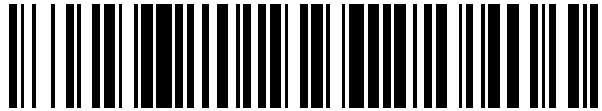


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 551 515**

51 Int. Cl.:

C12N 1/21 (2006.01)

C12N 1/15 (2006.01)

C12N 1/19 (2006.01)

C12N 9/92 (2006.01)

C12P 7/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.03.2009 E 09718191 (1)**

97 Fecha y número de publicación de la concesión europea: **29.07.2015 EP 2250262**

54 Título: **Una célula fermentadora de azúcar pentosa**

30 Prioridad:

07.03.2008 EP 08102407

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

19.11.2015

73 Titular/es:

**DSM IP ASSETS B.V. (100.0%)
Het Overloon 1
6411 TE Heerlen, NL**

72 Inventor/es:

**KLAASSEN, PAUL;
LAAN, VAN DER, JAN METSKE;
GIELESEN, BIANCA ELISABETH MARIA y
SUYLEKOM, VAN, GIJSBERDINA PIETERNELLA**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

ES 2 551 515 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

Una célula fermentadora de azúcar pentosa

Campo de la invención

5 La presente invención se relaciona con una célula que es capaz de isomerizar la xilosa a xilulosa. La invención también se relaciona con un proceso en el cual tales células son utilizadas para la producción de un producto de fermentación, tal como etanol.

Antecedentes de la invención

10 El consumo a gran escala de combustibles fósiles tradicionales (combustibles basados en petróleo) en las décadas recientes ha contribuido a altos niveles de contaminación. Esto, junto con la conclusión de que las reservas mundiales de combustibles fósiles no son limitadas y una preocupación ambiental creciente, ha estimulado nuevas iniciativas para investigar la factibilidad de combustibles alternativos tales como etanol, el cual es una fuente combustible de combustión libre de partículas que liberan menos CO₂ que la gasolina sin plomo por litro.

15 Aunque el etanol derivado de una biomasa puede producirse por la fermentación de azúcares de hexosa obtenidos a partir de muchas fuentes diferentes, los sustratos típicamente utilizados en producción a escala comercial de alcohol combustible, tales como caña de azúcar y almidón de maíz, son costosos. Los incrementos en la producción de etanol combustible requerirán por lo tanto el uso de materias primas menos costosas.

20 Actualmente, solamente las materias primas lignocelulósicas derivadas de biomasa vegetales están disponibles en cantidades suficientes para sustituir los cultivos utilizados actualmente para la producción de etanol. La mayor parte del material lignocelulósico, el segundo azúcar más común, después de la glucosa, es la xilosa. Así, para un proceso de producción de combustible económicamente factible, los azúcares tanto de hexosa como de pentosa deberían ser fermentados para formar etanol. La levadura *Saccharomyces cerevisiae* es robusta y bien adaptada para la producción de etanol, pero es incapaz de producir etanol utilizando xilosa como fuente de carbono. También, no se conocen organismos de origen natural que puedan fermentar la xilosa a etanol con tanto un alto rendimiento de etanol como con una alta productividad de etanol. El documento WO 2006/009434 divulga una célula huésped que comprende una secuencia de nucleótidos que codifica una xilosa isomerasa de *Bacteroides thetaiotaomicron* donde la célula es capaz de utilizar xilosa como fuente de carbono.

25 Por lo tanto hay necesidad de un organismo que posea estas propiedades de forma que permita la producción comercialmente viable de etanol a partir de materias primas lignocelulósicas.

Resumen de la invención

30 De acuerdo con la invención, se proporciona una célula que es capaz de realizar fermentación, tal como fermentación alcohólica, y de usar xilosa como fuente de carbono. Tal célula comprende una secuencia de nucleótidos que codifica una xilosa isomerasa, donde la secuencia de aminoácidos de la xilosa isomerasa tiene al menos aproximadamente un 96% de identidad en secuencia con la secuencia de aminoácidos definida en SEQ ID NO: 3 y donde la secuencia de nucleótidos es heteróloga al huésped, y donde la célula es capaz de isomerizar xilosa a xilulosa. Tal célula produce una cantidad mayor de etanol cuando utiliza xilosa como fuente de carbono en comparación con el hongo filamentoso de tipo silvestre.

La invención también proporciona:

- 40
- un proceso para producir un producto de fermentación proceso que comprende fermentar un medio que contiene una fuente de xilosa con una célula de la invención de tal manera que la célula fermente la xilosa al producto de fermentación;
 - un proceso para producir un producto de fermentación proceso que comprende fermentar un medio que contiene al menos una fuente de xilosa y una fuente de L- arabinosa con una célula tal como se define en la invención el cual es capaz de utilizar L-arabinosa de tal manera que la célula fermente la xilosa y la L-arabinosa hasta el producto de fermentación; y
 - 45 - un proceso para producir un producto de fermentación proceso que comprende fermentar un medio que contiene al menos una fuente de xilosa y una fuente de L-arabinosa con una célula de la invención y una célula capaz de utilizar L-arabinosa mediante lo cual cada célula fermenta xilosa y/o arabinosa hasta el producto de fermentación.

50 La invención proporciona adicionalmente el uso de una célula de la invención en un proceso para la producción de un producto de fermentación.

Breve descripción de los dibujos

La figura 1 define el mapa de plásmidos de pYISIT4-XKSI-xyIA (Baun CpO) que codifica la xilosa isomerasa a partir de *Bacteroides uniformis* ATCC 8492 para su expresión en *Saccharomyces cerevisiae*. La CpO denota el par de codones optimizados.

5 La figura 2 define un mapa físico de un plásmido pPWT080, cuya secuencia se da en SEQ ID No. 4.

La figura 3 define un mapa físico del locus GRE3 tipo silvestre (panel a) y una copia de integración de PWT080 en el locus GRE3 (panel b, que muestra donde los cebadores se enlazan y el panel c, que muestra donde la sonda *RKI1* se enlaza).

10 La figura 4 define un autoradiograma que muestra la integración correcta de una copia del plásmido pPWT080 en CEN.PK113-7D;

Panel a: Digestión de *XcmI* de las preparaciones de ADN cromosómico hibridizado con la sonda *RKI1*. Línea 1: CEN.PK113-7D; línea 2: BIE104F1; línea 3: BIE104P1.

Panel b: digestión de *PstI* de preparaciones de ADN cromosómico, hibridizado con la sonda *RKI1*. Línea 1: CEN.PK113-7D; línea 2: BIE104F1; línea 3: BIE104P1.

15 Δ GRE3::PPP significa el reemplazo de la región de codificación del gen *GRE3* por el casete que contiene los genes *TAL1*, *TKL1*, *RKI1* y *RPE1* bajo el control de promotores constitutivos fuertes, Δ GRE3::[*TPI1p-TAL1-ADH1p-TKL1-PGI1p-RPE1-ENO1p-RKI1*].

20 La figura 5 define un mapa físico del locus *GRE3*, donde la región de codificación del gen *GRE3* ha sido reemplazada por la integración de los genes PPP *TAL1*, *TKL1*, *RKI1* y *RPE1*. El panel a muestra donde los cebadores de SEQ ID 5 y 6 se enlazan, el panel b muestra donde se enlaza la sonda *RKI1*.

La figura 6 define un mapa físico del plásmido pYI#SIT4

La figura 7 define un mapa físico del plásmido pPWT007

La figura 8 define un mapa físico del plásmido pPWT042

25 La figura 9 define un mapa físico del locus *SIT4* tipo silvestre (panel a) y una copia de integración de PWT080 en el locus *SIT4* (panel b, que muestra donde se enlazan los cebadores).

30 La figura 10 representa una curva de crecimiento de BIE104P1Y9 sobre xilosa al 2% como única fuente de carbono, después de varios precultivos, y de la cepa de referencia sin una copia de pPWT042 integrada en el genoma. Los eventos indicados en la gráfica por números (1): transferencia a YNB 1% de glucosa + 1% de xilosa; (2): transferencia a YNB de 0.1% de glucosa + 2% de xilosa; (3) transferencia a YNB de 2% de xilosa; (4) transferencia de YNB de 2% de xilosa (solamente BIE104P1Y9).

La figura 11 representa una curva de crecimiento de la cepa de referencia BIE104P1 y una cepa metabolizante de xilosa, BIE104P1Y9.

35 La figura 12 representa el consumo de xilosa y glucosa y la producción de etanol con respecto al tiempo de las cepas BIE104P1 precultivadas sobre glucosa (panel a), BIE104P1Y9 precultivadas en glucosa (panel b) y BIE104P1Y9 precultivadas en xilosa (panel c).

La figura 13 representa un mapa físico del plásmido pPWT018.

La figura 14 define un mapa físico del plásmido pPWT006.

40 La figura 15 representa un autorradiograma Southern blot. EL ADN cromosómico de la cepa tipo silvestre CEN.PK113-7D (línea 1) y BIE104A2 (línea 2) fue digerido con Both *EcoRI* y *HindIII*. La siembra fue hibridizada con una sonda específica *SIT2*.

La figura 16 representa los mapas físicos del locus *SIT2* tipo silvestre (panel a) y después de la introducción de los genes ara por integración del plásmido pPWT018, seguido por la recombinación intramolecular que lleva a la pérdida del vector y a las secuencias de marcador seleccionable (panel b). La hibridización de la sonda está indicada.

45 La figura 17 muestra una representación gráfica de curvas de crecimiento de la cepa BIE104A2P1Y9 en medios diferentes. Panel a: cepa BIE104A2P1Y9 cultivada sobre galactosa, seguida por eventos indicados en la gráfica por números (1) transferencia a arabinosa 1% + xilosa al 1% y (2) transferencia a xilosa al 2% + arabinosa al 0.2%. Panel b: cepa BIE104A2P1Y9 cultivada sobre glucosa, seguida por (1) transferencia a arabinosa al 1% + xilosa al 1% y (2) transferencia a xilosa al 2% + arabinosa al 0.2%.

Breve descripción del listado de secuencias

- SEQ ID NO: 1 representa la secuencia de la xilosa isomerasa tipo silvestre de *Bacteroides uniformis* ATCC 8492. Acceso a Genbank no. AAYH02000036.
- SEQ ID NO: 2 representa una secuencia optimizada de codón derivada de SEQ ID NO. 1.
- 5 SEQ ID NO: 3 representa la secuencia de aminoácidos de la xilosa isomerasa a partir de *Bacteroides uniformis* ATCC 8492.
- SEQ ID NO. 4 representa la secuencia del plásmido pPWT080.
- SEQ ID NO: 5 representa la secuencia del cebador de avance.
- SEQ ID NO: 6 representa la secuencia del cebador reverso.
- 10 SEQ ID NO: 7 representa la secuencia del cebador multifuncional de avance para PCR de diagnóstico.
- SEQ ID NO: 8 representa la secuencia del cebador multifuncional reverso para PCR diagnóstico.
- SEQ ID NO: 9 representa la secuencia de la sonda RKI1 del cebador de avance.
- SEQ ID NO: 10 representa la secuencia de la sonda RKI1 del cebador reverso.
- SEQ ID NO: 11 representa la secuencia del casete kanMX del cebador de avance.
- 15 SEQ ID NO: 12 representa la secuencia del casete kanMX del cebador reverso.
- SEQ ID NO: 13 representa la secuencia del cebador de avance.
- SEQ ID NO: 14 representa la secuencia del cebador reverso.
- SEQ ID NO: 15 representa la secuencia del cebador multifuncional de avance para PCR diagnóstico.
- SEQ ID NO: 16 representa la secuencia del cebador multifuncional reverso para PCR diagnóstico.
- 20 SEQ ID NO: 17 representa la secuencia de la secuencia del plásmido pPWT018.
- SEQ ID NO: 18 representa la secuencia del cebador de integración pPWT018 de avance.
- SEQ ID NO: 19 representa la secuencia del cebador de integración pPWT018 reverso.
- SEQ ID NO: 20 representa la secuencia de la sonda SIT2 del cebador de avance.
- SEQ ID NO: 21 representa la secuencia del cebador SIT2 del cebador reverso.

25 **Descripción detallada de la invención**

- A lo largo de la presente especificación y en las reivindicaciones acompañantes las palabras “comprenden”, “incluyen” y variaciones tales como “comprende”, “que comprende”, “incluye” y “que incluye” deben interpretarse de forma inclusiva. Esto es, estas palabras pretenden abarcar la posible inclusión de otros elementos o enteros no específicamente descritos, donde el contexto lo permita.
- 30 La invención se relaciona con una célula que comprende una secuencia de nucleótidos que codifica una xilosa isomerasa, donde la secuencia de aminoácidos de la xilosa isomerasa tiene al menos aproximadamente 96% de identidad con respecto a la secuencia de aminoácidos representada en SEQ ID NO: 3 y donde la secuencia de nucleótidos es heteróloga al huésped.
- 35 La presencia de la secuencia de nucleótidos que codifica una xilosa isomerasa confiere a la célula la capacidad de isomerizar xilosa hasta xilulosa.
- Una “xilosa isomerasa” (EC 5.3.1.5) se define aquí como una enzima que cataliza la isomerización directa de D-xilosa en D-xilulosa y/o viceversa. La enzima también es conocida como D-xilosa cetoisomerasa. Una xilosa isomerasa aquí también puede ser capaz de catalizar la conversión entre el D-glucosa y D-fructosa (y de acuerdo con lo anterior también puede ser denominada como glucosa isomerasa). Una xilosa isomerasa aquí puede requerir un catión bivalente, tal como magnesio, manganeso o cobalto como cofactor.
- 40 De acuerdo con lo anterior, una célula de la invención es capaz de isomerizar xilosa a xilulosa. La capacidad de isomerizar xilosa a xilulosa es conferida sobre la célula huésped por transformación de la célula huésped con un constructo de ácido nucleico que comprende una secuencia de nucleótidos que codifican una xilosa isomerasa definida. Una célula de la invención isomeriza la xilosa en xilulosa mediante la isomerización directa de xilosa a

xilulosa. Se entiende que esto significa que la xilosa es isomerizada en xilulosa en una reacción sencilla catalizada por una xilosa isomerasa, en oposición a la conversión en dos etapas de la xilosa a xilulosa a través de un intermedio xilitol y catalizado por la xilosa reductasa y el xilitol deshidrogenasa, respectivamente.

5 Una unidad (U) de actividad de xilosa isomerasa puede definirse aquí como la cantidad de enzima que produce 1 nmol de xilulosa por minuto, bajo condiciones como las descritas por Kuyper et al. (2003, FEMS Yeast Res. 4: 69 – 78).

10 La célula de la invención se define con referencia a una xilosa isomerasa que tiene la secuencia de aminoácidos de SEQ ID NO: 3 o una secuencia que tiene al menos 96% de identidad en la secuencia a la misma. De la misma forma, una célula de la invención puede definirse con referencia a una xilosa isomerasa como una secuencia de nucleótidos que codifica tal secuencia de aminoácidos.

SEQ ID NO: 3 representa la secuencia de aminoácidos de la xilosa isomerasa de *Bacteroides uniformis* ATCC8492. Una célula de la invención comprende una secuencia de nucleótidos que codifica una xilosa isomerasa que tiene los aminoácidos de SEQ ID NO: 3 o uno que tiene al menos 96% de identidad en secuencia a la misma.

15 Preferiblemente, una célula de acuerdo con la presente invención es una célula que comprende una secuencia de nucleótidos que codifica una xilosa isomerasa que tienen una secuencia que tiene al menos 96%, al menos 97%, al menos 98% o al menos 99% de identidad en la secuencia con la secuencia de aminoácidos de SEQ ID NO: 3.

20 La identidad en secuencia (o similitud en secuencia) se define aquí como una relación entre dos o más secuencias de aminoácidos (polipéptidos o proteínas) o dos o más secuencias en ácidos nucleicos (polinucleótidos), según se determina por comparación de las secuencias. Usualmente, las identidades o similitudes en secuencias se comparan, típicamente en toda la longitud de las secuencias comparadas. Sin embargo, las secuencias pueden ser comparadas en ventanas de comparación más cortas. En la técnica, "identidad" también significa el grado de relación de secuencia entre secuencias de aminoácidos o ácidos nucleicos, según sea el caso, tal como se determina por la comparación entre segmentos de tales secuencias.

25 Los métodos preferidos para determinar la identidad están diseñados para dar la mayor coincidencia entre las secuencias probadas. Los métodos para determinar la identidad y similitud están codificados en programas de ordenador públicamente disponibles. Los métodos de programas de ordenador preferidos para determinar la identidad y similitud entre dos secuencias, incluyen por ejemplo, el BestFit, BLASTP, BLASTN, y FASTA (Altschul, S.F. et al., J. Mol. Biol. 215: 403-410 (1990), disponibles públicamente en NCBI y otras fuentes (BLAST Manual, Altschul, S., et al, NCBI NLM NIH Bethesda, MD 20894). Los parámetros preferidos para la comparación de secuencias de aminoácidos utilizando BLASTP son la abertura de espacios 11.0, extensión de espacios 1, matriz Blosum 62. Los parámetros preferidos para la comparación de secuencias de ácidos nucleicos utilizando BLASTP son abertura de espacio 11.0, extensión de espacio 1, matriz completa de ADN (matriz de identidad de ADN).

30 Opcionalmente, al determinar el grado de similitud de aminoácidos, la persona experimentada en la técnica también puede tener en cuenta las así llamadas sustituciones "conservadoras" de aminoácidos, como será evidente para la persona experimentada en la técnica.

35 Las sustituciones conservadoras de aminoácidos se refieren a la posibilidad de intercambio de residuos que tengan cadenas laterales similares. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas es glicina, alanina, valina, leucina, e isoleucina; un grupo de aminoácidos que tienen cadenas laterales con hidroxilo alifático es serina y treonina; un grupo de aminoácidos que tienen cadenas laterales que contienen amida es asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas es fenilalanina, tirosina y triptófano; un grupo de aminoácidos que tienen cadenas laterales básicas es lisina, arginina e histidina, y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina.

40 Grupos de sustitución conservadora de aminoácidos preferidos son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina y asparagina-glutamina. Variantes de sustitución de las secuencias de aminoácidos aquí divulgadas son aquellas en las cuales al menos un residuo en las secuencias divulgadas ha sido eliminado y se inserta un residuo diferente en su lugar. Preferiblemente, el cambio de aminoácido es conservador. Las sustituciones conservadoras preferidas para cada uno de los aminoácidos de origen natural es como sigue: Ala a ser; Arg a lys; Asn a gln o his; Asp a glu; Cys a ser o ala; Gln a asn; Glu a asp; Gly a pro; His a asn o gln; He a leur o val; Leu a ile o val; Lys a arg; gln o glu; Met a leu o ile; Phe a met, leu o tyr; Ser a thr; Thr a ser; Trp a tyr; Tyr a trp o phe; y, Val a ile o leu.

45 Una secuencia de nucleótidos que codifica una enzima que cataliza la conversión de xilosa a xilulosa también puede definirse por su capacidad de hibridizar con las secuencias de nucleótidos que codifican la enzima que tiene la secuencia definida en SEQ ID NO: 3 o una secuencia que tiene al menos aproximadamente 70% de identidad de secuencia con la misma, bajo condiciones de hibridización moderadas, o preferiblemente estrictas.

55 Formalmente, tales secuencias de nucleótidos hibridizan con el complemento reverso de las secuencias de nucleótidos que codifican la enzima que tiene la secuencia definida en SEQ ID NO: 3 o una secuencia que tiene al

menos aproximadamente 70% de identidad de secuencia con la misma, por ejemplo secuencias que hibridizan con el complemento reverso de SEQ ID NOs: 1 o 2.

5 Las condiciones de hibridización estrictas se definen aquí como condiciones que permiten que una secuencia de ácidos nucleicos de al menos aproximadamente 25, de forma preferible aproximadamente 50 nucleótidos, 75 o 100 y lo más preferiblemente de aproximadamente 200 o más nucleótidos, hibridice a una temperatura de aproximadamente 65°C en una solución que comprende aproximadamente una sal 1 M, preferiblemente 6 x SSC (cloruro de sodio, citrato de sodio) o cualquier otra solución que tenga una fuerza iónica comparable, y lavar a 65°C en una solución que comprende aproximadamente sal 0.1 M, o menos, preferiblemente 0,2 x SSC o cualquier otra solución que tenga fuerza iónica comparable. Preferiblemente, la hibridización se lleva a cabo durante la noche, esto es durante al menos 10 horas y preferiblemente el lavado se lleva a cabo durante al menos una hora con al menos dos cambios de la solución de lavado. Estas condiciones permitirán normalmente la hibridización específica de secuencias que tengan aproximadamente 90% o más de identidad de secuencia.

15 Condiciones moderadas se definen aquí como condiciones que permiten que secuencias de ácidos nucleicos de al menos 50 nucleótidos, preferiblemente de aproximadamente 200 o más nucleótidos, hibridicen a una temperatura de aproximadamente 45°C en una solución que comprende sal aproximadamente 1 M, preferiblemente 6 x SSC o cualquier otra solución que tenga una fuerza iónica comparable, y lavado a temperatura ambiente a una solución que comprende sal a aproximadamente 1 M, preferiblemente 6 x SSC o cualquier otra solución que tenga una fuerza iónica comparable. Preferiblemente, la hibridización se lleva a cabo durante la noche, esto es al menos durante 10 horas, y preferiblemente el lavado se lleva a cabo durante al menos una hora con al menos dos cambios de la solución de lavado. Estas condiciones permitirán usualmente la hibridización específica de secuencias que tengan hasta 50% de identidad de secuencia. La persona experimentada en la técnica será capaz de modificar estas condiciones de hibridización con el fin de identificar específicamente secuencias que varían en identidad entre 50% y 90%.

25 Para incrementar la probabilidad de que la enzima introducida se exprese en forma activa en una célula de la invención, la secuencia de nucleótidos codificante correspondiente puede ser adaptada para optimizar su uso de codón al de la célula de levadura escogida. Se conocen en la técnica varios métodos para optimización de codones. Un método preferido para optimizar el uso de codón de las secuencias de nucleótidos al de la levadura es una tecnología de optimización de pares de codones como se divulga en WO2006/077258 y/o WO2008/000632. La WO2008/000632 apunta a la optimización de pares de codones. La optimización de pares de codones es un método donde las secuencias de nucleótido que codifican un polipéptido se modifican con respecto al uso de codón, en particular a los pares de codones que son utilizados, para obtener una expresión mejorada de la secuencia de nucleótidos que codifica el polipéptido y/o una producción mejorada del polipéptido codificado. Los pares de codones se definen como un conjunto de dos tripletes subsecuentes (codones) en una secuencia de codificación.

30 Como una medida simple para la expresión de genes y la eficiencia de la traducción, aquí, se utiliza el Codon Adaptation Index (CAI), tal como se describe en Xuhua Xia, Evolutionary Bioinformatics 2007, 3: 53 – 58. El índice utiliza un conjunto de referencia de genes altamente expresados a partir de una especie para establecer los méritos relativos de cada codón, y una marcación para un gen se calcula a partir de una frecuencia de uso de todos los codones en ese gen. El índice establece el grado al cual la selección ha sido efectiva en el moldeo del patrón del uso de codón. En este respecto es útil predecir el nivel de expresión de un gen, para establecer la adaptación de genes virales a sus huéspedes, y para hacer comparaciones del uso de codón en los diferentes organismos. El índice también puede dar una indicación aproximada del éxito probable de la expresión del gen heterólogo. En los genes optimizados con par de codones de acuerdo con la invención el CAI es 0.6 o más, 0.7 o más, 0.8 o más, 0.85 o más, 0.87 o más, 0.90 o más, 0.95 o más, o aproximadamente 1.0.

45 En una célula de la invención, la xilosa isomerasa es típicamente heteróloga hacia la célula. Es decir, la xilosa isomerasa tiene una secuencia que no se presenta de forma natural en la célula en cuestión como parte del organismo, célula, genoma de ADN o secuencia de ARN en el cual está presente. Es decir, la xilosa isomerasa es exógena a la célula o no se presenta de forma natural en la célula. De acuerdo con lo anterior, una secuencia de nucleótidos que codifica una xilosa isomerasa se expresa típicamente o es capaz de ser expresada en forma activa en la célula huésped transformada.

50 Una célula de la invención es por lo tanto una célula que comprende, esto es que ha sido transformada con, un constructo de ácido nucleico que comprende la secuencia de nucleótidos que codifica la xilosa isomerasa como se definió anteriormente. El constructo de ácido nucleico que comprende la xilosa isomerasa que codifica la secuencia preferiblemente es capaz de la expresión de la xilosa isomerasa de la célula huésped.

55 Los métodos para expresar una secuencia heteróloga de xilosa isomerasa en una célula son bien conocidos para los expertos en la técnica.

De acuerdo con lo anterior, una célula de la invención es una célula recombinante. Es decir, una célula de la invención comprende, o se transforma con o está genéticamente modificada con una secuencia de nucleótidos que no se presenta en forma natural en la célula en cuestión.

- Las técnicas para la expresión recombinante de la xilosa isomerasa en una célula, así como las modificaciones genéticas adicionales de una célula de la invención son bien conocidas para las personas experimentadas en la técnica. Típicamente tales técnicas involucran la transformación de una célula con un constructo de ácido nucleico que comprende la secuencia relevante. Tales métodos son conocidos, por ejemplo, a partir de manuales estándar, tales como Sambrook and Russel (2001) "Molecular Cloning: A Laboratory Manual (3rd edition), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, o F. Ausubel et al., eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987). Los métodos para transformación y modificación genética de células huésped fúngicas son bien conocidos a partir por ejemplo de EP-A- 0635 574, WO 98/46772, WO 99/60102, WO 00/37671, WO 90/14423, EP-A-0481008, EP-A-0635574 y US 6,265,186.
- La mayor parte de los plásmidos episomales o 2μ son relativamente inestables, perdiéndose en aproximadamente 10^{-2} o más células después de cada generación. Incluso bajo condiciones de crecimiento selectivo, solamente el 60% al 95% de las células retienen el plásmido episomal. El número de copias de la mayoría de los plásmidos episomales varía de 10-40 por célula de huéspedes *cir⁺*. Sin embargo, los plásmidos no se distribuyen igualmente entre las células, y hay una alta varianza en el número de copias por célula en poblaciones. Las cepas transformadas con plásmidos integrantes son extremadamente estables, aún en ausencia de presión selectiva. Sin embargo la pérdida de plásmidos puede ocurrir a aproximadamente 10^{-3} hasta 10^{-4} frecuencias por recombinación homóloga entre ADN repetido en tándem, llevando a la apertura de circuito de la secuencia de vectores. Preferiblemente, el diseño del vector en el caso de la integración estable es de tal forma que por pérdida de los genes marcadores de selección (lo que también ocurre por recombinación homóloga intramolecular) la apertura del circuito constructo integrado no es posible. Preferiblemente los genes se integran así de forma estable. La integración estable se define aquí como la integración en el genoma, donde la apertura de circuito del constructo integrado no es posible. Preferiblemente los marcadores de selección están ausentes.
- Típicamente, el constructo de ácidos nucleicos puede ser un plásmido, por ejemplo un plásmido de copia baja o un plásmido de copia alta. La célula de acuerdo con la presente invención puede comprender copias sencillas o múltiples de la secuencia de nucleótidos que codifica una xilosa isomerasa, por ejemplo por copias múltiples de un constructo de nucleótidos o por uso de un constructo que tiene copias múltiples de la secuencia de xilosa isomerasa.
- El constructo de ácidos nucleicos puede ser mantenido de forma episomal y así comprender una secuencia de replicación autónoma, tal como una secuencia de replicación autosómica. Un constructo de ácido nucleico episomal adecuado puede por ejemplo, basarse en los plásmidos de levadura 2μ o pKD1 (Gleer et al., 1991, Biotechnology 9: 968-975), o el plásmido AMA (Fierro et al., 1995, Curr Genet 29: 482-489). Como alternativa, cada constructo de ácido nucleico puede integrarse en una o más copias en el genoma de la célula. La integración en el genoma de la célula puede ocurrir aleatoriamente por recombinación no homóloga pero preferiblemente, el constructo de ácido nucleico puede integrarse en el genoma de la célula mediante recombinación homóloga como es bien conocido en la técnica (véase por ejemplo WO 90/14423, EP-A-0481008, EP-A-0635 574 y US 6,265,186).
- Típicamente, la secuencia que codifica la xilosa isomerasa estará enlazada de forma operable a una o más secuencias de ácidos nucleicos, capaces de proveer o ayudar en la transcripción y/o traducción de la secuencia de xilosa isomerasa.
- El término "enlazado de forma operativa" se refiere a una yuxtaposición donde los componentes descritos están en una relación que les permite funcionar en la forma prevista. Por ejemplo, un promotor o potenciador está enlazado de forma operativa a una secuencia de codificación de dicho promotor o potenciador que afecta la transcripción de la secuencia de codificación.
- Tal como se usa aquí, el término "promotor" se refiere a un fragmento de ácido nucleico que funciona para controlar la transcripción de uno o más genes, localizado corriente arriba con respecto a la dirección de transcripción del sitio de iniciación de la transcripción del gen, y está identificado estructuralmente mediante la presencia de un sitio de enlace para polimerasa de ARN dependiente de ADN, sitio de iniciación de la transcripción y otras secuencias de ADN conocidas para los expertos en la técnica. Un promotor "constitutivo" es un promotor que es activo bajo la mayoría de las condiciones ambientales y de desarrollo. Un promotor "inducible" es un promotor que es activo bajo regulación ambiental o de desarrollo.
- El promotor que pudiera ser utilizado para alcanzar la expresión de una secuencia de nucleótidos que codifican una enzima de acuerdo con la presente invención pueden ser no nativo para la secuencia de nucleótidos que codifica la enzima que va a ser expresada, esto es, un promotor que es heterólogo a la secuencia de nucleótidos (secuencia de codificación) a la cual está enlazado operativamente. El promotor puede, sin embargo, ser homólogo, esto es endógeno a la célula huésped.
- Promotores adecuados en este contexto incluyen tanto promotores naturales constitutivos como inducibles así como promotores manipulados genéticamente, que son bien conocidos para las personas experimentadas en la técnica. Promotores adecuados en células huésped eucariotas pueden ser *GAL7*, *GAL10*, o *GAL1*, *CYC1*, *HIS3*, *ADH1*, *PGL*, *PH05*, *GAPDH*, *ADC1*, *TRP1*, *URA3*, *LEU2*, *ENO1*, *TPI1*, y *AOX1*. Otros promotores adecuados incluyen *PDC1*, *GPD1*, *PGK1*, *TEF1* y *TDH3*.

