

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 651 076**

51 Int. Cl.:

**C12N 1/36** (2006.01)

**C07K 14/395** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **19.04.2011 PCT/EP2011/056232**

87 Fecha y número de publicación internacional: **27.10.2011 WO11131667**

96 Fecha de presentación y número de la solicitud europea: **19.04.2011 E 11714788 (4)**

97 Fecha y número de publicación de la concesión europea: **18.10.2017 EP 2561064**

54 Título: **Célula adecuada para la fermentación de una composición de azúcares mixtos**

30 Prioridad:

**21.04.2010 US 326351 P**

**21.04.2010 EP 10160622**

**21.04.2010 US 326358 P**

**21.04.2010 EP 10160647**

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

**24.01.2018**

73 Titular/es:

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

**Het Overloon, 1**

**6411 TE Heerlen, NL**

72 Inventor/es:

**KLAASSEN, PAUL;**

**GIELESEN, BIANCA ELISABETH MARIA;**

**SUYLEKOM, VAN, GIJSBERDINA PIETERNELLA y**

**HEIJNE, WILBERT HERMAN MARIE**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**Observaciones:**

**Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes**

ES 2 651 076 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

Célula adecuada para la fermentación de una composición de azúcares mixtos

### Campo de la invención

5 La invención se refiere a una célula (también designada en el presente documento una célula de azúcares mixtos), adecuada para la fermentación de una composición de azúcares que comprende múltiples azúcares C5 y/o C6 (tal composición también se designa en el presente documento composición de azúcares mixtos). La composición de azúcares mixtos puede originarse a partir de material lignocelulósico. La invención también se refiere a un proceso para la producción de producto de fermentación a partir de la composición de azúcares mixtos usando la célula de azúcares mixtos.

### 10 Antecedentes de la invención

La mayor parte del etanol producido como alternativa a los combustibles fósiles es actualmente a partir de la fermentación de sacarosa basada en almidón de maíz y caña de azúcar. Con el fin de alcanzar los ambiciosos objetivos de producir combustibles renovables, están siendo desarrolladas nuevas tecnologías para convertir biomasa no alimentaria en productos de fermentación tales como etanol. *Saccharomyces cerevisiae* es el organismo de elección en la industria del etanol, pero no puede utilizar azúcares de cinco carbonos contenidos en el componente de hemicelulosa de las materias primas de biomasa. La hemicelulosa puede constituir hasta el 20-30 % de la biomasa, siendo la xilosa y arabinosa los azúcares C5 más abundantes. La expresión heteróloga de una xilosa isomerasa (XI) es una opción para permitir que las células de levadura se metabolicen y fermenten xilosa. Asimismo, la expresión de los genes bacterianos *araA*, *araB* y *araD* en cepas de *S. cerevisiae* produce la utilización y eficiente fermentación alcohólica de arabinosa. La galactosa es un azúcar C6 que también es un azúcar que está frecuentemente presente en lignocelulosa, frecuentemente en cantidades (~4 % de azúcares totales) que no deben ser rechazados por motivos económicos.

J. van den Brink et al., Microbiology (2009)155, 1340-1350, desvelan que la glucosa es la fuente de carbono favorecida para *Saccharomyces cerevisiae* y que tras el cambio de las condiciones de fermentación de glucosa limitada a condición de exceso de galactosa en condición anaerobia, no se consumió galactosa.

Hasta la fecha no se ha desvelado proceso para convertir glucosa, arabinosa, xilosa y galactosa en un producto de fermentación en el mismo proceso con glucosa.

### Sumario de la invención

30 Es un objeto de la invención proporcionar una célula capaz de convertir glucosa, xilosa y galactosa y manosa, es decir, una mezcla de cuatro azúcares, en un producto de fermentación. Es otro objeto de la invención tal célula que es capaz de convertir glucosa, xilosa, arabinosa, galactosa y manosa, es decir, una mezcla de cinco azúcares, en un producto de fermentación. Es un objeto adicional de la invención proporcionar tal célula anteriormente definida que es estable. Es un objeto adicional de la invención proporcionar tal cepa anteriormente definida que está libre de marcador. Es un objeto adicional de la invención proporcionar un proceso en el que la glucosa, xilosa, arabinosa y galactosa se convierten simultáneamente en un producto de fermentación. Uno o más de estos objetos se obtienen según la invención.

40 La presente invención proporciona una célula adecuada para la producción de uno o más productos de fermentación a partir de una composición de azúcares que comprende glucosa, galactosa, xilosa, arabinosa y manosa, en la que la célula comprende dos a quince copias de uno o más genes de xilosa isomerasa o dos a quince copias de uno o más genes de xilosa reductasa y xilitol deshidrogenasa, y dos a diez copias de cada uno de *araA*, *araB* y *araD*, en la que estos genes están integrados en el genoma de la célula.

