	FucokinasefwNcoI	18
P4	Fucokinase revBamHISall	19
P5	AldoB_BamHI_Fw	20
P6	AldoB_Hind3_Rv	21
P7	fbaB-Ndel-fw	22
P8	fbaB-BamHI-rev	23
P9	agaY-Ndel-fw	24
P10	agaY-BamHI-rev	25
P11	YC73-Fw YqhD	26
P12	YC74-Rev YqhD	27
P13	YC52-fwkhkfu	28
P14	YC53-revkhkfu	29
P15	YC-76 fw AldoB	30

ES 2 731 782 T3

P16	YC-75 fw AldoB	31
P17	aldA_KpnI_Fw	32
P18	aldA_Hind3_Rv	33
P19	pACT ycdW-Fw	34
P20	pACT ycdW-Rev	35
P21	Operón aceA-Fw	36
P22	pACT aceA-Rev	37
P23	Operón ycdW-Rev	38
P24	gldA_rbs_f	54
P25	gldA_rbs_r	55
P26	fucO_fw_inf	56
P27	fucO_rev_inf	57
P28	FbaB_oper_BamHI_F	58
P29	FbaB_oper_Sall_R	59
P30	khkrbsweak fw	60
P31	fbabrbsmax fw	61
P32	fbabrbsmax rev	62
P33	FbaB inf BamHI F	63
P34	FbaB inf Hind R	64
P40	pACTaceA-Rev	65
P41	Op aceA ghrA-Fw	66
P42	pACT3 ghrA_XbaI-Rev	67
P51	pACTghrA_XbaI-Rev	68
P52	glcDEFGB_fw	69
P53	glcDEFGB_rev	70
P54	edd eda_fw	71

P55	edd_eda_rev	72
P56	prom_galP*_fw	73
P57	prom_galP*_rev	74
P58	disruption_k7_fw	75
P59	disruption_k7_rev	76
P60	KHK-Cfw	77
P61	KHK-Crev	78
P62	AldoB_RBS_BaHI_F	79
P63	RevAldoB	80
P64	iclR-fw3	81
P65	iclR-rev2	82

Tabla 4. Iniciadores utilizados para la expresión de la vía en *S. cerevisiae*

AA	Iniciadores	SEQ ID
P1'	YC1_pTPI_fw	39
P2'	YC2_pTPI_Rev	40
P3'	YC3_khk-C_fw	41
P4'	YC4_khk-C_rev	42
P5'	YC5_ter_trki_fw	43
P6'	YC6_ter_trki_rev	44
P7'	YC7_pFBA1_fw	45
P8'	YC8_pFBA1_rev	46
P9'	YC9_aldoB_fw	47
P10'	YC10_aldoB_fw	48
P11'	YC11_Tlk1_terminador_fw	49
P12'	YC12_Tlk1_terminador_rev	50

AA	Iniciadores	SEQ ID
P13'	fw tpi ycp	83
P14'	Rev tpi	84
P15'	Fw khk-C opt	85
P16'	Rev khk-C opt	86
P17'	Trk1 fw	87
P18'	Term-trk1 rev	88
P19'	pFba1 fw	89
P20'	pfab1 rev	90
P21'	aldoB opt fw	91
P22'	aldoB opt rev	92
P23'	tlk1 fw	93
P24'	rev tlk1 ycp	94
P25'	fw deletion xks1	95
P26'	rev deletion xks1	96

Tabla 5: Lista de secuencias de los genes

Especie	Gen	SEQ ID N°
<i>H.sapiens</i>	<i>khkC</i>	1
<i>H.sapiens</i>	<i>aldoB</i>	2
<i>E.coli</i>	<i>aldA</i>	3
<i>E.coli</i>	<i>yqhD</i>	4
<i>E.coli</i>	<i>fucK</i>	5
<i>E.coli</i>	<i>rhaB</i>	6
<i>E.coli</i>	<i>glcD</i>	7
<i>E.coli</i>	<i>agaY</i>	8

Especie	Gen	SEQ ID N°
<i>E.coli</i>	<i>fbaB</i>	9
<i>E.coli</i>	<i>ghrA</i>	10
<i>E.coli</i>	<i>aceA</i>	11
<i>S.cerevisiae</i>	<i>trk1</i>	12
<i>S.cerevisiae</i>	<i>tlk1</i>	13
<i>S.cerevisiae</i>	<i>pFab</i>	14
<i>S.cerevisiae</i>	<i>pTpi</i>	15
<i>E.coli</i>	<i>gldA</i>	51
<i>E.coli</i>	<i>fucO</i>	52
<i>E.coli</i>	<i>xyIB</i>	53
<i>S.cerevisiae</i>	<i>Khk-C codon optimizado</i>	97
<i>S.cerevisiae</i>	<i>AldoB codon optimizado</i>	98
	<i>YCP-Tps1</i>	99
<i>E.coli</i>	<i>araB</i>	100
<i>S.cerevisiae</i>	<i>xks1</i>	101
<i>E. coli</i>	<i>proD</i>	102

Referencias

- Asipu, A., B. E. Hayward, J. O'Reilly, and D. T. Bonthron. 2003. "Properties of normal and mutant recombinant human ketohexokinases and implications for the pathogenesis of essential fructosuria." *Diabetes* 52 (9):2426-32.
- 5 Baba, T., T. Ara, M. Hasegawa, Y. Takai, Y. Okumura, M. Baba, K. A. Datsenko, M. Tomita, B. L. Wanner, and H. Mori. 2006. "Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection." *Mol Syst Biol* 2:2006.0008. doi: 10.1038/msb4100050.
- Bais, R., H. M. James, A. M. Rofe, and R. A. Conyers. 1985. "The purification and properties of human liver ketohexokinase. A role for ketohexokinase and fructose-bisphosphate aldolase in the metabolic production of oxalate from xylitol." *Biochem J* 230 (1):53-60.
- 10 Brat, D., E. Boles, and B. Wiedemann. 2009. "Functional expression of a bacterial xylose isomerase in *Saccharomyces cerevisiae*." *Appl Environ Microbiol* 75 (8):2304-11. doi: 10.1128/aem.02522-08.
- Cherepanov, P. P., and W. Wackernagel. 1995. "Gene disruption in *Escherichia coli*: TcR and KmR cassettes with the option of FLP-catalyzed excision of the antibiotic-resistance determinant." *Gene* 158 (1):9-14.

- Dagert, M., and S. D. Ehrlich. 1979. "Prolonged incubation in calcium chloride improves the competence of *Escherichia coli* cells." *Gene* 6 (1):23-8.
- Datsenko, K. A., and B. L. Wanner. 2000. "One-step inactivation of chromosomal genes in *Escherichia coli* K-12 using PCR products." *Proc Natl Acad Sci U S A* 97 (12):6640-5. doi: 10.1073/pnas.120163297.
- 5 Davis, J. H., A. J. Rubin, and R. T. Sauer. 2011. "Design, construction and characterization of a set of insulated bacterial promoters." *Nucleic Acids Res* 39 (3):1131-41. doi: 10.1093/nar/gkq810.
- Dischert, W., and P. Soucaille. 2012a. Method for producing high amount of glycolic acid by fermentation. Google Patents.
- 10 Dischert, W., and P. Soucaille. 2012b. Method for producing high amount of glycolic acid by fermentation. Google Patents.
- Dykxhoorn, D. M., R. St Pierre, and T. Linn. 1996. "A set of compatible tac promoter expression vectors." *Gene* 177 (1-2):133-6.
- Eliasson, A., C. Christensson, C. F. Wahlbom, and B. Hahn-Hägerdal. 2000. "Anaerobic xylose fermentation by recombinant *Saccharomyces cerevisiae* carrying XYL1, XYL2, and XKS1 in mineral medium chemostat cultures." *Appl Environ Microbiol* 66 (8):3381-6.
- 15 Gietz, R. D., and R. A. Woods. 2002. "Transformation of yeast by lithium acetate/single-stranded carrier DNA/polyethylene glycol method." *Methods Enzymol* 350:87-96.
- Klamt, S., J. Saez-Rodriguez, J. A. Lindquist, L. Simeoni, and E. D. Gilles. 2006. "A methodology for the structural and functional analysis of signaling and regulatory networks." *BMC Bioinformatics* 7:56. doi: 10.1186/1471-2105-7-56.
- 20 Lee, C., I. Kim, and C. Park. 2013. "Glyoxal detoxification in *Escherichia coli* K-12 by NADPH dependent aldo-keto reductases." *J Microbiol* 51 (4):527-30. doi: 10.1007/s12275-013-3087-8.
- Liu, H., K. R. Ramos, K. N. Valdehuesa, G. M. Nisola, W. K. Lee, and W. J. Chung. 2013. "Biosynthesis of ethylene glycol in *Escherichia coli*." *Appl Microbiol Biotechnol* 97 (8):3409-17. doi: 10.1007/s00253-012-4618-7.
- 25 Salis, H. M. 2011. "The ribosome binding site calculator." *Methods Enzymol* 498:19-42. doi: 10.1016/B978-0-12-385120-8.00002-4.
- Soucaille, P. 2009. Glycolic Acid Production by Fermentation from Renewable Resources. Google Patents.
- Stephanopoulos, Gregory (4 Russet Lane, Winchester, MA, 01890, US), Pereira, Brian (11 Everett Street, Apt. Ng4Cambridge, MA, 02138, US), DE MEY, Marjan (Ursulinen Straat 4/1, Gent, Gent, BE), Dugar, Deepak (69 Chestnut Street, Cambridge, MA, 02139, US), Avalos, Jose, Luis (65 Walnut Street, Arlington, MA, 02476, US). 2013. Engineering microbes and metabolic pathways for the production of ethylene glycol. Massachusetts Institute of Technology (77 Massachusetts Avenue, Cambridge, MA, 02139, US)
- 30 Winzeler, E. A., D. D. Shoemaker, A. Astromoff, H. Liang, K. Anderson, B. Andre, R. Bangham, R. Benito, J. D. Boeke, H. Bussey, A. M. Chu, C. Connelly, K. Davis, F. Dietrich, S. W. Dow, M. El Bakkoury, F. Foury, S. H. Friend, E. Gentalen, G. Giaever, J. H. Hegemann, T. Jones, M. Laub, H. Liao, N. Liebundguth, D. J. Lockhart, A. Lucau-Danila, M. Lussier, N. M'Rabet, P. Menard, M. Mittmann, C. Pai, C. Rebischung, J. L. Revuelta, L. Riles, C. J. Roberts, P. Ross-MacDonald, B. Scherens, M. Snyder, S. Sookhai-Mahadeo, R. K. Storms, S. Véronneau, M. Voet, G. Volckaert, T. R. Ward, R. Wysocki, G. S. Yen, K. Yu, K. Zimmermann, P. Philippsen, M. Johnston, and R. W. Davis. 1999. "Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis." *Science* 285 (5429):901-6.
- 40

Listado de secuencias

- <110> INRA INSA Toulouse CNRS
- 5 <120> PROCEDE POUR LA PRODUCTION D'AU MOINS UN METABOLITE D'INTERET PAR TRANSFORMATION D'UN PENTOSE DANS UN MICROORGANISME
- <130> Z376WO
- <160> 102
- <170> PatentIn version 3.5
- 10 <210> 1
 <211> 897
 <212> DNA
 <213> Homo sapiens
- <400> 1
 atggaagaga agcagatcct gtgcggtggg ctagtgggtgc tggacgtcat cagcctgggtg 60
 gacaagtacc ctaaggagga ctoggagata aggtgtttgt cccagagatg gcagcgcgga 120
 ggcaacgcgt ccaactcctg caccgttctc tcctgctcg gagccccctg tgccttcatg 180
 ggctcaatgg ctccctggcca tgttgctgac ttcctgggtgg ccgacttcag gcggcggggc 240
 gtggacgtgt ctccaggtggc ctggcagagc aagggggaca cccccagctc ctgctgcatc 300
 atcaacaact ccaatggcaa ccgtaccatt gtgctccatg acacgagcct gccagatgtg 360
 tctgctacag actttgagaa ggttgatctg acccagttca agtggatcca cattgagggc 420
 cggaacgcat cggagcaggt gaagatgctg cagcggatag acgcacacaa caccaggcag 480
 cctccagagc agaagatccg ggtgtccgtg gaggtggaga agccacgaga ggagctcttc 540
 cagctgtttg gctacggaga cgtgggtgtt gtcagcaaag atgtggccaa gcacttgggg 600
 ttccagtcag cagaggaagc cttgaggggc ttgtatggtc gtgtgaggaa aggggctgtg 660
 cttgtctgtg cctgggctga ggaggggcc gacgccctgg gccctgatgg caaattgctc 720
 cactcggatg ctttcccgcc accccgcgtg gtggatacac tgggagctgg agacaccttc 780
 aatgcctccg tcactttcag cctctcccag gggaggagcg tgcaggaagc actgagatcc 840
 ggggtgccagg tggccggcaa gaagtgtggc ctgcagggct ttgatggcat cgtgtga 897
- 15 <210> 2
 <211> 1095
 <212> DNA
 <213> Homo sapiens
- <400> 2

ES 2 731 782 T3

atggcacatc gctttccggc tctgacgcaa gaacagaaga aagaattgtc agagattgcg 60
cagtccatcg tggcgaatgg caaaggtatt ctggccgag atgagagcgt gggaaacctg 120
ggtaatcggc ttcaacgcat caaagtggag aacaccgaag aaaatcggcg tcagtttcgc 180
gaaatcctct ttagcgtcga cagtagcatc aatcagtcca ttggcggggt gattctgttt 240
cacgaaacgc tgtaccagaa agactcgcaa gggaaactgt tccgcaacat tctgaaagag 300
aaaggcattg tgggtggcat caaactggat cagggtggcg cacccttagc tggcactaat 360
aaggagacia cgattcaagg cctcgatggg ctttctgaac gttgtgccca gtataagaag 420
gatggtgttg actttggcaa atggcgcgcg gtactgcgta tcgctgatca gtgccctagc 480
agtctggcca ttcaggaaaa cgcgaaatgag ttagcacgct atgagagcat ttgccaacag 540
aacggcctgg ttcccattgt cgagccggaa gtaatcccgg atggcgatca tgacctggaa 600
cactgccagt atgtcaccga gaaagttctg gcggcgggtg ataaagccct gaatgacct 660
cacgtgtact tggaaagtac cctcttgaaa ccgaacatgg taactgccgg acatgcctgc 720
accaagaat acaccctga acaagtcgcc atggcaacag ttacggcttt acaccgcaact 780
gttccagcgg cggttccggg tatttgcttc ctgtcaggtg gcatgtcggg agaagatgcc 840
accctcaacc tgaatgcgat caacttgtgt ccaactgccga agccgtggaa actgtcattc 900
tcctatggac gtgcccttca ggccagtgcg ttagcggcat ggggtgggaa agcggcaaac 960
aaagaagcta cccaggaagc cttcatgaaa cgtgcaatgg cgaactgtca agctgcgaaa 1020
ggccagtaag tccataccgg ttctagcggg gctgcatcga cacaatctct gtttacggcc 1080
tgttacacgt attaa 1095