- Una célula de la invención, el extremo 3' de la secuencia de ácidos nucleicos que codifica la xilosa isomerasa esta preferiblemente enlazado de forma operativa a una secuencia terminadora de la transcripción. Preferiblemente la secuencia terminadora es operable en una célula huésped seleccionable, tal como por ejemplo la especie de levadura seleccionable. En cualquier caso la selección del terminador no es crítica; puede ser, por ejemplo de cualquier gen de levadura, aunque los terminadores algunas veces pueden trabajar si provienen de un gen eucariota no levadura. Usualmente, una secuencia de nucleótidos que codifican la xilosa isomerasa comprende un terminador. Preferiblemente, tales terminadores están combinados con mutaciones que evitan el no sentido mediado por la disminución de ARNm en la célula huésped de la invención (véase por ejemplo: Shirley et al., 2002, Genetics 161: 1465 – 1482).
- La secuencia de terminación de la transcripción comprende de manera preferible adicionalmente una señal de poliadenilación.
- Opcionalmente, puede estar presente un marcador seleccionable en un constructo de ácido nucleico adecuado para su uso en la invención. Tal como se usa aquí, el término "marcador" se refiere a un gen que codifica un trecho o un fenotipo que permite la selección de, o escoger, una célula huésped que contiene el marcador. El gen marcador puede ser un gen de resistencia a antibióticos donde puede utilizarse el antibiótico apropiado para seleccionar las células transformadas entre células que no están transformadas. Ejemplo de marcadores de resistencia a antibióticos adecuados incluye, por ejemplo, dihidrofolato reductasa, higromicin-B-fosfotransferasa, 3'-O-fosfotransferasa II (resistencia a kanamicina, neomicina y G418). Aunque los marcadores de resistencia a los antibióticos pueden ser los más convenientes para la transformación de células huésped polipoides, preferiblemente sin embargo se utilizan marcadores de no resistencia a antibióticos, tales como marcadores auxotróficos (*URA3*, *TRP1*, *LEU2*) o el gen *S. pombe* *TPI* (descrito por Russell P R, 1985, Gene 40: 125 – 130). En una realización preferida las células huésped transformadas con los constructos del ácido nucleico son libres de genes marcadores. Los métodos para construir las células huésped microbianas libres de genes marcadores recombinantes se describen en EP-A-O 635 574 y se basan en el uso de marcadores bidireccionales tales como el gen de *A. nidulans* *amdS* (acetamidasa) y los genes de levadura *URA3* y *LYS2*. Como alternativa, puede incorporarse un marcador seleccionable tal como Proteína Verde Fluorescente, *lacL*, luciferasa, cloranfenicol acetil transferasa, beta-glucoronidasa, en los constructos de ácido nucleico de la invención permitiendo seleccionar las células transformadas.
- Pueden estar presentes elementos adicionales opcionales en los constructos de ácido nucleico adecuados para su uso en la invención que incluyen, pero no se limitan a, uno o más secuenciadores de guía, potenciadores, factores de integración y/o genes reportadores, secuencias de intrones, centrómeros, telómeros y/o secuencias de unión a matrices (MAR). Los constructos de ácidos nucleicos de la invención pueden comprender adicionalmente una secuencia para replicación autónoma, tal como una secuencia ARS.
- Preferiblemente, la xilosa isomerasa se expresa en el citosol. La expresión citosólica puede ser alcanzada mediante eliminación o modificación de una señal de objetivo mitocondrial o peroxisomal.
- Una célula de la invención puede ser cualquier célula adecuada, tal como una célula procariota, tal como una bacteria, o una célula eucariota. Típicamente, la célula será una célula eucariota, por ejemplo una levadura o un hongo filamentoso.
- Las levaduras se definen aquí como microorganismos eucariotas e incluyen todas las especies de la subdivisión Eumicotina (Alexopoulos, C. J., 1962, In: Introductory Mycology, John Wiley & Sons, Inc., New York) que crecen predominantemente en forma unicelular.
- Las levaduras pueden crecer bien por fusión de un talus unicelular o pueden crecer por fisión del organismo. Una levadura preferida como célula de la invención puede pertenecer a los géneros *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces* o *Yarrowia*. Preferiblemente la levadura es capaz de fermentación anaeróbica, más preferiblemente capaz de fermentación anaeróbica alcohólica.
- Los hongos filamentosos se definen aquí como microorganismos eucariotas que incluyen todas las formas filamentosas de la subdivisión Eumicotina. Estos hongos son caracterizados por un micelio vegetativo compuesto de quitina, celulosa y otros polisacáridos complejos.
- Los hongos filamentosos adecuados para su uso como células de la presente invención son morfológica, fisiológica y genéticamente distintos de las levaduras. Las células de hongos filamentosos pueden ser utilizadas ventajosamente ya que la mayor parte de los hongos no requieren condiciones estériles para su propagación y son insensibles a infecciones de bacteriófagos. El crecimiento vegetativo por parte de los hongos filamentosos es por elongación del hifen y el catabolismo de carbono de la mayoría de los hongos filamentosos es obligatoriamente aeróbico. Los hongos filamentosos preferidos como células huésped de la invención pueden pertenecer a los géneros *Aspergillus*, *Trichoderma*, *Humicola*, *Acremonium*, *Fusarium* o *Penicillium*. Más preferiblemente, la célula fúngica filamentosa puede ser una célula de *Aspergillus niger*, *Aspergillus oryzae*, *Penicillium chrysogenum*, o *Rhizopus oryzae*.

A lo largo de los años se han hecho sugerencias para la introducción de diversos organismos para la producción de bioetanol a partir de azúcares de cultivo. En la práctica, sin embargo, todos los procesos de producción de bioetanol principales han continuado con el uso de levaduras del género *Saccharomyces* como productores de etanol. Esto se debe a los muchos rasgos atractivos de la especie *Saccharomyces* para los procesos industriales, esto es, una tolerancia alta a los ácidos, al etanol y a la osmosis, capacidad de crecimiento anaeróbico, y desde luego su alta capacidad fermentativa alcohólica. Las especies de levadura preferidas como células huésped incluyen *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus* o *K. fragilis*.

Una célula de la invención puede ser capaz de convertir biomásas de vegetales, celulosas, hemicelulosas, pectinas, ramnosa, galactosa, fucosa, maltosa, maltodextrinas, ribosa, ribulosa o almidón, derivados del almidón, sacarosa, lactosa y glicerol, por ejemplo en azúcares fermentables. De acuerdo con lo anterior, una célula de la invención puede expresar una o más enzimas tales como celulasa (una endocelulasa o una exocelulasa), una hemicelulasa (una endo o exo xilanaso o arabinasa) necesarias para la conversión de la celulosa en monómeros de glucosa y de la hemicelulosa en monómeros de xilosa y arabinosa, una peptinasa capaz de convertir las peptinas en ácido glucurónico y ácido galacturónico o una amilasa para convertir el almidón en monómeros de glucosa.

Una célula de la invención es preferiblemente un huésped capaz de un transporte activo o pasivo de xilosa hacia la célula.

Preferiblemente, una célula de la invención:

es capaz de glicolisis activa; y/o

muestra flujo a través de la ruta de la pentosa fosfato; y/o

despliega actividad de xilulosas quinasa de manera que la xilulosa isomerizada a partir de la xilosa puede ser metabolizada hasta piruvato.

La célula adicionalmente comprende de forma preferible aquellas actividades enzimáticas requeridas para la conversión de piruvato a un producto de fermentación deseado, tal como etanol, butanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, ácido fumárico, ácido málico, ácido itacónico, un aminoácido, 1,3-propano diol, etileno, glicerol, un antibiótico de β -lactama o cefalosporina.

Una célula preferida de la invención es una célula que sea capaz de forma natural de fermentación alcohólica, preferiblemente, fermentación alcohólica anaeróbica. Una célula de la invención preferiblemente tiene una alta tolerancia al etanol, una alta tolerancia a un pH bajo (por ejemplo es capaz de crecer a un pH inferior a aproximadamente 5, a aproximadamente 4, aproximadamente 3, o aproximadamente 2.5) y hacia ácidos orgánicos tales como ácido láctico, ácido acético o ácido fórmico y/o a productos de degradación del azúcar tales como furfural e hidroximetil furfural y/o una alta tolerancia a temperaturas elevadas.

Cualquiera de las características o actividades anteriores de una célula de la invención pueden estar presentes de forma natural en la célula o pueden ser introducidas o modificadas por modificación genética.

La secuencia de nucleótidos que codifica una xilosa isomerasa se expresa típicamente o es capaz de ser expresada en forma activa en la célula huésped transformada. Así, la expresión de la secuencia de nucleótidos en la célula huésped produce una xilosa isomerasa activa, típicamente con una actividad específica de al menos aproximadamente 10 U de actividad de xilosa isomerasa por mg de proteína a aproximadamente 30°C, preferiblemente al menos a aproximadamente 20, a al menos aproximadamente 25, a al menos aproximadamente 30, al menos aproximadamente 50, al menos aproximadamente 100, al menos aproximadamente 200, al menos aproximadamente 300, al menos aproximadamente 500, al menos aproximadamente 750 o al menos aproximadamente 1.000 U por mg a aproximadamente 30°C. La actividad específica de la xilosa isomerasa expresada en la célula huésped transformada se define aquí como la cantidad de actividad de xilosa isomerasa en unidades por mg de proteína de lisado libre de la célula huésped, por ejemplo un lisado libre de célula de levadura. La determinación de la actividad de la xilosa isomerasa, cantidad de proteína y preparación del lisado libre de células son como se describe aquí. Preferiblemente, la expresión de la secuencia de nucleótidos que codifican la xilosa isomerasa en la célula huésped produce una xilosa isomerasa con un K_m para la xilosa que es menor de 50, 40, 30 o 25 mM, más preferiblemente, el K_m para la xilosa es aproximadamente 20 mM o menos.

Una célula de la invención puede comprender una o más modificaciones genéticas que incrementan el flujo de la ruta del fosfato de pentosa. En particular, las modificaciones genéticas pueden llevar a un flujo incrementado a través de la parte de la ruta de fosfato de pentosa no oxidativa. Una modificación genética que produce un flujo incrementado de la parte no oxidativa de la ruta del fosfato de pentosa se entiende aquí como una modificación que incrementa el flujo en al menos un factor de aproximadamente 1.1, aproximadamente 1.2, aproximadamente 1.5, aproximadamente 2, aproximadamente 5, aproximadamente 10 o aproximadamente 20 en comparación con el flujo en una cepa que es genéticamente idéntica excepto por la modificación genética que causa el flujo incrementado. El flujo de la parte no oxidativa de la ruta del fosfato de pentosa puede medirse haciendo crecer el huésped modificado sobre xilosa como única fuente de carbono, determinando la tasa de consumo específica de xilosa y sustrayendo la tasa de producción de xilitol específica de la tasa de consumo de xilosa específica, si se produce

glicolaldehídotransferasa. Una transcetolasa puede definirse adicionalmente por sus aminoácidos. De la misma forma una transcetolasa puede definirse por una secuencia de nucleótidos que codifique la enzima así como por una secuencia de nucleótidos que hibridice a una secuencia de nucleótidos de referencia que codifiquen una transcetolasa. La secuencia de nucleótidos que codifica para una transcetolasa se designa aquí como *TKL1*.

5 La enzima "transaldolasa" (EC 2.2.1.2) se define aquí como una enzima que cataliza la reacción: sedoheptulosa 7-fosfato + D-gliceraldehído 3-fosfato <-> D-eritrosa 4-fosfato + D-fructosa 6-fosfato y viceversa. La enzima también es conocida como dihidroxiacetona transferasa; dihidroxiacetona sintasa; formaldehído transcetolasa; o sedoheptulosa-7-fosfato: D-gliceraldehído-3-fosfato gliceron transferasa. Una transaldolasa puede definirse adicionalmente por su secuencia de aminoácidos. De la misma forma una transaldolasa puede definirse mediante una secuencia de nucleótidos que codifiquen la enzima así como con una secuencia de nucleótidos que hibridicen a una secuencia de nucleótidos de referencia que codifiquen una transaldolasa. La secuencia de nucleótidos que codifica para la transcetolasa se designa aquí como *TAL1*.

15 Diversos medios son conocidos por los expertos en la técnica para la expresión y sobreexpresión de enzimas en una célula de la invención. En particular, una enzima puede ser sobreexpresada incrementando el número de copias del gen que codifica la enzima en la célula huésped, por ejemplo, integrando copias adicionales del gen en el genoma de la célula huésped, expresando el gen por un vector de expresión multicopias episomal o introduciendo un vector de expresión episomal que comprenda múltiples copias del gen.

20 Como alternativa, la sobreexpresión de las enzimas en las células huésped de la invención puede alcanzarse utilizando un promotor que no sea nativo a la secuencia que codifique para la enzima que va a ser sobreexpresada, esto es un promotor que sea heterólogo a la secuencia de codificación a la cual esta enlazado de forma operativa. Aunque el promotor preferiblemente es heterólogo a la secuencia de codificación a la cual está enlazado de forma operativa, también se prefiere que el promotor sea homólogo, esto es endógeno a la célula huésped. Preferiblemente el promotor heterólogo es capaz de producir un nivel de estado de equilibrio más alto del transcrito que comprende la secuencia de codificación (o es capaz de producir más moléculas transcritas, esto es, moléculas de ARNm, por unidad de tiempo) que el promotor que es nativo para la secuencia de codificación, preferiblemente bajo condiciones donde la xilosa o la xilosa y la glucosa están disponibles como fuentes de carbono, más preferiblemente como fuentes de carbono principales (esto es más del 50% de la fuente de carbono disponible consiste de xilosa o xilosa y glucosa), lo más preferiblemente como fuentes únicas de carbono. Los promotores adecuados en este contexto incluyen tanto promotores constitutivos como inducibles naturales así como promotores manipulados. Un promotor preferido para su uso en la presente invención será además insensible a la represión del catabolito (glucosa) y/o preferiblemente no requerirá xilosa para su inducción. Los promotores que tienen estas características están ampliamente disponibles y son conocidos para las personas experimentadas en la técnica. Ejemplos adecuados de tales promotores incluyen por ejemplo promotores de genes glicolíticos, tales como la fosfofructoquinasa (PFK), triosa fosfato isomerasa (TPI), gliceraldehído-3-fosfato deshidrogenasa (GPD, TDH3 o GAPDH), piruvato quinasa (PYK), fosfogliceratoquinasa (PGK) como promotores de levaduras u hongos filamentosos; más detalles acerca de tales promotores provenientes de levaduras pueden ser encontrados en (WO 93/03159). Otros promotores útiles son proteínas ribosomales que codifican promotores de genes, el promotor de gen lactasa (LAC4), promotores de alcohol deshidrogenasa (ADHI, ADH4, y similares), y el promotor enolasa (ENO). Otros promotores, tanto constitutivos como inducibles, y potenciadores o secuencias de activación de corriente arriba serán conocidos para los expertos en la técnica. Los promotores utilizados en las células huésped de la invención pueden ser modificados, si se desea, para afectar sus características de control.

La secuencia de codificación utilizada para la sobreexpresión de las enzimas mencionadas más arriba puede ser preferiblemente homóloga a la célula huésped de la invención. Sin embargo, pueden usarse secuencias de codificación que sean heterólogas a la célula huésped de la invención.

45 La sobreexpresión de una enzima, cuando se refiere a la producción de la enzima en una célula huésped genéticamente modificada, significa que la enzima se produce en un nivel más alto de actividad enzimática específica en comparación con la célula huésped no modificada bajo condiciones idénticas. Usualmente esto significa que la proteína activa enzimáticamente (o proteínas en el caso de enzimas multisubunidades) se producen cantidades mayores, o mejor a un nivel de estado de equilibrio más alto en comparación con la célula huésped no modificada bajo condiciones idénticas. De la misma forma esto significa usualmente que el ARNm que codifica para la proteína enzimáticamente activa se produce en cantidades mayores, o de nuevo en un nivel de estado de equilibrio más alto en comparación con la célula huésped no modificada bajo condiciones idénticas. La sobreexpresión de una enzima se determina así preferiblemente midiendo el nivel de la actividad específica de la enzima en la célula huésped utilizando ensayos enzimáticos apropiados tal como se describen aquí. Como alternativa, la sobreexpresión de la enzima puede determinarse indirectamente cuantificando el nivel de estado de equilibrio específico de la proteína enzimática, por ejemplo, utilizando anticuerpos específicos para la enzima o cuantificando el nivel de estado de equilibrio específico del ARNm que codifica para la enzima. Esto último puede ser particularmente adecuado para enzimas de la ruta del fosfato de pentosa para la cual los ensayos enzimáticos no son fácilmente factibles como sustratos para las enzimas y no son comercialmente disponibles. Preferiblemente, en una célula huésped de la invención, una enzima que va a ser sobreexpresada es sobreexpresada por al menos un factor de aproximadamente 1.1, aproximadamente 1.2, aproximadamente 1.5, aproximadamente 2, aproximadamente 5, aproximadamente 10 o aproximadamente 20 en comparación con una cepa que sea

genéticamente idéntica excepto por la modificación genética que causa la sobreexpresión. Debe entenderse que estos niveles de sobreexpresión pueden aplicarse al nivel de estado de equilibrio de la actividad de la enzima, el nivel de estado de equilibrio de la proteína de la enzima así como el nivel de estado de equilibrio del transcripto que codifica para la enzima.

5 Una célula de la invención puede comprender una o más modificaciones genéticas para incrementar la actividad específica de la xilulosa quinasa. Preferiblemente la modificación o modificaciones genéticas producen sobreexpresión de una xilulosa quinasa, por ejemplo por sobreexpresión de una secuencia de nucleótidos que codifican una xilulosa quinasa. El gen que codifica la xilulosa quinasa puede ser endógeno a la célula huésped o puede ser una xilulosa quinasa que sea heteróloga a la célula huésped. Una secuencia de nucleótidos usada para
10 sobreexpresión de xilulosa quinasa en la célula huésped de la invención es una secuencia de nucleótidos y codifica un polipéptido con actividades de xilulosa quinasa.

La enzima "xilulosa quinasa" (EC 2.7.1.17) se define aquí como una enzima que cataliza la reacción $ATP + D\text{-xilulosa} = ADP + D\text{-xilulosa } 5 \text{ Fosfato}$. La enzima también es conocida como xiluloquinasa fosforilante, D-xiluloquinasa o ATP: D-xilulosa 5-fosfotransferasa. Una xilulosa quinasa de la invención puede ser definida
15 adicionalmente por su secuencia de aminoácidos. Probablemente una xilulosa quinasa puede ser definida por una secuencia de nucleótidos que codifica la enzima así como una secuencia de nucleótidos que hibridiza a una secuencia de nucleótidos de referencia que codifica una xilulosa quinasa.

En una célula de la invención, una modificación o modificaciones genéticas que incrementen la actividad específica de la xilulosa quinasa pueden ser combinadas con cualquiera de las modificaciones que incrementen el flujo de la
20 ruta del fosfato de pentosa como se describió más arriba. Sin embargo esto no es esencial.

Así, una célula huésped de la invención puede comprender solamente una modificación o modificaciones genéticas que incrementen la actividad específica de xilulosa quinasa. Los diversos medios disponibles en la técnica para alcanzar y analizar la sobreexpresión de una xilulosa quinasa en la célula huésped de la invención son los mismos que se describieron más arriba para enzimas de la ruta del fosfato de pentosa. Preferiblemente en las células
25 huésped de la invención, una xilulosa quinasa que va a ser sobreexpresada se sobreexpresa por al menos un factor de aproximadamente 1.1, aproximadamente 1.2, aproximadamente 1.5, aproximadamente 2, aproximadamente 5, aproximadamente 10 o aproximadamente 20 en comparación con una cepa que es genéticamente idéntica excepto por las modificaciones genéticas que causan la sobreexpresión. Debe entenderse que estos niveles de sobreexpresión pueden aplicarse al nivel de estado de equilibrio de la actividad de la enzima, el nivel de estado de equilibrio de la proteína de la enzima así como el nivel de estado de equilibrio del transcripto que codifica para la
30 enzima.

Una célula de la invención puede comprender una o más modificaciones genéticas que reduzcan la actividad no específica de la aldosa reductasa en la célula huésped. Preferiblemente, la actividad no específica de la aldosa reductasa es reducir a la célula huésped por una o más modificaciones genéticas que reducen la expresión o
35 inactivan un gen que codifica una aldosa reductasa no específica. Preferiblemente, las modificaciones genéticas reducen o inactivan la expresión de cada copia endógena de un gen que codifica una aldosa reductasa no específica en la célula huésped. Las células huésped pueden comprender copias múltiples de genes que codifican aldosas reductasas no específicas como resultado de un di, poli o aneoploide, y/o la célula huésped puede contener diferentes (iso) enzimas con actividades de aldosa reductasa que difiere en la secuencia de aminoácidos y que son
40 cada una codificada por un gen diferente. También en tales instancias preferiblemente la expresión de cada gen que codifica una aldosa reductasa no específica se reduce o inactiva. Preferiblemente, el gen es inactivado por eliminación de al menos una parte del gen o por interrupción del gen, por lo cual en este contexto el término gen también incluye cualquier secuencia no codificante corriente arriba o corriente debajo de la secuencia de codificación, la eliminación (parcial) o inactivación de las cuales resulta en la reducción de la expresión de la
45 actividad de aldosa reductasa no específica en la célula huésped.

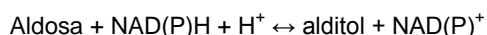
Una secuencia de nucleótidos que codifica una aldosa reductasa cuya actividad va a ser reducida en la célula huésped de la invención es una secuencia de nucleótidos que codifica un polipéptido con actividad de aldosa reductasa.

En las células huésped de la invención, la modificación genética que reduce la actividad no específica de aldosa reductasa en la célula huésped puede combinarse con cualquiera de las modificaciones incrementando el flujo de la
50 ruta del fosfato de pentosa y/o con cualquiera de las modificaciones que incrementan la actividad específica de xilulosa quinasa en la célula huésped como se describió más arriba. Sin embargo, esto no es esencial.

Así, una célula huésped de la invención que comprende solamente una modificación o modificaciones genéticas que
55 reduzcan la actividad no específica de la aldosa reductasa en la célula huésped se incluye específicamente en la invención.

La enzima "aldosa reductasa" (EC 1.1.1.21) se define aquí como cualquier enzima que sea capaz de reducir la xilosa o xilulosa a xilitol. En el contexto de la presente invención una aldosa reductasa puede ser una aldosa

reductasa no específica que sea nativa (endógena) para una célula huésped de la invención y que sea capaz de reducir la xilosa o xilulosa a xilitol. Las aldosas reductasas no específicas catalizan la reacción:



5 La enzima tiene una alta especificidad y también es conocida como aldosa reductasa; poliol dehidrogenasa (NADP⁺); alditol: NADP óxidorreductasa; alditol: NADP⁺ 1-óxidorreductasa; NADPH-aldopentosa reductasa; o NADPH- aldosa reductasa.

10 Un ejemplo particular de tal aldosa reductasa no específica que es endógena a *S. cerevisiae* y que es codificada por el gen GRE3 (Traff et al.; 2001, Appl. Environ. Microbiol. 67: 5668-74). Así, una aldosa reductasa de la invención puede ser definida adicionalmente por su secuencia de aminoácidos. De la misma forma una aldosa reductasa puede ser definida por la secuencia de nucleótidos que codifican la enzima así como una secuencia de nucleótidos que hibridiza a una secuencia de nucleótidos de referencia que codifica una aldosa reductasa.

15 Una célula de la invención puede adaptarse a la utilización de la xilosa por selección de mutantes, bien espontánea o inducida (por ejemplo por radiación o productos químicos), para el crecimiento sobre xilosa, preferiblemente sobre xilosa como única fuente de carbono, y más preferiblemente bajo condiciones anaeróbicas. La selección de mutantes puede ser llevada a cabo por técnicas que incluyen el paso en serie de cultivos, tal como lo describe por ejemplo, Kuyper et al. (2004, FEMS Yeast Res. 4: 655-664) o por cultivo bajo presión selectiva en un cultivo quimiostático. Una célula huésped preferida de la invención al menos una de las modificaciones genéticas descritas más arriba, incluyendo modificaciones obtenidas por selección de mutantes, confiere a la célula huésped la capacidad de crecer sobre xilosa como fuente de carbono, preferiblemente como única fuente de carbono, y preferiblemente bajo condiciones anaeróbicas. Preferiblemente la célula huésped modificada no produce esencialmente xilitol, por ejemplo, el xilitol producido está por debajo del límite de detección o por ejemplo menos de aproximadamente 5, aproximadamente 2, aproximadamente 1, aproximadamente 0.5, o aproximadamente 0.3 del carbono consumido sobre una base molar.

25 Una célula de la invención puede tener la capacidad de crecer sobre xilosa como única fuente de carbono a una tasa de al menos aproximadamente 0.05, aproximadamente 0.1, aproximadamente 0.2, aproximadamente 0.25 o aproximadamente 0.3 h⁻¹ en condiciones aeróbicas, o, si es aplicable, a una tasa de al menos aproximadamente 0.03, aproximadamente 0.05, aproximadamente 0.07, aproximadamente 0.08, aproximadamente 0.09, aproximadamente 0.1, aproximadamente 0.12, aproximadamente 0.15 o aproximadamente 0.2 h⁻¹ bajo condiciones anaeróbicas. Preferiblemente, la célula huésped modificada tiene la capacidad de crecer sobre una mezcla de glucosa y xilosa (en una relación en peso 1:1) como única fuente de carbono a una tasa de al menos aproximadamente 0.05, aproximadamente 0.1, aproximadamente 0.2, aproximadamente 0.25 o aproximadamente 0.3 h⁻¹ bajo condiciones aeróbicas, o si es aplicable, a una tasa de al menos aproximadamente 0.03, aproximadamente 0.05, aproximadamente 0.1, aproximadamente 0.2, aproximadamente 0.15, o aproximadamente 0.2 h⁻¹ bajo condiciones anaeróbicas.

35 Una célula de la invención puede tener una tasa de consumo de xilosa específica de al menos aproximadamente 200, aproximadamente 250, aproximadamente 300, aproximadamente 346, aproximadamente 350, aproximadamente 400, aproximadamente 500, aproximadamente 600, aproximadamente 750 o aproximadamente 1,000 mg de xilosa/g célula/h. Una célula de la invención puede tener un rendimiento de producto de fermentación (tal como etanol) sobre xilosa que es al menos aproximadamente 40, aproximadamente 50, aproximadamente 55, aproximadamente 55, aproximadamente 60, aproximadamente 70, aproximadamente 80, aproximadamente 85, aproximadamente 90, aproximadamente 95, aproximadamente 98 o aproximadamente 99% del rendimiento de la célula huésped de producto de fermentación (tal como etanol) sobre glucosa. Más preferiblemente, el rendimiento de un producto de fermentación (tal como etanol) de una célula de la invención sobre xilosa puede ser igual al rendimiento de la célula en producto de fermentación (tal como etanol) sobre glucosa. De la misma forma, el rendimiento en biomasa de la célula sobre xilosa puede ser al menos aproximadamente 40, aproximadamente 50, aproximadamente 55, aproximadamente 60, aproximadamente 70, aproximadamente 80, aproximadamente 85, aproximadamente 90, aproximadamente 95, aproximadamente 98 o aproximadamente 99% del rendimiento de biomasa de la célula huésped sobre glucosa. Más preferiblemente, el rendimiento de biomasa de la célula sobre xilosa puede ser igual al rendimiento de biomasa de la célula huésped sobre glucosa. Se entiende que en comparación con los rendimientos de glucosa y xilosa ambos rendimientos se comparan bajo condiciones aeróbicas o ambos bajo condiciones anaeróbicas.

Una célula de la invención puede ser capaz de utilizar arabinosa. Una célula de la invención puede, por lo tanto, ser capaz de convertir la L-arabinosa en L-ribulosa y/o xilulosa 5-fosfato y/o en un producto de fermentación deseado, por ejemplo uno de los mencionados aquí.

50 Los organismos, por ejemplo cepas de *S. cerevisiae*, capaces de producir etanol a partir de L-arabinosa pueden ser producidos modificando una célula que introduzca los genes *araA* (L-arabinosa isomerasa), *araB* (L-ribuloquinasa) y *araD* (L-ribulosa-5-P4-epimerasa) a partir de una fuente adecuada. Tales genes pueden ser introducidos en una célula de la invención con el fin de que sea capaz de utilizar la arabinosa. Tal modalidad se describe en WO 2003/095627.

Una célula de la invención puede ser una célula adecuada para la producción de etanol. Una célula de la invención sin embargo, puede ser adecuada para la producción de productos de fermentación o diferentes al etanol. Tales productos de fermentación no etanólica incluyen en principio cualquier producto químico a granel o fino que sea producible por microorganismos eucarióticos tales como una levadura o un hongo filamentoso.

5 Tales productos de fermentación pueden ser, por ejemplo, butanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, ácido málico, ácido fumárico, ácido itacónico, un aminoácido, 1,3-propanodiol, etileno, glicerol, un antibiótico de β -lactama o una cefalosporina. Una célula huésped modificada preferida de la invención para la producción de productos de fermentación no etanólicos es una célula huésped que contiene una modificación genética que da como resultado una actividad de alcohol deshidrogenasa disminuida.

10 En un aspecto adicional la invención se relaciona con procesos de fermentación en los cuales las células huésped modificadas de la invención se utilizan para la fermentación de una fuente de carbono que comprende una fuente de xilosa, tal como xilosa. Además de una fuente de xilosa la fuente de carbono en el medio de fermentación también pueden comprender una fuente de glucosa. La fuente de xilosa o glucosa puede ser xilosa o glucosa como tales o puede ser cualquier carbohidrato oligo o polimérico que comprende unidades de xilosa o glucosa, tal como, por
15 ejemplo, linocelulosa, xilanos, celulosa, almidón y similares. Para la liberación de unidades de xilosa o glucosa a partir de tales carbohidratos, pueden añadirse carbohidrasas adecuadas (tales como xilanasas, glucanasas, amilasas y similares) al medio de fermentación o pueden ser producidos por la célula huésped modificada. En este último caso la célula huésped modificada puede ser manipulada genéticamente para producir y excretar tales carbohidrasas. Una ventaja adicional del uso de fuentes oligo o poliméricas de glucosa es que permite mantener una
20 concentración baja (más baja) de glucosa libre durante la fermentación, por ejemplo, utilizando cantidades que limitan la tasa de las carbohidrasas. Así, a su vez, se prevendrá la represión de sistemas requerida para el metabolismo y transporte de azúcares diferentes a la glucosa tales como la xilosa.