5

<210> 3
<211> 1440
<212> DNA
<213> Escherichia coli

<400> 3

ES 2 731 782 T3

atgtcagtac ccgttcaaca tcoctatgtat atcgatggac agtttggttac ctggcgtgga 60
 gacgcatgga ttgatgtggt aaaccctgct acagaggctg tcatttcccg catacccgat 120
 ggtcaggccg aggatgcccg taaggcaatc gatgcagcag aacgtgcaca accagaatgg 180
 gaagcgttgc ctgctattga acgcgccagt tggttgcgca aaatctccgc cgggatccgc 240
 gaacgcgcca gtgaaatcag tgcgctgatt gttgaagaag ggggcaagat ccagcagctg 300
 gctgaagtgc aagtggcttt tactgccgac tatatcgatt acatggcggga gtgggcacgg 360
 cgttacgagg gcgagattat tcaaagcgat cgtccaggag aaaatattct tttgtttaa 420
 cgtgcgcttg gtgtgactac cggcattctg ccgtggaact tcccgttctt cctcattgcc 480
 cgcaaaatgg ctcccgtctt tttgaccggt aataccatcg tcattaaacc tagtgaattt 540
 acgccaaca atgcgattgc attcgcaaaa atcgtcgatg aaataggcct tccgcgcggc 600
 gtgtttaacc ttgtactggg gcgtggtgaa accgttgggc aagaactggc gggtaaccga 660
 aaggtcgcaa tggcagtat gacaggcagc gtctctgcag gtgagaagat catggcgact 720
 gcggcgaaaa acatcaccaa agtgtgtctg gaattggggg gtaaagcacc agctatcgta 780
 atggacgatg ccgatcttga actggcagtc aaagccatcg ttgattcacg cgtcattaat 840
 agtgggcaag tgtgtaactg tgcagaacgt gtttatgtac agaaaggcat ttatgatcag 900
 ttcgtcaatc ggctgggtga agcgatgcag gcggttcaat ttggtaaccg cgctgaacgc 960
 aacgacattg cgatggggcc gttgattaac gccgcggcgc tggaaagggt cgagcaaaaa 1020
 gtggcgcgcg cagtagaaga aggggcgaga gtggcgttcg gtggcaaagc ggtagagggg 1080
 aaagatatt attatccgcc gacattgctg ctggatgttc gccaggaaat gtcgattatg 1140
 catgaggaaa cctttggccc ggtgctgcca gttgtcgcat ttgacacgct ggaagatgct 1200
 atctcaatgg ctaatgacag tgattacggc ctgacctcat caatctatac ccaaaatctg 1260
 aacgtcgcga tgaagccat taaaggctg aagtttggtg aaacttacat caaccgtgaa 1320
 aacttcgaag ctatgcaagg cttccacgcc ggatggcgta aatccggtat tggcggcgca 1380
 gatggtaaac atggcttgca tgaatatctg cagaccagc tggtttattt acagtcttaa 1440

<210> 4
 <211> 1164
 <212> DNA
 <213> Escherichia coli

<400> 4

5

ES 2 731 782 T3

```

atgaacaact ttaatctgca caccccaacc cgcattctgt ttggtaaagg cgcaatcgct      60
ggtttacgcg aacaaattcc tcacgatgct cgcgtattga ttacctacgg cggcggcagc      120
gtgaaaaaaaa cggcggttct cgatcaagtt ctggatgcc tgaaggcat ggacgtgctg      180
gaatttgcg gtattgagcc aaaccggct tatgaaacgc tgatgaacgc cgtgaaactg      240
gttcgcgaac agaaagtgac tttcctgctg gcggttgcg gcggttctgt actggacggc      300
accaaattta tcgccgagc ggctaactat ccgaaaata tcgatccgtg gcacattctg      360
caaacggcg gtaaagagat taaaagccc atccgatgg gctgtgtgct gacgctgcca      420
gcaaccggtt cagaatccaa cgcaggcgcg gtgatctccc gtaaaaccac aggcgacaag      480
caggcgttcc attctgccca tggtcagccg gtatttgccg tgctcgatcc ggtttatacc      540
tacaccctgc cgccgctca ggtggctaac ggcgtagtgg acgcctttgt acacaccgtg      600
gaacagtatg ttaccaaac ggttgatgcc aaaattcagg accgtttcgc agaaggcatt      660
ttgctgacgc taatcgaaga tgggccgaaa gccctgaaag agccagaaaa ctacgatgtg      720
cgcgccaacg tcatgtgggc ggcgactcag gcgctgaacg gtttgattgg cgctggcgta      780
ccgcaggact gggcaacgca tatgctgggc cacgaaactga ctgcgatgca cggctctgat      840
cacgcgcaaa cactggctat cgtcctgcct gcactgtgga atgaaaaacg cgataccaag      900
cgcgctaagc tgctgcaata tgctgaacgc gtctggaaca tcaactgaagg ttccgatgat      960
gagcgtattg acgccgcat tgccgcaacc cgcaatttct ttgagcaatt aggcgtgccg     1020
accacctct ccgactacgg tctggacggc agctccatcc cgctttgct gaaaaaactg     1080
gaagagcacg gcatgaccca actgggcgaa aatcatgaca ttacggttga tgtcagccgc     1140
cgtatatacg aagccgcccg ctaa                                           1164

```

5

```

<210> 5
<211> 1419
<212> DNA
<213> Escherichia coli

<400> 5

```

ES 2 731 782 T3

atgaaacaag aagttatcct ggtactcgac tgtggcgcgga ccaatgtcag ggccatcgcg 60
 gttaatcggc agggcaaaat tgttgcccgc gcctcaacgc ctaatgccag cgatatcgcg 120
 atgaaaaaca acacctggca ccagtgtct ttagacgcca ttttgcaacg ctttgctgat 180
 tgctgtcggc aaatcaatag tgaactgact gaatgccaca tccgcggtat cgccgtcacc 240
 acctttggtg tggatggcgc tctggtagat aagcaaggca atctgctcta tccgattatt 300
 agctggaat gtccgcgaac agcagcggtt atggacaata ttgaacggtt aatctccgca 360
 cagcggttgc aggctatttc tggcgtcgga gccttagtt tcaatacgtt atataagttg 420
 gtgtggttga aagaaatca tccacaactg ctggaacgcg cgcacgcctg gctctttatt 480
 tcgtcgctga ttaaccaccg ttttaaccggc gaattcacta ctgatatcac gatggccgga 540
 accagccaga tgctggatat ccagcaacgc gatttcagtc cgcaaatttt acaagccacc 600
 ggtattccac gccgactcct ccctcgtctg gtggaagcgg gtgaacagat tggtagccta 660
 cagaacagcg ccgcagcaat gctcggctta cccgttgga taccggtgat ttccgcaggt 720
 cacgataccc agttcgccct ttttggcgct ggtgctgaac aaaatgaacc cgtgctctct 780
 tccggtacat gggaaatttt aatggttcgc agcgcaccag ttgatacttc gctgttaagt 840
 cagtacgccg gttccacctg cgaactggat agccaggcag gttgtataa cccaggtatg 900
 caatggctgg catccggcgt gctggaatgg gtgagaaaac tgttctggac ggctgaaaca 960
 ccctggcaaa tgttgattga agaagctcgt ctgatcgcgc ctggcgcgga tggcgtaaaa 1020
 atgcagtgtg atttattgtc gtgtcagaac gctggctggc aaggagtgc gcttaatacc 1080
 acgcgggggc atttctatcg cgcgcgctg gaagggttaa ctgcgcaatt acagcgaat 1140
 ctacagatgc tggaaaaaat cgggcacttt aaggcctctg aattattgtt agtcggtgga 1200
 ggaagtcgca acacattgtg gaatcagatt aaagccaata tgcttgatat tccggtaaaa 1260
 gttctcgacg acgccgaaac gaccgtcgca ggagctgcgc tgttcggttg gtatggcgta 1320
 ggggaattta acagcccgga agaagcccgc gcacagattc attatcagta ccgttatttc 1380
 taccgcaaa ctgaacctga atttatagag gaagtgtga 1419

<210> 6

<211> 1470

<212> DNA

<213> Escherichia coli

5

<400> 6

ES 2 731 782 T3

atgacctttc gcaattgtgt cgccgtcgat ctcggcgat ccagtggcg cgtgatgctg 60
 gcgcgttacg agcgtgaatg ccgcagcctg acgctgcgagc aaatccatcg ttttaacaat 120
 gggctgcata gtcagaacgg ctatgtcacc tgggatgtgg atagccttga aagtgccatt 180
 cgccttggat taaacaagggt gtgagaggaa gggattcgta tcgatagcat tgggatgcat 240
 acctggggcg tggactttgt gctgctcgac caacagggtc agcgtgtggg cctgcccgtt 300
 gcttatcgcg atagccgcac caatggccta atggcgcagg cacaacaaca actcggcaaa 360
 cgcgatattt atcaacgtag cggcatccag tttctgccct tcaatacgct ttatcagttg 420
 cgtgcgctga cggagcaaca acctgaactt attccacaca ttgctcacgc tctgctgatg 480
 ccggattact tcagttatcg cctgaccggc aagatgaact gggaatatac caacgccacg 540
 accacgcaac tggtaaatat caatagcgac gactgggacg agtcgctact ggcgtggagc 600
 ggggccaaca aagcctgggt tggtcgcccg acgcatccgg gtaatgtcat aggtcactgg 660
 atttgcccg caggtaatga gattccagtg gtcgcccgtt ccagccatga taccgccagc 720
 gcggttatcg cctcgccgtt aaacggctca cgtgctgctt atctctcttc tggcacctgg 780
 tcattgatgg gcttcgaaag ccagacgcca ttaccaatg acacggcact ggcagccaac 840
 atcaccaatg aaggcggggc ggaaggtcgc tatcgggtgc tgaaaaatat tatgggctta 900
 tggctgcttc agcagagtgc tcaggagcag caaatcaacg atcttccggc gcttatctcc 960
 gcgacacagg cacttccggc ttgccgcttc attatcaatc ccaatgacga tcgctttatt 1020
 aatcctgaga cgatgtgcag cgaaattcag gctgcgtgtc gggaaacggc gcaaccgatc 1080
 ccgaaaagtg atgctgaact ggcgcgctgc attttcgaca gtctggcgct gctgtatgcc 1140
 gatgtgttgc atgagctggc gcagctgcgc ggtgaagatt tctcgcaact gcatattgtc 1200
 ggcggaggct gccagaacac gctgctcaac cagctatgag ccgatgcctg cggatttcgg 1260
 gtgatcgccg ggcctgttga agcctcgacg ctcggaata tcggcatcca gttaatgacg 1320
 ctggatgaac tcaacaatgt ggatgatttc cgtcaggtcg tcagcaccac cgcgaatctg 1380
 accaccttta cccctaattc tgacagtga attgcccact atgtggcgca gattcactct 1440
 acacgacaga caaaggagct ttgcatga 1470

<210> 7
 <211> 1500
 <212> DNA
 <213> Escherichia coli

5

<400> 7

ES 2 731 782 T3

atgagcatct tgtacgaaga gcgtcttgat ggcgctttac ccgatgtcga ccgcacatcg 60
 gtactgatgg cactgcgtga gcatgtccct ggacttgaga tcctgcatac cgatgaggag 120
 atcattcctt acgagtgtga cgggttgagc gcgtatcgca cgcgtccatt actggttgtt 180
 ctgcctaagc aaatggaaca ggtgacagcg attctggctg tctgccatcg cctgcgtgta 240
 ccggtggtga cccgtggtgc aggcaccggg ctttctggtg gcgcgctgcc gctggaaaaa 300
 ggtgtgttgt tggatgatgg gcgctttaa gagatcctcg acattaacc cgttggtcgc 360
 cgcgcgcgcg tgcagccagg cgtgcgtaac ctggcgatct cccaggccgt tgcaccgcat 420
 aatctctact acgcaccgga cccttcctca caaatcgct gttccattgg cggcaatgtg 480
 gctgaaaatg ccggcggcgt ccaactgcctg aaatatggtc tgaccgtaca taacctgctg 540
 aaaattgaag tgcaaacgct ggacggcgag gcaactgacgc ttggatcggg cgcgctggat 600
 tcacctggtt ttgacctgct ggcgctgttc accggatcgg aaggtatgct cggcgtgacc 660
 accgaagtga cggtaaaact gctgccgaag ccgcccgctg cgcgggttct gttagccagc 720
 tttgactcgg tagaaaaagc cggacttgcg gttggtgaca tcatcgcaa tggcattatc 780
 cccggcgggc tggagatgat ggataacctg tcatccgcg cggcgggaaga ttttattcat 840
 gccggttata ccgtcgacgc cgaagcgatt ttgttatgcg agctggacgg cgtggagtct 900
 gacgtacagg aagactgcga gcgggttaac gacatcttgt tgaaagcggg cgcgactgac 960
 gtccgtctgg cacaggacga agcagagcgc gtacgtttct gggccggctc caaaaatgcg 1020
 ttcccggcgg taggacgtat ctccccgat tactactgca tggatggcac catcccgcgt 1080
 cgcgccctgc ctggcgtact ggaaggcatt gcccgttat cgcagcaata tgatttactg 1140
 gttgccaaag tctttcatgc cggagatggc aacatgcacc cgtaaatcct tttcgatgcc 1200
 aacgaaccgg gtgaatttgc ccgcgcggaa gagctgggcg ggaagatcct cgaactctgc 1260
 gttgaagtgt gcggcagcat cagtggcgaa catggcatcg ggcgagaaaa aatcaatcaa 1320
 atgtgcgcc agttcaacag cgatgaaatc acgaccttcc atgcggtcaa ggcggcgttt 1380
 gaccccgatg gtttgctgaa ccctgggaaa aacattccca cgctacaccg ctgtgctgaa 1440
 tttggtgcca tgcattgtga tcacggatcat ttaccttcc ctgaactgga gcgtttctga 1500

<210> 8
 <211> 861
 <212> DNA
 <213> Escherichia coli

5

<400> 8
 atgagcatta tctccactaa atatctgtta caggacgcc aggccaatgg ctacgcggtg 60
 cctgctttta acattcataa cgccgagacg atccaagcga tcctcgaagt gtgcagtgaa 120

ES 2 731 782 T3

atgcatcgcc cggtgatcct cgccggaacg ccggggacct ttaaaccacat cgcgctggaa 180
gagatctacg cctgtgttag cgcctattcc acaacctaca acatgccact ggcgctgcac 240
ctcgaccacc acgaatcgct ggatgatatt cgccgtaaag tccacgcagg tgtgcgcagt 300
gcatgatcg acggcagcca cttcccgttt gccgagaacg tgaagctggt gaaatcggtt 360
gttgacttct gccactcaca agattgcagc gtggaagcag aactgggccg cctgggcggt 420
gttgaagatg acatgagcgt tgacgccgaa agtgattcc tgaccgatcc acaagaagct 480
aaacgctttg tcgaactgac tggcgtcgac agcctggcgg tagcgattgg tacggcgcac 540
ggcttatata gcaaaacgcc gaagattgat ttccagcggc tggcggaaat tcgtgaagtg 600
gtggatgttc ctctggtgct gcatggtgcc agcagatgttc cggatgaatt tgtccgtcgc 660
actattgaac ttggcgtcac aaaagtgaac gttgccacag aattaataat agccttcgct 720
ggcgcggtta aagcctgggt tgcggaaaat ccgcagggta atgacacctg ttattatatg 780
cgcgtcgaa tggatcgcat gaaagaagt gtcagaaata aaattaatgt ctgtggttca 840
gcgaatcgaa tttcagcata a 861

<210> 9
<211> 1053
<212> DNA
<213> Escherichia coli

5

<400> 9
atgacagata ttgctgagtt gcttggcaaa gacgccgaca accttttaca gcaccgttgt 60
atgacaattc cttctgacca gctttatctc cccggacatg actacgtaga ccgcgtaatg 120
attgacaata atcggcccgc agcgggtgta cgtaatatgc agacgttcta caacaccggg 180
cgtctggctg gcacaggata tctttctatt ctgccggttg accagggcgt tgagcactct 240
gccggagctt catttgctgc taaccgcctc tactttgacc cgaaaaacat tgttgaactg 300
gcatcgaag cgggctgtaa ctgtgtgctg tcaacttacg gcgtgctggc gtcggtatcg 360
cggcgttatg cgcacatcctc tccattcctc gtcaaaacta atcacaacga gacgctaagt 420
taccgaata cctacgatca aacgctgtat gccagcgtgg agcagggcgt caacatgggc 480
gctgttgctg ttggtgctgc tatctatctt ggctcggaa agtcacgtcg ccagattgaa 540
gaaatttctg cggcttttga acgtgctcac gagctgggta tgggtgacagt gctgtgggcc 600
tatttgctga actccgcctt taagaaagat ggcgttgatt accatgtttc cgccgacctg 660
accggtcagg caaacctctc gggggcaacc atcgtgctcag atatcgtcaa acaaaaaatg 720
gctggaaaata acggcggcta taaagcaatt aattacggtt acaccgacga tcgtgtttac 780
agcaaatgta ccagcgaaaa ccgattgat ctggtgctgt atcagttagc taactgctat 840
atgggtcggg ctgggttgat aaactccggc ggtgctgctg gcggtgaaac tgacctcagc 900

ES 2 731 782 T3

gatgcagtgc gtactgCGgt tatcaacaaa cgCGcaggcg gaatggggct gattcttggg 960
 cgtaaagcgt tcaagaaatc gatggctgac ggcgtgaaac tgattaacgc cgtgcaggac 1020
 gtttatctcg atagcaaaat tactatcgcc tga 1053

5 <210> 10
 <211> 939
 <212> DNA
 <213> Escherichia coli

<400> 10
 atggatatca tcttttatca cccaacgttc gataccaat ggtggattga ggcactgCGc 60
 aaagctattc ctCaggcaag agtcagagca tggaaaagcg gagataatga ctctgctgat 120
 tatgctttag tctggcatcc tctgttgaa atgctggcag gCGcGatct taaagcggcg 180
 ttCGcactcg gggccgggtg tGattctatt ttgagcaagc tacaggcaca ccctgaaatg 240
 ctgaaccctt ctgttccact ttttCGcctg gaagataccg gtatgggCGa gcaaatgCag 300
 gaatatgctg tcagtcaggT gctgCattgg tttCGacgTt ttgacgatta tcgcatccag 360
 caaaatagTt cGcattggca accgctgCct gaatatcAtc gggaagattt taccatcggc 420
 attttggcg caggcgtact gggcagTaaa gttgctcaga gtctgcaaac ctggcGcttt 480
 ccgctgCGTt gctggagTcg aaccCGTaaa tCGTggcctg gCGTgcaag ctttCGcggg 540
 cgggaagaac tGctgCatt tctgagCca tGctCGgTat tgattaattt gttaccgaaT 600
 acccctgaaa cGtCGgcat tattaatCaa caattactcg aaaaattacc ggatggcCGc 660
 tatctcctca acctggcCGc Tggtgttcat gttgtggaag atgacctgct cGcggcGctg 720
 gatagcggca aagttaaagG cGcaatgttG gatgttttTa atcgtgaacc cttaccgCct 780
 gaaagtCGc tctggcaaca tccacGcTg acgataacac cacatgTcg cGcGattacc 840
 cgtccGctg aagctgtgga gtacatttct cGcaccattg cccagctCGa aaaaggggag 900
 agggTctgCG gGcaagTcga cCGcGcagCG gGctactaa 939

10 <210> 11
 <211> 1305
 <212> DNA
 <213> Escherichia coli

<400> 11
 atgaaaaccC gtacacaaca aattgaagaa ttacagaaag agtggactca accgCgttgg 60
 gaaggcatta ctCGccata cagTcgGaa gatgtgTga aattacGcgG ttCagtcaat 120
 cctgaaTgca cGctggCGca actgggCGca gCGaaaatgt gCGtctgct gCacggTgag 180
 tcgaaaaaag gctacatCaa cagcctCGc gCactgactg gCGTcagGc gctgcaacag 240
 gCGaaagCGg gtattgaagc agtctatctg tCGgagTgC aggtagCGc ggacGtaac 300
 ctggcGgcca gcatgtatcc ggatcagTcg ctctatCGg caaactCGgt gCcagctgTg 360

ES 2 731 782 T3

gtggagcggg tcaacaacac cttccgctcgt gccgatcaga tccaatggtc cgcgggcatt 420
gagccggggc atcccgcgcta tgtcgattac ttctcgccga tcggtgccga tgcggaagcc 480
ggttttggcg gtgtcctgaa tgcctttgaa ctgatgaaag cgatgattga agccggtgca 540
gcggcagttc acttcgaaga tcagctggcg tcagtgaaga aatgcgggtca catgggcggc 600
aaagttttag tgccaactca ggaagctatt cagaaactgg tcgcgggcgcg tctggcagct 660
gacgtgacgg gcgttccaac cctgctgggt gcccgtagcg atgctgatgc ggcgtagctg 720
atcacctccg attgcgaccc gtatgacagc gaatttatta ccggcgagcg taccagtga 780
ggcttcttcc gtactcatgc gggcattgag caagcgatca gccgtggcct ggcgtagctg 840
ccatagctg acctggctctg gtgtgaaacc tccacgcccg atctggaact ggcgctcgc 900
tttgacaag ctatccagc gaaatatccg ggcaaactgc tggcttataa ctgctcgcg 960
tcggtcaact ggcaaaaaa cctcgacgac aaaactattg ccagcttcca gcagcagctg 1020
tcggatatgg gctacaagtt ccagttcatc accctggcag gtatccacag catgtggttc 1080
aacatgtttg acctggcaaa cgcctatgcc cagggcgagg gtatgaagca ctacgttgag 1140
aaagtgcagc agccggaatt tgccgcccgc aaagatggct atacctcgt atctcaccag 1200
caggaagtgg gtacaggtta cttcgataaa gtgacgacta ttattcaggc cggcacgtct 1260
tcagtcaccg cgtgaccgg ctccactgaa gaatcgcagt tctaa 1305

<210> 12
<211> 323
<212> DNA
<213> Saccharomyces cerevisiae

5

<400> 12
tgcagatcaa aggcaaagac agaaaccgta gtaaaggttg acttttcaca acagtgtctc 60
cattttttat attgtattat taaagctatt tagttatttg gatactgttt tttttccaga 120
agttttcttt ttagtaaagt acaatccagt aaaaatgaag gatgaacaat cgggtgatgc 180
agattcaaca ccaataaatg caatgtttat ttctttggaa cgtttgtgtt gttcgaatc 240
caggataatc cttcaacaag accctgtccg gataaggcgt tactaccgat gacacaccaa 300
gaaccacaca tctatttcct tat 323

<210> 13
<211> 311
<212> DNA
<213> Saccharomyces cerevisiae

10

<400> 13
tattctgata gtagatcatc agatttgata tgatattatt tgtgaaaaaa tgaaataaaa 60
ctttatacaa cttaaataca acttttttta taacgatta agcaaaaaaa tagtttcaaa 120

ES 2 731 782 T3

cttttaacaa tattccaaac actcagtcct tttccttctt atattatagg tgtacgtatt 180
 atagaaaaat ttcaatgatt actttttctt tctttttcct tgtaccagca catggccgag 240
 cttgaatggt aaacccttcg agagaatcac accattcaag tataaagcca ataaagaata 300
 taaaaaatta t 311
 <210> 14
 <211> 514
 <212> DNA
 <213> Saccharomyces cerevisiae
 5
 <400> 14
 atgggtcatt acgtaaataa tgataggaat gggattcttc tatttttctt ttttccattc 60
 tagcagccgt cgggaaaacg tggcatcctc tctttcgggc tcaattggag tcacgctgcc 120
 gtgagcatcc tctctttcca tatctaacaa ctgagcacgt aaccaatgga aaagcatgag 180
 cttagcgttg ctccaaaaaa gtattggatg gttaatacca tttgtctggt ctcttctgac 240
 tttgactcct caaaaaaaaa aaatctacaa tcaacagatc gttcaatta cgcctcaca 300
 aaaacttttt tctttcttct tgcgccacgt taaattttat ccctcatggt gtctaacgga 360
 tttctgcact tgatttatta taaaagaca aagacataat acttctctat caatttcagt 420
 tattgttctt ccttgcggtt ttcttctggt cttctttttc ttttgtcata tataaccata 480
 accaagtaat acatattcaa aaagaaacat aaat 514
 <210> 15
 <211> 391
 <212> DNA
 <213> Saccharomyces cerevisiae
 10
 <400> 15
 tatttaaact gtgaggacct taatacattc agacacttct gcggtatcac cctacttatt 60
 cccttcgaga ttatatctag gaacccatca ggttggtgga agattaccg ttctaagact 120
 tttcagcttc ctctattgat gttacacctg gacaccctt ttctggcatc cagtttttaa 180
 tcttcagtgg catgtgagat tctccgaaat taattaaagc aatcacacaa ttctctcgga 240
 taccacctcg gttgaaactg acaggtggtt tgttacgcat gctaatgcaa aggagcctat 300
 atacctttgg ctcggtgct gtaacagga atataaaggg cagcataatt taggagtta 360
 gtgaacttgc aacatttact attttccctt c 391
 <210> 16
 <211> 23
 <212> DNA
 <213> Artificial Sequence
 15
 <220>
 <223> Amorce/Primer

<400> 16
 catatgacct ttcgcaattg tgt 23

5
 <210> 17
 <211> 26
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> Amorce/Primer

10
 <400> 17
 ggatcctcat gcgcaaagct cctttg 26

<210> 18
 <211> 52
 <212> DNA
 <213> Artificial Sequence

15
 <220>
 <223> AMORCE/PRIMER

<400> 18
 ccatggatgc accatcacca tcaccatatg ttatccggct atattgcagg ag 52

20
 <210> 19
 <211> 38
 <212> DNA
 <213> Artificial Sequence

<220>
 <223> AMORCE/PRIMER

25
 <400> 19
 ggatccgtcg acattaacgg cgaaattggt tcagcatt 38

<210> 20
 <211> 40
 <212> DNA
 30 <213> Artificial Sequence

<220>
 <223> AMORCE/PRIMER

<400> 20
 ttgatccag gaggattcat atggcacatc gcttccggc 40

35
 <210> 21
 <211> 28
 <212> DNA
 <213> Artificial Sequence

40
 <220>
 <223> AMORCE/PRIMER

<400> 21

ttaagctttt aatacgtgta acaggccg 28
 <210> 22
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 5
 <220>
 <223> AMORCE/PRIMER
 <400> 22
 catatgacag atattgcgca gttgcttg 29
 10
 <210> 23
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 15
 <220>
 <223> AMORCE/PRIMER
 <400> 23
 ggatcctcag gcgatagtaa tttgctat 29
 20
 <210> 24
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 25
 <220>
 <223> AMORCE/PRIMER
 <400> 24
 catatgagca ttatctccac taaatatct 29
 30
 <210> 25
 <211> 29
 <212> DNA
 <213> Artificial Sequence
 35
 <220>
 <223> AMORCE/PRIMER
 <400> 25
 ggatccttat gctgaaattc gattcgctg 29
 40
 <210> 26
 <211> 56
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> AMORCE/PRIMER
 <400> 26
 cctctagagt cgactgcag aggaggattc atatgaacaa cttaaatctg cacacc 56

ES 2 731 782 T3

<210> 27
<211> 34
<212> DNA
<213> Artificial Sequence

5 <220>
<223> AMORCE/PRIMER

<400> 27
gccaaaacag aagctttag cgggCGGctt cgta 34

10 <210> 28
<211> 42
<212> DNA
<213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

15 <400> 28
cgtaccCGG gGatcaggag gCacacgatg gaagaagcag at 42

<210> 29
<211> 25
<212> DNA
20 <213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

<400> 29
tcacacgatg ccatcaaagc cctgc 25

25 <210> 30
<211> 54
<212> DNA
<213> Artificial Sequence

30 <220>
<223> AMORCE / PRIMER

<400> 30
cctctagagt cgacctgCag agGaggattc atatgCaca tCGctttcCG gctc 54

35 <210> 31
<211> 36
<212> DNA
<213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

40 <400> 31
gccaaaacag aagctttaa tacgtgtaac aggccg 36

<210> 32

ES 2 731 782 T3

<211> 44
<212> DNA
<213> Artificial Sequence

5 <220>
<223> AMORCE / PRIMER

<400> 32
ttggtaccaa gaaggagaat tcatatgtca gtaccggtc aaca 44

10 <210> 33
<211> 28
<212> DNA
<213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

15 <400> 33
ttaagctttt aagactgtaa ataaacca 28

<210> 34
<211> 55
<212> DNA
<213> Artificial Sequence

20 <220>
<223> AMORCE / PRIMER

<400> 34
gagctcggta cccggggatc caggaggcac acgatggata tcatctttta tcacc 55

25 <210> 35
<211> 49
<212> DNA
<213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

30 <400> 35
ctcatccgcc aaaacagaag ctttagtag ccgctgctgcg ggtcgactt 49

<210> 36
<211> 52
<212> DNA
<213> Artificial Sequence

35 <220>
<223> AMORCE / PRIMER

<400> 36
gcacgcggct actaaaggag gaaccgtatg aaaaccgta cacaacaat tg 52

40 <210> 37
<211> 47

ES 2 731 782 T3

<212> DNA
<213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

5 <400> 42
tgtctttgcc ttgatctgc tcacacgatg ccatcaaagc cc 42

<210> 43
<211> 45
<212> DNA
10 <213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

<400> 43
gatggcatcg tgtgagcaga tcaaaggcaa agacagaaac cgtag 45

15 <210> 44
<211> 56
<212> DNA
<213> Artificial Sequence

<220>
20 <223> AMORCE / PRIMER

<400> 44
attcctatca ttattacgt aatgacccat ctggtgtgt catcggtagt aacgcc 56

<210> 45
<211> 49
25 <212> DNA
<213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