En un proceso preferido la célula huésped modificada fermenta tanto la xilosa como la glucosa, preferiblemente en forma simultánea en cuyo caso preferiblemente se utiliza una célula huésped modificada que sea insensible a la
25 represión de la glucosa para prevenir el crecimiento diauxico. Además de una fuente de xilosa (y glucosa) como fuente de carbono, el medio de fermentación comprenderá adicionalmente el ingrediente apropiado requerido para el crecimiento de la célula huésped modificada. Las composiciones de medios de fermentación para el crecimiento de microorganismos tales como levaduras son bien conocidas en la técnica. El proceso de fermentación es un proceso para la producción de un producto de fermentación tal como por ejemplo etanol, butanol, ácido láctico, ácido 3-
30 hidroxipropiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, ácido málico, ácido fumárico, ácido itacónico y un aminoácido, 1,3-propanodiol etileno, glicerol, un antibiótico de β -lactama, tal como penicilina G o penicilina V y derivados de fermentación de los mismos, y una cefalosporina.

El proceso de fermentación puede ser un proceso de fermentación aeróbico o anaeróbico. Un proceso de fermentación anaeróbico se define aquí como un proceso de fermentación llevado a cabo en la ausencia de oxígeno
35 o en el cual sustancialmente no se consume oxígeno, preferiblemente menos de aproximadamente 5, aproximadamente 2.5 o aproximadamente 1 mmol/L/h, más preferiblemente se consume 0 mmol/L/h (esto es el consumo de oxígeno no es detectable), y donde las moléculas orgánicas sirven tanto como donores de electrones como aceptores de electrones. En ausencia de oxígeno, la NADH producida en la glicolisis y formación de biomasa, no puede ser oxidada por fosforilación oxidativa. Para resolver este problema muchos microorganismos usan
40 piruvato o uno de sus derivados como aceptor de electrones e hidrógeno regenerándose por lo tanto NAD^+ .

Así, en un proceso de fermentación anaeróbico preferido se utiliza piruvato como aceptor de electrones (y de hidrógeno) y se reduce a productos de fermentación tales como etanol, butanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, ácido málico, ácido fumárico, y un aminoácido, 1,3-propanodiol, etileno, glicerol, un antibiótico de β -lactama, y una cefalosporina.

45 El proceso de fermentación se lleva a cabo preferiblemente a una temperatura que sea óptima para la célula huésped modificada. Así, para la mayoría de las células huésped de levadura o fúngicas, el proceso de fermentación se lleva a cabo a una temperatura que es menor de aproximadamente 42°C, preferiblemente menor de aproximadamente 38°C. Para células huésped de levadura o fúngicas filamentosas, el proceso de fermentación se
50 lleva a cabo preferiblemente a una temperatura que es inferior a aproximadamente 35, aproximadamente 33, aproximadamente 30 o aproximadamente 28°C y a una temperatura que es superior a aproximadamente 20, aproximadamente 22 o, aproximadamente 25°C.

Un proceso preferido es un proceso para la producción de etanol, por lo cual el proceso comprende las etapas de:
(a) fermentar un medio que contiene una fuente de xilosa con una célula huésped modificada como se definió más
55 arriba, por lo cual la célula huésped fermenta la xilosa a etanol; y opcionalmente, (b) recuperar el etanol. El medio de fermentación también puede comprender una fuente de glucosa que también es fermentada a etanol. En el proceso la productividad de etanol volumétrica es preferiblemente al menos aproximadamente 0.5, aproximadamente 1.0, aproximadamente 1.5, aproximadamente 2.0, aproximadamente 2.5, aproximadamente 3.0, aproximadamente 5.0 o aproximadamente 10.0 g de etanol por litro por hora. El rendimiento en etanol sobre xilosa y/o glucosa en el proceso preferiblemente es aproximadamente al menos 50, aproximadamente 60, aproximadamente 70, aproximadamente

80, aproximadamente 90, aproximadamente 95 o aproximadamente 98%. El rendimiento en etanol se define aquí como un porcentaje de rendimiento máximo teórico.

La invención también se relaciona con un proceso para producir un producto de fermentación, siendo seleccionado tal producto del grupo consistente de butanol, ácido láctico, ácido 3-hidroxi propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, ácido málico, ácido fumárico, ácido itacónico, un aminoácido, 1,3-propanodiol, etileno, glicerol, un antibiótico de β -lactama y una cefalosporina. El proceso comprende preferiblemente fermentar un medio que contiene una fuente de xilosa con una célula huésped modificada como se definió más arriba, mediante lo cual la célula huésped fermenta la xilosa al producto de fermentación.

La invención también provee un proceso para producir un producto de fermentación, tal como un producto seleccionado del grupo consistente de etanol, butanol, ácido láctico, ácido 3-hidroxi propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, ácido málico, ácido fumárico, ácido itacónico, un aminoácido, 1,3-propanodiol, etileno, glicerol, un antibiótico de β -lactama y una cefalosporina. El proceso comprende preferiblemente la fermentación de un medio que contiene al menos una fuente de xilosa y una fuente de L-arabinosa con una célula como se definió más arriba la cual es capaz de utilizar tanto xilosa como L-arabinosa de forma tal que la célula fermenta la xilosa y la L-arabinosa hasta el producto de fermentación.

La invención también proporciona un proceso para producir un producto de fermentación, tal como un producto seleccionado del grupo consistente de etanol, butanol, ácido láctico, ácido 3-hidroxi propiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, ácido málico, ácido fumárico, ácido itacónico, un aminoácido, 1,3-propanodiol, etileno, glicerol, un antibiótico de β -lactama y una cefalosporina. El proceso comprende preferiblemente la fermentación de un medio que contiene al menos una fuente de xilosa y una fuente de L-arabinosa con una célula como se definió más arriba y una célula capaz de usar L-arabinosa, mediante el cual cada célula fermenta xilosa y/o arabinosa hasta el producto de fermentación.

Un proceso de la invención puede comprender también la recuperación del producto de fermentación. El medio con el cual el proceso se lleva a cabo también puede contener una fuente de glucosa.

El proceso de acuerdo con la presente invención puede transcurrir bajo condiciones aeróbicas y anaeróbicas. Preferiblemente, el proceso se lleva a cabo bajo condiciones microaerófilas o limitadas en oxígeno.

Un proceso de fermentación anaeróbico se define aquí como un proceso de fermentación que transcurre en ausencia de oxígeno o en el cual no se consume sustancialmente oxígeno, preferiblemente menos de aproximadamente 5, aproximadamente 2.5 o aproximadamente 1 mmol/L/h, y donde las moléculas orgánicas sirven tanto como donores de electrones como aceptores de electrones.

Un proceso de fermentación limitado en oxígeno es un proceso en el cual el consumo de oxígeno se limita por la transferencia de oxígeno desde el gas hasta el líquido. El grado de limitación del oxígeno se determina por la cantidad y composición del flujo de gas entrante así como las propiedades de mezclado/transferencia de masa reales del equipo de fermentación usado. Preferiblemente, en un proceso bajo condiciones limitadas en oxígeno, la tasa de consumo de oxígeno es al menos aproximadamente de 5.5, más preferiblemente al menos 6, tal como al menos 7 mmol/L/h.

Los siguientes ejemplos ilustran la invención:

Ejemplos

A menos que se indique otra cosa, los métodos utilizados son técnicas bioquímicas estándar. Ejemplos de libros de texto de metodología general adecuados incluyen Sambrook et al., Molecular Cloning, a Laboratory Manual (1989) y Ausubel et al., Current Protocols in Molecular Biology (1995), John Wiley & Sons, Inc.

Actividad de la xilosa isomerasa (según se determina en los ejemplos 1 y 2)

La actividad de la xilosa isomerasa puede probarse a 37°C en una mezcla de reacción que contiene tampón de fosfato 50 mM (pH 7.0), xilosa 10 mM, MgCl₂ 10 mM y una cantidad adecuada de extracto libre de células. La cantidad de xilulosa formada puede determinarse mediante el método de cisteína-carbazole (Goldstein y McCusker, Yeast 15, 1541-1553, 1999). Como alternativa, la actividad de la xilosa isomerasa se prueba a 30°C utilizando la prueba enzimática de Kesters-Hildersson et al. (Kinetic characterization of D-xylose isomerases by enzymatic assays using D-sorbitol dehydrogenase. Enz. Microb. Technol. 9 (1987) 145-148. La actividad in vitro de la xilosa isomerasa en los extractos libres de células de cepas de *S. cerevisiae* transformadas dependen de los cationes bivalentes (Mg²⁺ o Co²⁺).

Transformación de *S. cerevisiae*

La transformación de *S. cerevisiae* se llevó a cabo como lo describe Gietz y Woods (2002; Transformation of the yeast by the LiAc/SS carrier DNA/PEG method. Methods in Enzymology 350: 87-96).

PCR de colonia

Un aislado de colonia individual fue tomado con un palillo plástico y resuspendido en 50 µl de agua milliQ. La muestra fue incubada durante 10 minutos a 99°C. Se utilizaron 5µl de la muestra incubada como patrón para la reacción de PCR, utilizando Phusion® DNA polimerasa (Finnzymes) de acuerdo con las instrucciones suministradas por el proveedor.

Condiciones de reacción de PCR

Etapa 1	3´	98°C	
Etapa 2	10´´	98°C	
Etapa 3	15´´	58°C	repetir la etapa 2 a 4 durante 30 ciclos
Etapa 4	30´´	72°C	
Etapa 5	4´	72°C	
Etapa 6	30´´	20°C	

Pretratamiento de muestra para determinaciones de la actividad de la xilosa isomerasa (en general aquí y en el ejemplo 3).

Se añadieron 0.5 ml de tampón MOPS 0.1 M (pH 7.5) a la pella de células de un cultivo llevado a cabo durante la noche. Las células fueron resuspendidas y transferidas a un tubo Eppendorf de 2 ml que ya contenía 0.5 g de perlas de vidrio con un diámetro de 0.4 – 0.5 mm. Todas las muestras fueron agitadas vigorosamente en el agitador de tubos Eppendorf (IKA VIBRAX-VXR) durante 20 minutos a 4°C, a velocidad máxima. El extracto fue centrifugado durante 5 minutos a 14.000 rpm y 4°C. El sobrenadante, que es el extracto libre de células, fue transferido a un tubo de Eppendorf nuevo.

Condiciones de ensayo de la prueba de actividad de la xilosa isomerasa (en general aquí y como se determina en el ejemplo 3).

El siguiente método es una versión modificada del método descrito por Dische-Borenfreud (J. Biol. Chem. (1951) 192, 2, 583 – 587). Un (1.0) ml de la mezcla de sustrato (MOPS 100 mM pH 7.5, MgCl₂ 10 Mm, D-xilosa 10 mM) se mezcló con 50 µl (diluidos) de extracto libre de células, en duplicado, sobre hielo. A continuación los tubos de reacción fueron colocados en un baño de agua a 50°C durante 30 minutos. Además, las reacciones fueron llevadas a cabo a 30°C, también en duplicado. La reacción fue detenida colocando los tubos de reacción sobre hielo con agua, seguido por la adición de 0.2 ml de clorhidrato monohidratado de L-cisteína al 1.67% (Merck) en solución. La mezcla se mezcló bien mediante sometimiento a vórtex. A continuación, se añadieron 6 ml de solución de H₂SO₄ (190 ml de agua con 450 ml de H₂SO₄ concentrado al 95 – 97%), seguido inmediatamente por 0.2 ml de carbazol al 0.12% (p/v) (Merck), disuelto en etanol. Esta mezcla final fue mezclada bien mediante vórtex y se dejó a temperatura ambiente durante 60 minutos. Se midió la absorción a 560 nm utilizando cubetas plásticas.

La D(+)-fructosa, que también es una cetosa, se usó como referencia. Con este fin, se pesaron aproximadamente 1.000 mg de D-fructosa exactamente y se disolvieron en tampón MOPS 0.1 M, pH 7.5 en un matraz volumétrico de 50 ml. Se llevó a cabo una serie de diluciones que variaron desde aproximadamente 2 hasta 20 µmol/ml. Se utilizaron 50 µl de estas soluciones de fructosa en la prueba como se describió más arriba y se usó la absorción a 560 nm para hacer una curva de calibración. La actividad de las muestras fue calculada relacionando la absorbancia a 560 nm con la curva de calibración.

La concentración en proteína de la muestra se determinó de acuerdo con un protocolo modificado del método de Bradford, utilizando la prueba de Coomassie Plus Proten (Thermo Scientific). La actividad específica de la xilosa isomerasa se expresa como nmol/mg de proteína.min.

Ejemplo 1**Expresión de xilosa isomerasa a partir de *Bacteroides uniformis* ATCC 8492 en *Saccharomyces cerevisiae*****1.1.1 Construcción del vector de expresión de xilosa isomerasa.**

La xilosa isomerasa [E.C. 4.2.1.2], número de acceso GenBank AAYH02000036 (SEQ ID NO: 1) de *Bacteroides uniformis* ATCC 8492 fue analizado en cuanto al uso de codón. El uso de codón fue optimizado como se describió en WO 2006/077258 y WO 2008/000632 (SEQ ID NO: 2).

El gen de acuerdo con SEQ ID NO: 2 fue clonado frente al promotor *TPI1* de *S. cerevisiae*. Con el fin de prevenir una expresión potencial ineficiente de la xilosa isomerasa, se colocó la siguiente secuencia al frente de la secuencia de codificación:

ACTAGTAAAAACACATACATAAACTAAAAATG,

Mostrando subrayado el codón de inicio.

Se introdujo un sitio de restricción SpeI ACTAGT en el promotor fuerte constitutivo TPI1, cambiando la secuencia

TCTTGCTTAAATCTATAACTACAAAAACACATACATAAACTAAAAATG.

5 (promotor *TPI1* original) por

TCTTGCTTAAATCTATA**ACTAGT**AAAAACACATACATAAACTAAAAATG.

Esto permitir enlazar de manera operativa la secuencia que codifica la xilosa isomerasa optimizada mediante el codón al promotor *TPI1*.

10 Además, el codón de terminación TAA fue cambiado en TAAG, el cual es el codón de terminación más eficiente en levaduras. Los sitios de restricción convenientes fueron añadidos para facilitar la clonación. La secuencia es sintetizada por GeneArt AG (Regensburg, Alemania).

El constructo de expresión final de levadura pYISIT4-XKS1-xylA (Baun CpO) se define como en la Figura 1.

1.2 Transformación de levaduras

15 La cepa *S. cerevisiae* CEN.PK113-7D (MATa URA3 HIS3 LEU2 TRP1 MAL2-8 SUC2) y un derivado de CEN.PK113-7D, en el cual el gen GRE3 fue reemplazado por los genes de la parte no oxidativa de la ruta del fosfato de pentosa (véase más arriba) (MATa URA3 HIS3 LEU2 TRP1 MAL2-8 SUC2 GRE3::[TPI1p-TAL1_ADH1p-TKL1_PGI1p-RPE1_ENO1p-RKI1]) se transforman con el constructo pYISIT4-XKS1-xylA (Baun CpO). Las mezclas de transformación se colocan sobre una Base de Levadura Carbono (YCB) sin sulfato de amonio (Difco), KPi 40 mM (pH 6.8) y acetamida 5 mM. Las células no transformadas no pueden crecer en este medio.

20 Los transformantes se caracterizan utilizando técnicas de PCR y/o técnicas de Southern blotting.

Ejemplo 2

Crecimiento de cepas de levadura transformadas sobre xilosa

2.1 Composición del medio

25 Experimentos de crecimiento: las cepas de *Saccharomyces cerevisiae* se cultivan sobre un medio que tiene la siguiente composición: 0.67% (p/v) de base de nitrógeno de levadura y glucosa, galactosa o xilosa, o una combinación de estos sustratos (véase más abajo). Para placas de agar el medio se suplementa con agar bacteriológico al 2% (p/v).

30 Producción de etanol: se llevaron a cabo cultivos en matraces con agitación a 30°C en un medio sintético (Verduyn et al., *Yeast* 8: 501-517, 1992). El pH del medio se ajustó a 6.0 con KOH con 2 M antes de la esterilización. Para el medio sintético sólido se añadió 1.5% de agar.

35 Los precultivos fueron preparados inoculando 100 ml de medio que contenía el azúcar apropiado en un matraz de agitación de 500 ml con un cultivo de reserva congelado. Después de la incubación a 30°C en un agitador orbital (200 rpm), este cultivo fue utilizado para inocular cultivos en matraces de agitación. El medio sintético para cultivo anaeróbico fue suplementado con 0.01 g l⁻¹ de ergosterol y 0.42 g l⁻¹ de Tween 80 disuelto en etanol (Andreasen and Stier. *J. Cell Physiol.* 41: 23-36, 1953; and Andreasen and Stier. *J. Cell Physiol.* 43: 271-281, 1954).

2.2 Experimentos de crecimiento

40 La cepa de *Saccharomyces cerevisiae* CEN.PK113-7D o el derivado que constitutivamente expresa el PPP (véase ejemplo 1), transformado con pYISIT4-XKS1-xylA (Baun CpO), se hizo crecer sobre placas de agar con glucosa al 2% como fuente de carbono. Cuando las colonias se hicieron visibles, las colonias individuales se utilizaron para inocular un medio líquido con xilosa 100mM, glucosa 100mM y galactosa 100mM como fuentes de carbono, o combinaciones de las mismas. El crecimiento se monitorizó por medición del incremento en la densidad óptica a 600 nm en un espectrofotómetro LKB Ultrospec K.

2.3. Producción de etanol

45 La cepa de *Saccharomyces cerevisiae* CEN.PK113-7D o el derivado que expresa constitutivamente el PPP (véase ejemplo 1), transformada con pYISIT4-XKS1-xylA (Baun CpO), se hace crecer sobre placa de agar con glucosa al 2% como fuente de carbono. Cuando las colonias se hicieron visibles, se utilizaron colonias individuales para inocular un medio sintético (Verduyn et al., supra). Se añadieron mezclas de glucosa, xilosa y/o galactosa al medio como fuente de carbono, variando de 0 a 50 gramos por litro. El crecimiento se monitorizó midiendo el incremento

en densidad óptica a 600 nm en un espectrofotómetro LKB Ultrospec K. La producción de etanol y el consumo de azúcar en el tiempo se monitorizaron por análisis de HPLC y/o RMN.

Ejemplo 3

3.1 Introducción de cuatro genes expresados constitutivamente de la ruta no oxidativa del fosfato de pentosa

Se obtuvo *Saccharomyces cerevisiae* BIE104P1, que expresaba los genes *TAL1*, *TKL1*, *RK11* y *RPE1* constitutivamente, transformando CEN.PK113-7D (*MATa URA3 HIS3 LEU2 TRP1 MAL2-8 SUC2*) con el plásmido pPWT080 (figura 2). En un alto grado, el plásmido pPWT080 fue construido utilizando ADN sintético, sintetizado por GeneArt AG (Regensburg, Alemania). La secuencia del plásmido pPWT080 se define en SEQ ID 4. Para resumir, el plásmido pPWT080 consiste en la región promotora del gen GRE3, seguida por los cuatro genes PPP, *TAL1*, *TKL1*, *RK11* y *RPE1* bajo control de promotores fuertemente constitutivos, y las secuencias 3' no codificantes del gen GRE3, como se representa en la figura 2. Como marcadores seleccionables, el gen *kanMX* que confiere resistencia a G418 y el gen *Aspergillus amdS* que permite que los transformantes crezcan en acetamida como fuente única de nitrógeno están presentes en este plásmido. Por integración, seguida por recombinación intramolecular, los marcadores se pierden y la integración de este constructo lleva a la inactivación de la región codificante del gen GRE3 y la sobreexpresión de los genes *TAL1*, *TKL1*, *RPE1* y *RK11*.

Antes de la transformación de CEN.PK113-7D, el pPWT080 fue linealizado usando la enzima de restricción Sfi1 (New England Biolabs), de acuerdo con las instrucciones suministradas por el proveedor. Las mezclas de transformación se colocaron sobre placa en YPD (por litro: 10 gramos de extracto de levadura, 20 gramos por litro de peptona, 20 gramos por litro de dextrosa, 20 gramos de agar) que contenían 100 µg de G418 (Sigma Aldrich) por ml.

Después de dos a cuatro días, aparecieron las colonias sobre las placas, mientras que el control negativo (esto es sin adición de ADN en el experimento de transformación) dio como resultado placas de YPD/G418 en blanco.

La integración del plásmido pPWT080 está dirigida al locus GRE3. Los transformantes fueron caracterizados utilizando técnicas de PCR y Southern blotting.

Las reacciones de PCR, que son indicadoras de la integración correcta de una copia del plásmido pPWT080, se llevaron a cabo con los cebadores indicados por SEQ ID 5 y 6, y 6 y 7 (véase figura 3). Con el primer par de SEQ ID 5 y 6, se revisó la integración correcta en el locus GRE3. Si el plásmido pPWT080 estaba integrado en copias múltiples (integración de cabeza a cola), el primer par de SEQ ID 6 y 7 dará un producto PCR. Si el último producto de PCR está ausente, esto es indicativo de una integración en una copia.

Con el fin de verificar la integración correcta en una copia en los transformantes identificados como tales utilizando la técnica de PCR antes descrita, se llevó a cabo un análisis de Southern blot. Con este fin, el ADN de cromosómico fue aislado de la cepa tipo silvestre CEN.PK113-7D y con técnicas transformantes utilizando técnicas biología molecular estándar. El ADN cromosómico fue digerido con las enzimas de restricción XcmI y PstI, por electroforesis sobre un gel de agarosa al 0.7% y el ADN fue transferido a una membrana de nailon (Hybond N+, Amersham Pharmacia Biotech) de acuerdo con las instrucciones del fabricante.

Como sonda para detectar la integración correcta del plásmido pPWT080, se usó una sonda derivada del gen *RK11* presente en el plásmido pPWT080. La sonda fue hecha utilizando los cebadores de SEQ ID 9 y 10 y el plásmido pPWT080 como patrón. La marcación de la sonda y la hibridización y los procedimientos de lavado posteriores fueron llevados a cabo de acuerdo a como sugiere el proveedor del ECL Direct Labeling and Detection System (GE Life Sciences).

El autorradiograma, como se presenta en la figura 4, muestra una integración correcta de una copia del plásmido pPWT080, de acuerdo con el patrón de hibridización esperado como puede deducirse a partir de la figura 3 (panel c). La cepa fue designada como BIE104F1.

Con el fin de ser capaces de introducir los genes que codifican la xilosa isomerasa y la xiluloquinasa (sección 3.2), es necesario eliminar los marcadores de selección introducidos por la integración del plásmido pPWT080. El diseño del plásmido pPWT080 fue tal, que por integración del pPWT080 en el cromosoma, las secuencias homólogas están en cercana proximidad una de otra. Este diseño permite que los marcadores seleccionables se pierdan por recombinación intramolecular espontánea de estas regiones homólogas. La eliminación de los marcadores a partir de la cepa da como resultado una cepa libre de marcador que es más estable en su uso, que una cepa que contenga marcadores. Más específicamente, la región promotora del gen GRE3 y la región no codificadora 3' del gen GRE3 duplican después de la integración de una copia de pPWT080 y el locus GRE3 de *S. cerevisiae*. Por crecimiento vegetativo, tomará lugar la recombinación intramolecular, aunque a baja frecuencia. La frecuencia de esta recombinación depende de la longitud de la homología y el locus en el genoma (resultados no publicados). Por transferencia secuencial de una subfracción del cultivo a medio fresco se acumularán recombinantes intramoleculares con el tiempo.

Con este propósito la cepa BIE104F1 fue cultivada en YPD con glucosa al 2%, partiendo de un aislado de colonias. Se usaron 25 µl de un cultivo durante la noche para inocular un medio fresco de YPD con glucosa al 2%. Después de cinco transferencias en serie, se determinó la densidad óptica del cultivo y las células fueron diluidas a una concentración de aproximadamente 5.000 por ml. Se colocaron 100 µl de la suspensión de células sobre una placa sobre medio Yeast Carbon Base (Difco) que contenían KPi 30 mM (pH 6.8), (NH₄)₂SO₄ al 0.1%, fluoroacetamida 40 nM (Amersham) y agar al 1.8% (Difco). Las células idénticas a las células de las cepas BIE104F1, es decir sin recombinación intracelular, contenían aún el gen amdS. Para estas células la fluoroacetamida es tóxica. Estas células no serán capaces de crecer y no formarán colonias sobre un medio que contenga fluoroacetamida. Sin embargo, si se ha presentado la recombinación intramolecular, las variantes de BIE104F1 que hayan perdido sus marcadores seleccionables serán capaces de crecer sobre el medio de fluoroacetamida, puesto que son incapaces de convertir la fluoroacetamida en compuestos inhibidores del crecimiento. Estas células formarán colonias sobre este medio de agar.

Las colonias resistentes a la fluoroacetamida así obtenidas fueron sometidas a un análisis de PCR utilizando cebadores de SEQ ID 5 y 6, y 7 y 8. Los cebadores de SEQ ID 5 y 6 darán una banda si la recombinación de los marcadores seleccionables ha tomado lugar como se pretendía, como se representan en la figura 5. Como resultado, la región codificadora del gen GRE3 es reemplazada por los cuatro genes TKL1, TAL1, RKI1, y RPE1. En este caso, una reacción de PCR que utiliza los cebadores de SEQ ID 7 y 8 no daría como resultado un producto de PCR, puesto que el cebador 7 actúa como iniciador en una región que debería ser recombinada hacia afuera (véase figura 3, panel b). Si se obtiene una banda con estos cebadores, esto es indicativo de la presencia del plásmido pPWT080 completo en el genoma, de manera que no ha tenido lugar la recombinación.

Si los cebadores de SEQ ID 5 y 6 no dan como resultado un producto de PCR, la recombinación ha tenido lugar, pero de tal manera que el plásmido completo pPWT080 se ha recombinado fuera del genoma. No solo donde los marcadores seleccionables se pierden, sino también los cuatro genes PPP. En efecto, se ha recuperado la levadura tipo silvestre.

Los aislados que exhibían los resultados de PCR esperados, fueron sometidos a análisis por Southern blot (vide supra). El resultado se presenta en la figura 4. Una de las cepas que mostró el patrón correcto de bandas sobre el Southern blot (como puede deducirse a partir de la figura 3) es la cepa designada como BIE104P1.

3.2. Introducción de genes expresados constitutivamente que codifican xilosa isomerasa y xiluloquinasa

El plásmido pYISIT4-XKS1-xylA (Baun CpO), como se representa en la figura 1, fue mejorado con el fin de permitir la selección de G418 de los transformantes. Con este fin, un inserto 4630 bp que contenía el gen xylA bajo el control del promotor TPI1 y el gen XKS1 bajo control del promotor TDH1 fue escindido a partir del plásmido pYISIT4-XKS1-xylA (Baun) (figura 1), utilizando las enzimas de restricción MluI y SacII.

El plásmido pYI#SIT4, como se representa en la figura 6, fue digerido con la enzima de restricción Acc65I.

El marcador de resistencia a la kanamicina (kanMX) presente en el plásmido p427TEF (Dualsystems Biotech AG), que permite la selección en *E. coli* (kanamicina) y *S. cerevisiae* (G418) fue aislado por PCR utilizando cebadores de SEQ ID 11 y 12. La secuencia del cebador de SEQ ID 12 fue diseñada de tal manera que el sitio MluI en el fragmento kanMX se perdió, lo que mantiene el sitio MluI en el plásmido resultante (pPWT007, véase más abajo) único. El producto de PCR fue subclonado en el vector PCRII-TOPO usando el kit de clonación Zero Blunt® TOPO PCR para subclonación (Invitrogen). Se utilizaron los clones correctos para escindir el marcador de resistencia kanMX utilizando la enzima de restricción Acc65I. La ligación de este fragmento con el plásmido digerido pYI#SIT4 dio como resultado pPWT007, el cual se representa en la figura 7.

El plásmido pPWT007 fue escindido con las enzimas de restricción MluI y SacII. Después de la limpieza de este vector, el fragmento antes descrito 4630 bp MluI-SacII del pYISIT4-XKS1-xylA (Baun) fue ligado. El plásmido resultante se denomina pPWT042, el cual se representa en la figura 8.

La cepa BIE104P1 (*MATa URA3 HIS3 LEU2 TRP1 MAL2-8 SUC2 ΔGRE3::[TPI1p-TAL1-ADH1p-TKL1-PGI1p-RPE1-ENO1p-RKI1]*) (véase sección 3.1) fue transformada con el plásmido pPWT042. Antes de la transformación de BIE104P1, se linealizó el pPWT042 usando la enzima de restricción SfiI, de acuerdo con las instrucciones suministradas por el proveedor. Las mezclas de transformación fueron colocadas sobre placas de YPD (por litro: 10 gramos de extracto de levadura, 20 gramos por litro de peptona, 20 gramos por litro de dextrosa, 20 gramos de agar) que contenían 100 µg G418 (Sigma Aldrich) por ml.

Después de 2 a 4 días, aparecieron las colonias sobre las placas, mientras que el control negativo (esto es, sin adición de ADN en el experimento de transformación) dio como resultado placas de YPD/G418 en blanco.

Por digestión del plásmido pPWT042 con SfiI, su integración es dirigida al locus SIT4 (Gottlin Ninfa and Kaback (1986) Molecular and Cellular Biology Vol. 6, No. 6, 2185-2197) en el genoma. Los transformantes fueron caracterizados utilizando PCR y técnicas de Southernblotting.

Las reacciones de PCR utilizando polimerasa Phusion® DNA (Finnzymes), que son indicativas de la integración correcta de una copia de plásmido pPWT042, se llevaron a cabo con los cebadores indicados por SEQ ID 13 y 14, y 14 y 15.