<400> 45
30 ccgatgacac accaagatgg gtcattacgt aaataatgat aggaatggg 49

<210> 46
<211> 55
<212> DNA
<213> Artificial Sequence

35 <220>
<223> AMORCE / PRIMER

<400> 46
tgctgcagag ccgaaagcg atgtgccatt tgaatatgt attacttggg tatgg 55

<210> 47
40 <211> 44
<212> DNA

ES 2 731 782 T3

<213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

<400> 47
5 ccaagtaata catattcaaa atggcacatc gcttccggc tctg 44

<210> 48
<211> 46
<212> DNA
<213> Artificial Sequence

10 <220>
<223> AMORCE / PRIMER

<400> 48
ctgatgatct acgatcagaa ttaatacgt gtaacaggcc gtaaac 46

15 <210> 49
<211> 45
<212> DNA
<213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

20 <400> 49
gttacacgta ttaaattctg atcgtagatc atcagattg atatg 45

<210> 50
<211> 60
<212> DNA
25 <213> Artificial Sequence

<220>
<223> AMORCE / PRIMER

<400> 50
cgcgcgtaat acgactcact atagggcgaa ttgatattc ttattggct ttacttga 60

30 <210> 51
<211> 1104
<212> DNA
<213> Escherichia coli

<400> 51

ES 2 731 782 T3

atggaccgca ttattcaatc accgggtaaa tacatccagg gcgctgatgt gattaatcgt 60
ctgggcgaat acctgaagcc gctggcagaa cgctggtag tggtaggtga caaatttgtt 120
ttaggttttg ctcaatccac tgtcgagaaa agctttaaag atgctggact ggtagtagaa 180
attgcgccgt ttggcgggtga atgttcgcaa aatgagatcg accgtctgcg tggcatcgcg 240
gagactgcbc agtgtggcgc aattctcggc atcggtaggc gaaaaaccct cgatactgcc 300
aaagcactgg cacatttcat ggggtgtccg gtagcgatcg caccgactat cgcctctacc 360
gatgcaccgt gcagcgcatt gtctgttacc tacaccgatg agggtaggtt tgaccgctat 420
ctgctgttgc caaataaacc gaatatggc attgtcgaca ccaaaatcgt cgctggcgca 480
cctgcacgct tgttagcggc gggtagcggc gatgcgctgg caacctggt tgaagcgcgt 540
gcctgctctc gtagcggcgc gaccaccatg gcggcggca agtgcacca ggctgcgctg 600
gcactggctg aactgtgcta caacaccctg ctggaagaag gcgaaaaagc gatgcttgcct 660
gccgaacagc atgtagtgc tccggcgctg gagcgcgtga ttgaagcga cacctatttg 720
agcgggtgtg gttttgaaag tggtagctg gctgcggcgc acgcagtgc taacggcctg 780
accgctatcc cggacgcgca tcaactattat cacggtgaaa aagtggcatt cggtagcctg 840
acgcagctgg ttctggaaaa tgcgccggtg gaggaatcg aaaccgtagc tgccttagc 900
catgcggtag gtttgccaat aactctcgt caactggata ttaaagaaga tgtcccggcg 960
aaaatgcgaa ttgtggcaga agcggcatgt gcagaagggtg aaaccattca caacatgcct 1020
ggcggcgcga cggcagatca ggtttacgcc gctctgctgg tagccgacca gtacggtcag 1080
cgtttctcgc aagagtggga ataa 1104

<210> 52
<211> 1149
<212> DNA
<213> Escherichia coli
<400> 52

5

ES 2 731 782 T3

```

atggctaaca gaatgattct gaacgaaacg gcatggtttg gtcgggggtgc tgttggggct      60
ttaaccgatg agtgaaaacg ccgtggttat cagaaggcgc tgatcgtcac cgataaaacg      120
ctggtgcaat gcggcgtggt ggcgaaaagt accgataaga tggatgctgc agggctggca      180
tgggcgattt acgacggcgt agtgcccaac ccaacaatta ctgtcgtcaa agaagggtc      240
gggtgattcc agaatagcgg cgcggttac ctgatcgcta ttggtggtg ttctccacag      300
gatacttgta aagcgattgg cattatcagc aacaaccgag agtttgccga tgtgcgtagc      360
ctggaagggc tttccccgac caataaacc agtgtaccga ttctggcaat tcctaccaca      420
gcaggtactg cggcagaagt gaccattaac tacgtgatca ctgacgaaga gaaacggcgc      480
aagtttgttt gcgttgatcc gcatgatatc ccgcagggtg cgtttattga cgctgacatg      540
atggatggta tgcctccagc gctgaaagct gcgacgggtg tcgatgcgct cactcatgct      600
attgaggggt atattaccg tggcgcgtgg gcgctaaccg atgcactgca cattaagcg      660
attgaaatca ttgctggggc gctgcgagga tcggttgctg gtgataagga tgccggagaa      720
gaaatggcgc tcgggcagta tgttgccgggt atgggcttct cgaatggtg gttagggtt      780
gtgcatggta tggcgcaccc actgggcgcg ttttataaca ctccacacgg tgttgccaac      840
gccatcctgt taccgcatgt catgcgttat aacgctgact ttaccggtga gaagtaccgc      900
gatatcgcgc gcgttatggg cgtgaaagtg gaaggtatga gcctggaaga ggcgcgtaat      960
gccgctgttg aagcgggtgt tgctctcaac cgtgatgctg gtattccgcc acatttgcgt     1020
gatgttggtg tacgcaagga agacattccg gcaactggcgc aggcggcact ggatgatgtt     1080
tgtaccggtg gcaaccgctg tgaagcaacg cttgaggata ttgtagagct ttaccatacc     1140
gcctggtaa                                     1149

```

<210> 53

<211> 1455

<212> DNA

<213> Escherichia coli

5

<400> 53

ES 2 731 782 T3

atgtatatcg ggatagatct tggcacctcg ggcgtaaaag ttatTTTgct caacgagcag 60
 ggtgaggtgg ttgctgCGca aacggaaaag ctgaccgTTT cgcgcccGca tccactctgg 120
 tcggaacaag acccggaaCa gtggtggcag gcaactgatc gcgcaatgaa agctctgggc 180
 gatcagcatt ctctgcagga cgTtaaagca ttggtattg ccggccagat gcacggagca 240
 accttGctgg atgctcagca acgggtgTta cgccctgcca tTTTgtggaa cgacgggCGc 300
 tgtgcgcaag agTgcactTT gctggaagcg cgagttccgc aatcgCGgt gattaccggc 360
 aacctgatga tgcccggatt tactgCGcct aaattgctat gggttcagcg gcatgagccg 420
 gagatattcc gtcaaatcga caaagtatta ttaccgaaag attacttgCG tctgcgtatg 480
 acgggggagT ttgccagcga tatgtctgac gcagctggca ccatgtggct ggatgtCGca 540
 aagcgtgact ggagtgacgt catgctgcag gcttgcgact tatctCGtga ccagatgccc 600
 gcattatacg aagcagcga aattactggt gcttTgttac ctgaagttgc gaaagcgtgg 660
 ggtatggcga cggtgccagT tgtcgCagc ggtggcGaca atgcagctgg tgCagttggt 720
 gtgggaatgg ttgatgctaa tcaggcaatg ttatcGctgg ggacgtcggg ggtctatTTT 780
 gctgtcagcg aagggttctt aagcaagcca gaaagcCGc tacatagctt ttgccatgcg 840
 ctaccgcaac gttggcattt aatgtctgtg atgctgagtg cagcgtcgtg tctggattgg 900
 gccgcgaaat taaccggcct gagcaatgtc ccagctTtaa tcgctgcagc tcaacaggct 960
 gatgaaagtG ccgagccagT ttggttctg ccttatcttT ccggcgagcg tacgccacac 1020
 aataatcccc aggcgaaggg ggttttcttT ggtttgactc atcaacatgg ccccaatgaa 1080
 ctggcgcgag cagtgctgga aggcgtgggt tatgcgctgg cagatggcat ggatgtcgtg 1140
 catgcctgcg gtattaaacc gcaaagtgtt acgttgattg gggcggggc gcgtagtgag 1200
 tactggcgtc agatgctggc ggatatcagc ggtcagcagc tcgattaccg tacggggggg 1260
 gatgtggggc cagcactggg cgcagcaagg ctggcgcaga tcgCGcgaa tccagagaaa 1320
 tcgctcattg aattgttgcc gcaactaccg ttagaacagT cgcatctacc agatgcgag 1380
 cgttatgccg cttatcagcc acgacgagaa acgttccgtc gcctctatca gcaacttctg 1440
 ccattaatgg cgtaa 1455

<210> 54

<211> 45

<212> DNA

<213> Artificial Sequence

5

<220>

<223> amorce/primer

<400> 54

cctctagagt cgacctgagc agggaggttc atatggaccg catta 45

10

<210> 55

ES 2 731 782 T3

<211> 35
<212> DNA
<213> artificial sequence

5 <220>
<223> amorce/primer

<400> 55
gccaaaacag aagctttat tccactct gcagg 35

10 <210> 56
<211> 40
<212> DNA
<213> artificial sequence

<220>
<223> amorce/primer

15 <400> 56
ttggatccag gaggattcat atggctaaca gaatgattct 40

<210> 57
<211> 27
<212> DNA
<213> artificial sequence

20 <220>
<223> amorce/primer

<400> 57
ttaagctttt accaggcggg atggtaa 27

25 <210> 58
<211> 52
<212> DNA
<213> artificial sequence

<400> 58
gatggcatcg tgtgaaggag gaaccgatg acagatattg cgcagttgct tg 52

30 <210> 59
<211> 44
<212> DNA
<213> artificial sequence

35 <220>
<223> amorce/primer

<400> 59
atgcctgcag gtcgacaatt gtcaggcgat agtaatttg ctat 44

40 <210> 60
<211> 68
<212> DNA
<213> artificial sequence

ES 2 731 782 T3

<220>
<223> amorce/primer

<400> 60
cggtagccgg ggcgagaaa tactataaac ggttaataa acaaaccata atggaagaga 60

5 agcagatc 68

<210> 61
<211> 70
<212> DNA
<213> artificial sequence

10 <220>
<223> amorce/primer

<400> 61
gatggcatcg tgtgaacagc ttttaattat acactttaag gaggacagac atgacagata 60
ttgctgagtt 70

15 <210> 62
<211> 35
<212> DNA
<213> artificial sequence

<220>
<223> amorce/primer

20 <400> 62
atgcctgcag gtcgagcggat agtaatttg ctatc 35

<210> 63
<211> 46
<212> DNA
25 <213> artificial sequence

<220>
<223> amorce/primer

<400> 63
cggtagccgg ggcgagcgg ggcgagcgg ggcgagcgg ggcgagcgg 46

30 <210> 64
<211> 39
<212> DNA
<213> artificial sequence

<220>
<223> amorce/primer

35 <400> 64
tcgcaaaaa cagaagcttt caggcgatag taatttgc 39

<210> 65
<211> 47

ES 2 731 782 T3

<212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 5 <400> 65
 ctcatccgcc aaaacagaag cttttagaac tgcgattctt cagtgga 47

 <210> 66
 <211> 49
 <212> DNA
 10 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 66
 gaagaatcgc agttctaaag gaggcacacg atggatatca tctttatc 49

 15 <210> 67
 <211> 53
 <212> DNA
 <213> artificial sequence

 <220>
 20 <223> amorce/primer

 <400> 67
 catccgcaa aacagaagct tctagatta gtagccgctg gcgcggtcga ctt 53

 <210> 68
 <211> 53
 25 <212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 68
 30 catccgcaa aacagaagct tctagatta gtagccgctg gcgcggtcga ctt 53

 <210> 69
 <211> 102
 <212> DNA
 <213> artificial sequence

 35 <220>
 <223> amorce/primer

 <400> 69
 gCGTcttgat ggcgctttac ccgatgtcga ccgcacatcg gtactgatgg cactgcgtga 60

 gcatgtccct ggacttgaga tcgtgtaggc tggagctgct tc 102

 <210> 70

ES 2 731 782 T3

<211> 99
 <212> DNA
 <213> artificial sequence

 <220>
 5 <223> amorce/primer

 <400> 70
 cgcgtaaacg ccaggcgtgt aataacggtt cggatatagcc gtttggctgt ttcacgccga 60
 ggaagattaa atcgtggcc atatgaatat cctccttag 99

 <210> 71
 10 <211> 102
 <212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 71
 cgcgcgagac tcgctctgct tatctcgccc ggatagaaca agcgaaaact tcgaccgttc 60
 15 atcgttcgca gttggcatgc ggggtgtaggc tggagctgct tc 102

 <210> 72
 <211> 100
 <212> DNA
 <213> artificial sequence

 20 <220>
 <223> amorce/primer

 <400> 72
 gcttagcgcc ttctacagct tcacgcgcca gcttagtaat gcggtcgtaa tcgcccgtt 60
 ccagcgcatac tgccggaacc catatgaata tcctccttag 100

 <210> 73
 25 <211> 20
 <212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 30 <400> 73
 cacagctaac accacgtcgt 20

 <210> 74
 <211> 82
 <212> DNA
 35 <213> artificial sequence

 <220>
 <223> amorce/primer

ES 2 731 782 T3

<400> 74
 acgtcattgc cttgtttgac cgcccctggt ttttagcgtc aggcataataa tacctcctaa 60
 agttaaacia aattatttgt ag 82

 <210> 75
 <211> 70
 <212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 75
 ccgcccgcac aataacatca ttcttcctga tcacgtttca ccgcagatta gtgtaggctg 60
 gagctgcttc 70

 <210> 76
 <211> 47
 <212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 76
 gatagggacg acgtggtgt agctgtgcat atgaatatcc tccttag 47

 <210> 77
 <211> 47
 <212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 77
 ggatccgtcg acaggaggta acatatggaa gagaagcaga tcctgtg 47

 <210> 78
 <211> 35
 <212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 78
 ctgcagtcta gaagatctca cacgatgcca tcaaa 35

 <210> 79
 <211> 44
 <212> DNA
 <213> artificial sequence

 <220>

ES 2 731 782 T3

<223> amorce/primer

<400> 79
 ttgatccaa gaaggagaat tcatatggca catcgcttc cggc 44

5 <210> 80
 <211> 21
 <212> DNA
 <213> artificial sequence

<220>
 <223> amorce/primer

10 <400> 80
 ttaatacgtg taacaggccg t 21

<210> 81
 <211> 101
 <212> DNA
 15 <213> artificial sequence

<220>
 <223> amorce/primer

<400> 81
 cgcaccatt cccgcgaaac gcggcagaaa acccgccggt gccaccgcac cagcgactgg 60
 acaggttcag tctttaacgc gtgtaggctg gagctgcttc g 101

20 <210> 82
 <211> 100
 <212> DNA
 <213> artificial sequence

<220>
 25 <223> amorce/primer

<400> 82
 gcgcattcca cgtacgcca gcgtcacttc cttcgcgct ttaatcacca tcgcgcaaaa 60
 ctcggtcacg cgtcatcgg catatgaata tcctccttag 100

<210> 83
 <211> 63
 <212> DNA
 30 <213> artificial sequence

<220>
 <223> amorce/primer

<400> 83
 gaattgctgg cccgctgcgg cataaaaatt aattggtacc gttatttaa actgtgagga 60

35 cct 63

<210> 84
 <211> 46

ES 2 731 782 T3

<212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 5 <400> 84
 gattgcttt tctccattt ttagttatg tatgtgttt tttgta 46

 <210> 85
 <211> 40
 <212> DNA
 10 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 85
 cacatacata aactaaaaat ggaagaaaag caaatcttgt 40

 15 <210> 86
 <211> 45
 <212> DNA
 <213> artificial sequence

 <220>
 20 <223> amorce/primer

 <400> 86
 gtttctgtct ttgccttga tctgcttaa cgataccgtc gaaac 45

 <210> 87
 <211> 42
 25 <212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 87
 30 ggtatcgttt aagcagatca aaggcaaaga cagaaaccgt ag 42

 <210> 88
 <211> 56
 <212> DNA
 <213> artificial sequence

 <220>
 <223> amorce/primer

 <400> 88
 attcctatca ttatttacgt aatgacccat ctgggtgtgt catcggtagt aacgcc 56

 <210> 89
 <211> 49
 40 <212> DNA

<213> artificial sequence
 <220>
 <223> amorce/primer
 <400> 89
 5 cccgatgacac accaagatgg gtcattacgt aaataatgat aggaatggg 49
 <210> 90
 <211> 45
 <212> DNA
 <213> artificial sequence
 <220>
 <223> amorce/primer
 <400> 90
 10 gctgggaatc tgtgagccat ttgaatatg tattactgg ttatg 45
 <210> 91
 <211> 42
 <212> DNA
 <213> artificial sequence
 <220>
 <223> amorce/primer
 <400> 91
 15 tgatctacga tcagaattta gtaagtgtaa caagcagtga ac 42
 <210> 92
 <211> 42
 <212> DNA
 <213> artificial sequence
 <220>
 <223> amorce/primer
 <400> 92
 20 gtaatacata ttcaaatgg ctacagatt cccagctttg ac 42
 <210> 93
 <211> 49
 <212> DNA
 <213> artificial sequence
 <220>
 <223> amorce/primer
 <400> 93
 25 gcttggtaca ctactaaat tctgacgta gatcatcaga ttgatatg 49
 <210> 94
 <211> 66
 <212> DNA
 <213> artificial sequence
 <220>
 <223> amorce/primer
 <400> 93
 30 gcttggtaca ctactaaat tctgacgta gatcatcaga ttgatatg 49
 <210> 94
 <211> 66
 <212> DNA
 <213> artificial sequence
 <220>
 <223> amorce/primer
 <400> 93
 35 gcttggtaca ctactaaat tctgacgta gatcatcaga ttgatatg 49
 <210> 94
 <211> 66
 <212> DNA
 <213> artificial sequence
 <220>
 <223> amorce/primer
 <400> 93
 40 gcttggtaca ctactaaat tctgacgta gatcatcaga ttgatatg 49
 <210> 94
 <211> 66
 <212> DNA
 <213> artificial sequence

ES 2 731 782 T3

<220>
<223> amorce/primer

<400> 94
ggtgacccgg cggggacgag gcaagctaaa cagatctcta gaataattt ttatttctt 60

5 tattgg 66

<210> 95
<211> 25
<212> DNA
<213> artificial sequence

10 <220>
<223> amorce/primer

<400> 95
gagatttgt tcttctgagc ttctg 25

15 <210> 96
<211> 25
<212> DNA
<213> artificial sequence

<220>
<223> amorce/primer

20 <400> 96
gaaattgcac ttagcatttt tgaat 25

<210> 97
<211> 897
<212> DNA
25 <213> *Saccharomyces cerevisiae*

<400> 97

ES 2 731 782 T3

atggaagaaa agcaaatcct gtgtggttgg ttggttgttt tggacgttat ctctttggtt 60
 gacaagtacc caaaggaaga ctctgaaatc agatgtttgt ctcaaagatg gcaaagaggt 120
 ggtaacgctt ctaactcttg tactgttttg tctttgttgg gtgctccatg tgctttcatg 180
 ggttctatgg ctccagggtca cgttgctgac ttcttggttg ctgacttcag aaggagaggt 240
 gttgacgttt ctcaagttgc ttggcaatct aagggtgaca ctccatcttc ttgttgatc 300
 atcaacaact ctaacggtaa cagaactatc gttttgcacg acacttcttt gccagacgtt 360
 tctgctactg acttcgaaaa gtttgacttg actcaattca agtggatcca catcgaaggt 420
 agaaacgctt ctgaacaagt taagatgttg caaagaatcg acgctcacia cactagacia 480
 ccaccagaac aaaagatcag agtttctggt gaagttgaaa agccaagaga agaattgttc 540
 caattgttcg gttacgggtga cgttgttttc gtttctaagg acgttgctaa gcacttgggt 600
 ttccaatctg ctgaagaagc tttgagaggt ttgtacggta gagttagaaa ggggtgctgtt 660
 ttggtttgtg cttgggctga agaagggtct gacgctttgg gtccagacgg taagttgttg 720
 cactctgacg ctttcccacc gccaaagatt gttgacactt tgggtgctgg tgacactttc 780
 aacgctctcg ttatcttctc tttgtctcaa ggtagatctg ttcaagaagc tttgagattc 840
 ggttgtcaag ttgctggtaa gaagtgtggt ttgcaagggt tcgacggtat cgtttaa 897

5

<210> 98
 <211> 1095
 <212> DNA
 <213> *saccharomyces cerevisiae*

<400> 98

ES 2 731 782 T3

atggctcaca gattcccagc tttgactcaa gaacaaaaga aggaattgtc tgaaatcgct 60
 caatctatcg ttgctaacgg taagggtatc ttggctgctg acgaatctgt tggactatg 120
 ggtaacagat tgcaaagaat caaggttgaa aacctgaag aaaacagaag acaattcaga 180
 gaaatcttgt tctctgttga ctcttctatc aaccaatcta tcggtggtgt tatcttgttc 240
 cacgaaactt tgtacaaaa ggactctcaa ggtaagttgt tcagaaacat cttgaaggaa 300
 aagggtatcg ttgttggat caagttggac caagtggtg ctccattggc tggactaac 360
 aaggaaacta ctatccaagg tttggacggt ttgtctgaaa gatgtgctca atacaagaag 420
 gacggtgttg acttcggtaa gtggagagct gttttgagaa tcgctgacca atgtccatct 480
 tctttggcta tccaagaaaa cgctaacgct ttggctagat acgcttctat ctgtcaacia 540
 aacggtttgg ttccaatcgt tgaaccagaa gttatcccag acggtgacca cgacttgaa 600
 cactgtcaat acgttactga aaaggttttg gctgctgttt acaaggcttt gaacgaccac 660
 cacgtttact tggaaaggtac tttgttgaag ccaaactatg ttactgctgg tcacgcttgt 720
 actaagaagt aactccaga acaagttgct atggctactg ttactgcttt gcacagaact 780
 gttccagctg ctgttccagg tatctgtttc ttgtctggtg gtatgtctga agaagacgct 840
 actttgaact tgaacgctat caacttgtgt ccattgcaa agccatggaa gttgtctttc 900
 tcttacggta gagctttgca agcttctgct ttggctgctt ggggtggtaa ggctgctaac 960
 aaggaagcta ctcaagaagc tttcatgaag agagctatgg ctaactgtca agctgctaag 1020
 ggtcaatacg ttcacactgg ttcttctggt gctgcttcta ctcaatcttt gttcactgct 1080
 tgttacactt actaa 1095

<210> 99

<211> 10023

<212> DNA

<213> *Saccharomyces cerevisiae*

5

<400> 99

ggtaccttta cgatagagcc agagacgatg cgcacgaaga tgaagaagat cccgacacia 60
 gaagctccgg taagaagaag gacgccagcc aagaagaatc tctaacttaa gaggacggtt 120
 gctgaagaaa aaggcttttt ttattttgtc cgtttttttt ttgtaaaacc caaagatctg 180
 aatctaaagc ttttttaaac gtatatagat gtctacatgt gtgtttttgt ttttttacgt 240
 acgtataccc acctatatat gcataatccg taattgaaaa aaaaaaaga aaaagatcaa 300

ES 2 731 782 T3

ggaacacatc accctgggca catcaagcgt gaggaatgcc gtccaactgg tggagacgct 360
 tgatttgctc tttttgttct gggccaacc cggctctgaa gaacatcagc accacgcccg 420
 caacgacaaa gaacattgca atacacttgc atatgtgagc atagtcgagc ggtccgttct 480
 gtggttgatg ctggtgttct ttcttctggt tgtcaggggt gatagccata tcttcgtgct 540
 cttgttgoga ttgttctggt ccatctgcac cagaacaaaag aacaaaagaa caaggaacaa 600
 agtccaagca cgtcagcgcct gtttataagg ggattgacga gggatcgggc ctagagtgcc 660
 agcgcgccag ggagagggag cccctgggc cctcatccgc aggctgatag gggtcacccc 720
 gctgggcagg tcagggcagg ggctctcagg ggggcgccat ggacaaaactg cactgaggtt 780
 ctaagacaca tgtattattg tgagtatgta tatatagaga gagattaagg cgtacagccg 840
 ggtggtagag attgattaac ttggtagtct tatcttgca attgagtttc tgtcagtttc 900
 ttcttgaaca agcacgcagc taagtaagca acaaagcagg ctaacaaaact aggtactcac 960
 atacagactt attaagacat agaactatga ctacggataa cgctaaggcg caactgacct 1020
 cgtcttcagg gggtaacatt attgtggtgt ccaacaggct tcccgtgaca atcactaaaa 1080
 acagcagtac gggacagtac gagtacgcaa tgtcgtccgg agggctggtc acggcgttgg 1140
 aagggttgaa gaagacgtac actttcaagt ggctcggatg gcctgggcta gagattcctg 1200
 acgatgagaa ggatcagggtg aggaaggact tgctggaaaa gtttaatgcc gtaccatct 1260
 tcctgagcga tgaaatcgca gacttacct acaacgggtt cagtaattct attctatggc 1320
 cgttattcca ttaccatcct ggtgagatca atttcgacga gaatgcgtgg ttggcataca 1380
 acgaggcaaa ccagacgttc accaacgaga ttgctaagac tatgaacat aacgatttaa 1440
 tctgggtgca tgattacat ttgatggtg ttccgaaat gttgagagtc aagattcacg 1500
 agaagcaact gcaaacgtt aaggtcgggt ggctcctgca cacaccattc ccttcgagtg 1560
 aaatttacag aatcttacct gtcagacaag agattttgaa ggtgttttg agttgtgatt 1620
 tagtcgggtt ccacacatac gattatgcaa gacatttctt gtcttcctg caaagagtgc 1680
 ttaacgtgaa cacattgcct aatgggtggt aataccagg cagattcgtt aacgtagggg 1740
 ccttccctat cggtatcgac gtggacaagt tcaccgatgg gttgaaaaag gaatccgtac 1800
 aaaagagaat ccaacaattg aaggaaactt tcaagggtg caagatcata gttggtgctg 1860
 acaggtgga ttacatcaaa ggtgtgcctc agaagttgca cgccatggaa gtgtttctga 1920
 acgagcatcc agaatggagg ggcaagggtt ttctggtaca ggttgacgtg ccaagtctg 1980
 gagatgtgga agagtaccaa tatttaagat ctgtggtcaa tgagttggtc ggtagaatca 2040
 acggtcagtt cggfactgtg gaattcgtcc ccatccattt catgcacaag tctataccat 2100
 ttgaagagct gatttcgtta tatgctgtga gcgatgttg tttggtctcg tccaccctg 2160

ES 2 731 782 T3

atggtatgaa cttggtttcc tacgaatata ttgcttgcca agaagaaaag aaaggttcct 2220
taatcctgag tgagttcaca ggtgcccac aatccttgaa tgggtgctatt attgtaaadc 2280
cttggaaacac cgatgatcct tctgatgcca tcaacgaggc cttgactttg cccgatgtaa 2340
agaaaagaagt taactgggaa aaactttaca aatacatctc taaatacact tctgccttct 2400
ggggtgaaaa tttcgtccat gaattataca gtacatcatc aagctcaaca agtcctctg 2460
ccacaaaaaa ctgatgaacc cgatgcaaat gagacgatcg tctattcctg gtcgggtttt 2520
ctctgccctc tcttctattc acttttttta tactttatat aaaattatat aaatgacata 2580
actgaaacgc cacacgtcct ctcctattcg ttaacgcctg tctgtagcgc tgttactgaa 2640
gctgcgcaag tagttttttc accgtatagg atgacaacc ttaatataac ttcgtataat 2700
gtatgctata cgaagttatt aggtctagag atctgtttag cttgcctcgt ccccgccggg 2760
tcacccggcc agcgacatgg aggccagaa taccctcctt gacagtcttg acgtgcgag 2820
ctcaggggca tgatgtgact gtcgcccga catttagccc atacatcccc atgtataatc 2880
atgtgatcc atacattttg atggccgcac ggcgcgaagc aaaaattacg gctcctcgt 2940
gcagacctgc gagcagggaa acgctcccct cacagacgcg ttgaattgtc cccacgccgc 3000
gccctgtag agaaatataa aaggtagga tttgccactg aggttcttct ttcataact 3060
tccttttaaa atcttgctag gatacagttc tcacatcaca tccgaacata aacaacctg 3120
ggtaaggaaa agactcacgt ttcgaggccg cgattaaatt ccaacatgga tgctgattta 3180
tatgggtata aatgggctcg cgataatgtc gggcaatcag gtgcgacaat ctatcgattg 3240
tatgggaagc ccgatgccc agagttggtt ctgaaacatg gcaaaggtag cgttgccaat 3300
gatgttacag atgagatggc cagactaaac tggctgacgg aatztatgcc tcttccgacc 3360
atcaagcatt ttatccgtac tcctgatgat gcatggttac tcaccactgc gatccccggc 3420
aaaacagcat tccaggtatt agaagaatat cctgattcag gtgaaaatat tgttgatgcg 3480
ctggcagtgt tcctgcgccg gttgcattcg attcctgttt gtaattgtcc ttttaacagc 3540
gatcgcgtat ttcgtctcgc tcaggcgcaa tcacgaatga ataacggttt ggttgatgcg 3600
agtgatthtg atgacgagcg taatggctgg cctggtgaac aagtctggaa agaaatgcat 3660
aagctthtgc cattctcacc ggattcagtc gtcactcatg gtgatttctc acttgataac 3720
cttathhthg acgaggggaa attaataggt tgtattgatg ttggacgagt cggaaatcgca 3780
gaccgatacc aggatcttgc catcctatgg aactgcctcg gtgagthhth tccttcatta 3840
cagaaacggc thhthcaaaa atatggtatt gataatcctg atatgaataa atgacagthh 3900
cathhthgct tcgatgagth tthctaatca gtactgacaa taaaagatt cthhthhthca 3960
agaacththc atthhthtag thhthhthata thhthhthgt thhthhthaa thhthhthta 4020
gchhthhthta thhthhthhth chhthhthgaca thhthhthccc agathhthgag thhthhthhth 4080