5 Como se representa en la figura 9, con el par de cebadores de SEQ ID 13 y 14, se revisó la integración correcta en el locus SIT4. La integración correcta del plásmido en el locus SIT4 puede también revisarse con el par de cebadores SEQ ID 15 y 16 (figura 9). Si se integró el plásmido pPWT042 en copias múltiples (integración cabeza a cola), el par de cebadores SEQ ID 14 y 15 dará un producto de PCR. Si este último producto de PCR está ausente, esto es indicativo de una integración en una copia del plásmido pPWT042.

Una cepa con una copia de plásmido pPWT042 integrada en el genoma se designó como BIE104P1Y9.

10 3.3 Experimentos de crecimiento

Se utilizaron aislados de colonia individual de cepas BIE104P1 y BIE104P1Y9 para inocular medio YNB (Difco) suplementado con 2% de glucosa. Los matraces inoculados fueron incubados durante aproximadamente 16 horas a 30°C y 280 rpm. Se determinó la densidad óptica a 600nm de los cultivos mantenidos durante la noche. El medio YNB suplementado con 1% de glucosa y 1% de xilosa fue inoculado con los cultivos durante la noche con un OD600 de partida de 0.2. Las células fueron cultivadas durante la noche a 30°C y 280 rpm. Subsecuentemente, se inoculó medio YNB que contenía 2% de xilosa y 0.1% de glucosa en un OD600 de partida de 0.2.

La diminuta cantidad de glucosa presente en este último medio fue consumida rápidamente por ambas cepas. Por transferencia al YNB con xilosa al 2% como única fuente de carbono, a un OD600 de partida de 0.2, solamente el BIE104P1Y9 fue capaz de crecer en este medio después de una fase de pausa muy larga de aproximadamente 4 semanas. Si la densidad óptica a 600 nm alcanzó un valor de al menos 2.0, las células fueron transferidas a un matraz con medio YNB fresco que contenía 2% de xilosa, con OD600 de partida de 0.2.

Esto fue repetido un cierto número de veces, como se establece en la figura 10. La gráfica muestra claramente que la cepa BIE104P1Y9 crece rápidamente y en forma eficiente sobre un medio mineral que contiene 2% de xilosa como única fuente de carbono, mientras que una cepa de referencia, que carece del plásmido integrado pPWT042, no es capaz de hacerlo así.

3.4 Actividad de la xilosa isomerasa

Aislado de colonias individuales de cepas BIE104P1 y BIE104P1Y9 fueron utilizados para inocular YPD con glucosa al 2%. Los matraces inoculados fueron incubados durante aproximadamente 16 horas a 30°C y 280 rpm. Se determinó la densidad óptica a 600nm de los cultivos mantenidos durante la noche. Las células fueron recolectadas por centrifugación. La pella fue lavada una vez con tampón de MOPS (ácido 3-(N-morfolino) propanosulfónico; Sigma) 0.1 M, pH 7.5 y congelado a -20°C hasta que se llevó a cabo el análisis.

Los resultados de los análisis se resumen en la tabla más abajo.

Cepa	Actividad XI a 30°C (nmol/mg proteína.min)	Actividad XI a 50°C (nmol/mg proteína.min)
35 Cepa de referencia	<20	<20
BIE104P1		
BIE104P1Y9	110	640

Los valores son el promedio de dos experimentos independientes.

3.5 Producción de etanol

40 Los aislados de colonias individuales de cepas BIE104P1 y BIE104P1Y9 fueron utilizados para inocular medio Verduyn (Verduyn et al., Yeast 8: 501-517, 1992) suplementados con glucosa al 2% como fuente única de carbono. Además, la cepa BIE104P1Y9 fue inoculada en medio Verduyn con xilosa al 2% como fuente única de carbono. Los matraces inoculados fueron incubados durante aproximadamente 64 horas a 30°C y 280 rpm. Se determinó la densidad óptica a 600 nm de los cultivos. Las células fueron recolectadas por centrifugación y la pella de células fue lavada con agua estéril milliQ (Millipore).

Medio de Verduyn fresco suplementado con glucosa al 2% y xilosa al 2% fue inoculado con los tres precultivos descritos más arriba. La cantidad de células inoculadas fue tal que el OD600 inicial fue 0.2. Los matraces fueron cerrados a prueba de agua, asegurando condiciones de crecimiento anaeróbico después que el oxígeno fue extraído del medio y del espacio de cabeza.

Los matraces fueron incubados durante 72 horas a 30°C y 280 rpm. Se tomaron muestras a 23, 47 y 71 horas para su análisis. Se llevaron a cabo los siguientes análisis: determinación OD600, análisis de RMN (xilosa, glucosa, etanol, ácido acético y glicerol). Los resultados se muestran en las figuras 11 y 12 y en la tabla más abajo. Los datos representan la cantidad de azúcares residuales en lo indicado (glucosa y xilosa en gramos por litro) y la formación de (sub)productos (etanol, glicerol y ácido acético).

En la figura 11, se muestra el desarrollo de la densidad óptica a 600 nm (OD600) en tiempo. La cepa de referencia BIE104P1 alcanza su máximo OD600 antes o a las 23 horas después del inicio del experimento. Aparentemente, en o antes del punto de tiempo de 23 horas, la glucosa ha sido extraída del medio (figura 12, panel a). También, la producción de etanol alcanzó su máximo en el momento en que la glucosa había sido consumida por esta cepa de levadura. Tanto el crecimiento como la producción de etanol se detiene, por esta cepa no puede utilizar y fermentar xilosa puesto que pierde las proteínas activas necesarias (esto es, una xilosa isomerasa y la xiluloquinasa sobreexpresada).

La cepa BIE104P1Y9 sin embargo, en la cual la xilosa isomerasa derivada del *Bacteroides uniformis* y la xiluloquinasa nativa son sobre expresadas, es capaz de crecer y fermentar la xilosa hasta etanol (figuras 11 y 12). Después de aproximadamente un día de cultivo anaeróbico, la cepa BIE104P1Y9 ya consumió algo de xilosa, mientras que ya se había consumido toda la glucosa. Subsecuentemente, la cantidad residual de glucosa fue fermentada en etanol como es evidente en la figura 12 (panel b y c) y la tabla más abajo. La formación de subproducto (ácido acético y glicerol) es baja, como es evidente a partir de la tabla de más abajo.

Los resultados no son (significativamente) influenciados por los precultivos (glucosa o xilosa) como es evidente a partir de los resultados presentados en las figuras 11 y 12.

Precrecimiento de BIE104P1 sobre glucosa

Tiempo (h)	Glucosa	Xilosa	Glicerol	Ácido acético	Etanol
0	19,5	19,6	0,0	0,0	0,7
23	0,0	20,8	0,5	0,3	9,0
47	0,0	20,7	0,7	0,7	8,8
71	0,0	19,9	0,7	0,9	8,4

Precrecimiento de BIE104P1Y9 sobre glucosa

Tiempo (h)	Glucosa	Xilosa	Glicerol	Ácido acético	Etanol
0	19,5	19,6	0,0	0,0	0,7
23	0,0	16,6	0,6	0,5	11,4
47	0,0	6,3	0,7	0,8	14,3
71	0,0	1,7	0,4	1,1	16,8

Precrecimiento de BIE104P1Y9 sobre xilosa

Tiempo (h)	Glucosa	Xilosa	Glicerol	Ácido acético	Etanol
0	19,5	19,6	0,0	0,0	0,7
23	0,7	17,7	0,0	0,5	11,3
47	0,0	6,1	0,7	0,8	14,9
71	0,0	1,1	0,5	1,1	16,7

Todos los valores se dan en gramos por litro

Con base en estos resultados, se calculó un Qs de 363 mg de xilosa por gramo de biomasa, por hora (intervalo de tiempo 23-47 horas; densidad óptica de 30 igual a 6 gramos de materia seca por litro), en caso de precrecimiento de la cepa BIE104P1Y9 sobre xilosa.

Ejemplo 4

4.1 Introducción de los genes araA, araB y araD en el genoma de *S. cerevisiae*

El plásmido pPWT018, tal como se representa en la figura 13, fue construido como sigue: el vector pPWT006 (figura 14), consistente de un locus SIT2 (Gottlin Ninfa and Kaback (1986) Molecular and Cell Biology vol. 6, no. 6, 2185-2197) y los marcadores que permiten la selección de los transformantes sobre el antibiótico G418 y la capacidad de crecer sobre acetamida (vide supra), fue digerido con las enzimas de restricción BsiWI y MluI. Los genes que codifican arabinosa isomerasa (araA), L-ribuloquinasa (araB) y L-ribulosa-5-fosfato-4-epimerasa (araD) del *Lactobacillus plantarum* como se describió en la solicitud de patente WO 2008/041840, fueron sintetizados por GeneArt AG (Regensburg, Alemania). Se sintetizó un fragmento grande, rodeando los tres genes ara antes mencionados, bajo control de (o enlazados operativamente a) promotores fuertes de *S. cerevisiae*, esto es, el promotor TDH3 que controla la expresión del gen araA, el promotor ENO1 que controla el gen araB y el promotor PGI1 que controla el gen araD. Este fragmento fue rodeado por las enzimas de restricción única Acc65I y MluI. La clonación de este fragmento en pPWT006 digerido con MluI y BsiWI produjo el plásmido pPWT018 (figura 13). La secuencia del plásmido pPWT018 se representa en SEQ ID 17.

La CEN.PK113-7D (MATa URA3 HIS3 LEU2 TRP1 MAL2-8 SUC2) se transformó con el plásmido pPWT018, el cual fue linealizado previamente con SfiI (New England Biolabs), de acuerdo con las instrucciones del proveedor. Se diseñó un sitio sintético SfiI en el flanco 5' del gen SIT2 (véase figura 13). Las mezclas de transformación fueron colocadas en placas sobre agar YPD (por litro: 10 gramos de extracto de levadura, 20 gramos por litro de peptona, 20 gramos por litro de dextrosa, 20 gramos de agar) que contenían 100 µg de G418 (Sigma Aldrich) por ml.

Después de 2 a 4 días, aparecieron colonias sobre las placas, mientras que el control negativo (esto es sin adición de ADN en el experimento de transformación) dio como resultado placas YPD/G418 de blanco.

La integración del plásmido pPWT018 está dirigida al locus SIT2. Los transformantes fueron caracterizados utilizando técnicas de PCR y Southern blotting.

Las reacciones de PCR, que son indicativas de la integración correcta de una copia del plásmido pPWT018, se llevaron a cabo con los cebadores indicados por SEQ ID 18 y 15, y 15 y 14 (véase figura 3). Con el par de cebadores de SEQ ID 18 y 15, se revisó la integración correcta del locus SIT2. Si se integró el plásmido pPWT018 en copias múltiples (integración cabeza a cola), el par de cebadores de SEQ ID 15 y 14 dará un producto de PCR. Si este último producto de PCR está ausente, es indicativo de una integración de una copia de pPWT018. Una cepa en la cual una copia del plásmido pPWT018 se integró en el locus SIT2, se designa BIE104R2.

Con el fin de ser capaces de transformar las cepas de levadura con otros constructos, es necesario eliminar los marcadores seleccionables. El diseño del plásmido pPWT018 fue tal, que por integración del pPWT018 en el cromosoma, las secuencias homólogas están en cercana proximidad una de otra. Este diseño permite que los marcadores seleccionables se pierdan por recombinación intramolecular espontánea de estas regiones homólogas.

Por crecimiento vegetativo, tendrá lugar la recombinación intramolecular, aunque a baja frecuencia. La frecuencia de esta recombinación depende de la longitud de la homología y del locus en el genoma (resultados sin publicar). Por transferencia secuencial de una subfracción del cultivo a medio fresco, los recombinantes intramoleculares se acumularán en el tiempo.

Con este fin, la cepa BIE104R2 fue cultivada en medio YPD (por litro: 10 gramos de extracto de levadura, 20 gramos de peptona por litro, 20 gramos de dextrosa por litro), partiendo de un aislado de colonia individual. Se usaron 25 µl de un cultivo mantenido durante la noche para inocular medio YPD fresco. Después de al menos cinco de tales transferencias en serie, se determinó la densidad óptica del cultivo y las células fueron diluidas hasta una concentración de aproximadamente 5.000 por ml. Se colocaron sobre placa 100 µl de la suspensión de células sobre medio Yeast Carbon Base (Difco) que contenía KPi 30 Mm (pH 6.8), (NH₄)₂SO₄ al 0.1%, fluoroacetamida 40 Mm (Amersham) y agar al 1.8% (Difco). Las células idénticas a las células de la cepa BIE104R2, esto es sin recombinación intracelular, contenían aún el gen amdS. Para estas células, la fluoroacetamida es tóxica. Estas células no serán capaces de crecer y no formarán colonias en un medio que contiene fluoroacetamida. Sin embargo, si hubiera presentado la recombinación intramolecular, las variantes de BIE104R2 que han perdido los marcadores seleccionables serán capaces de crecer sobre el medio de fluoroacetamida, puesto que son incapaces de convertir la fluoroacetamida en compuestos inhibidores del crecimiento. Estas células formarán colonias sobre este medio de agar.

Las colonias resistentes a la fluoroacetamida así obtenidas fueron sometidas a análisis de PCR utilizando cebadores de SEQ ID 18 y 15, y 14 y 19. Los cebadores de SEQ ID 18 y 5 darán una banda si ha tenido lugar la recombinación de los marcadores seleccionables como se pretendía. Como resultado, el casete con los genes araA, araB y araD bajo control de los promotores fuertes de la levadura habrán sido integrados en el locus SIT2 del genoma de la cepa huésped. En este caso, una reacción de PCR que utiliza cebadores de SEQ ID 14 y 19 no dará como resultado un producto de PCR, puesto que el cebador 14 inicia en una región que debía ser recombinada fuera. Si se obtiene una banda con los anteriores cebadores, esto es indicativo de la presencia del plásmido completo pPWT018 en el genoma, de forma que no ha habido recombinación.

Si los cebadores de SEQ ID 18 y 15 no resultan en un producto de PCR, ha tenido lugar la recombinación, pero de tal forma que el plásmido completo pPWT018 se ha recombinado fuera del genoma. No solamente se han perdido los marcadores seleccionables, sino también los genes *ara*. En efecto, se ha recuperado la levadura tipo silvestre.

5 Los aislados que muestran resultados de PCR de acuerdo con una integración de copia de pPWT018 fueron sometidos a análisis por Southern blot. El ADN cromosómico de la cepa CEN.PK113-7D y los recombinantes correctos fueron digeridos con EcoRI y HindIII (doble digestión). Se preparó una sonda SIT2 con los cebadores SEQ ID 20 y 21, utilizando pPW018 como patrón. El resultado del experimento de hibridización se muestra en la figura 15. El patrón de hibridización esperado puede deducirse a partir de los mapas físicos como se representa en la figura 16 (paneles a y b).

10 En la cepa tipo silvestre, se observa una banda de 2.35 kb, la cual está de acuerdo con el tamaño esperado (figura 16, panel a). Por integración y pérdida parcial por recombinación del plásmido pPWT018, se esperaba una banda de 1.06 kb (figura 16, panel b). En efecto, se observa esta banda, como se muestra en la figura 15 (línea 2).

Una de las cepas que mostró el patrón correcto de bandas sobre el Southern blot (como puede deducirse a partir de la figura 15) es la cepa designada como BIE104A2.

15 **4.2 Introducción de cuatro genes consecutivamente expresados de la ruta no oxidativa del fosfato de pentosa**

20 La *Saccharomyces cerevisiae* BIE104A2, que expresa los genes *araA*, *araB* y *araD* constitutivamente fue transformada con el plásmido pPWT080 (figura 2). El procedimiento y los resultados fueron descritos ya en el ejemplo 3 (sección 3.1). Para resumir, la BIE104A2 fue transformada con pPWT080 digerido con SfiI. Las mezclas de transformación fueron colocadas sobre placas sobre agar YPD (por litro: 10 gramos de extracto de levadura, 20 gramos de peptona por litro, 20 gramos de dextrosa por litro, 20 gramos de agar) que contenían 100 µg de G418 (Sigma Aldrich) por ml.

Después de dos a cuatro días, aparecieron colonias sobre las placas, mientras que el control negativo (esto es sin adición de ADN en el experimento de transformación), dio como resultado placas YPD/G418 en blanco.

25 La integración del plásmido pPWT080 es dirigida al locus *GRE3*. Los transformantes fueron caracterizados utilizando técnicas de PCR y Southern blotting, como se describe en el ejemplo 3, sección 3.1.

Un transformante que muestra una integración correcta de una copia de plásmido pPWT080, de acuerdo con el patrón de hibridización esperado, fue designado como BIE104A2F1.

30 Con el fin de ser capaces de introducir los genes que codifican la xilosa isomerasa y la xiluloquinasa (sección 3.2), es necesario eliminar los marcadores de selección introducidos por la integración del plásmido pPWT080. Con este fin, la cepa BIE104A2F1 fue cultivada en medio YPD, a partir de un aislado de colonia. Se utilizaron 25 µl de un cultivo mantenido durante la noche para inocular el medio YPD fresco. Después de cinco transferencias en serie, se determinó la densidad óptica del cultivo y las células fueron diluidas hasta una concentración de aproximadamente 5.000 por ml. Se colocaron 100 µl de la suspensión celular sobre placas sobre medio Yeasts Carbon Base (Difco) que contenían KPi 30 mM (pH 6.8), (NH₄)₂SO₄ al 0.1%, fluoroacetamida 40 mM (Amersham) y agar al 1.8% (Difco). Las colonias resistentes a la fluoroacetamida fueron sometidas a análisis por PCR, y en el caso de perfiles de PCR correctos, al análisis por Southern blot (sección 3.1 del ejemplo 3). Una de las cepas que mostró el patrón correcto de bandas sobre el Southern blot es la cepa designada como BIE104A2P1.

35 **4.3. Introducción de genes constitutivamente expresados que codifican xilosa isomerasa y xiluloquinasa**

40 La cepa BIE104A2P1 (MATa URA3 HIS3 LEU2 TRP1 MAL2-8 SUC2 SIT2::[TDH3-*araA*, ENO1-*araB*, PGI1-ARAd] Δ*GRE3*::[TPI1p-TAL1, ADH1p-TKL1, PGI1-RPE1, ENO1p-RK11]) fue transformada con el plásmido pPWT042. Antes de la transformación de BIE104A2P1, el pPWT042 fue linealizado utilizando la enzima de restricción SfiI, de acuerdo con las instrucciones suministradas por el proveedor. Las mezclas de transformación fueron colocadas sobre placas sobre agar YPD (por litro: 10 gramos de extracto de levadura, 20 gramos de peptona por litro, 20 gramos de dextrosa por litro, 20 gramos de agar) que contenían 100 µg de G418 (Sigma Aldrich) por ml.

Después de dos a cuatro días, aparecieron colonias sobre las placas, donde el control negativo (esto es sin adición de ADN en el experimento de transformación) dio como resultado placas YPD/G418 en blanco.

50 Por digestión del plásmido pPWT042 con SfiI, su integración es dirigida al locus SIT4 (Gottlin Ninfa and Kaback (1986) *Molecular and Cellular Biology* Vol. 6, No.6, 2185-2197) en el genoma. Los transformantes fueron caracterizados utilizando técnicas de PCR y Southern blotting, como se describió en el ejemplo 3 (sección 3.2).

Una cepa con una copia de plásmido pPWT042 integrada en el genoma fue designada como BIE104A2P1Y9.

4.4. Experimentos de crecimiento

Los aislados de colonia individual de las cepas BIE104A2P1Y9 fueron utilizados para inocular medio YNB (Difco) suplementado con glucosa al 2% o galactosa al 2%. Los matraces inoculados fueron incubados a 30°C y 280 rpm hasta que la densidad óptica a 600nm alcanzó un valor de al menos 2.0.

- 5 El medio YNB suplementado con arabinosa al 1% y xilosa al 1% fue inoculado con los cultivos durante la noche con un OD600 de partida de 0.2. Las células fueron cultivadas a 30°C Y 280 rpm. La densidad óptica a 600 nm fue monitorizada regularmente. Cuando la densidad óptica alcanzó un valor superior a 2.0, se transfirió un alícuota del cultivo a medio YNB que contenía 2% de xilosa y 0.2% de arabinosa. La cantidad de células añadidas fue tal que el OD600 de partida del cultivo era 0.2.
- 10 La densidad óptica fue monitorizada de forma regular. Los resultados se muestran en la figura 17, panel a (precultivos sobre galactosa) y panel b (precultivos sobre glucosa).

Los resultados muestran claramente que las cepas son capaces de utilizar tanto arabinosa como xilosa.

ES 2 551 515 T3

LISTADO DE SECUENCIAS

<110> DSM IP Assets B.V.

<120> UNA CÉLULA FERMENTADORA DE AZÚCAR PENTOSA

<130> 26577WO

<150> 081024077

<151> 07-03-2008

<160> 21

<170> PatentIn versión 3.2

<210> 1

<211> 1317

<212> ADN

<213> Bacteroides uniformis

<400> 1

```

atggcaaca aagagtattt tcccggaata ggaaagatta aattcgaagg taaagagagc      60
aagaaccgca tggcattccg ttattacgat gccgataaag taatcatggg taagaaaatg      120
agcgaatggc tgaagtccgc catggcatgg tggcacactc tttgcgcaga aggtggtgac      180
caattcggtg gcggaacaaa gaaattcccc tggaacgggtg aggctgacaa ggttcaggct      240
gccaagaaca aaatggacgc cggctttgaa ttcattgcaga aaatgggtat cgaatactac      300
tgcttecacg atgtagacct ctgcgaagaa gccgagacca ttgaagaata cgaagccaac      360
ttgaaggaaa tcgtagcgta tgccaagcag aaacaagcag aaaccggcat caaactggtg      420
tggggtactg ccaacgtatt cggccatgcc cgctacatga atggtgcagc caccaatccc      480
gatttcgatg ttgtggcacg tgccgccatc caaatcaaaa acgccatcga cgctactatc      540
gaactgggag gctcaacta tgtattctgg ggcggtcgcg aaggctacat gtcattgctg      600
aatacagacc agaagcgtga gaaagagcac ctgcacaga tgttgacat cgcgccgac      660
tatgcacgtg cocgcggtt caaaggtacc ttcttgattg aaccgaaacc gatggaacct      720
acaaaacacc agtatgatgt agacaccgaa accgttatcg gcttcttgaa ggctcacaat      780
ctggacaaaag atttcaaggt gaacatcgaa gtgaaccacg ctactttggc gggccacacc      840
ttcgagcaag aactcgcagt agccgtagac aacggtatgc tcggctccat cgacgccaac      900
cgtggtgact accagaacgg ctgggataca gaccagttcc ccattgacaa cttcgaactg      960
accaggcaa tgatgcaaat catccgtaac ggaggctttg gcaatggcgg taaaaacttc     1020
gatgccaaga ccgctcgcaa ctccaccgac ctggaagaca ttttcattgc ccacatcgcc     1080
ggtatggacg tgatggcacg tgcaactgaa agtgcagcca aactgcttga agagtctcct     1140
tacaagaaga tgctggccga ccgctatgct tccttcgaca gtggtaaagg caaggaattt     1200
gaagatggca aactgacgct ggaggatttg gtagcttacg caaaagcaa cggtgagccg     1260
aacagacca gcggcaagca ggaattgtat gaggcaatcg tgaatatgta ctgctaa      1317

```

<210> 2

ES 2 551 515 T3

<211> 1318
 <212> ADN
 <213> Secuencia optimizada de codón de Bacteroides uniformis

<220>
 <221> CDS
 <222> (1)..(1314)

<400> 2
 atg gct acc aag gaa tac ttc cca ggt att ggt aag atc aaa ttc gaa 48
 Met Ala Thr Lys Glu Tyr Phe Pro Gly Ile Gly Lys Ile Lys Phe Glu
 1 5 10 15
 ggt aag gaa tcc aag aac cca atg gcc ttc aga tac tac gat gct gac 96
 Gly Lys Glu Ser Lys Asn Pro Met Ala Phe Arg Tyr Tyr Asp Ala Asp
 20 25 30
 aag gtt atc atg ggt aag aag atg tct gaa tgg tta aag ttc gct atg 144
 Lys Val Ile Met Gly Lys Lys Met Ser Glu Trp Leu Lys Phe Ala Met
 35 40 45
 gct tgg tgg cat acc ttg tgt gct gaa ggt ggt gac caa ttc ggt ggt 192
 Ala Trp Trp His Thr Leu Cys Ala Glu Gly Gly Asp Gln Phe Gly Gly
 50 55 60
 ggt acc aag aaa ttc cca tgg aac ggt gaa gct gac aag gtc caa gct 240
 Gly Thr Lys Lys Phe Pro Trp Asn Gly Glu Ala Asp Lys Val Gln Ala
 65 70 75 80
 gct aag aac aag atg gac gct ggt ttc gaa ttt atg caa aag atg ggt 288
 Ala Lys Asn Lys Met Asp Ala Gly Phe Glu Phe Met Gln Lys Met Gly
 85 90 95
 att gaa tac tac tgt ttc cac gat gtt gac ttg tgt gaa gaa gct gaa 336
 Ile Glu Tyr Tyr Cys Phe His Asp Val Asp Leu Cys Glu Glu Ala Glu
 100 105 110
 acc atc gaa gaa tac gaa gct aac ttg aag gaa att gtt gct tac gct 384
 Thr Ile Glu Glu Tyr Glu Ala Asn Leu Lys Glu Ile Val Ala Tyr Ala
 115 120 125
 aag caa aag caa gct gaa act ggt atc aag cta tta tgg ggt act gct 432
 Lys Gln Lys Gln Ala Glu Thr Gly Ile Lys Leu Leu Trp Gly Thr Ala
 130 135 140
 aac gtc ttt ggt cat gcc aga tac atg aac ggt gcc gct acc aac cca 480

ES 2 551 515 T3

Asn Val Phe Gly His Ala Arg Tyr Met Asn Gly Ala Ala Thr Asn Pro 145 150 155 160	
gat ttc gat gtt gtt gcc aga gct gcc atc caa atc aag aac gcc atc Asp Phe Asp Val Val Ala Arg Ala Ala Ile Gln Ile Lys Asn Ala Ile 165 170 175	528
gat gct acc att gaa tta ggt ggt tcc aac tac gtt ttc tgg ggt ggt Asp Ala Thr Ile Glu Leu Gly Gly Ser Asn Tyr Val Phe Trp Gly Gly 180 185 190	576
aga gaa ggt tac atg tcc ttg ttg aac act gac caa aag aga gaa aag Arg Glu Gly Tyr Met Ser Leu Leu Asn Thr Asp Gln Lys Arg Glu Lys 195 200 205	624
gaa cac ttg gct caa atg ttg acc att gct cgt gac tac gct cgt gcc Glu His Leu Ala Gln Met Leu Thr Ile Ala Arg Asp Tyr Ala Arg Ala 210 215 220	672
aga ggt ttc aag ggt act ttc ttg att gaa cca aag cca atg gaa cca Arg Gly Phe Lys Gly Thr Phe Leu Ile Glu Pro Lys Pro Met Glu Pro 225 230 235 240	720
acc aag cac caa tac gat gtt gac acc gaa act gtc atc ggt ttc ttg Thr Lys His Gln Tyr Asp Val Asp Thr Glu Thr Val Ile Gly Phe Leu 245 250 255	768
aag gct cac aac ttg gac aag gac ttc aag gtc aac atc gaa gtc aac Lys Ala His Asn Leu Asp Lys Asp Phe Lys Val Asn Ile Glu Val Asn 260 265 270	816
cac gct act ttg gcc ggt cac act ttc gaa cac gaa ttg gct gtt gct His Ala Thr Leu Ala Gly His Thr Phe Glu His Glu Leu Ala Val Ala 275 280 285	864
gtc gac aac ggt atg ttg ggt tcc att gat gct aac aga ggt gac tac Val Asp Asn Gly Met Leu Gly Ser Ile Asp Ala Asn Arg Gly Asp Tyr 290 295 300	912
caa aac ggt tgg gac acc gac caa ttc cca atc gac aac ttt gaa ttg Gln Asn Gly Trp Asp Thr Asp Gln Phe Pro Ile Asp Asn Phe Glu Leu 305 310 315 320	960
act caa gct atg atg caa atc atc aga aac ggt ggt ttc ggt aac ggt Thr Gln Ala Met Met Gln Ile Ile Arg Asn Gly Gly Phe Gly Asn Gly 325 330 335	1008
ggt acc aac ttc gat gct aag acc aga aga aac tct act gac ttg gaa Gly Thr Asn Phe Asp Ala Lys Thr Arg Arg Asn Ser Thr Asp Leu Glu 340 345 350	1056
gat atc ttc atc gct cac att gcc ggt atg gat gtc atg gcc aga gct Asp Ile Phe Ile Ala His Ile Ala Gly Met Asp Val Met Ala Arg Ala 355 360 365	1104
ttg gaa tct gct gct aaa tta ttg gaa gaa tct cct tac aag aag atg Leu Glu Ser Ala Ala Lys Leu Leu Glu Glu Ser Pro Tyr Lys Lys Met	1152

ES 2 551 515 T3

370 375 380

ttg gct gac aga tac gct tct ttc gac tct ggt aag ggt aag gaa ttt 1200
 Leu Ala Asp Arg Tyr Ala Ser Phe Asp Ser Gly Lys Gly Lys Glu Phe
 385 390 395 400

gaa gat ggt aag ttg act ttg gaa gat ttg gtt gct tac gcc aag gct 1248
 Glu Asp Gly Lys Leu Thr Leu Glu Asp Leu Val Ala Tyr Ala Lys Ala
 405 410 415

aac ggt gaa cca aag caa act tct ggt aag caa gaa ttg tac gaa gcc 1296
 Asn Gly Glu Pro Lys Gln Thr Ser Gly Lys Gln Glu Leu Tyr Glu Ala
 420 425 430

att gtc aac atg tac tgt taag 1318
 Ile Val Asn Met Tyr Cys
 435

<210> 3
 <211> 438
 <212> PRT
 <213> Secuencia optimizada de codón de Bacteroides uniformis