ES 2 731 782 T3

agaaagtaat atcatgcgtc aatcgtatgt gaatgctggt cgctatactg ctgtcgaattc 4140
 gatactaacg ccgccatcca gtgtcgaaaa cgagctctcg agaaccctta atataacttc 4200
 gtataatgta tgctatacga agttattagg tgatatacga tccggatcct cagcagcgaa 4260
 acgttgttct gggaaacccat ggagaaactg aaagcgcttg gcgccagctc gattctggta 4320
 ctgccgatcg agaagatgat ggagtgagca tgcaagcttg gcgtaatcat ggtcatagct 4380
 gtttcctgtg tgaattgtt atccgctcac aattccacac aacatacgag ccggaagcat 4440
 aaagtgtaaa gcctggggtg cctaatagagt gagctaactc acattaattg cgttgcgctc 4500
 actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg 4560
 cgcggggaga ggcggtttgc gtattgggcg ctcttccgct tcctcgcctca ctgactcgct 4620
 gcgctcggtc gttcggctgc ggcgagcgg atcagctcac tcaaaggcgg taatacggtt 4680
 atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 4740
 caggaaccgt aaaaggccg cgttgctggc gtttttccat aggctccgcc cccctgacga 4800
 gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 4860
 ccaggcgttt cccoctgga gctccctcgt gcgctctcct gttccgacct tgccgcttac 4920
 cggatacctg tccgccttcc tcccttggg aagcgtggcg ctttctcata gctcacgctg 4980
 taggtatctc agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 5040
 cgttcagccc gaccgctcgc cttatccgg taactatcgt cttgagtcca acccgtaag 5100
 acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 5160
 aggcggtgct acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt 5220
 atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 5280
 atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 5340
 gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca 5400
 gtggaacgaa aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 5460
 ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 5520
 ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 5580
 tcgttcatcc atagttgcct gactccccgt cgtgtagata actacgatac gggagggctt 5640
 accatctggc cccagtgtcg caatgatacc gcgagacca cgctcaccgg ctccagattt 5700
 atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc 5760
 cgcctccatc cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa 5820
 tagtttgccg aacgttggtg ccattgctac aggcacgtg gtgtcacgct cgtcgtttgg 5880
 tatggcttca ttcagctccg gttcccaacg atcaaggcga gttacatgat ccccatggtt 5940

ES 2 731 782 T3

gtgcaaaaaa gcggttagct ccttcggtcc tccgatcggt gtcagaagta agttggccgc 6000
 agtgttatca ctcatggta tggcagcact gcataattct cttactgtca tgccatccgt 6060
 aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcy 6120
 gcgaccgagt tgctcttgcc cggcgtaaat acgggataat accgcgccac atagcagaac 6180
 tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 6240
 gctgttgaga tccagttcga tgtaaccocac tctgtcacc aactgatctt cagcatcttt 6300
 tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 6360
 aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 6420
 catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 6480
 acaaataggg gttccgcgca ctttccccc aaaagtgcc aactgacgtct aagaaacat 6540
 tattatcatg acattaacct ataaaaatag gcgtatcacg aggccctttc gtcttcaaga 6600
 attagctttt caattcaatt catcattttt tttttattct tttttttgat ttcggtttct 6660
 ttgaaatfff tttgattcgg taatctccga acagaaggaa gaacgaagga aggagcacag 6720
 acttagattg gtatatatac gcatatgtag tgttgaagaa acatgaaatt gccagatt 6780
 cttaacccaa ctgcacagaa caaaaacatg caggaaacga agataaatca tgtcgaaagc 6840
 tacatataag gaacgtgctg ctactcatcc tagtctgtt gctgccaagc tatttaatat 6900
 catgcacgaa aagcaaaaca acttgtgtgc ttcattggat gttcgtacca ccaaggaatt 6960
 actggagtta gttgaagcat taggtcccaa aatttgttta ctaaaaacac atgtggatat 7020
 cttgactgat ttttccatgg agggcacagt taagccgcta aaggcattat ccgccaagta 7080
 caatttttta ctcttogaag acagaaaatt tgctgacatt ggtaatacag tcaaattgca 7140
 gtactctgcy ggtgtataca gaatagcaga atgggcagac attacgaatg cacacggtgt 7200
 ggtgggcccc ggtattgtta cgggtttgaa gcaggcggca gaagaagtaa caaaggaacc 7260
 tagaggcctt ttgatgttag cagaattgtc atgcaagggc tccctatcta ctggagaata 7320
 tactaagggt actgttgaca ttgcgaagag cgacaaagat tttgttatcg gctttattgc 7380
 tcaaagagac atgggtggaa gagatgaagg ttacgattgg ttgattatga caccgggtgt 7440
 gggtttagat gacaaggag acgcattggg tcaacagtat agaaccgtgg atgatgtggt 7500
 ctctacagga tctgacatta ttattgttg aagaggacta tttgcaaagg gaagggatgc 7560
 taaggtagag ggtgaacggt acagaaaagc aggctgggaa gcatatttga gaagatgcgg 7620
 ccagcaaac taaaaaactg tattataagt aatgcatgt atactaaact cacaaattag 7680
 agcttcaatt taattatatac agttattacc caattctcat gtttgacagc ttatcatcga 7740
 tctccaact gcatggagat gagtctggc aagaatacca agagttcctc ggtttgccag 7800
 ttattaaaag actcgtatftt ccaaaagact gcaacatact actcagtga gcttcacaga 7860

ES 2 731 782 T3

aacctcattc gtttattccc ttgtttgatt cagaagcagg tgggacaggt gaacttttgg 7920
attggaactc gatttctgac tgggttgaa ggcaagagag ccccgagagc ttacatttta 7980
tgtagactgg tggactgacg ccagaaaatg ttggtgatgc gcttagatta aatggcgтта 8040
ttggtgttga tgtaagcggg ggtgtggaga caaatggtgt aaaagactct aacaaaatag 8100
caaatctcgt caaaaatgct aagaaatagg ttattactga gtagtattta ttttaagtatt 8160
gtttgtgcac ttgcctgcaa gccttttgaa aagcaagcat aaaagatcta aacataaaat 8220
ctgtaaaata acaagatgta aagataatgc taaatcattt ggctttttga ttgattgtac 8280
aggaaaaat acatcgcagg ggggtgactt ttaccatttc accgcaatgg aatcaaaactt 8340
gttgaagaga atgttcacag gcgcatacgc tacaatgacc cgattcttgc tagccttttc 8400
tcggtcttgc aaacaaccgc cggcagctta gtatataaat acacatgtac atacctctct 8460
ccgtatcctc gtaatcattt tcttgtattt atcgtctttt cgctgtaaaa actttatcac 8520
acttatctca aatacactta ttaaccgctt ttactattat cttctacgct gacagtaata 8580
tcaaacagt acacatatta aacacagtgg tttctttgca taaacacat cagcctcaag 8640
tcgtcaagta aagatttcgt gttcatgcag atagataaca atctatatgt tgataattag 8700
cgttgctca tcaatgcgag atccgtttaa cgggacccta gtgcacttac cccacgttgc 8760
gtccactgtg tgccgaacat gtccttcac tattttaaca tgtggaatta attctcatgt 8820
ttgacagctt atcatcgaac tctaagaggt gatacttatt tactgtaaaa ctgtgacgat 8880
aaaaccggaa ggaagaataa gaaaactcga actgatctat aatgcctatt ttctgtaaag 8940
agttaaagct atgaaagcct cggcattttg gccgctccta ggtagtgcct tttttccaag 9000
gacaaaaacag tttcttttct ttgagcaggt tttatgtttc ggtaatcata aacaataaat 9060
aaattatttc atttatgttt aaaaataaaa aataaaaaag tattttaaat ttttaaaaa 9120
gttgattata agcatgtgac cttttgcaag caattaaatt ttgcaatttg tgattttagg 9180
caaaagttac aatttctggc tcgtgtaata tatgtatgct aaagtgaact tttacaaagt 9240
cgatatggac ttagtcaaaa gaaatcttct taaaaatata tagcactagc caatttagca 9300
cttctttatg agatatatta tagactttat taagccagat ttgtgtatta tatgtattta 9360
cccggcgaat catggacata cattctgaaa taggtaatat tctctatggt gagacagcat 9420
agataaccta ggatacaagt taaaagctag tactgttttg cagtaatttt tttcttttt 9480
ataagaatgt taccacctaa ataagttata aagtcaatag ttaagtttga tatttgattg 9540
taaaataccg taatatattt gcatgatcaa aaggctcaat gttgactagc cagcatgtca 9600
accactatat tgatcaccga tatatggact tccacaccaa ctagtaatat gacaataaat 9660
tcaagatatt cttcatgaga atggcccagc gatatatgcg gtgtgaaata ccgacagat 9720

ES 2 731 782 T3

gcgtaaggag aaaataccgc atcaggcgcc attcgccatt caggctgcmc aactgttggg 9780
aagggcgatc ggtgcccccc tcttcgctat tacgccagct ggcgaaagg ggatgtgctg 9840
caaggcgatt aagttgggta acgccagggt tttcccagtc acgacgttgt aaaacgacgg 9900
ccagtgaatt cgagctcatg ttagacaaca cccgcttacg catagctatt cagaaatcag 9960
gccgtttaag cgatgattca cgagaattgc tggcccgtg cggcataaaa attaattggt 10020
acc 10023

<210> 100

<211> 1701

<212> DNA

<213> *Saccharomyces cerevisiae*

5

<400> 100

atggcgattg caattggcct cgatthttggc agtgattctg tgcgagcttt ggcggtggac 60
tgcgctaccg gtgaagagat cgccaccagc gtagagtggg atccccgttg gcagaaaggg 120
caatthttgtg atgccccgaa taaccagttc cgtcatcatc cgcgtgacta cattgagtca 180
atggaagcgg cactgaaaac cgtgcttgca gagcttagcg tcgaacagcg cgcagctgtg 240
gtcgggattg gcgttgacag taccggctcg acgcccgcac cgattgatgc cgacggaaac 300
gtgctggcgc tgcgcccgga gtttgccgaa aaccggaacg cgatgttcgt attgtggaaa 360
gaccacactg cggttgaaga agcggaaagag attaccogtt tgtgccacgc gccgggcaac 420
gttgactact cccgctacat tgggtgtatt tattccagcg aatggttctg ggcaaaaatc 480
ctgcatgtga ctgcgcaaga cagcgccgtg gcgcaatctg ccgcatcgtg gattgagctg 540
tgcgactggg tgccagctct gctttccggt accaccogcc cgcaggatat tcgtcgcgga 600
cgttgcagcg ccgggcataa atctctgtgg cagcaaaagct gggcgggcct gccgccagcc 660
agtttctttg atgagctgga cccgatcctc aatcgccatt tgccttcccc gctgttcaact 720
gacacttgga ctgcccgatat tccggtgggc accttatgcc cggaatgggc gcagcgtctc 780
ggcctgcctg aaagcgtggt gatttccggc ggcgcgtttg actgccatat gggcgcagtt 840
ggcgcagggc cacagcctaa cgcactggtg aaagtatcgt gtacttccac ctgcgacatt 900
ctgattgccg acaaacagag cggtggcgag cgggcagttg aaggtatthg cggtcaggtt 960
gatggcagcg tgggtgcctg atthtatcggg ctggaagcag gccaatcggc gthttggtgat 1020
atctacgcct ggtthgtgctg cgtactcggc tggccgctgg aacagcttgc cggccagcat 1080
ccggaactga aaacgcaaat caacgccagc cagaaacaac tgcttccggc gctgaccgaa 1140
gcatgggcca aaaatccgtc tctggatcac ctgccgggtg tgctcgactg gthtaacggc 1200
cgccgcacac cgaacgctaa ccaacgcctg aaaggggtga thaccgatct taacctcgct 1260
accgacgctc cgctgctgth cggcggtthg atthctgcca ccgcctthg cgcacgcgca 1320

ES 2 731 782 T3

atcatggagt gctttaccga tcaggggatc gccgtaata acgtgatggc actgggcggc 1380
atcgcgcgga aaaaccaggt cattatgcag gcctgctgcg acgtgctgaa tcgcccgctg 1440
caaattggtg cctctgacca gtgctgtgcg ctcggtgctg cgatTTTTTgc tgccgtcgcc 1500
gcgaaaagtgc acgcagacat cccatcagct cagcaaaaaa tggccagtgc ggtagagaaa 1560
accctgcaac cgtgcagcga gcaggcacia cgctttgaac agctttatcg ccgctatcag 1620
caatgggcga tgagcgccga acaacactat cttccaactt ccgccccggc acaggctgcc 1680
caggccgttg cgactctata a 1701

<210> 101

<211> 1803

<212> DNA

<213> *Saccharomyces cerevisiae*

5

<400> 101

atgttgtgtt cagtaattca gagacagaca agagaggttt ccaacacaat gtctttagac 60
tcatactatc ttgggtttga tctttcgacc caacaactga aatgtctcgc cattaaccag 120
gacctaaaaa ttgtccattc agaaacagtg gaatttgaaa aggatcttcc gcattatcac 180
acaaagaagg gtgtctatat acacggcgac actatcgaat gtcccgtagc catgtggtta 240
gaggctctag atctggttct ctcgaaatat cgcgaggcta aatttccatt gaacaaagtt 300
atggccgtct cagggctctg ccagcagcac gggctctgtct actggctctc ccaagccgaa 360
tctctgtag agcaattgaa taagaaaccg gaaaaagatt tattgcacta cgtgagctct 420
gtagcatttg caaggcaaac cgcccccaat tggcaagacc acagtactgc aaagcaatgt 480
caagagtttg aagagtgcac aggtgggcct gaaaaaatgg ctcaattaac agggccaga 540
gccattttta gatttactgg tctctcaaatt ctgaaaattg cacaattaga accagaagct 600
tacgaaaaaa caaagacat ttcttttagtg tctaattttt tgacttctat cttagtgggc 660
catcttgttg aattagagga ggcagatgcc tgtggtatga acctttatga tatacgtgaa 720
agaaaattca gtgatgagct actacatcta attgatagtt cttctaagga taaaactatc 780
agacaaaaat taatgagagc acccatgaaa aattgatag cgggtacat ctgtaaatat 840
tttattgaga agtacggttt caatacaaac tgcaaggctct ctcccatgac tggggataat 900
ttagccacta tatgttcttt acccctgcgg aagaatgacg ttctcgtttc cctaggaaca 960
agtactacag ttcttctggt caccgataag tatcaccctt ctccgaacta tcatcttttc 1020
attcatccaa ctctgcaaaa ccattatatg ggtatgattt gttattgtaa tggttctttg 1080
gcaagggaga ggataagaga cgagttaaac aaagaacggg aaaataatta tgagaagact 1140
aacgattgga ctctttttta tcaagctgtg ctagatgact cagaaagtag tgaaaatgaa 1200
ttaggtgtat attttctctt gggggagatc gttcctagcg taaaagccat aaacaaaagg 1260

ES 2 731 782 T3

gttatcttca atccaaaaac gggatgatt gaaagagagg tggccaagtt caaagacaag 1320
 aggcacgatg ccaaaaatat tgtagaatca caggctttaa gttgcagggt aagaatatct 1380
 cccctgcttt cggattcaaa cgcaagctca caacagagac tgaacgaaga tacaatcgtg 1440
 aagtttgatt acgatgaatc tccgctgcgg gactacctaa ataaaaggcc agaaaggact 1500
 tttttgtag gtggggcttc taaaaacgat gctattgtga agaagtttgc tcaagtcatt 1560
 ggtgctacaa aggtaattt taggctagaa acaccaaact catgtgccct tgggtggtgt 1620
 tataaggcca tgtggtcatt gttatatgac tctaataaaa ttgcagttcc ttttgataaa 1680
 tttctgaatg acaatthttcc atggcatgta atggaaagca tatccgatgt ggataatgaa 1740
 aattgggatc gctataattc caagattgtc cccttaagcg aactggaaaa gactctcatc 1800
 taa 1803

<210> 102

<211> 160

<212> DNA

<213> *Saccharomyces cerevisiae*

5

<400> 102

ttctagagca cagctaacac cacgtcgtcc ctatctgctg ccctaggtct atgagtgggt 60

gctggataac tttacgggca tgcataaggc tcgtataata tattcagga gaccacaacg 120

gtttccctct acaataaatt ttgtttaact ttactagag 160

REIVINDICACIONES

1. Un procedimiento para la producción de al menos un metabolito de interés mediante la transformación de una pentosa en un microorganismo, caracterizado porque dicho procedimiento comprende al menos:

5 (i) una operación de cultivo un microorganismo recombinante que expresa una vía sintética de asimilación de pentosas catalizada por un conjunto de enzimas:

- una enzima recombinante que consiste en una quinasa adecuada para fosforilar en posición 1 una pentosa escogida entre (D)-xilulosa y/o (L)-ribulosa,

- una enzima que consiste en una aldolasa adecuada de escindir la(s) pentosas-1-fosfonato para la obtención de glicolaldehído y dihidroxiacetona fosfato (DHAP),

10 y en que la vía sintética de asimilación de pentosas comprende al menos las etapas siguientes:

a) fosforilación en posición 1 de una pentosa escogida de (D)-xilulosa y/o (L)-ribulosa, caracterizada por dicha quinasa,

b) escisión de pentosa-1-fosfato obtenido al final de la etapa a), para la obtención de glicolaldehído y dihidroxiacetona fosfato (DHAP), catalizada por dicha aldolasa,

15 (ii) una operación de recuperación de dicho al menos un metabolito de interés obtenido al final de la operación de cultivo (i).