<400> 3

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

Gly Lys Glu Ser Lys Asn Pro Met Ala Phe Arg Tyr Tyr Asp Ala Asp
 20 25 30

Lys Val Ile Met Gly Lys Lys Met Ser Glu Trp Leu Lys Phe Ala Met
 35 40 45

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

Gly Thr Lys Lys Phe Pro Trp Asn Gly Glu Ala Asp Lys Val Gln Ala
 65 70 75 80

Ala Lys Asn Lys Met Asp Ala Gly Phe Glu Phe Met Gln Lys Met Gly
 85 90 95

Ile Glu Tyr Tyr Cys Phe His Asp Val Asp Leu Cys Glu Glu Ala Glu
 100 105 110

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

ES 2 551 515 T3

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

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

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

Asp Ala Thr Ile Glu Leu Gly Gly Ser Asn Tyr Val Phe Trp Gly Gly
 180 185 190

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

Glu His Leu Ala Gln Met Leu Thr Ile Ala Arg Asp Tyr Ala Arg Ala
 210 215 220

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

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

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

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

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

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

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

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

ES 2 551 515 T3

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

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

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

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

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

Ile Val Asn Met Tyr Cys
435

<210> 4

<211> 16176

<212> ADN

<213> Secuencia artificial

<220>

<223> ADN sintético

<400> 4

ES 2 551 515 T3

tcgcgcgttt cggatgatgac ggtgaaaacc tcttgacaca tgcagctccc ggagacggtc	60
acagcttgtc tgtaagcgga tgccgggagc agacaagccc gtcagggcgc gtcagcgggt	120
gttggcgggt gtcggggctg gcttaactat gcggcatcag agcagattgt actgagagtg	180
caccatatgc ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcaggcgc	240
cattcgccat tcaggctgcg caactgttgg gaagggcgat cggtgcgggc ctcttcgcta	300
ttacgccagc tggcgaaagg gggatgtgct gcaaggcgat taagttgggt aacgccaggg	360
ttttcccagt cacgacgttg taaaacgacg gccagtaagc ttgcatgcct gcaggtcgac	420
gcggccgcat attttttgta actgtaattt cactcatgca caagaaaaaa aaaactggat	480
taaaaggag cccaaggaaa actcctcagc atatatttag aagtctctc agcatatagt	540
tgtttgttt ctttacacat tcaactgtta ataaaacttt tataatattt cattatcgga	600
actctagatt ctatacttgt ttccaattg ggccgatcgg gccttgctgg tagtaaacgt	660
atacgtcata aaagggaaaa gccacatgcg gaagaatttt atggaaaaaa aaaaaacctc	720

ES 2 551 515 T3

gaagttacta cttctagggg gcctatcaag taaattactc ctggtacact gaagtatata 780
 agggatatag aagcaaatag ttgtcagtgc aatccttcaa gacgattggg aaaatactgt 840
 aggtaccgga gacctaacta catagtgttt aaagattacg gatatttaac ttacttagaa 900
 taatgccatt tttttgagtt ataataatcc tacgttagtg tgagcgggat ttaaactgtg 960
 aggaccttaa tacattcaga cacttctgcg gtatcacctt acttattccc ttcgagatta 1020
 tatctagga cccatcaggt tgggtgaaga ttaccggtc taagactttt cagcttcctc 1080
 tattgatgtt acacctggac accccttttc tggcatccag tttttaatct tcagtggcat 1140
 gtgagattct ccgaaattaa ttaaagcaat cacacaattc tctcggatac cacctcggtt 1200
 gaaactgaca ggtggtttgt tacgcatgct aatgcaaagg agcctatata cctttggctc 1260
 ggctgctgta acaggaata taaagggcag cataatttag gagtttagtg aacttgcaac 1320
 atttactatt ttcccttctt acgtaaatat ttttcttttt aattctaaat caatcttttt 1380
 caattttttg tttgtattct tttcttgctt aaatctataa ctacaaaaaa cacatacata 1440
 aactaaaaat gtctgaacca gctcaaaaga aacaaaagggt tgctaacaac tctctagaac 1500
 aattgaaagc ctccggcact gtcggtgttg ccgacactgg tgatttcggc tctattgcca 1560
 agtttcaacc tcaagactcc acaactaacc catcattgat cttggctgct gccaaagcaac 1620
 caacttacgc caagttgatc gatggtgccc tggaatacgg taagaagcat ggtaagacca 1680
 ccgaagaaca agtcgaaaat gctgtggaca gattgttagt cgaattcggg aaggagatct 1740
 taaagattgt tccaggcaga gtctccaccg aagttgatgc tagattgtct tttgacactc 1800
 aagctacat tgaaaaggct agacatatca ttaaattggt tgaacaagaa ggtgtctcca 1860
 aggaaagagt ccttattaaa attgcttcca cttgggaagg tattcaagct gccaaagaat 1920
 tggaagaaaa ggacggtatc cactgtaatt tgactctatt attctccttc gttcaagcag 1980
 ttgctgtgc cgaggcccaa gttactttga tttccccatt tgttggtaga attctagact 2040
 ggtacaaatc cagcactggt aaagattaca aggggtgaagc cgaccaggt gttatttccg 2100
 tcaagaaaat ctacaactac tacaagaagt acggttacaa gactattggt atgggtgctt 2160
 ctttcagaag cactgacgaa atcaaaaact tggctggtgt tgactatcta acaatttctc 2220
 cagctttatt ggacaagttg atgaacagta ctgaaccttt cccaagagtt ttggacctg 2280
 tctccgctaa gaaggaagcc ggcgacaaga tttcttacat cagcgacgaa tctaaattca 2340
 gattcgactt gaatgaagac gctatggcca ctgaaaaatt gtccgaaggt atcagaaaat 2400
 tctctgccga tattgttact ctattcgact tgattgaaaa gaaagttacc gcttaaggaa 2460

ES 2 551 515 T3

gtatctcgga aatattaatt taggccatgt cttatgcac gtttcttttg atacttacgg 2520
 gtacatgtac acaagtatat ctatatatat aaattaatga aatccccta tttatatata 2580
 tgactttaac gagacagaac agttttttat tttttatcct atttgatgaa tgatacagtt 2640
 tcttattcac gtgttatacc cacaccaaat ccaatagcaa taccggccat cacaatcact 2700
 gtttcggcag cccctaagat cagacaaaac atccggaacc acctaaatc aacgtcccat 2760
 atgaatcctt gcagcaaagc cgctcgtacc ggagatatac aatagaacag ataccagaca 2820
 agacataatg ggctaaacaa gactacacca attacactgc ctcatatgat gtggtacata 2880
 acgaactaat actgtagccc tagacttgat agccatcatc atatcgaagt ttcactaccc 2940
 ttttccatt tgccatctat tgaagtaata ataggcgcat gcaacttctt ttcttttttt 3000
 ttcttttctc tctccccgt tgttgtctca ccatatccgc aatgacaaaa aatgatgga 3060
 agacactaaa ggaaaaaatt aacgacaaag acagcaccaa cagatgtcgt tgttccagag 3120
 ctgatgaggg gtatctcgaa gcacacgaaa ctttttctt cttcattca cgcacactac 3180
 tctctaata gcaacggat acggccttcc ttccagttac ttgaatttga aataaaaaaa 3240
 agtttgctgt cttgctatca agtataaata gacctgcaat tattaatctt ttgtttctc 3300
 gtcattgttc tcgttccctt tcttcttgt ttctttttct gcacaatatt tcaagctata 3360
 ccaagcatal aatcaactat ctcatataca atgactcaat tcaactgacat tgataagcta 3420
 gccgtctcca ccataagaat tttggctgtg gacaccgtat ccaaggcca ctcaggtcac 3480
 ccaggtgctc cattgggtat ggcaccagct gcacacgttc tatggagtca aatgcgcatg 3540
 aaccaacca acccagactg gatcaacaga gatagatttg tcttgtctaa cggtcacgcy 3600
 gtcgctttgt tgtattctat gctacatttg actggttacg atctgtctat tgaagacttg 3660
 aaacagttca gacagttggg ttccagaaca ccaggtcac ctgaatttga gttgccaggt 3720
 gttgaagtta ctaccgtcc attagggtcaa ggtatctcca acgctgttg tatggccatg 3780
 gctcaagcta acctggctgc cacttacaac aagccgggct ttaccttgc tgacaactac 3840
 acctatgttt tcttgggtga cggttggttg caagaaggta tttcttcaga agcttctcc 3900
 ttggctggtc atttgaaatt gggttaacttg attgccatct acgatgaca caagatcact 3960
 atcgatggtg ctaccagtat ctcatcogat gaagatgttg ctaagagata cgaagcctac 4020
 ggttgggaag ttttgtacgt agaaaatggg aacgaagatc tagccggtat tgccaaggct 4080
 attgctcaag ctaagttatc caaggacaaa ccaactttga tcaaatgac cacaaccatt 4140

ES 2 551 515 T3

ggttacggtt ccttgcacgc cggctctcac tctgtgcacg gtgccccatt gaaagcagat 4200
 gatgttaaac aactaaagag caaatcgggt ttcaaccag acaagtcctt tgttgttcca 4260
 caagaagttt acgaccacta ccaaaagaca attttaaagc caggtgtcga agccaacaac 4320
 aagtggaaca agttgttcag cgaataccea aagaaattcc cagaattagg tgctgaattg 4380
 gctagaagat tgagcggcca actaccgca aattgggaat ctaagttgcc aacttacacc 4440
 gccaaaggact ctgccgtggc cactagaaaa ttatcagaaa ctgttcttga ggatgtttac 4500
 aatcaattgc cagagttgat tgggtggttct gccgatttaa caccttctaa cttgaccaga 4560
 tggaaggaag cccttgactt ccaacctcct tcttccggtt caggtaaacta ctctggtaga 4620
 tacattaggt acggtattag agaacacgct atgggtgcc taatgaacgg tatttcagct 4680
 ttcggtgcc actacaaacc atacggtggt actttcttga acttcgtttc ttatgctgct 4740
 ggtgccgta gattgtccgc tttgtctggc caccagta tttgggttgc tacacatgac 4800
 tctatcgtg tcggtgaaga tggccaaca catcaaccta ttgaaacttt agcacacttc 4860
 agatccctac caaacattca agtttgaga ccagctgatg gtaacgaagt ttctgccgcc 4920
 tacaagaact ctttagaatc caagcactt ccaagtatca ttgctttgtc cagacaaaac 4980
 ttgccacaat tggaaggtag ctctattgaa agcgcttcta aggggtggtta cgtactaaa 5040
 gatgttgcta acccagatat tatttttagtg gctactggtt ccgaagtgtc tttgagtgtt 5100
 gaagctgcta agactttggc cgcaaagaac atcaaggctc gtgttgtttc tctaccagat 5160
 ttcttcaact ttgacaaaaca acccctagaa tacagactat cagtcttacc agacaacggt 5220
 ccaatcatgt ctgttgaagt tttggctacc acatgttggg gcaaatacgc tcatcaatcc 5280
 ttcggtattg acagatttgg tgcctccggt aaggcaccag aagtcttcaa gttcttcggt 5340
 ttcaccccag aagggtttgc tgaaagagct caaaagacca ttgcattcta taagggtgac 5400
 aagctaattt ctctttgaa aaaagctttc taaattctga tcgtagatca tcagatttga 5460
 tatgatatta tttgtgaaaa aatgaaataa aactttatac aacttaaata caactttttt 5520
 tataaacgat taagcaaaaa aatagtttca aacttttaac aatattcaa aactcagtc 5580
 ctttccttc ttatattata ggtgtacgta ttatagaaaa atttcaatga ttactttttc 5640
 tttctttttc cttgtaccag cacatggccg agcttgaatg ttaaaccctt cgagagaatc 5700
 acaccattca agtataaagc caataaagaa tatcgtacca gagaattttg ccatcggaca 5760
 tgctacctta cgcttatatc tctcattgga atatcgtttt ctgattaaaa cacggaagta 5820
 agaacttaat tcgtttttcg ttgaactatg ttgtgccagc gtaacattaa aaaagagtgt 5880

ES 2 551 515 T3

acaaggccac gttctgtcac cgtcagaaaa atatgtcaat gaggcaagaa ccgggatggt 5940
 aacaaaaatc acgatctggg tgggtgtggg tgtattggat tataggaagc cacgcgctca 6000
 acctggaatt acaggaagct ggtaatTTTT tgggtttgca atcatcacca tctgcacggt 6060
 gttataatgt cccgtgtcta tatatatcca ttgacgggat tctatTTTT tgctattgaa 6120
 atgagcgttt tttgttacta caattggttt tacagacgga attttccta tttgtttcgt 6180
 cccatTTTT cttttctcat tgtttcata tcttaaaaag gtcctttctt cataatcaat 6240
 gctttctttt acttaatatt ttacttgcac tcagtgaatt ttaatacata ttctcttagt 6300
 cttgcaaaat cgatttagaa tcaagatacc agcctaaaaa tggTcaaacc aattatagct 6360
 ccagtatcc ttgcttctga cttcgccaac ttgggttgcg aatgtcataa ggtcatcaac 6420
 gccggcgcag attggttaca tatcgatgtc atggacggcc attttgttcc aaacattact 6480
 ctgggccaac caattgttac ctccctacgt cgttctgtgc cacgccctgg cgatgctagc 6540
 aacacagaaa agaagcccac tgcgttcttc gattgtcaca tgatggttga aaatcctgaa 6600
 aaatgggtcg acgattttgc taaatgtggt gctgaccaat ttacgttcca ctacgaggcc 6660
 acacaagacc ctttgcattt agttaagttg attaagtcta agggcatcaa agctgcatgc 6720
 gccatcaaac ctggtacttc tgttgacggt ttatttgaac tagctcctca tttggatatg 6780
 gctcttgтта tgactgtgga acctgggttt ggaggccaaa aattcatgga agacatgatg 6840
 ccaaaagtgg aaactttgag agccaagttc cccatttga atatccaagt cgatggtggt 6900
 ttgggcaagg agaccatccc gaaagccgcc aaagccggtg ccaacgttat tgctcgtgga 6960
 accagtgttt tcaactgcagc tgaccgcgac gatgttatct ccttcatgaa agaagaagtc 7020
 tcgaaggaat tgcgttctag agatttgcta gattagttgt acatatgagg catttcttat 7080
 atttatactc tctatactat acgatatggt attttttct cgttttgatc tctaatata 7140
 cataaaccga gccattccta ctatacaaga tacgtaagtг cctaactcat gggaaaaatg 7200
 ggccgcccag ggtggtgcct tgtcggtttt cgatgatcaa tccctgggat gcagtatcgt 7260
 caatgacact ccataaggct tcttaacca aagtcaaaga actcttcttt tcattctctt 7320
 tcactttctt accgccatct agatcaatat ccatttcgta ccccgcgga cgcagata 7380
 ttcattactt gacgcaaaag cgtttgaaat aatgacgaaa aagaaggaag aaaaaaaag 7440
 aaaaataccg cttctaggcg ggttatctac tgatccgagc ttccactagg atagcaccca 7500
 aacacctgca tatttggagc acctttactt acaccaccaa aaaccacttt cgcctctccc 7560

ES 2 551 515 T3

gccctgata acgtccacta attgagcgat tacctgagcg gtcctctttt gtttgcagca 7620
 tgagacttgc atactgcaaa tcgtaagtag caacgtctca aggtcaaaac tgtatggaaa 7680
 ccttgtcacc tcaactaatt ctagctagcc taccctgcaa gtcaagaggt ctccgtgatt 7740
 cctagccacc tcaaggtatg cctctccccc gaaactgtgg ccttttctgg cacacatgat 7800
 ctccacgatt tcaacatata aatagctttt gataatggca atattaatca aatttatttt 7860
 acttctttct tgtaacatct ctcttgtaat cccttattcc ttctagctat ttttcataaa 7920
 aaaccaagca actgcttata aacacacaaa cactaaatca aatggctgc cgggtgcccc 7980
 aaaattgatg cgttagaatc tttgggcaat cctttggagg atgccaagag agctgcagca 8040
 tacagagcag ttgatgaaaa tttaaaattt gatgatcaca aaattattgg aattggtagt 8100
 ggtagcacag tggtttatgt tgccgaaaga attggacaat atttgcatac ccctaaattt 8160
 tatgaagtag cgtctaaatt catttgcatt ccaacaggat tccaatcaag aaacttgatt 8220
 ttggataaca agttgcaatt aggctccatt gaacagtatc ctgcgattga tatagcgttt 8280
 gacggtgctg atgaagtgga tgagaattta caattaatta aaggtggtgg tgcttgtcta 8340
 tttcaagaaa aattggttag tactagtgct aaaaccttca ttgtcgttgc tgattcaaga 8400
 aaaaagtcac caaaacattt aggtaagaac tggaggcaag gtgttcccat tgaaattgta 8460
 ccttcctcat acgtgagggg caagaatgat ctattagaac aattgcatgc tgaaaaagtt 8520
 gacatcagac aaggagggtc tgctaaagca ggtcctgttg taactgacaa taataacttc 8580
 attatcgatg cggatttcgg tgaaatttcc gatccaagaa aattgcatag agaaatcaaa 8640
 ctgttagtgg gcgtggtgga aacaggttta ttcacgaca acgcttcaaa agcctacttc 8700
 ggtaattctg acggtagtgt tgaagttacc gaaaagtgag cagatcaaag gcaaagacag 8760
 aaaccgtagt aaaggttgac ttttcacaac agtgtctcca tttttatat tgtattatta 8820
 aagctattta gttatttggg tactgttttt tttccagaag ttttcttttt agtaaagtac 8880
 aatccagtaa aatgaagga tgaacaatcg gtgatgcag attcaacacc aataaatgca 8940
 atgtttattt ctttggaaac tttgtgttgt tcgaaatcca ggataatcct tcaacaagac 9000
 cctgtccgga taaggcgta ctaccgatga cacaccaagc tcgagtaacg gagcaagaat 9060
 tgaaggatat ttctgcacta aatgccaaca tcagatttaa tgatccatgg acctggttgg 9120
 atggtaaatt ccccactttt gcctgatcca gccagtaaaa tccatactca acgacgatat 9180
 gaacaaattt cctcattcc gatgctgtat atgtgtataa atttttacat gctcttctgt 9240
 ttagacacag aacagcttta aataaaatgt tggatatact ttttctgcct gtgggtgcat 9300

ES 2 551 515 T3

ccacgctttt aattcatctc ttgtatgggt gacaatttgg ctatttttta acagaaccca 9360
 acggtaattg aaattaaaag ggaaacgagt gggggcgatg agtgagtgat actaaaatag 9420
 acaccaagag agcaaagcgg tcccagcggc cgcgaaattcg gcgtaatcat ggtcatagct 9480
 gtttcctgtg tgaaattggt atccgctcac aattccacac aacatacgag ccggaagcat 9540
 aaagtgtaaa gcctgggggt cctaatgagt gagctaactc acattaattg cgttgcgctc 9600
 actgcccgct ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg 9660
 cgcggggaga ggcggtttgc gtattgggag ctcttcgctc tcctcgctca ctgactcgct 9720
 gcgctcggtc gttcggctgc ggcgagcggg atcagctcac tcaaaggcgg taatacgggt 9780
 atccacagaa tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc 9840
 caggaaccgt aaaaaggcgg cgttgcctggc gtttttccat aggctccgcc ccctgacga 9900
 gcatcacaaa aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata 9960
 ccaggcgttt cccctggaa gctccctcgt gcgctctcct gttccgacct tgccgcttac 10020
 cggatacctg tccgcotttc tcccttcggg aagcgtggcg ctttctcaat gctcacgctg 10080
 taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc 10140
 cgttcagccc gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaa 10200
 acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt 10260
 aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta gaaggacagt 10320
 atttggtatc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg 10380
 atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 10440
 gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggg ctgacgctca 10500
 gtggaacgaa aactcacggt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac 10560
 ctagatcctt ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 10620
 ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 10680
 tcgttcatcc atagttgctt gactccccgt cgtgtagata actacgatac gggagggtt 10740
 accatctggc cccagtgctg caatgatacc gcgagacca cgctcaccgg ctccagatt 10800
 atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggctctg caactttatc 10860
 cgctccatc cagtctatta attggtgccc ggaagctaga gtaagtagtt cgccagttaa 10920
 tagtttgccg aacgttggtt ccattgctac aggcacgctg gtgtcacgct cgtcgtttgg 10980

ES 2 551 515 T3

tatggcttca ttcagctcog gttcccaacg atcaaggcga gttacatgat ccccatggt 11040
 gtgcaaaaaa gcggttagct ccttcgggcc tccgatcggt gtcagaagta agttggccgc 11100
 agtgttatca ctcatgggta tggcagcact gcataattct cttactgtca tgccatccgt 11160
 aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 11220
 gcgaccgagt tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac 11280
 tttaaaagtg ctcatcattg gaaaacggtc ttcggggcga aaactctcaa ggatcttacc 11340
 gctgttgaga tccagttoga tgtaaccac tctgtcacc aactgatctt cagcatcttt 11400
 tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg 11460
 aataagggcg acacggaaat gttgaatact catactcttc ctttttcaat attattgaag 11520
 catttatcag ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa 11580
 acaaataggg gttccgcgca ctttccccg aaaagtgcc cctgacgtca actatacaaa 11640
 tgacaagttc ttgaaaacaa gaatcttttt attgtcagta ctgattagaa aaactcatcg 11700
 agcatcaaat gaaactgcaa tttattcata tcaggattat caataccata tttttgaaaa 11760
 agccgtttct gtaatgaagg agaaaactca ccgaggcagt tccataggat ggcaagatcc 11820
 tggatcgggt ctgcgattcc gactcgtcca acatcaatac aacctattaa tttcccctcg 11880
 tcaaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat 11940
 ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca 12000
 tcaaaatcac tcgcatcaac caaacggtt ttcattcgtg attgcgctg agcgagacga 12060
 aatacgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa ccggcgcagg 12120
 aacactgcc aacatcaaac aatattttca cctgaatcag gatattcttc taatacctgg 12180
 aatgctgttt tgccggggat cgcagtggtg agtaaccatg catcatcagg agtacggata 12240
 aatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct gaccatctca 12300
 tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc tggcgcacgc 12360
 ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc gcgagcccat 12420
 ttataccat ataaatcagc atccatggtg gaatttaac gcggcctcga aacgtgagtc 12480
 ttttccttac ccatggttgt ttatgttcgg atgtgatgtg agaactgtat cctagcaaga 12540
 ttttaaaagg aagtatatga aagaagaacc tcagtggcaa atcctaacct tttatatttc 12600
 tctacagggg cgcggcgtgg ggacaattca acgcgactgt gacgcgttct agaacacaca 12660
 atatgcatgt aatcgctgat ttttttgggt ttagaagctc tatcttcagg taaaaatgag 12720

ES 2 551 515 T3

tagagaaaaa aaaacatact ggatcgatgc agaattaggg ggttattatc ctgcaggtac 12780
atgattttca gtgggaacat tgcttttttag tagtccgggt ctcaacaact tgtctaagtg 12840
ttgaaaacaa aagaaatggc gtagaaacaa agtagtgtaa gtaaactctgc caatgttcta 12900
tgtataaaaa gtaaaggcaa gaagaggttc tatgcatatt tctgaaaata tctaatacac 12960
tattataatg catcaagaaa ctgtcgtatg atgaagtgcc tatgagtttt tgtgtacgtg 13020
cttctctagt atgtagccgg ttttctcttt ttacctcttt ttactactta tactactact 13080
tttactacct ttcttccacg taatctagat ctcaagccac aattcttgcc ctatgctcca 13140
acgtatacaa catcgaagaa gagtctttct ttagggagtc attggaaaag atagtatgat 13200
ggatttcgat ttacctatgt cgcaaaagaa agtccggggc aacaccacag aatgctttct 13260
ctgtactaat aacctgttgt gcgcttaacg gtctaactcg taatcagcgg tggttaaatt 13320
tttgtaaate taatgttcca tgattttctt tcttcaaaag gaacatgtag cgaaaatctt 13380
ttttttactt tgatacactg caattgtttc tgagcatgct gaaattttct cgatgttttt 13440
tttttttatt ggcatccaag taattaatcc ttatgctacg aaaaagttgt aggaatgaat 13500
catgcataat ctaacggata tcatcatata ctctgtgcta atattctaaa caagttcgaa 13560
aatattttct tggcccatgt aataggtggt aagtgtattg ctttgatagg aacgtcatta 13620
tcgcacaaga caatcggcac taataaccgt ttaaataatta tcatgcatgt atacatcagt 13680
atctcataga aatatacctg taagtacata cttatctaag tataaattct cgacctatgg 13740
agtcaccaca ttcccagca acttccccac ttctctgca atcgccaacg tctctcttc 13800
actgagtctc cgccgataa cctgcactgc aaccggtgcc ccatggtacg cctccgcatc 13860
atactcttcc tgcaagaggg catcaagctc actaaccgcc ttgaaactct cattcttctt 13920
atcgatgttc ttatccgcaa aggttaaccgg aacaaccacg ctogtgaat ccagcaggtt 13980
gatcacagag gcatacccat agtaccggaa ctggtcatgc cgtaccgcag cggtaggcgt 14040
aatcggcgcg atgatggcgt ccagttcctt ccgggccttt tcttcagcct cccgccattt 14100
ctcaaggtag tccatctggt aattccactt ctggagatgc gtgtcccaga gctcgttcat 14160
gttaacagct ttgatgttcg ggttcagtag gtctttgata tttggaatcg ccggctcgcc 14220
ggatgcactg atatcgcgca ttacgtcggc gctgccgtca gccgcgtaga tatgggagat 14280
gagatcgtgg ccgaaatcgt gcttgtagtg cgtccacggg gtcacgggtg gaccggcttt 14340
ggcagagtgc gcgacggtag tttccacgcc gcgcagata ggagggtgtg gaaggacatt 14400

ES 2 551 515 T3

gccgtcgaag ttgtagtagc cgatattgag cccgccgttc ttgatcttgg aggcaataat 14460
 gtccgactcg gactggcgcc agggcatggg gatgacctg gagtcgtatt tccatggctc 14520
 ctgaccgagg acggatttgg tgaagaggcg gaggtcctca acagagtgcg taatcggccc 14580
 gacaacgctg tgcaccgtct cctgaccctc catgctgttc gccatctttg catacggcag 14640
 ccgcccataga ctccggcctta gaccgtacag gaagttgaac gcggccggca ctcgaatcga 14700
 gccaccgata tccgttccta caccgatgac gccaccacga atcccaacga tcgcaccctc 14760
 accaccagaa ctgccgccc acgaccagtt cttgttgctt gggttgacgg tgcgcccgat 14820
 gatgttgttg actgtctcgc agaccatcag ggtctgcggg acagaggtct tgacgtagaa 14880
 gacggcaccg gctttgcgga gcatggttgt cagaaccgag tccccttcgt cgtacttggt 14940
 tagccatgag atgtagccca ttgatgtttc gtagcccttg actcgaagct ggtctttgag 15000
 agagatgggg aggccatgga gtggaccaac gggctctctg tgctttgctt agtattcatc 15060
 gagttccctt gcctgcgcga gagcggcgtc aggaagaac tcgtgggccc agtttgtaa 15120
 ctgctggggc attgctgcc gtttacagaa tgetagcgt acttccaccg aggtcaactc 15180
 tccggccgcc agcttgaca caagatctgc agcggaggcc tctgtgatct tcagttcggc 15240
 ctctgaaagg atccccgatt tctttgggaa atcaataacg ctgtcttccg caggcagcgt 15300
 ctggactttc cattcatcag ggatggtttt tgcgaggcgg gcgcgcttat cagcggccag 15360
 ttcttcccag gattgaggca ttgtatatga gatagttgat tgtatgcttg gtatagcttg 15420
 aatatatttg cagaaaaaga aacaaggaag aaagggaaac agaacaatga cgaggaaaca 15480
 aaagattaat aattgcaggt ctatttatac ttgatagcaa agcggcaaac tttttttatt 15540
 tcaaattcaa gtaactggaa ggaaggccgt ataccgttg tcattagaga gtagtgtgcg 15600
 tgaatgaagg aaggaaaaag tttcgtgtgt tcgaagatac ccctcatcag ctctggaaca 15660
 acgacatctg ttggtgctgt ctttgcgtt aatttttcc ttagtgtct tccatcattt 15720
 tttttgtcat tgcggatatg gtgagacaac aacgggggag agagaaaaga aaaaaaaga 15780
 aaagaagttg catgcgccta ttattacttc aatagatggc aatggaaaa agggtagtga 15840
 aacttcgata tgatgatggc tatcaagtct agggctacag tattagtctg ttatgtacca 15900
 ccatcaatga ggcagtgtaa tttgtgtagt cttgtttagc ccattatgtc ttgtctggta 15960
 tctgttctat tgtatatctc ccctccgcca cctacatgtt aggagacca acgaaggat 16020
 tataggaatc ccgatgatg ggtttggtt ccagaaaaga ggaagtccat attgtacacc 16080
 cggaaacaac aaaaggatgg gcccatgacg tctaagaaac cattattatc atgacattaa 16140

cctataaaaa taggcgtatc acgaggccct ttcgtc

16176

ES 2 551 515 T3

<210> 5	
<211> 22	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 5	
gaaatgggcg cactactaca ag	22
<210> 6	
<211> 22	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 6	
caccaacctg atgggtcct ag	22
<210> 7	
<211> 22	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 7	
acgccagggt ttcccagtc ac	22
<210> 8	
<211> 22	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 8	
ccagcaccct aagccgacta gg	22
<210> 9	
<211> 21	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 9	
acgggtgctga tgaagtgat g	21
<210> 10	
<211> 21	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 10	
accacgccca ctaacagttt g	21

<210> 11	
<211> 32	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 11	
gggggtacc ctggatggcg gcgtagtat cg	32
<210> 12	
<211> 32	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 12	
gggggtacc tcacagtcgc gttgaattgt cc	32
<210> 13	
<211> 24	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 13	
ccaaggcagc ggtacatcaa gtag	24
<210> 14	
<211> 23	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 14	
tgcacatgtt gtccatcaag atg	23
<210> 15	
<211> 24	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 15	
ggaaacagct atgacatgat tacg	24
<210> 16	
<211> 23	
<212> ADN	
<213> Secuencia artificial	
<220>	
<223> ADN sintético	
<400> 16	
gtagcgaaat catgtattgc acc	23

ES 2 551 515 T3

<210> 17
<211> 18215
<212> ADN
<213> Secuencia artificial

<220>
<223> ADN sintético

<400> 17
ggccaagatg gccgatctgc atttttcata ataatcctcg gtactttcta caagatcaat 60
taaattccaa tcaaaaatcg tcttttgcaa gattttgaag tcacagtact tttcattttc 120
aatgtcaaca gcgccccatt tgtattgtct tcctttaact ttttcgcct tttcattaaa 180
aatgtactca ttagatgcaa ttatactgaa tggatatttt tgaaaaatat cttgtgttgc 240
attcaaaact tcatcgccga aaaagaaca tacagggata tcttgtactc ttattatttc 300
tctaacttgt gttttgaagt ttttcaattc ctctttcgtt agcaaactcg atttagcaat 360
aaccgggatt aaattcactc tcttcgctaa ttttttcatt gttacgacgt ctāaagtatc 420

ES 2 551 515 T3

aattccctta tttgaaggtc tcagaaagta caaacaacaa tggactctat tatcaaccat 480
ttttgtccta tcaggttgtt cttcttgga aatgtacgat cttatttctt catcaatata 540
gtttctagac tgcagcccg gatccgtoga caagcttgtg gagaggtgac ttcataaacc 600
aagtgtctgt cgatatacaa caaaaaggaa ccattttcat cttgatggac aacatgtgca 660
tcaaaaacct tatcgtaaag agttcttggga cccttggatg gagtgtaaac catgatttaa 720
aacagcaaat aataaaaatc gatagcgaca aaaactgtca atttcaatat tctttatatt 780
tgttgactgc ttagatattt tgagaaaatt cagcggaaac agcgtgatga gtgagttaag 840
ttctgctgtt taaataagta ttcaactact attgaagccg actcatgaag ccggttacgg 900
acaaaaccgg gcaaatctcg ccggtcccgg aattttcgtt tccgcaataa aagaaccgct 960
catcatcata gcgccagggt agtatactat agaaggtcag actaaactga gtcactaga 1020
gtaatgacgc cttagtagct tttacatctt cataagaaaa ggaaacttgt agaatggcct 1080
ggcgatttgt ttgctttctt gtgatgaaga aatttcgatg cgattaaccg gcaaaatcag 1140
taaaggtatt tcgcggaggc ggccttcaat catcgaatac tacgtcttaa tatgatgtac 1200
tgtggttcat attttcaagt agtgtagta aatttgata cgttcatgta agtgtgtatc 1260
ttgagtgtct gtatggcgc ataaacgtaa gcgagacttc caaatggagc aaacgagaag 1320
agatctttaa agtattatag aagagctggg caggaactat tatgacgtaa agccttgacc 1380
ataataaaga cgattctttg tcctctata caaacatctt gcaaagatac caaatatttt 1440
caaatcctac tcaataaaaa attaataaat aaattagtgt gtgtgcatta tatatattaa 1500
aaattaagaa ttagactaaa taaagtgtt ctaaaaaat attaaagttg aaatgtgcgt 1560
gttgtaatt gtgctctatt agaataatta tgacttgtgt gcgtttcata ttttaaata 1620
ggaaataacc aagaaagaaa aagtaccatc cagagaaacc aattatatca aatcaataa 1680
aacaaccagc ttcggtgtgt gtgtgtgtgt gaagctaaga gttgatgcca tttaatctaa 1740
aaattttaag gtgtgtgtgt ggataaaata ttagaatgac aattcgaatt gcgtacctta 1800
gtcaaaaaat tagcctttta attctgctgt aaccctgaca tgcccaaat agggggcggg 1860
ttacacagaa tatataacat cgtaggtgtc tgggtgaaca gtttattcct ggcattccact 1920
aatataatg gagcccgtt tttaagctgg catccagaaa aaaaaagaat cccagcacca 1980
aatattgtt ttcttacca accatcagtt cataggtcca ttctcttagc gcaactacag 2040
agaacagggg cacaaacagg caaaaaacgg gcacaacctc aatggagtga tgcaacctgc 2100
ctggagtaaa tgatgacaca aggcaattga cccagcatg tatctatctc attttcttac 2160

ES 2 551 515 T3

accttotatt accttctgct ctctctgatt tggaaaaagc tgaaaaaaa ggttgaacc 2220
 agttccctga aattattccc ctacttgact aataagtata taaagacggt aggtattgat 2280
 tgtaattctg taaatctatt tcttaaactt cttaaattct acttttatag ttagtctttt 2340
 ttttagtttt aaaacaccaa gaacttagtt tcgaataaac acacataaac aaacaaaatg 2400
 ttatcagtac ctgattatga gttttggttt gttaccggtt cacaacacct ttatggtgaa 2460
 gaacaattga agtctgttgc taaggatgcg caagatattg cggataaatt gaatgcaagc 2520
 ggcaagttac cttataaagt agtctttaag gatgttatga cgacggctga aagtatcacc 2580
 aactttatga aagaagttaa ttacaatgat aaggtagccg gtggtattac ttggatgcac 2640
 acattctcac cagctaagaa ctggattcgt ggaactgaac tgttacaaaa accattatta 2700
 cacttagcaa cgcaatattt gaataatatt ccatatgcag acattgactt tgattacatg 2760
 aaccttaacc aaagtgccca tggcgaccgc gagtatgcct acattaacgc ccggttgacg 2820
 aacataata agattgttta cggctattgg ggcgatgaag atgtgcaaga gcagattgca 2880
 cgttgggaag acgtcgccgt agcgtacaat gagagcttta aagttaaggt tgctcgcttt 2940
 ggcgacacaa tgcgtaatgt ggccgttact gaaggtgaca aggttgaggc tcaaattaag 3000
 atgggctgga cagttgacta ttatggtatc ggtgacttag ttgaagagat caataaggtt 3060
 tcggatgctg atgttgataa ggaatacgtc gacttgaggt ctcggtatga aatggtccaa 3120
 ggtgataacg atgcggacac gtataaacat tcagttcggg ttcaattggc acaatatctg 3180
 ggtattaagc ggttcttaga aagaggcggg tacacagcct ttaccacgaa ctttgaagat 3240
 ctttggggga tggagcaatt acctggtcta gcttcacaat tattaattcg tgatgggtat 3300
 ggttttggtg ctgaaggtga ctggaagacg gctgctttag gacgggttat gaagattatg 3360
 tctcacaaca agcaaaccgc ctttatggaa gactacacgt tagacttgcg tcatggtcat 3420
 gaagcgatct taggttcaca catgttggaa gttgatccgt ctatcgcaag tgataaacca 3480
 cgggtcgaag ttcattcatt ggatattggg ggtaagatg atcctgctcg cctagtattt 3540
 actggttcag aaggtgaagc aattgatgtc accggtgccc atttccgtga tgggttcaag 3600
 atgattagct acgcggtaga tgcgaataag ccagaagccg aaacacctaa tttaccagtt 3660
 gctaagcaat tatggacccc aaagatgggc ttaaagaaag gtgcactaga atggatgcaa 3720
 gctggtggtg gtcaccacac gatgctgtcc ttctcgttaa ctgaagaaca aatggaagac 3780
 tatgcaacca tggttggcat gactaaggca ttcttaaagt aagtgaattt actttaaatc 3840

ES 2 551 515 T3

ttgcatttaa ataaatthtc tttttatagc tttatgactt agtttcaatt tatatactat 3900
 tttaatgaca ttttcgattc attgattgaa agctttgtgt tttttcttga tgcgctattg 3960
 cattgttctt gtctttttcg ccacatgtaa tatctgtagt agatacctga tacattgtgg 4020
 atgctgagtg aaatthtagt taataatgga ggcgctctta ataatthtg g gatattggc 4080
 tttttttttt aaagthtaca aatgaattht ttcgcccagg atcgtacgcc gcggaaccgc 4140
 cagatattca ttacttgacg caaaagcgtt tgaataatg acgaaaaaga aggaagaaaa 4200
 aaaaagaaaa ataccgcttc taggcgggtt atctactgat ccgagcttcc actaggatag 4260
 cacccaaaca cctgcatatt tggacgacct ttacttacac caccaaaaac cactttcgcc 4320
 tctcccgcgc ctgataacgt ccactaattg agcgttacc tgagcggctc tcttttgtht 4380
 gcagcatgag acttgcatc tgcaaatcgt aagtagcaac gtctcaaggc caaaactgta 4440
 tggaaacctt gtcacctcac ttaattctag ctacgctacc ctgcaagtca agaggctcctc 4500
 gtgattctca gccacctcaa ggtatgcctc tcccggaaa ctgtggcctt ttctggcaca 4560
 catgatctcc acgatttcaa catataaata gctthtgata atggcaatat taatcaaatt 4620
 tattttactt ctttcttga acatctctct tgtaatccct tttcttct agctatthtt 4680
 cataaaaaac caagcaactg cttatcaaca cacaaacct aatcaaaat gaatthtagtt 4740
 gaaacagccc aagcgattaa aactggcaaa gthtctthtag gaattgagct tggctcaact 4800
 cgaattaaag ccgthtthgat cacggacgat ttaatacga ttgcttcggg aagttacgtht 4860
 tgggaaaacc aatthgttga tggthcttg acttacgcac ttgaagatgt ctggaccgga 4920
 attcaacaaa gttatacgc attagcagca gatgtccgca gtaaatatca catgagthtg 4980
 aagcatatca atgctatthg cattagthgc atgatgcaag gatacctagc atthgatcaa 5040
 caagcgaat tattagthcc gthtccgact tggcgtaata acattacggg gcaagcagca 5100
 gatgaattga ccgaattatt tgatthcaac attccacaac ggtggagat cgcacactta 5160
 taccaggcaa tcttaataa tgaagcgcac gthaaacagg tggacttcat aacaacgctg 5220
 gctggctatg taacctggaa atthtccggg gagaaagtht taggaatcgg tgatgctct 5280
 ggcgthtthcc caatthgatg aacgactgac acatacaatc agacgatgtht aaccaagtht 5340
 agccaacttg acaaagthta accgthtca tgggatatcc ggcaththtt accgcccgtt 5400
 ttaccagcgg gagccatthc tggaaagtht acggtgccc gggcgagctt actthgatcag 5460
 agcggcacgc tcgacgctgg cagthttht gcaccgccag aaggggatgc tggaacagga 5520
 atggtcggta cgaacagcgt ccgthaaacgc accggthaca tctcggthgg aacctcagca 5580

ES 2 551 515 T3

ttttcgatga acgttctaga taaaccattg tctaaagtct atcgcgatat tgatattggt 5640
 atgacgccag atgggtcacc agttgcaatg gtgcatgtta ataattgttc atcagatatt 5700
 aatgogtggg caacgatttt tcatgagttt gcagcccggg tgggaatgga attgaaaccg 5760
 gatcgattat atgaaacggt attcttgga tcaactcgcg ctgatgcgga tgctggaggg 5820
 ttggctaatt atagttatca atccggtgag aatattacta agattcaagc tggtcggccg 5880
 ctatttgtac ggacacccaa cagtaaattt agtttaccga actttatggt gactcaatta 5940
 tatgcccgtc tcgcaccctt ccaacttggt atggatattc ttgttaacga agaacatggt 6000
 caaacggacg ttatgattgc acagggtgga ttgttccgaa cgccggtaat tggccaacaa 6060
 gtattggcca acgcactgaa cattccgatt actgtaatga gtactgctgg tgaaggcggc 6120
 ccatggggga tggcagtgtt agccaacttt gcttgcggc aaactgcaat gaacctagaa 6180
 gatttcttag atcaagaagt ctttaaagag ccagaaagta tgacgttgag tccagaaccg 6240
 gaacgggtgg ccggatatcg tgaatttatt caacgttacc aagctggctt accagttgaa 6300
 gcagcggctg ggcaagcaat caaatattag agcttttgat taagccttct agtccaaaaa 6360
 acacgttttt ttgtcattta tttcattttc ttagaatagt ttagtttatt cattttatag 6420
 tcacgaatgt tttatgattc tatatagggt tgcaaacaag cttttttcat tttatgttaa 6480
 aacaatttca ggtttacctt ttattctgct tgtggtgacg cgggtatccg cccgctcttt 6540
 tggtcacca tgtatttaat tgcataaata attcttaaaa gtggagctag tctatttcta 6600
 tttacatacc tctcatttct catttcctcc actagtagag aattttgcca tcggacatgc 6660
 taccttacgc ttatatctct cattggaata tcgttttctg attaaaacac ggaagtaaga 6720
 acttaattcg tttttcgttg aactatgttg tgccagcgta acattaaaaa agagtgtaca 6780
 aggccacggt ctgtcaccgt cagaaaaata tgtcaatgag gcaagaaccg ggatggtaac 6840
 aaaaatcacg atctgggtgg gtgtgggtgt attggattat aggaagccac gcgctcaacc 6900
 tgggaattaca ggaagctggt aattttttgg gtttgcaatc atcaccatct gcacgttggt 6960
 ataatgtccc gtgtctatat atatccattg acgggtattct atttttttgc tattgaaatg 7020
 agcgtttttt gttactacaa ttggttttac agacggaatt ttccctattt gtttcgtccc 7080
 atttttcoct ttctcattgt tctcatalct taaaaaggtc ctttcttcat aatcaatgct 7140
 ttcttttact taatatttta cttgcattca gtgaatttta atacatattc ctctagtctt 7200
 gcaaaaatcga tttagaatca agataccagc ctaaaaatgc tagaagcatt aaaacaagaa 7260

ES 2 551 515 T3

gtttatgagg ctaacatgca gcttccaaag ctgggcctgg ttacttttac ctggggcaat 7320
 gtctcgggca ttgaccggga aaaaggccta ttcgtgatca agccatctgg tgttgattat 7380
 ggtgaattaa aaccaagcga tttagtcggt gttaacttac aggggtgaagt ggttgaaggt 7440
 aaactaaatc cgtctagtga tacgccgact catacggtgt tatataacgc ttttcctaata 7500
 attggcggaa ttgtccatac tcattcgcca tgggcagttg cctatgcagc tgctcaaatg 7560
 gatgtgccag ctatgaacac gacccatgct gatacgttct atggtgacgt gccggccgcg 7620
 gatgcgctga ctaaggaaga aattgaagca gattatgaag gcaacacggg taaaaccatt 7680
 gtgaagacgt tccaagaacg gggcctcgat tatgaagctg taccagcctc attagtcagc 7740
 cagcacggcc catttgcttg gggaccaacg ccagctaaag ccgtttacaa tgctaaagtg 7800
 ttggaagtgg ttgccgaaga agattatcat actgcgcaat tgaccctgac aagtagcgaa 7860
 ttaccacaat atttattaga taagcattat ttacgtaagc atggtgcaag tgcctattat 7920
 ggtcaaaaata atgcgcattc taaggatcat gcagttcgca agtaaacaaa tcgctcttaa 7980
 atatatacct aaagaacatt aaagctatat tataagcaaa gatacgtaaa ttttgcttat 8040
 attattatac acatatcata tttctatatt ttttaagattt ggttatataa tgtacgtaat 8100
 gcaaaggaaa taaattttat acattattga acagcgtcca agtaactaca ttatgtgcac 8160
 taatagttta gcgtcgtgaa gactttattg tgtcgcgaaa agtaaaaatt ttaaaaatta 8220
 gagcaccttg aacttgcgaa aaaggttctc atcaactggt taaaaacgcg tgtcttctgt 8280
 gtttcagttc agggcttttc ggaggatgtg aatcgacggc gtactgtcct tgggaacttt 8340
 gtctacgtat tttcacttcc tcagcgaatc cagagactat cttgggaaat tcgacaggac 8400
 agtctgttga caaccgactc ccttttgact tcataataaa aattcaatga cgcaaaagga 8460
 attttaggtt tttattattt atttatttat ttctgttaat tgatcctttt ctttccacta 8520
 ccaacaacaa aaaagggggg aaaaagatgt ataatctaaa agacactaat ctgctcttga 8580
 tatocttatt atgtaatgga ataactcata taaatgtaaa atagaacttc aaattaatat 8640
 tataatgata gtcgaggtca gacacactta taatacatta agtaaagaaa aaaaaatgct 8700
 tgtcatcgag gtctcttttg tgtcgctaac aaaacatcac taaatacgaa gacactttgc 8760
 atgggaagga tgcagcaaat ggcaactaa cgggccattg attggtttac ctcttctatt 8820
 tgtattacga ccagaaagaa cgaatggtt tcatcaatga ggtaggaaac gacctaaata 8880
 taatgtagca tagataaaat ctttgtactg tatggttgca atgccttctt gattagtatc 8940
 gaatttctg aataattttg ttaatctcat tagccaaact aacgcctcaa cgaatttatc 9000

ES 2 551 515 T3

aaactttagt tcttttcctg ttccatttct gtttataaac tcagcatatt ggtcaaatgt 9060
 tttctcgccta acttcaaaag gtattagata tcctagttct tgaagtgagt tatgaaattc 9120
 gcttacagaa atggtgagcg atccggtgat atcattgtcc acataaaactt ttctccaact 9180
 tttcactctt ttgtataggg cgatgaattc tgcttggtg acagtgcaa acctggaagc 9240
 accaaataaa tttatcagcg catctactga tgatatacaa aaatgggagt tgcctgctt 9300
 ttgtagtaag ttctgtagtt cctcagctgt cagtcggttt ttgcccttta catcatggtt 9360
 atgaaatagc tgtgtggcca cttgcatgtc tcgtacatct tctctgctat cgaacgaagc 9420
 aggtgcaact ttcttcaaga gttgtgcagg cactgcttga ttgtgaatta ggggaggagg 9480
 agaggaagct atccggtgag cggaaagtgt caagttgta taatgggtt ggcctggagg 9540
 tataggcctg cctgctggtt tctgtgcgat aacattatat ctaggatcca caggtgtttt 9600
 cgtatgtctt ggagaataac tttggggaga accataggag tggtgaccgt tttctgctct 9660
 gtttttgta tattgagttt gtaagggaaat tggagctgag tggactctag tgttgggagt 9720
 ttgtgcttga gtaaccgta ccaaggctcc tcgctgcaga cctgcgagca gggaaacgct 9780
 cccctcacag tcgcgttgaa ttgtccccac gccgcgcccc thtagagaaa tataaaaggt 9840
 taggatttgc cactgaggtt cttctttcat atacttctt ttaaaatctt gctaggatac 9900
 agttctcaca tcacatccga acataaaca ccatgggtaa ggaaaagact cacgtttcga 9960
 ggccgcgatt aaattccaac atggatgctg atttatatgg gtataaatgg gctcgcgata 10020
 atgtcgggca atcaggtgcg acaatctatc gattgtatgg gaagccgat gcgccagagt 10080
 tgtttctgaa acatggcaaa ggtagcgttg ccaatgatgt tacagatgag atggtcagac 10140
 taaactggct gacggaattt atgcctcttc cgaccatcaa gcattttatc cgtactcctg 10200
 atgatgcatg gttactcacc actgcgatcc ccggcaaac agcattccag gtattagaag 10260
 aatctctga ttcaggtgaa aatattggtg atgcgctggc agtgttctg cgccggttgc 10320
 attcgattcc tgtttgtaat tgtcctttta acagcgatcg cgtatttctg ctgcctcagg 10380
 cgcaatcacg aatgaataac ggtttggtt atgcgagtga ttttgatgac gagcgtaatg 10440
 gctggcctgt tgaacaagtc tggaaagaaa tgcataagct tttgccattc tcaccggatt 10500
 cagtcgtcac tcatggtgat ttctcacttg ataacctat ttttgacgag gggaaattaa 10560
 taggttgat tgatgttga cgagtcggaa tcgcagaccg ataccaggat cttgccatcc 10620
 tatggaactg cctcggtgag tttctcctt cattacagaa acggcttttt caaaaatag 10680

ES 2 551 515 T3

gtattgataa tcctgatatg aataaattgc agtttcattt gatgctcgat gagtttttct 10740
 aatcagtact gacaataaaa agattcttgt tttcaagaac ttgtcatttg tatagttttt 10800
 ttatattgta gttgttctat tttaatcaaa tgtttagcgtg atttatattt tttttcgcct 10860
 cgacatcatc tgcccagatg cgaagttaag tgcgcagaaa gtaatatcat gcgtcaatcg 10920
 tatgtgaatg ctggtecgta tactgctgtc gattcgatac taacgccgcc atccagggta 10980
 ccatcctttt gttgtttccg ggtgtacaat atggacttcc tcttttctgg caaccaaacc 11040
 catacatcgg gattcctata ataccttcgt tgggtctcct aacatgtagg tggcggaggg 11100
 gagatataca atagaacaga taccagacaa gacataatgg gctaaacaag actacaccaa 11160
 ttacactgcc tcattgatgg tggtagataa cgaactaata ctgtagccct agacttgata 11220
 gccatcatca tatcgaagtt tcaactacct ttttccattt gccatctatt gaagtaataa 11280
 taggcgcatg caacttcttt tctttttttt tcttttctct ctcccccggt gttgtctcac 11340
 catatccgca atgacaaaaa aatgatgga agacactaaa ggaaaaaatt aacgacaaag 11400
 acagcaccaa cagatgtcgt tgttccagag ctgatgaggg gtatcttcga acacacgaaa 11460
 ctttttctct ccttcattca cgcacactac tctctaata gcaacgggat acggccttcc 11520
 ttccagttac ttgaatttga aataaaaaaa gtttgccgct ttgctatcaa gtataaatag 11580
 acctgcaatt attaactctt tgtttcctcg tcattgttct cgttcccttt ctcccttgtt 11640
 tctttttctg cacaatattt caagctatac caagcataca atcaactatc tcatatacaa 11700
 tgccctcaatc ctgggaagaa ctggccgctg ataagcgcgc ccgcctcgca aaaaccatcc 11760
 ctgatgaatg gaaagtccag acgctgcctg cggaagacag cgttattgat ttcccaaaga 11820
 aatcggggat cctttcagag gccgaactga agatcacaga ggcctccgct gcagatcttg 11880
 tgtccaagct gggggccgga gagttgacct cggtggaagt tacgctagca ttctgtaaac 11940
 gggcagcaat cgcccagcag ttaacaaact gcgcccacga gttcttcctt gacgccgctc 12000
 tcgcgcaggc aaggaactc gatgaatact acgcaaagca caagagacc gttggtccac 12060
 tccatggcct ccccatctct ctcaaagacc agcttcgagt caagggctac gaaacatcaa 12120
 tgggctacat ctcatggcta aacaagtacg acgaagggga ctcggttctg acaaccatgc 12180
 tccgcaaagc cggtgccgct ttctacgtca agacctctgt cccgcagacc ctgatggtct 12240
 gcgagacagt caacaacatc atcggggcga ccgtcaacc acgcaacaag aactggtcgt 12300
 gcggcgccag ttctggtggt gaggggtcga tcgttgggat tcgtggtggc gtcacgggtg 12360
 taggaacgga tatcgggtggc tcgattcgag tgccggccgc gttcaacttc ctgtacggtc 12420

ES 2 551 515 T3

taaggccgag tcatgggagg ctgccgtatg caaagatggc gaacagcatg gagggtcagg 12480
 agacgggtgca cagcgttgctc gggccgatta cgcactctgt tgaggacctc cgcctcttca 12540
 ccaaatccgt cctcgggtcag gagccatgga aatacgactc caaggtcatc cccatgccct 12600
 ggcgccagtc cgagtcggac attattgcct ccaagatcaa gaacggcggg ctcaatatcg 12660
 gctactacaa cttcgacggc aatgtccttc cacaccctcc tatcctgagc ggctggaaa 12720
 ccaccgtgac cgcactcgcc aaagccggtc acaccgtgac cccgtggacg ccatacaagc 12780
 acgatttcgg ccacgatctc atctccata tctacggcgc tgacggcagc gccgacgtaa 12840
 tgcgagatag cagtgcattc ggcgagccgg cgattccaaa tatcaaagac ctactgaacc 12900
 cgaacatcaa agctgttaac atgaacgagc tctgggacac gcatctccag aagtggatt 12960
 accagatgga gtacctgag aatggcggg aggctgaaga aaaggccggg aaggaaactg 13020
 acgccatcat cgcgccgatt acgcctaccg ctgcggtagc gcatgaccag ttccggtact 13080
 atgggtatgc ctctgtgatc aacctgctgg atttcacgag cgtggttgtt ccggttacct 13140
 ttgcggataa gaacatcgat aagaagaatg agagtttcaa ggccggttagt gagcttgatg 13200
 ccctcgtgca ggaagagtat gatccggagg cgtaccatgg ggcaccggtt gcagtgcagg 13260
 ttatcggacg gagactcagt gaagagagga cgttggcagat tgcagaggaa gtggggaagt 13320
 tgctgggaaa tgtggtgact ccataggtcg agaatttata cttagataag tatgtactta 13380
 caggatatt tctatgagat actgatgat acatgcatga taatatttaa acggttatta 13440
 gtgccgattg tcttgtgca taatgacgtt cctatcaaag caatacactt accacctatt 13500
 acatgggcca agaaaatatt ttcgaacttg tttagaatat tagcacagag tatatgatga 13560
 tatccgtagg attatgcatg attcattcct acaacttttt cgtagcataa ggattaatta 13620
 cttggatgcc aataaaaaaa aaaaacatcg agaaaatttc agcatgctca gaaacaattg 13680
 cagtgtatca aagtaaaaaa aagattttcg ctacatgttc cttttgaaga aagaaaatca 13740
 tggaacatta gatttacaaa aatttaacca ccgctgatta acgattagac cgttaagcgc 13800
 acaacagggtt attagtacag agaaagcatt ctgtggtggt gccccggact ttcttttgcg 13860
 acataggtaa atcgaatacc atcactat cttttccaat gactccctaa agaaagactc 13920
 ttcttcgatg ttgtatacgt tggagcatag ggcaagaatt gtggcttgag atctagatta 13980
 cgtggaagaa aggtagtaaa agtagtagta taagtagtaa aaagagtaa aaagagaaaa 14040
 ccggctacat actagagaag cacgtacaca aaaactcata ggcacttcat catacgacag 14100

ES 2 551 515 T3

tttcttgatg cattataata gtgtattaga tattttcaga aatatgcata gaacctcttc 14160
 ttgcctttac tttttataca tagaacattg gcagatttac ttacactact ttgtttctac 14220
 gccatttctt ttgttttcaa cacttagaca agttgttgag aaccggacta ctaaaaagca 14280
 atgttcccac tgaaaatcat gtacctgcag gataataacc ccctaattct gcatcgatcc 14340
 agtatgtttt tttttctcta ctcatTTTTa cctgaagata gagcttctaa aacaaaaaaaa 14400
 atcagcgatt acatgcatat tgtgtgttct agaattgCGG atcaccagat cgccattaca 14460
 atgtatgcag gcaaatatTTt ctCagaatga aaaatagaga aaaggaaacg aaaattctgt 14520
 aagatgcctt cgaagagatt tctcgatatg caaggCGTgc atcagggTga tccaaaggaa 14580
 CTCgagagag agggcgaaag gcaatttaat gcattgcttc tccattgact tctagtTgag 14640
 cggataagtt cggaaatgta agtcacagct aatgacaaat ccactttagg tttcgaggca 14700
 ctatttaggc aaaaagacga gtggggaaat aacaaacgct caaacatatt agcatatacc 14760
 ttcaaaaaat gggaaatagta tataaccttc cggttcgTta ataaatcaa tctttcatct 14820
 agttctctta agatttcaat attttgcttt ctTgaagaaa gaatctactc tctccccca 14880
 ttcgcactgc aaagctagct tggcactggc cgtcgTTTTa caacgTcgTg actgggaaaa 14940
 ccoTggcctt acccaactta atcgCctTgc agcacatccc cttttcgcca gctggcgtaa 15000
 tagcgaagag gcccgCaccg atcgcccttc ccaacagTtg cgcagcctga atggcgaatg 15060
 ggaaattgta aacgttaata ttttgTtaa attcgcgTta aatttttgtt aaatcagctc 15120
 attttttaac caatagccg aaatcgGcaa aatcccttat aaatcaaaag aatagaccga 15180
 gatagggTtg agtgtTgttc cagtttgGaa caagagTcca ctattaaaga acgtggactc 15240
 caacgtcaaa gggcgaaaaa cGtctatca gggcgatggc ccactacgtg aaccatcacc 15300
 ctaatcaagt tttttggggt cgaggtgCcg taaagcacta aatcggaacc ctaaaggag 15360
 cccccgattt agagctTgac ggggaaagcc ggcgaacgtg gcgagaaagg aagggaagaa 15420
 agcgaaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac 15480
 cacaccgccc gcgcttaatg cgccgctaca gggcgcgTca ggtggcactt ttcggggaaa 15540
 tgtgcgcgga acccctattt gtttattttt ctaaatacat tcaaataTgt atccgctcat 15600
 gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca 15660
 acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttctg tttttgctca 15720
 occagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta 15780
 catcgaactg gatctcaaca gcggtaaat ccttgagagt tttcgccccg aagaacgttt 15840

ES 2 551 515 T3

tccaatgatg agcactttta aagttctgct atgtggcgcg gtattatccc gtattgacgc 15900
 cgggcaagac caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtactc 15960
 accagtcaca gaaaagcadc ttacggatgg catgacagta agagaattat gcagtgtctgc 16020
 cataacccatg agtgataaca ctgcggccaa cttacttctg acaacgatcg gaggaccgaa 16080
 ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg atcgttggga 16140
 accggagctg aatgaagcca taccaaacga cgagcgtgac accacgatgc ctgtagcaat 16200
 ggcaacaacg ttgcgcaaac tattaactgg cgaactactt agtctagctt cccggcaaca 16260
 attaatagac tggatggagg cggataaagt tgcaggacca cttctgcgct cggcccttcc 16320
 ggctggctgg tttattgctg ataaatctgg agccgggtgag cgtgggtctc gcggtatcat 16380
 tgcagcactg gggccagatg gtaagccctc ccgatcgtg gttatctaca cgacggggag 16440
 tcaggcaact atggatgaac gaaatagaca gatcgtgag ataggtgcct cactgattaa 16500
 gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt taaaacttca 16560
 tttttaatth aaaaggatct aggtgaagat cttttttgat aatctcatga ccaaatccc 16620
 ttaacgtgag ttttcgttcc actgagcgtc agaccccgta gaaaagatca aaggatcttc 16680
 ttgagatcct ttttttctgc gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc 16740
 agcggtggtt tgtttgccgg atcaagagct accacctctt tttccgaagg taactggctt 16800
 cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt 16860
 caagaactct gtagcaccgc ctacatacct cgctctgcta atcctgttac cagtggctgc 16920
 tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa 16980
 ggcgcagcgg tcgggctgaa cgggggggtc gtgcacacag cccagcttgg agcgaacgac 17040
 ctacaccgaa ctgagatacc tacagcgtga gcattgagaa agcggccacgc ttcccgaagg 17100
 gagaaaggcg gacaggtatc cggtaagcgg cagggtcggg acaggagagc gcacgagggg 17160
 gcttccaggg ggaaacgcct ggtatcttta tagtctgtc gggtttcgcc acctctgact 17220
 tgagcgtcga tttttgtgat gctcgtcagg ggggcgagc ctatggaaaa acgccagcaa 17280
 cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt tctttcctgc 17340
 gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg ataccgctcg 17400
 ccgcagccga acgaccgagc gcagcgagtc agtgagcgag gaagcggaag agcgccaat 17460
 acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc acgacaggtt 17520