2. Procedimiento para la producción de al menos un metabolito de interés según la reivindicación 1, caracterizado porque:

20 - la etapa de fosforilación a) de la vía sintética de asimilación de pentosas está catalizada por una enzima expresada de forma recombinante escogida entre el grupo constituido por cetohecoquinasa C (KhcC), ramnulosa quinasa (rhaB) y fuculosa quinasa (fucK) y/o

- la etapa de escisión b) de la vía sintética de asimilación de pentosas está catalizada por una aldolasa escogida entre el grupo constituido por aldolasa B (Aldo-B) y fructosa-1,6 bisfosfato aldolasa (FbaB).

25 3. Procedimiento para la producción de al menos un metabolito de interés según la reivindicación 2, caracterizado porque:

- la etapa a) es catalizada por la cetohecoquinasa C, y

- la etapa b) es catalizada por aldolasa B (Aldo-B) y/o la fructosa -1,6-bisfosfato aldolasa B.

30 4. Procedimiento para la producción de al menos un metabolito de interés según una de las reivindicaciones 1 a 3, siendo escogido dicho metabolito entre etilenglicol y/o ácido glicólico, caracterizado porque dicho microorganismo recombinante expresa una vía sintética de asimilación de pentosas catalizada además por las enzimas siguientes:

- una enzima que consiste en una glicolaldehído reductasa, adecuada para catalizar la reducción del glicolaldehído obtenido al finas de la etapa b) a etilenglicol, y/o

- una enzima que consiste en una glicolaldehído deshidrogenasa, adecuada para catalizar una oxidación del glicolaldehído obtenido al fina de la etapa b), a ácido glicólico

35 y porque dicha vía sintética de asimilación de pentosas comprende además las etapas siguientes

c) reducción del glicolaldehído obtenido al final de la etapa b) a etilenglicol, catalizada por dicha glicolaldehído reductasa y/o

c') oxidación del glicolaldehído obtenido al final de la etapa b) a ácido glicólico, catalizada por dicha glicolaldehído deshidrogenasa.

5. Procedimiento para la producción de al menos un metabolito de interés según la reivindicación 4, para la producción de etilenglicol y/o de ácido glicólico, caracterizado porque:

5 - la etapa de reducción c) está catalizada por una glicolaldehído reductasa escogida entre el grupo constituido por aldehído reductasa (YqhD), glicerol deshidrogenasa (GldA) y propano-1,2-diol oxidorreductasa (FucO), y/o

- la etapa de oxidación c') está catalizada por una glicolaldehído deshidrogenasa que consiste en lactaldehído deshidrogenasa (AldA).

10 6. Procedimiento para la producción de al menos un metabolito de interés según una cualquiera de las reivindicaciones anteriores, caracterizado porque el microorganismo se pone en cultivo sobre un medio carbonado que contiene (D)-xilosa y/o (L)-arabinosa.

7. Procedimiento para la producción de al menos un metabolito de interés según la reivindicación 6, caracterizado porque el medio de cultivo comprende un hidrolizado de biomasa que comprende hemicelulosa.

15 8. Procedimiento para la producción de al menos un metabolito de interés según una cualquiera de las reivindicaciones anteriores, caracterizado porque dicha vía sintética de asimilación de pentosas comprende, previamente a la etapa a):

- una etapa de conversión de (D)-xilosa en (D)-xilulosa y/o,

- una etapa de la conversión de (L)-arabinosa en (L)-ribulosa.

20 9. Procedimiento para la producción de al menos un metabolito de interés según una cualquiera de las reivindicaciones anteriores, caracterizado porque el microorganismo recombinante es escogido entre *E. coli*, *S. cerevisiae* y *Scheffersomyces stipitis*.

10. Procedimiento para la producción de al menos un metabolito de interés según una cualquiera de las reivindicaciones anteriores, caracterizado porque el microorganismo recombinante consiste en un microorganismo cuyas actividades de (D)-xilulosa-quinasa y/o (L)-ribulosa-5-quinasa han sido suprimidas.

25 11. Microorganismo recombinante que expresa una vía sintética de asimilación de pentosas y que comprende al menos ácidos nucleicos que codifican las enzimas siguientes:

i) una enzima que consiste en una quinasa adecuada de fosforilar en posición 1 una pentosa escogida entre (D)-xilulosa y/o (L)-ribulosa,

30 ii) una enzima que consiste en una aldolasa adecuada para escindir la(s) pentosa(s)-1 fosfato obtenidas al final de la etapa (i), para la obtención de glicolaldehído y de dihidroxiacetona fosfato (DHAP),

caracterizado porque las actividades de (D)-xilulosa-5-quinasa y/o (L)-ribulosa-5-quinasa de dicho microorganismo han sido suprimidas y/o el microorganismo es portador de al menos una de las modificaciones siguientes:

- sobreexpresión de un gen que codifica una glioxilato reductasa;

35 - sobreexpresión de un gen que codifica una isocitrato liasa acompañada o no de la delección de su represor transcripcional;

- delección de los genes que codifican malato sintasas;

- delección de los genes que codifican glioxilato carboligasas;

- delección de los genes que codifican glicolato oxidasas y/o glicolato deshidrogenasas;
 - delección de los genes que codifican 2-ceto-4-hidroxi-glutarato aldolasas, particularmente Entner-Doudouroff aldolasa y fosfogluconato deshidratasas;
 - 5 - delección de un gen que codifica un represor de genes implicados en el metabolismo respiratorio, particularmente el gen *arcA*;
 - atenuación y, particularmente, delección de la expresión de isocitrato deshidrogenasa;
 - delección de los genes que codifican sistemas de internalización de ácido glicólico;
 - atenuación de vías metabólicas que conducen a la producción de productos secundarios como acetato, lactato o etanol;
 - 10 - sobreexpresión de al menos un gen que codifica un transportador de azúcares.
12. Microorganismo recombinante según la reivindicación 11, caracterizado porque comprende:
- al menos un ácido nucleico exógeno que codifica cetohexoquinasa C, y
 - al menos un ácido nucleico exógeno que codifica aldolasa B.
13. Microorganismo según una de las reivindicaciones 11 o 12, que comprende además al menos una de las modificaciones siguientes:
- 15 - sobreexpresión del gen que codifica glicolaldehído reductasa principal;
- delección del gen que codifica glicolaldehído deshidrogenasa.
14. Microorganismo según una de las reivindicaciones 11 o 12, que comprende además al menos una de las modificaciones siguientes:
- 20 - sobreexpresión del gen que codifica glicolaldehído deshidrogenasa;
- inactivación de al menos uno de los genes que codifican al menos una de las subunidades de glicolato oxidasa.
15. Microorganismo según la reivindicación 14, que comprende las siguientes modificaciones:
- 25 - sobreexpresión de un gen que codifica una glioxilato reductasa;
- sobreexpresión de un gen que codifica una isocitrato liasa acompañada o no de la delección de su represor transcripcional;
- delección de los genes que codifican malato sintasas;
- delección de los genes que codifican glioxilato carboligasas;
- delección de los genes que codifican glicolato oxidasas y/o glicolato deshidrogenasas;
- 30 - delección de los genes que codifican 2-ceto-4-hidroxi-glutarato aldolasas, particularmente Entner-Doudouroff aldolasa y fosfogluconato deshidratasas;
- delección de un gen que codifica un represor de genes implicados en el metabolismo respiratorio, particularmente el gen *arcA*;

- atenuación y, particularmente, delección de la expresión de isocitrato deshidrogenasa;
- ocasionalmente, la sobreexpresión de un gen que codifica un transportador de azúcares.

16. Microorganismo según la reivindicación 15, que comprende las modificaciones siguientes:

- la sobreexpresión del gen *aldA* que codifica glicolaldehído deshidrogenasa;
- 5
- la sobreexpresión del gen *ghrA*;
 - la sobreexpresión del gen *aceA*;
 - la delección del gen *iclR*;
 - la delección del gen *gcl*;
 - la delección de los genes *aceB* y *glcB*.
- 10
- la delección de al menos un gen escogido entre *glcD*, *glcE*, *glcF* o *glcG*;
 - la delección de los genes *edd-eda*;
 - la delección del gen *arcA*;
 - la delección del gen *icd*;
 - la sobreexpresión del gen *galP*.