ES 2 551 515 T3

tccccgactgg aaagcgggca gtgagcgcaa cgcaattaat gtgagttagc tcaactcatta 17580
 ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg 17640
 ataacaattt cacacaggaa acagctatga catgattacg aatttaatac gactcacaat 17700
 agggaattag cttgcgcgaa attattggct tttttttttt ttttaattaat actacctttt 17760
 gatgtgaacg tttactaaag tagcactatc tgtggaatgg ctggttggaaac tttttccgat 17820
 taacagcttg tattccaagt cctgacattc cagttgtaag tttccaact tgtgattcaa 17880
 ttgttcaatc tcttggttaa aattctcttg ttccatgaat aggctctttt tccagtctcg 17940
 aaattttgaa atttctctgt tggacagctc gttgaatttt ttcttagctt ctaattgtct 18000
 agttataaat tcaggatccc attctgtagc caccttatcc atgaccgttt tattaattat 18060
 ttcatagcac ttgtaatttt tgagtttggt ttccctcgatt tcatcgaagt tcatttcttc 18120
 ctccaaaaat ttcttttggt cttccgttat gtcaaacatt ttcgttgta agcaatctct 18180
 ggcctttaat agcctagttc ttagcatttc agatc 18215

<210> 18
 <211> 23
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> ADN sintético

<400> 18
 tgatcttga gaaagtaccg agg 23

<210> 19
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> ADN sintético

<400> 19
 cttgttctt ccggtatgc aacac 25

<210> 20
 <211> 23
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> ADN sintético

<400> 20
 ttccaagaag aacaacctga tag 23

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

<220>

ES 2 551 515 T3

<223> ADN sintético

<400> 21

tgatgtgaac gttactaaa g

21

REIVINDICACIONES

- 5 1. Una célula que comprende una secuencia de nucleótidos que codifica una xilosa isomerasa, donde la secuencia de aminoácidos de la xilosa isomerasa tiene al menos aproximadamente 96% de identidad de secuencia con la secuencia de aminoácidos representada en SEQ ID NO 3 donde la secuencia de nucleótidos es heteróloga al huésped y donde la célula es capaz de utilizar xilosa como fuente de carbono.
2. Una célula de acuerdo a la reivindicación 1 la cual es una célula eucariótica.
3. Una célula de acuerdo con la reivindicación 1 o 2 la cual es una célula de levadura.
- 10 4. Una célula de acuerdo con la reivindicación 3 la cual es una célula de levadura del género *Saccharomyce*, *Kluyveromyces*, *Cándida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Klockera*, *Schwanniomyces* o *Yarrowia*.
5. Una célula de acuerdo con la reivindicación 4, donde la célula de levadura es de la especie *S. cerevisiae*, *S. bulderi*, *S barnetti*, *S.exiguus*, *S.uvarum*, *S.diastaticus*, *K. lactis*, *K. marxianus* o *K. fragilis*.
6. Una célula de acuerdo con la reivindicación 1 o 2 la cual es una célula fúngica filamentosa.
- 15 7. Una célula de acuerdo con la reivindicación 6 donde la célula fúngica filamentosa es del género *Aspergillus*, *Penicillium*, *Rhizopus*, *Trichoderma*, *Humicola*, *Acremonium* o *Fusarium*.
8. Una célula de acuerdo con la reivindicación 7, donde la célula fúngica filamentosa es de la especie *Aspergillus niger*, *Aspergillus oryzae*, *Penicillium chrysogenum*, o *Rhizopus oryzae*.
9. Una célula de acuerdo con una cualquiera de las reivindicaciones precedentes, donde la célula comprende una o más modificaciones genéticas que dan como resultado:
 - 20 a. un incremento en el transporte de xilosa en la célula;
 - b. un incremento en la actividad de xilulosa quinasa;
 - c. un incremento en el flujo a través de la ruta del fosfato de pentosa;
 - d. una disminución en la actividad de la aldosa reductasa;
 - e. una disminución en la sensibilidad a la represión de catabolitos;
 - 25 f. un incremento en la tolerancia al etanol, osmolaridad u ácidos orgánicos; o
 - g. una reducción reducida de subproductos.
10. Una célula de acuerdo con la reivindicación 9, donde la una o más modificaciones genéticas dan como resultado la sobreexpresión de al menos un gen que codifica una enzima de la parte no oxidativa de la ruta del fosfato de pentosa.
- 30 11. Una célula de acuerdo con la reivindicación 10, donde el gen es un gen que codifica una ribulosa-5-fosfato isomerasa, una ribulosa-5-fosfato epimerasa, una transcetolasa o una transaldolasa.
12. Una célula de acuerdo con la reivindicación 10 u 11, donde la una o más modificaciones genéticas dan como resultado una sobreexpresión de al menos los genes que codifican una transcetolasa y una transaldolasa.
- 35 13. Una célula de acuerdo con una cualquiera de las reivindicaciones 9 a 12, donde la una o más modificaciones genéticas dan como resultado la sobreexpresión de un gen que codifica una xilulosa quinasa.
14. Una célula de acuerdo con una cualquiera de las reivindicaciones 8 a 13, donde el gen que es sobreexpresado es un gen que es endógeno a la célula.
15. Una célula de acuerdo con una cualquiera de las reivindicaciones 9 a 14, donde la una o más modificaciones genéticas dan como resultado un descenso en la actividad no especificada en la aldosa reductasa en la célula.
- 40 16. Una célula de acuerdo con la reivindicación 15, donde la una o más modificaciones genéticas reducen la expresión de un gen endógeno que codifica una aldosa reductasa no específica o reduce la actividad de dicha aldosa reductasa no específica.
- 45 17. Una célula de acuerdo con la reivindicación 16, donde el gen es inactivado por eliminación de al menos parte del gen o por interrupción del gen.

18. Una célula de acuerdo con la reivindicación 16 o 17, donde la expresión de cada gen en la célula que codifica una aldosa reductasa no específica es reducida.
19. Una célula de acuerdo con una cualquiera de las reivindicaciones precedentes que tiene la capacidad de utilizar L-arabinosa.
- 5 20. Una célula de acuerdo con cualquiera de las reivindicaciones 9 a 19, donde los genes TAL1, TKL1, RPE1 y RKI1 están sobreexpresados.
21. Una célula de acuerdo con cualquiera de las reivindicaciones 9 a 20, donde la región codificante del gen GRE3 es inactivado por reemplazo de la región codificante con una secuencia de nucleótidos que comprende los genes TAL1, TKL1, RPE1 y RKI1.
- 10 22. Una célula de acuerdo con una cualquiera de las reivindicaciones 9 a 21, donde los genes araA, araB y araD de *Lactobacillus plantarum* son expresados.
23. Una célula de acuerdo con cualquiera de las reivindicaciones 9 a 22, donde todos los genes expresados son expresados constitutivamente o sobreexpresados constitutivamente.
- 15 24. Una célula de acuerdo con la reivindicación 23 donde uno o más genes expresados constitutivamente o sobreexpresados constitutivamente están integrados de forma estable en el genoma de la célula.
25. Una célula de acuerdo con la reivindicación 24, donde todos los genes expresados constitutivamente o sobreexpresados constitutivamente están integrados de forma estable en el genoma de la célula.
- 20 26. Un proceso para producir un producto de fermentación proceso que comprende fermentar un medio que contiene una fuente de xilosa con una célula de acuerdo con una cualquiera de las reivindicaciones precedentes de forma que la célula fermenta la xilosa hasta el producto de fermentación.
27. Un proceso para producir un producto de fermentación proceso que comprende fermentar un medio que contiene al menos una fuente de xilosa y una fuente de L-arabinosa con una célula como se define en cualquiera de las reivindicaciones 19 a 23 de tal forma que la célula fermenta la xilosa y la L-arabinosa hasta el producto de fermentación.
- 25 28. Un proceso para producir un producto de fermentación proceso que comprende fermentar un medio que contiene al menos una fuente de xilosa y una fuente de L-arabinosa con una célula como se define en cualquiera de las reivindicaciones 1 a 18 y una célula capaz de usar L-arabinosa, por lo cual cada célula fermenta xilosa y/o arabinosa hasta el producto de fermentación.
- 30 29. Un proceso de acuerdo con una cualquiera de las reivindicaciones 26 a 28, que comprende recuperar el producto de fermentación.
- 30 30. Un proceso de acuerdo con cualquiera de las reivindicaciones 26 a 29 donde el medio también contiene una fuente de glucosa.
- 35 31. Un proceso de acuerdo con cualquiera de las reivindicaciones 28 a 30, donde el producto de fermentación es etanol, butanol, ácido láctico, ácido 3-hidroxipropiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, ácido málico, ácido fumárico, ácido itacónico, un aminoácido, 1,3-propanodiol, etileno, glicerol, butanol, un antibiótico de β -lactama y una cefalosporina.
32. Un proceso de acuerdo con una cualquiera de las reivindicaciones 26 a 31, donde el proceso es anaeróbico.
- 40 33. Un proceso de acuerdo con una cualquiera de las reivindicaciones 26 a 32, donde el proceso es aeróbico, preferiblemente llevado a cabo bajo condiciones limitadas de oxígeno.
34. Uso de una célula de acuerdo con las reivindicaciones 1 a 25 en un proceso para la producción de un producto de fermentación.

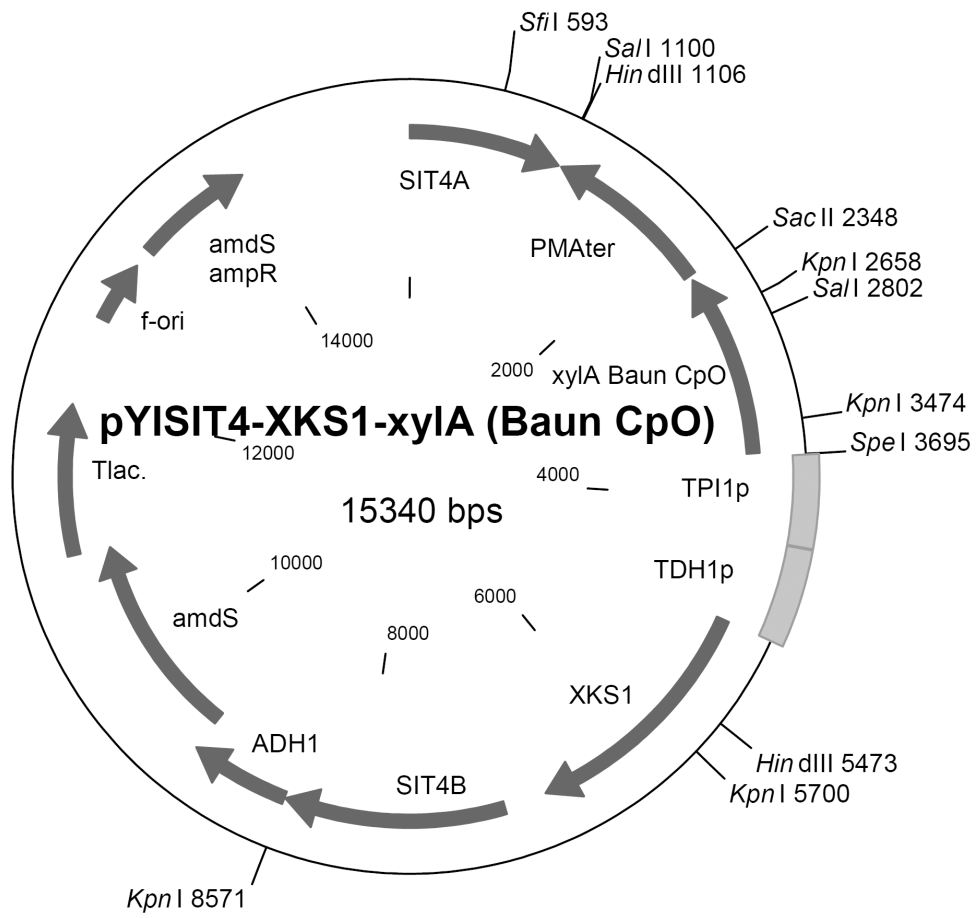


Figura 1

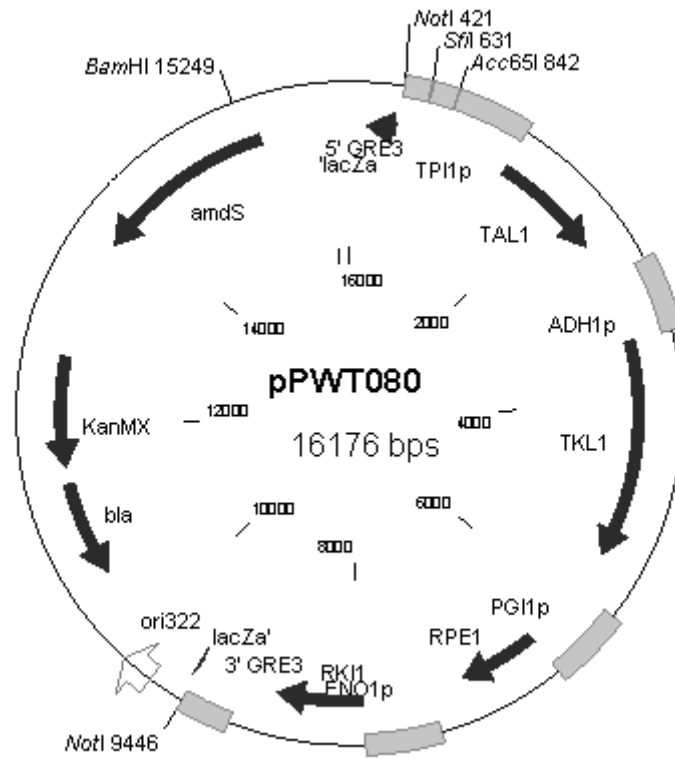
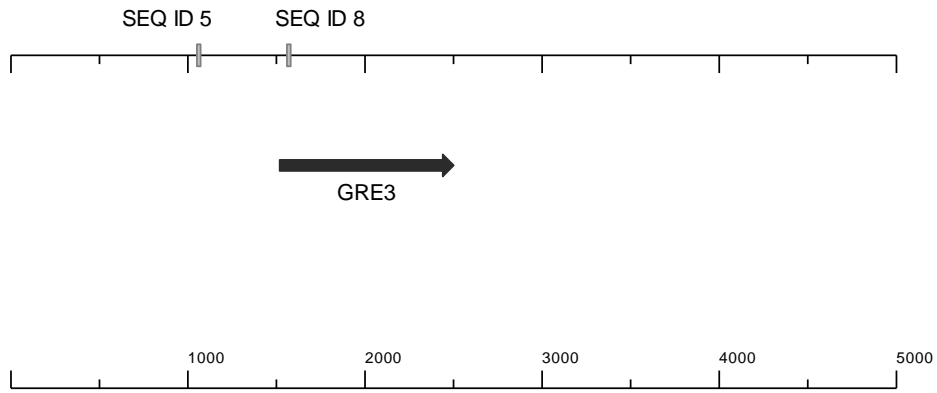
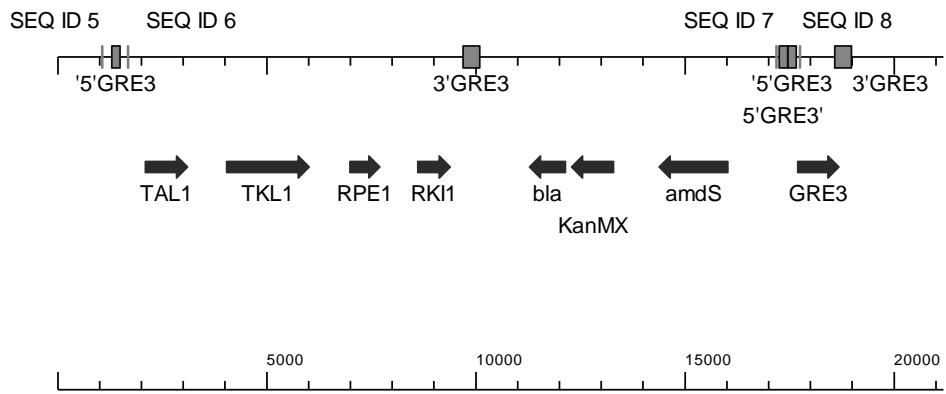


Figura 2

ES 2 551 515 T3

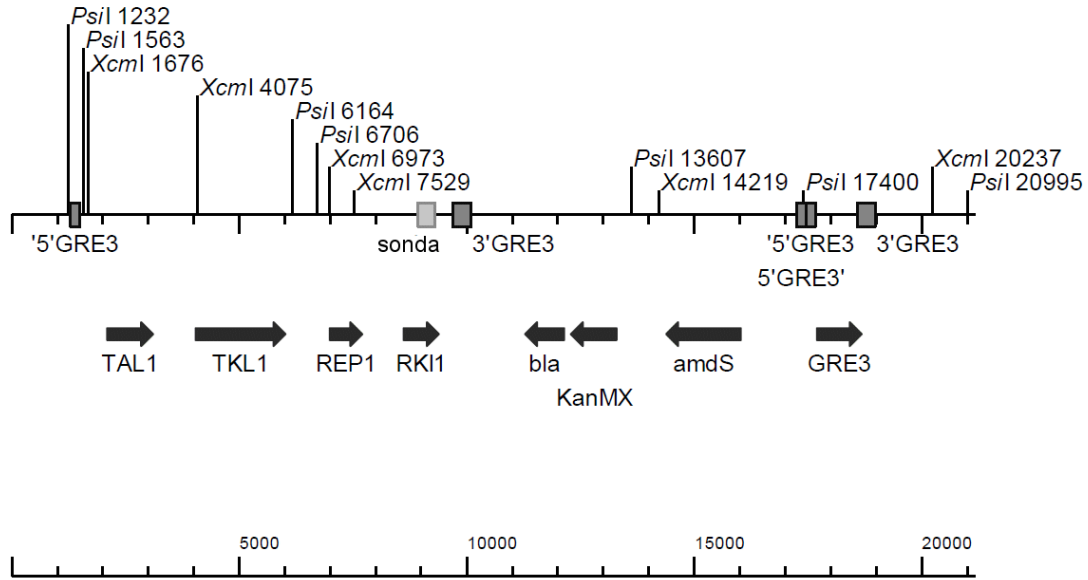


Panel a



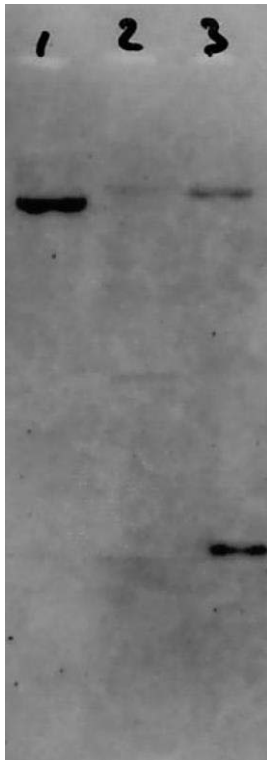
Panel b

Figura 3

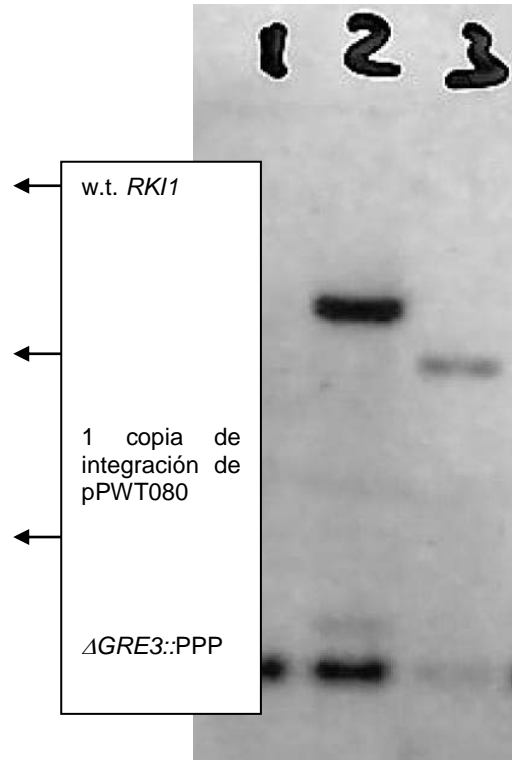


Panel c

Figura 3



Panel a



Panel b

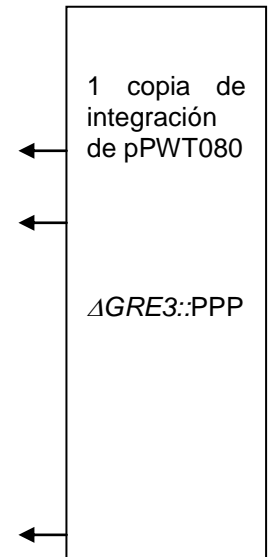
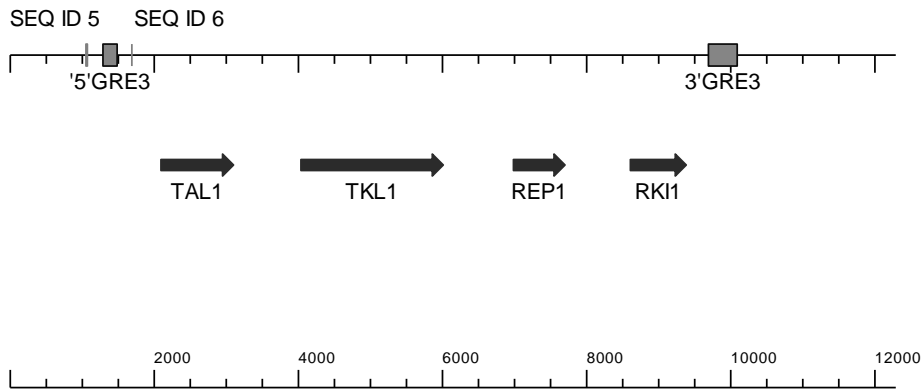
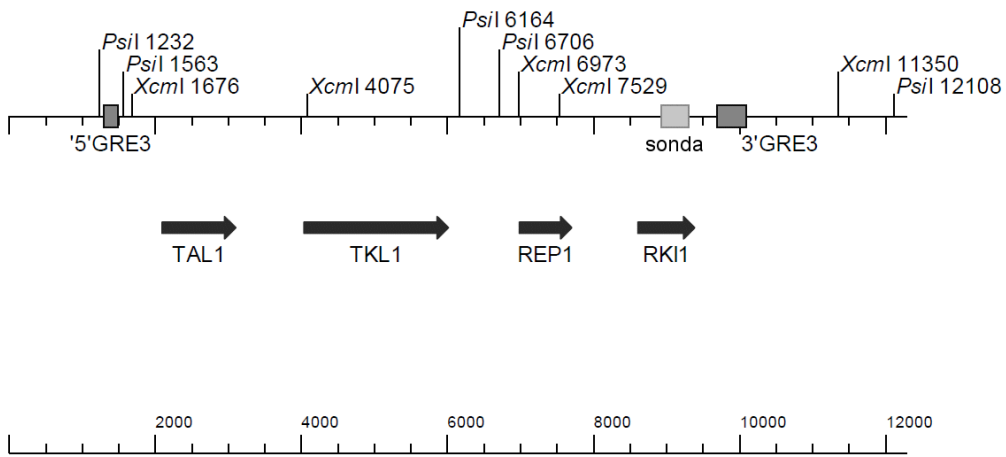


Figura 4

ES 2 551 515 T3



Panel a



Panel b

Figura 5

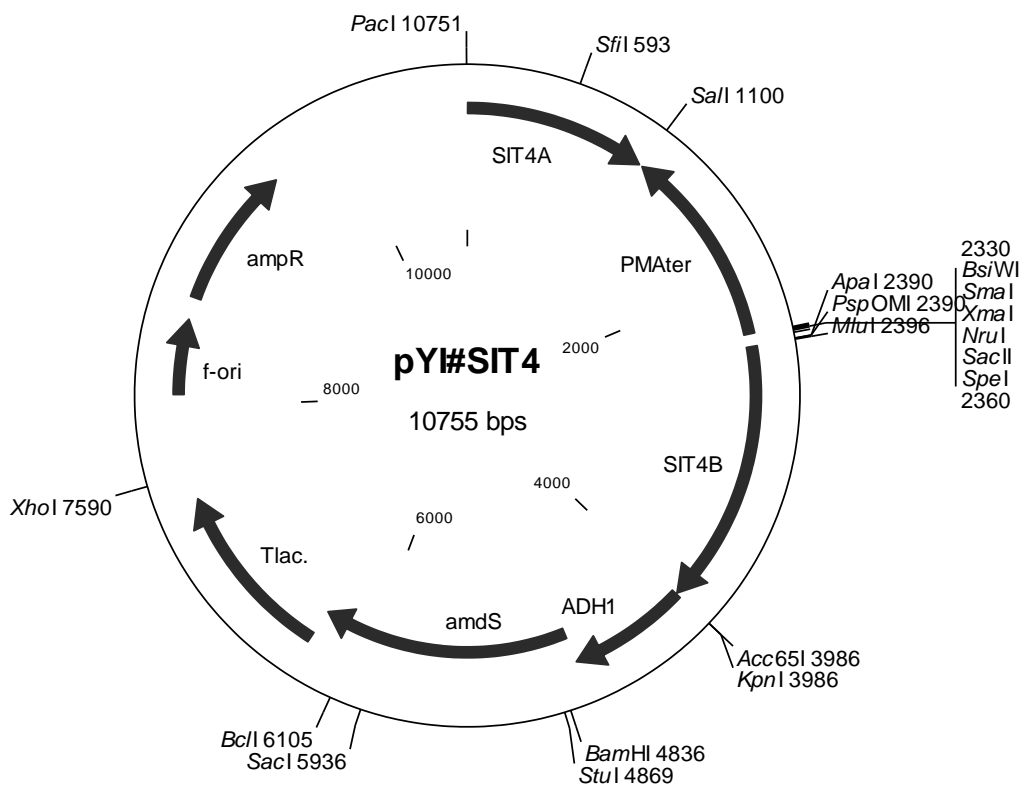


Figura 6

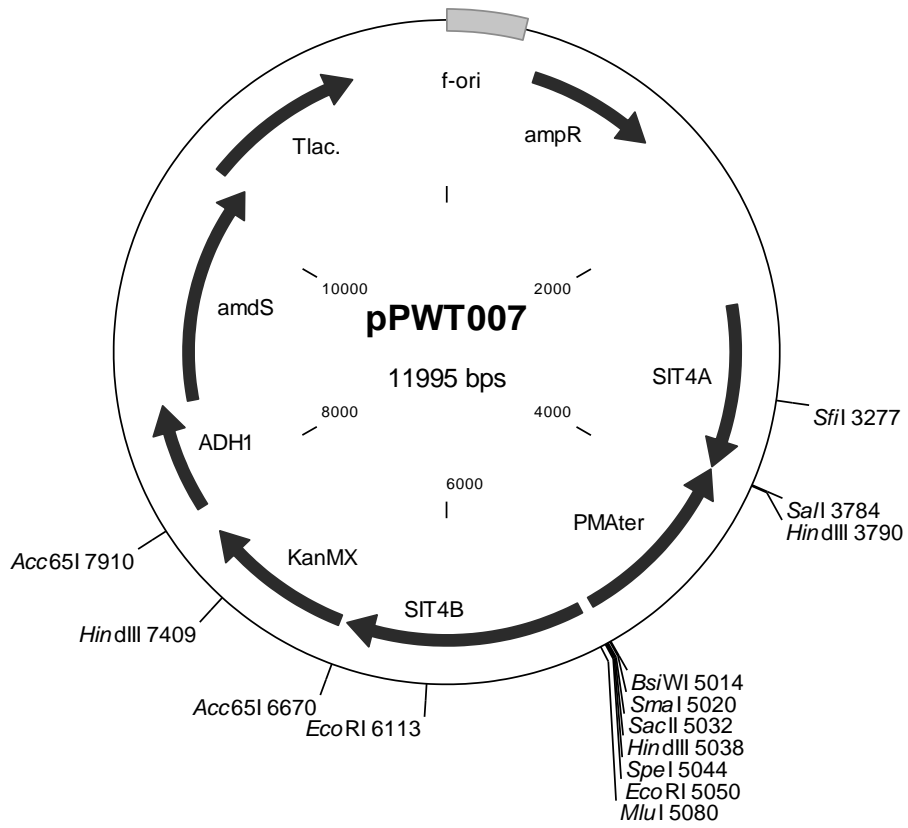


Figura 7

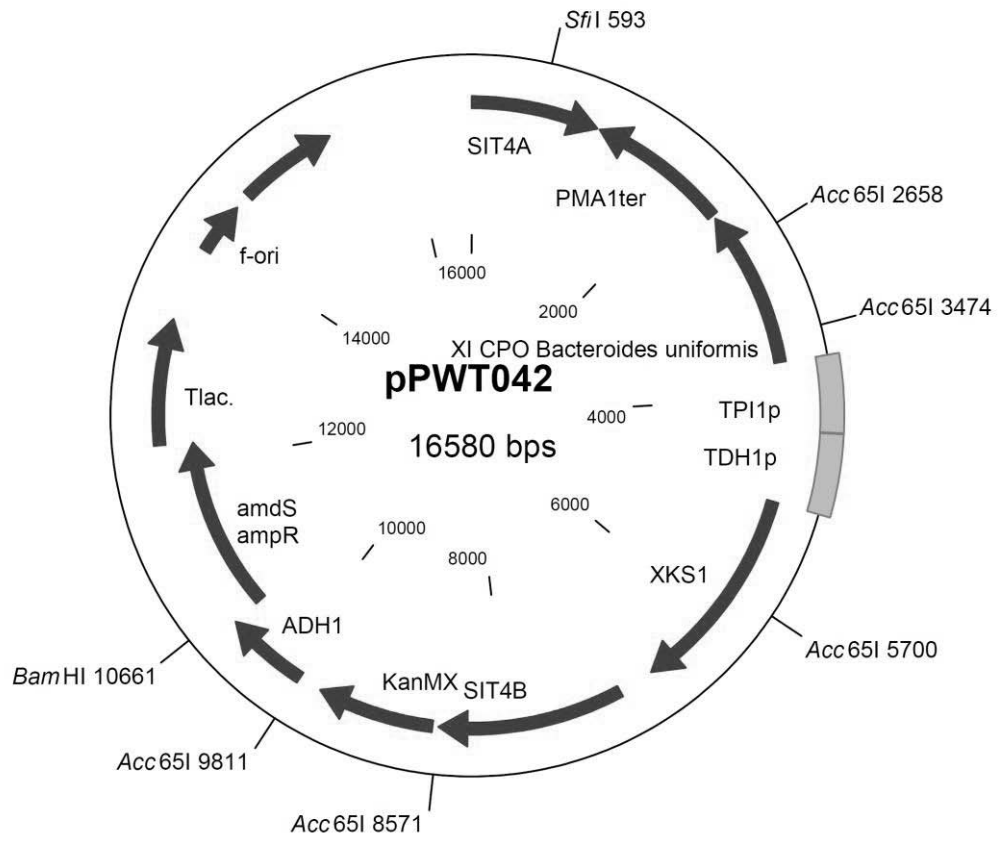
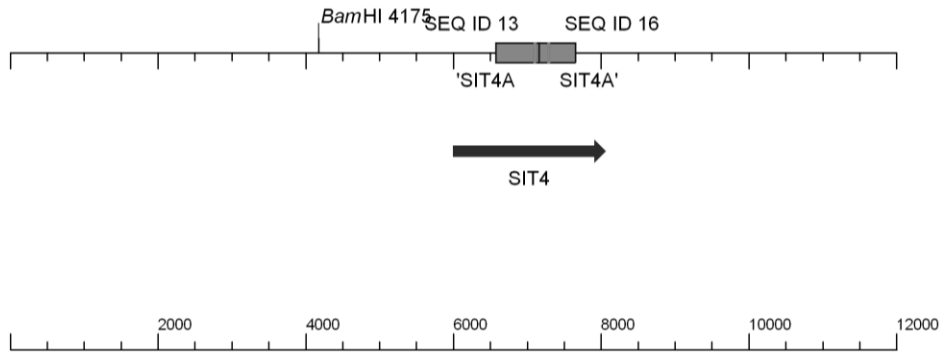
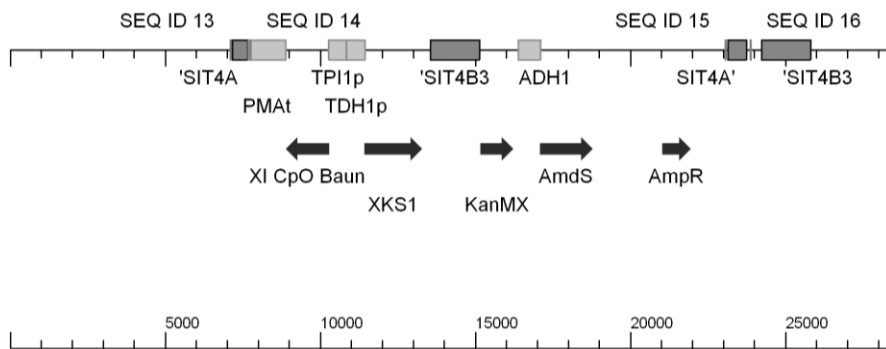