15

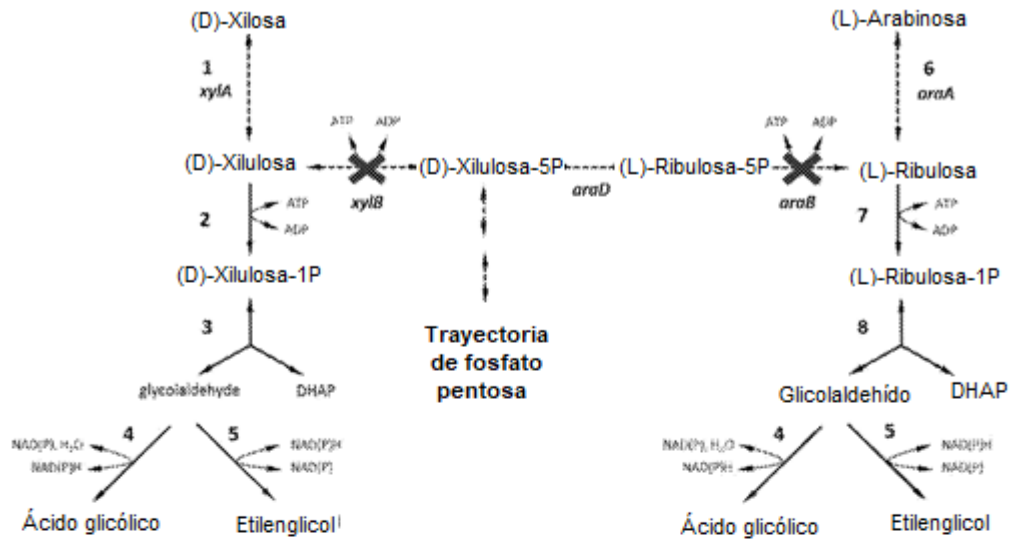


Fig. 1

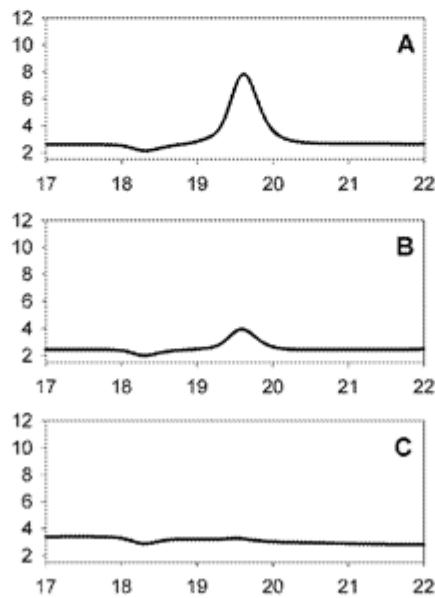


Fig. 2

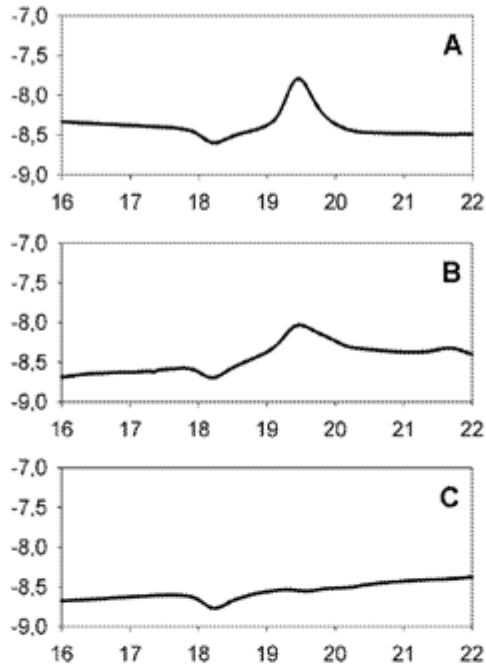


Fig. 3

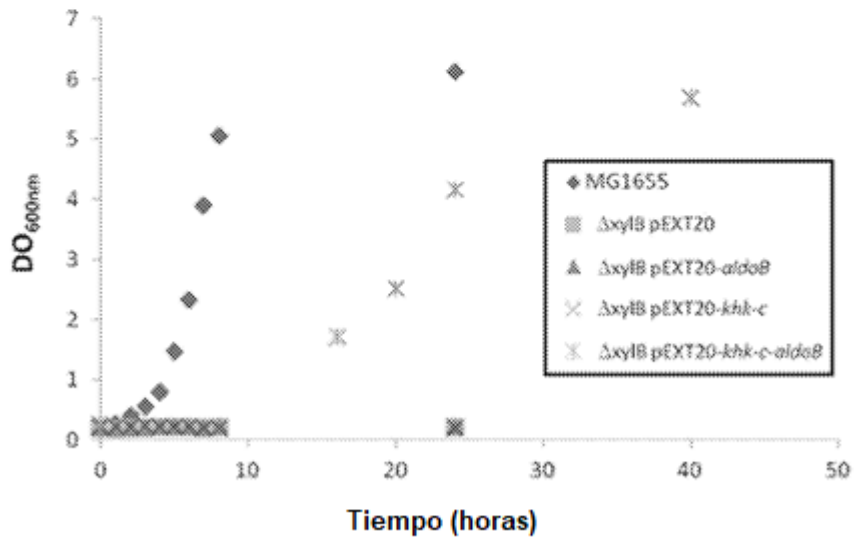


Fig. 4

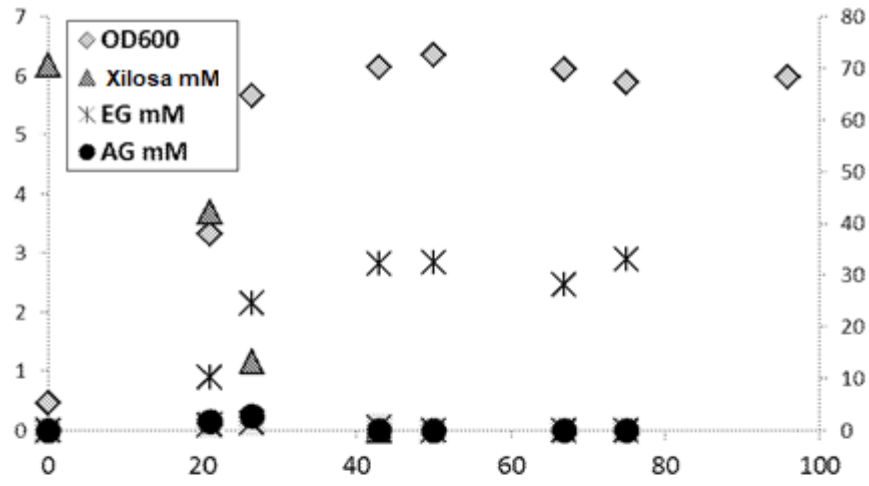


Fig. 5

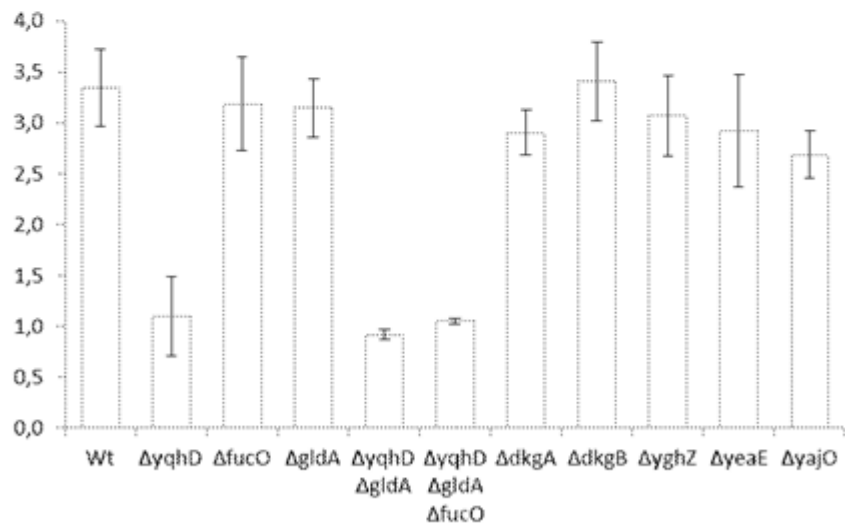


Fig. 6

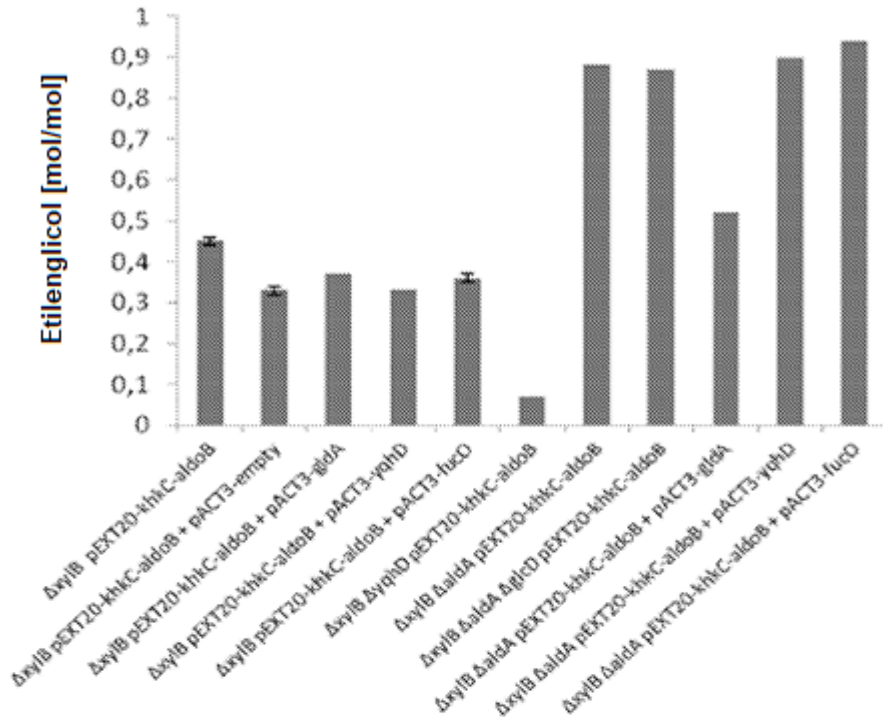


Fig. 7

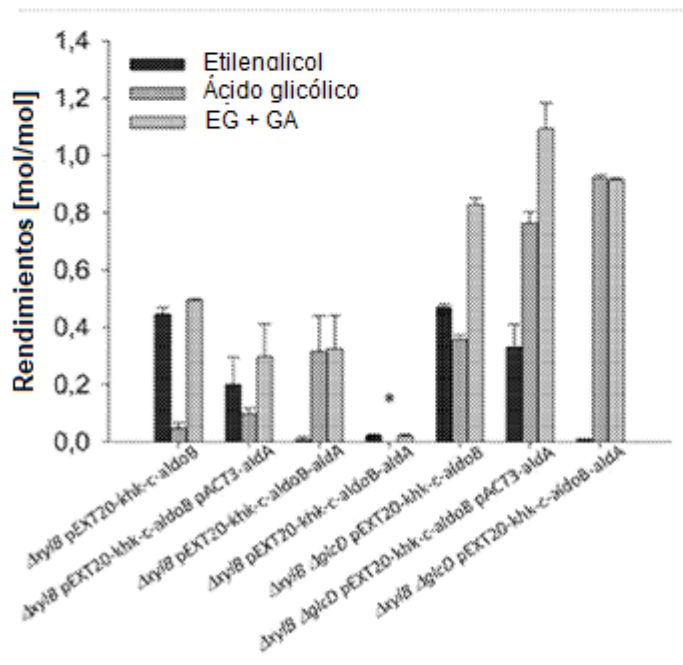


Fig. 8

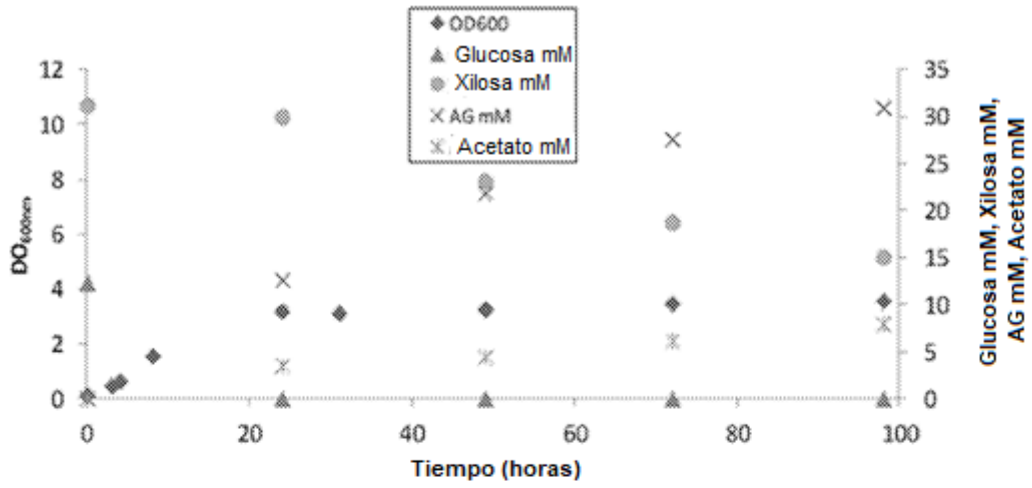


Fig. 9

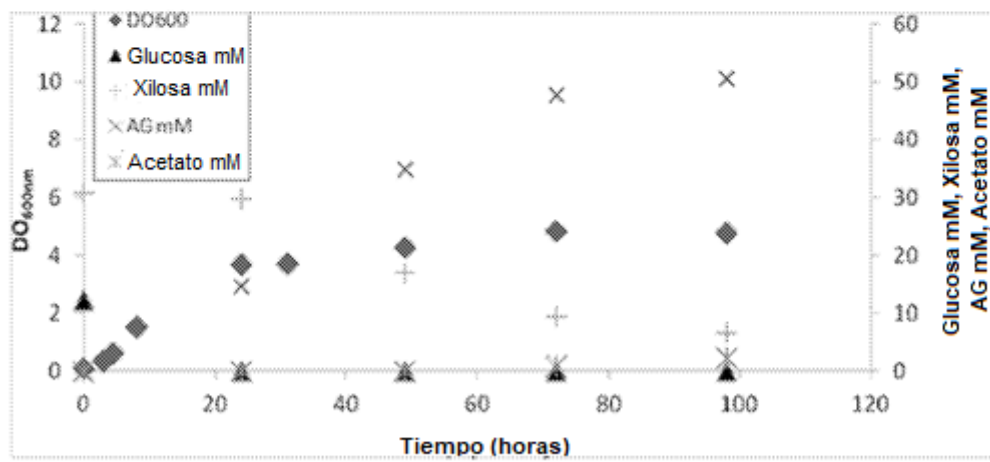


Fig. 10

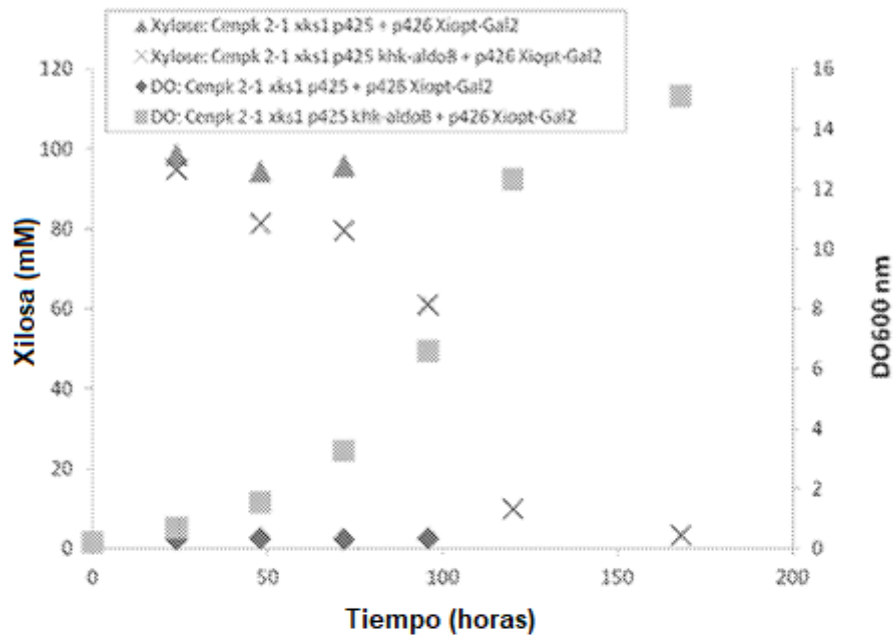


Fig. 11

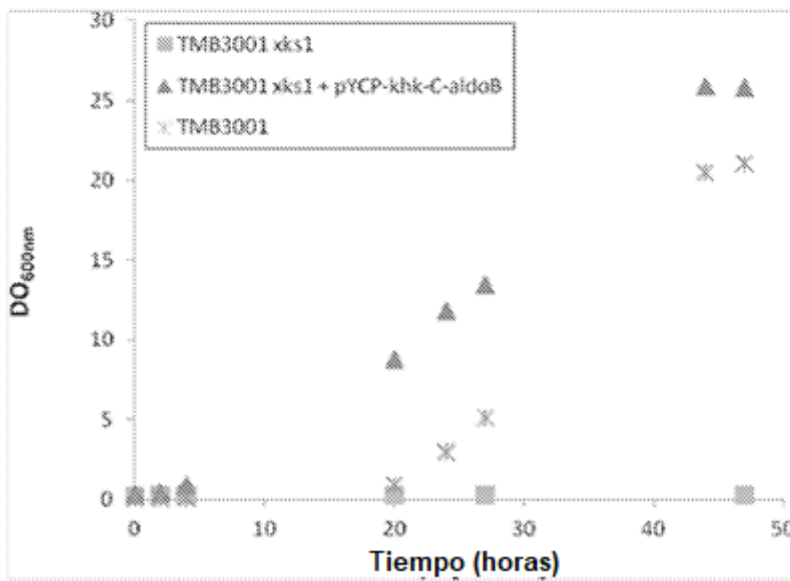


Fig. 12

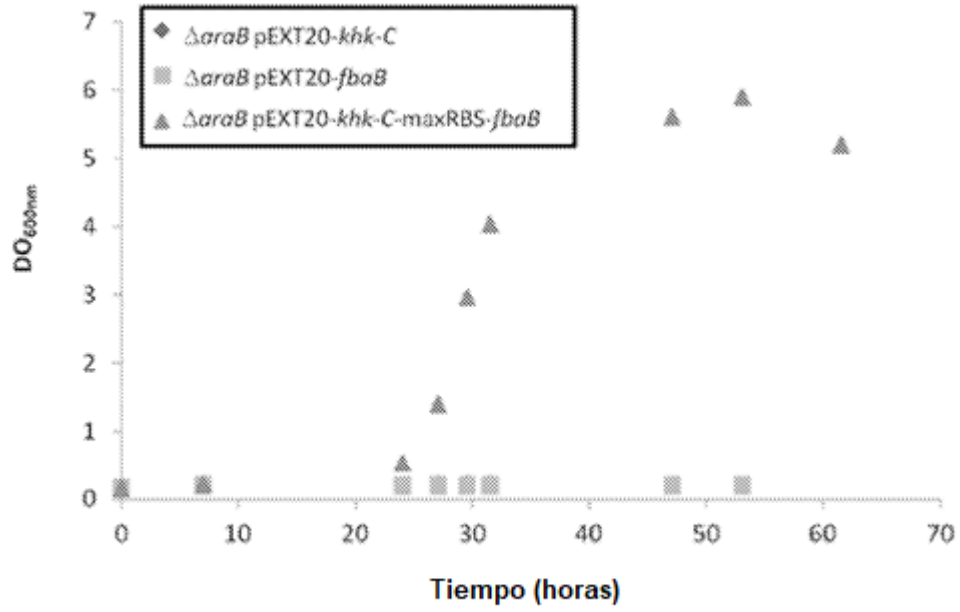


Fig. 13