Figura 8

ES 2 551 515 T3



Panel a



Panel b

Figura 9

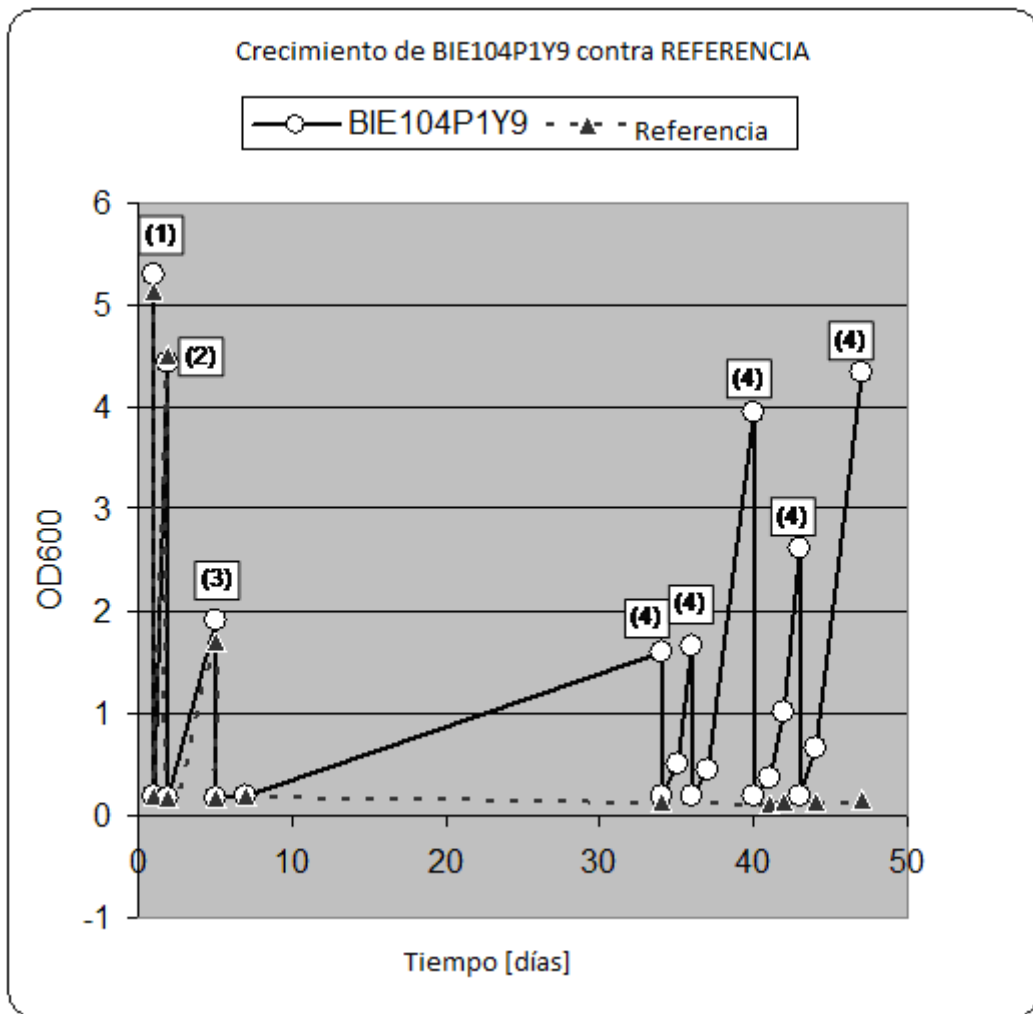


Figura 10

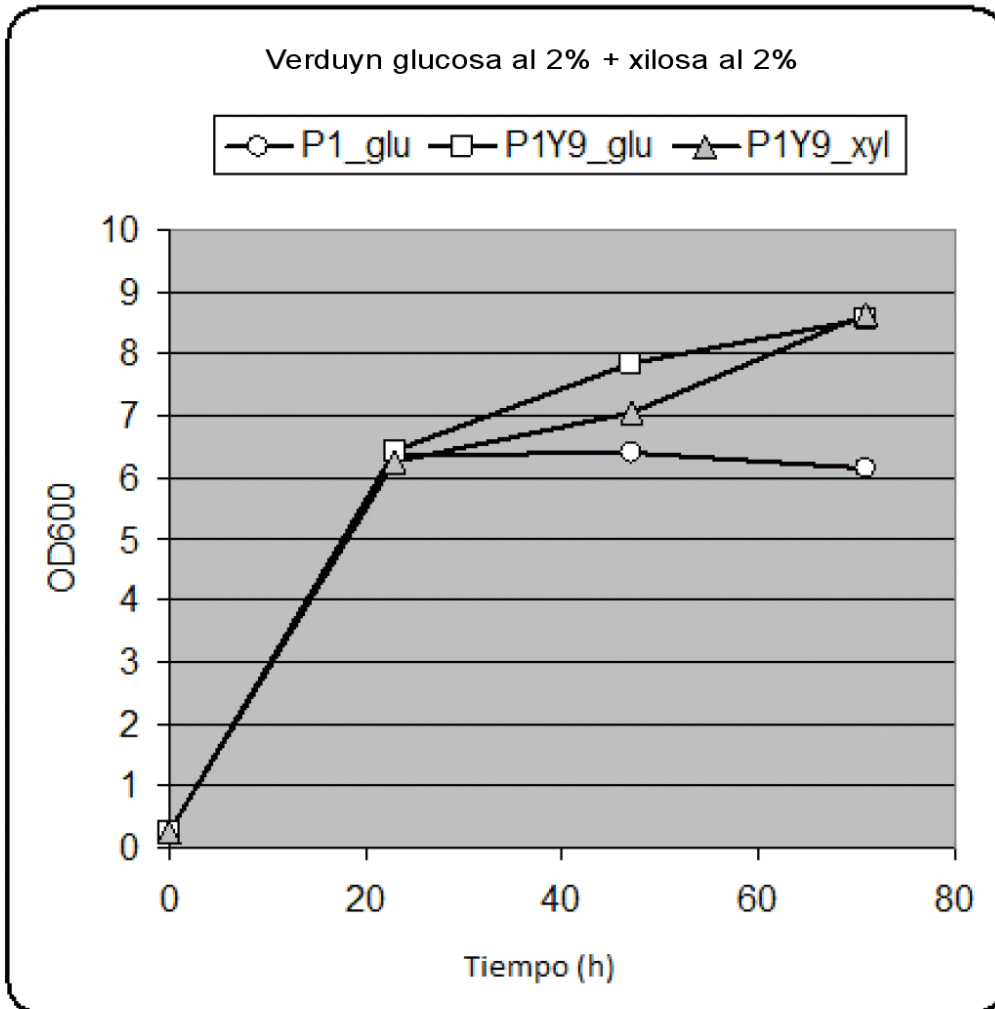
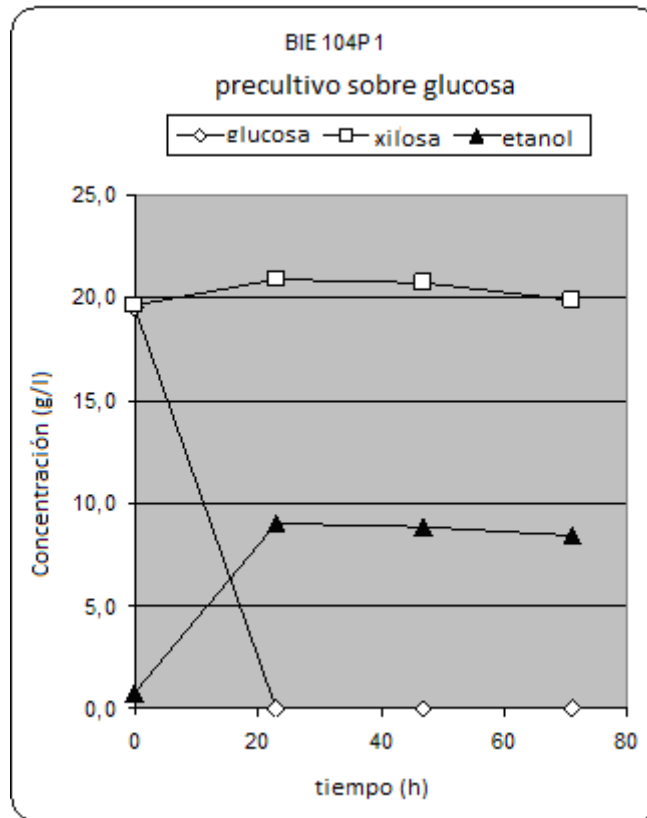
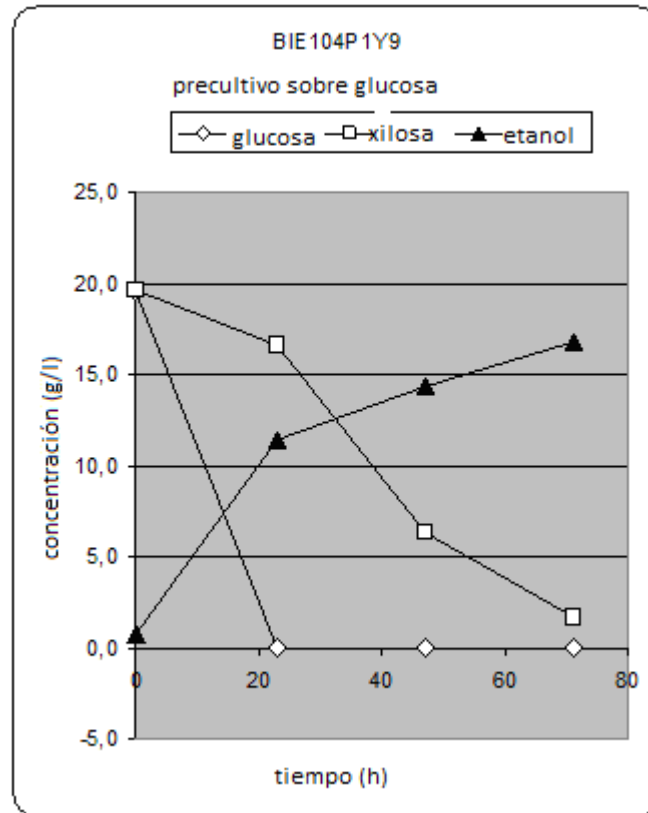


Figura 11



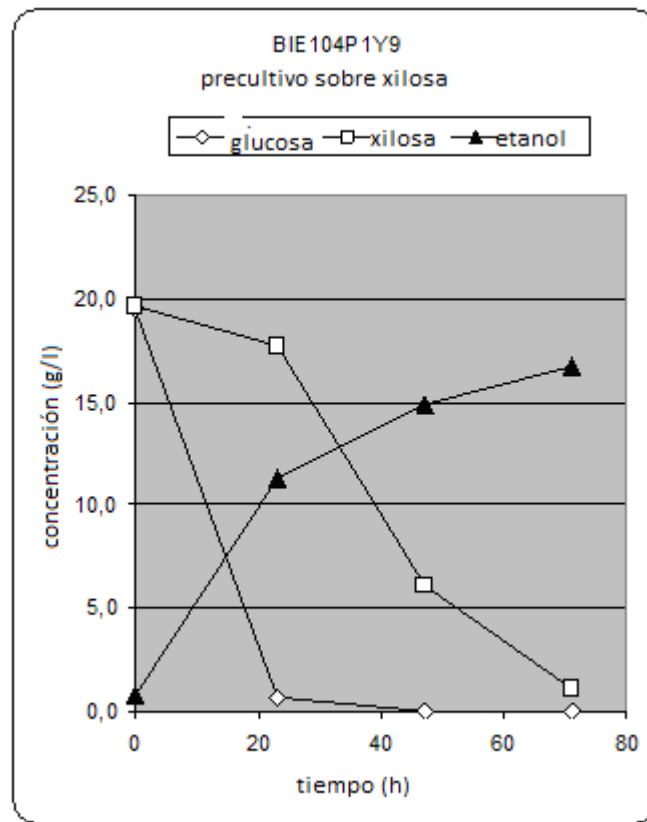
Panel a

Figura 12



Panel b

Figura 12



Panel c

Figura 12

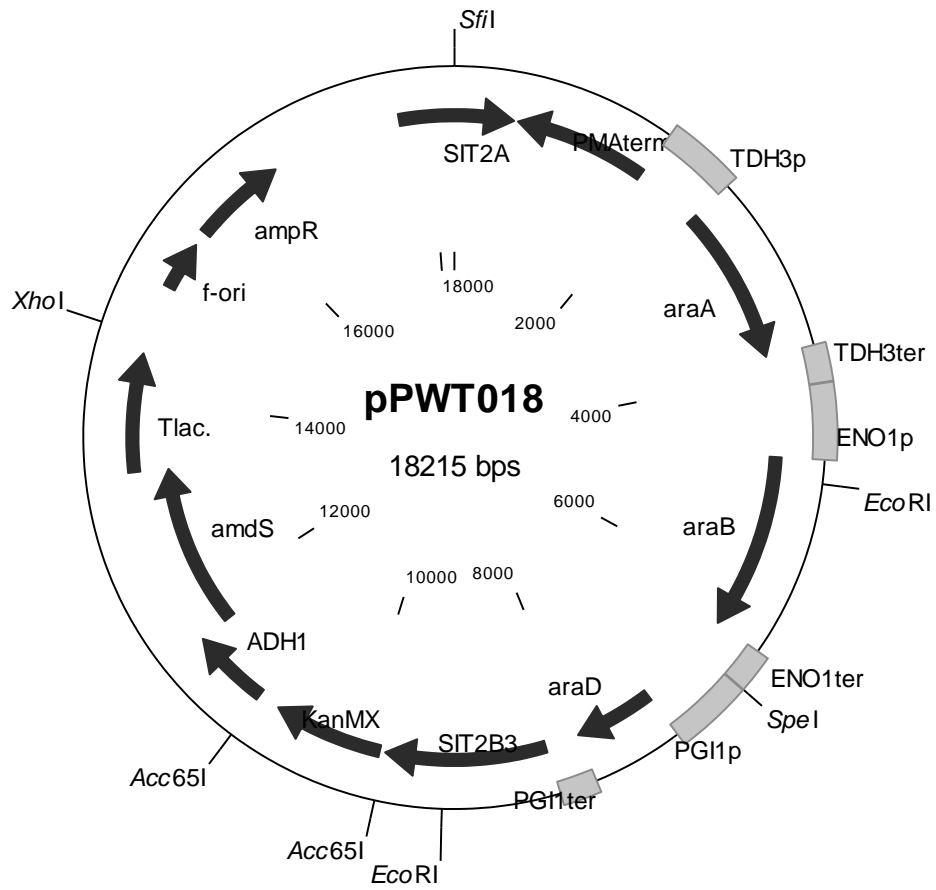


Figura 13

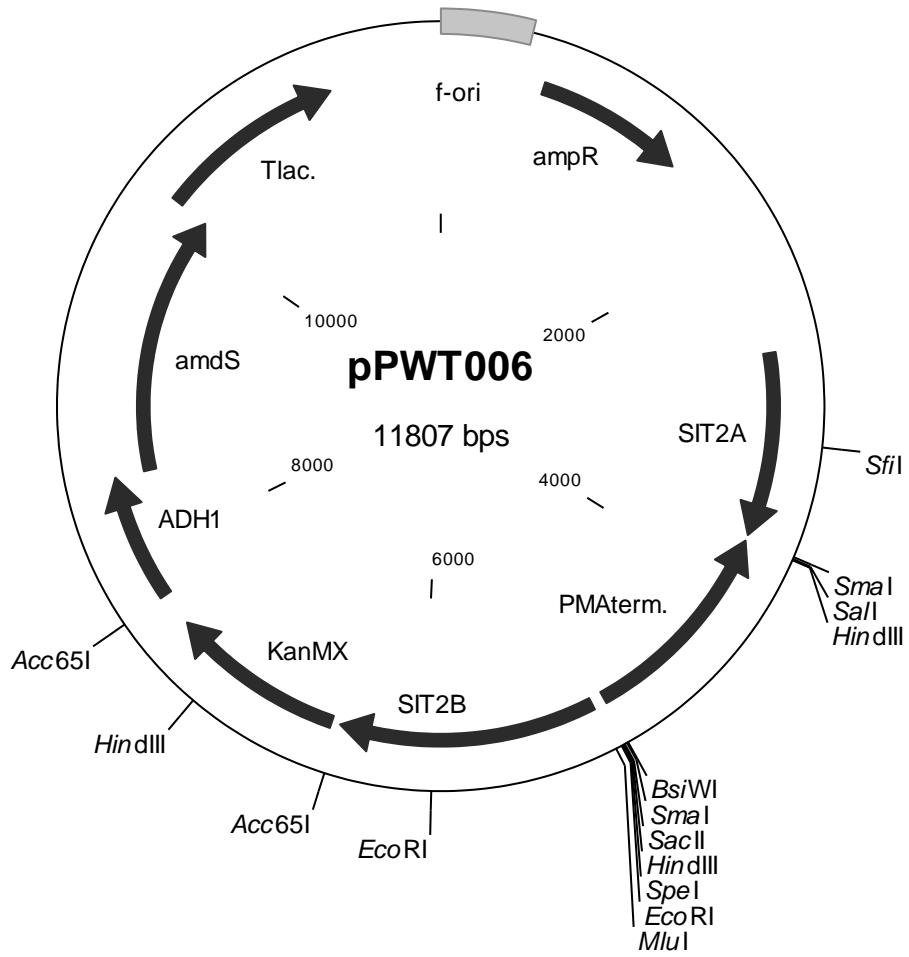


Figura 14

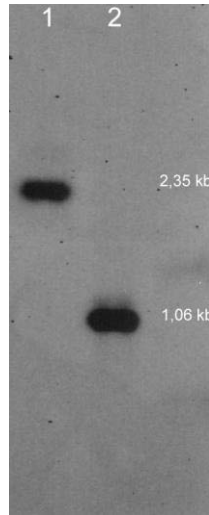
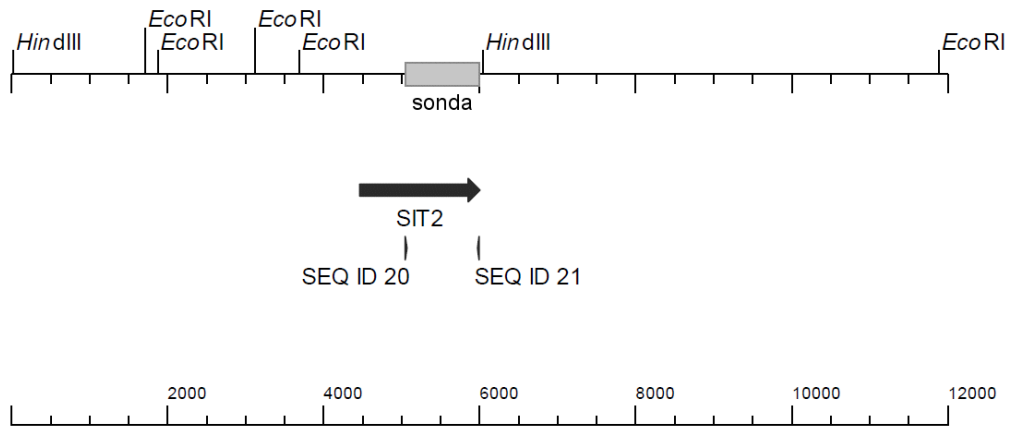
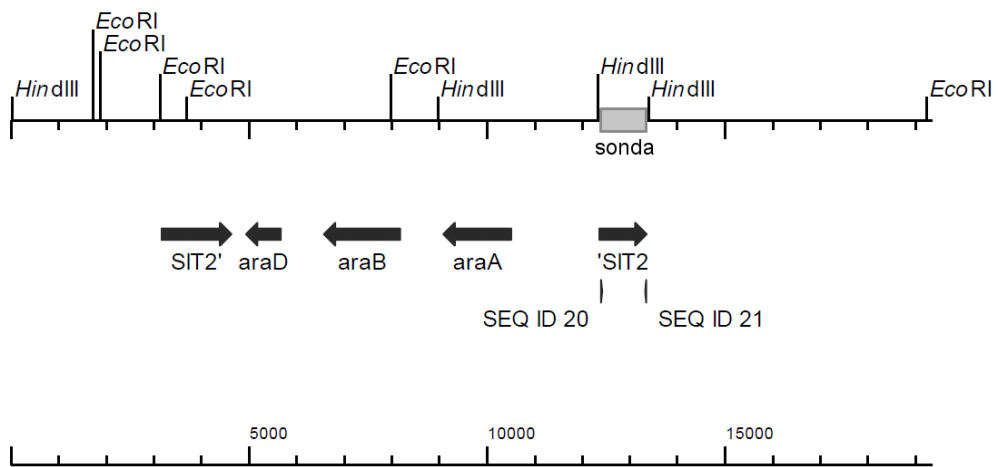


Figura 15

ES 2 551 515 T3

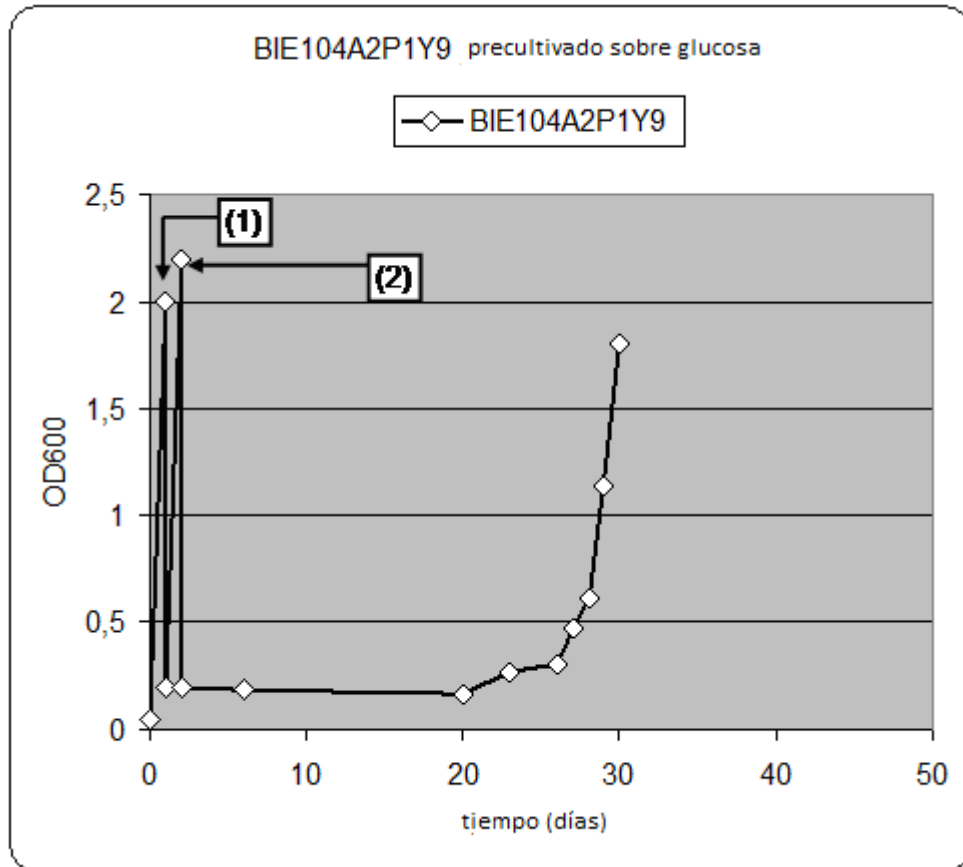


Panel a



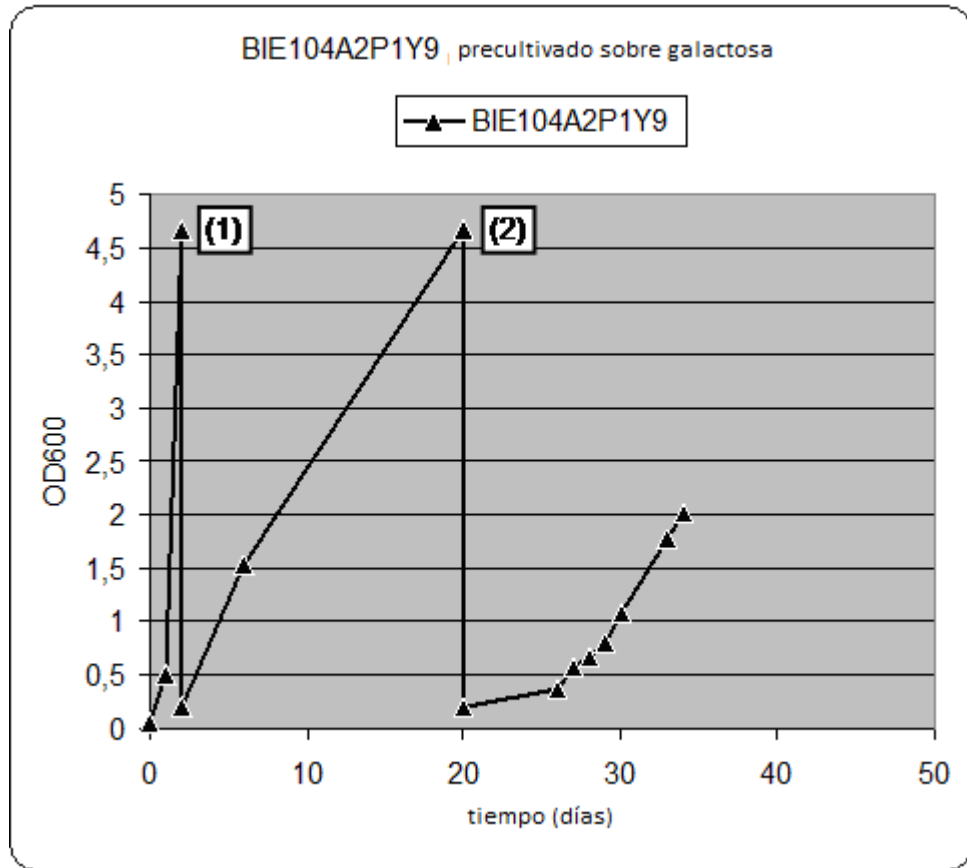
Panel b

Figura 16



Panel a

Figura 17



Panel b

Figura 17