La invención proporciona además un proceso para la producción de uno o más productos de fermentación a partir de una composición de azúcares que comprende glucosa, arabinosa y xilosa y opcionalmente galactosa y/o manosa que comprende las siguientes etapas:

45 a) fermentación de la composición de azúcares en presencia de una célula adecuada para la producción de uno o más productos de fermentación a partir de una composición de azúcares que comprende glucosa, galactosa, xilosa, arabinosa y manosa, en la que la célula comprende dos a quince copias de uno o más genes de xilosa isomerasa o dos a quince copias de uno o más genes de xilosa reductasa y xilitol deshidrogenasa, y dos a diez copias de cada uno de *araA*, *araB* y *araD*, en la que estos genes están integrados en el genoma de la célula; y

b) recuperación del producto de fermentación.

En una realización, la célula de azúcares mixtos es del género *Saccharomyces*. En una realización, la célula de azúcares mixtos es de la especie *Saccharomyces cerevisiae*. En una realización, el producto de fermentación es etanol.

**Breve descripción de los dibujos**

- La Figura 1 expone un mapa físico del plásmido pPWT006.
- La Figura 2 expone un mapa físico del plásmido pPWT018.
- 5 La Figura 3 expone una autorradiograma de transferencia Southern. Se digirió ADN cromosómico de la cepa no mutada CEN.PK113-7D (carril 1) y BIE104A2 (carril 2) con tanto *EcoRI* como *HindIII*. La transferencia se hibridó con una sonda SIT2 específica.
- La Figura 4 expone un mapa físico del plásmido pPWT080, cuya secuencia de da en SEQ ID NO: 8
- 10 La Figura 5 expone una curva de crecimiento en condiciones aerobias de BIE104P1A2 sobre diferentes medios. Se precultivó la cepa BIE104A2P1 sobre YNB con 2 % de galactosa. La curva de crecimiento empezó con 2 % de galactosa y 1 % de arabinosa, seguido del evento indicado en el gráfico por el número (1) transferencia a YNB con 2 % de arabinosa como única fuente de carbono. Después de alcanzar una DO 600 superior a 1, el cultivo se transfirió a medio fresco con una DO 600 inicial de 0,2. Tras tres transferencias sobre medio de arabinosa puro, la cepa resultante se designó BIE104P1A2c.
- 15 La Figura 6 expone una curva de crecimiento en condiciones anaerobias de BIE104P1A2c sobre YNB con 2 % de arabinosa como única fuente de carbono. Después de alcanzar una DO 600 superior a 1, el cultivo se transfirió a medio fresco con una DO 600 inicial de 0,2. Después de varias transferencias, la cepa resultante se llamó BIE104P1A2d (= BIE201).
- La Figura 7 expone un mapa físico del plásmido pPWT042, cuya secuencia de da en SEQ ID NO: 14.
- 20 La Figura 8 expone una curva de crecimiento en condiciones aerobias de una mezcla de transformación. La cepa BIE201X9 se transformó con el fragmento X18-Ty1 y la mezcla de transformación se usó para inocular medio Verduyn complementado con 2 % de xilosa. Después de alcanzar una DO 600 de aproximadamente 35, se usó una alícuota del cultivo para inocular un segundo matraz con medio fresco. A partir del segundo matraz, se aislaron colonias para análisis adicional.
- 25 La Figura 9 expone el crecimiento de cepas aisladas individuales sobre medio Verduyn complementado con ya sea 2 % de glucosa, o 2 % de xilosa, o 2 % de arabinosa o 2 % de galactosa.
- La Figura 10 expone el crecimiento de cepas aisladas individuales seleccionadas sobre medio Verduyn complementado con ya sea 2 % de xilosa o 2 % de arabinosa.
- 30 La Figura 11 expone la conversión de azúcar y formación de producto de un cultivo mixto de cepas BIE104P1Y9mc y BIE201 sobre medio sintético, en un experimento de lotes alimentados. Se midió la producción de CO<sub>2</sub> constantemente. El crecimiento se monitorizó siguiendo la densidad óptica del cultivo. El precultivo se cultivó sobre 2 % de glucosa.
- 35 La Figura 12 expone la conversión de azúcar y formación de producto de la cepa BIE252 sobre medio sintético, en un experimento de lotes alimentados. Se midió la producción de CO<sub>2</sub> constantemente. El crecimiento se monitorizó siguiendo la densidad óptica del cultivo. El precultivo se cultivó sobre 2 % de glucosa.
- La Figura 13 expone el CO<sub>2</sub> total que se produjo en los experimentos de lotes alimentados.
- 40 La Figura 14 expone la conversión de azúcar y formación de producto de un cultivo mixto de cepas BIE104P1Y9mc y BIE201 sobre medio sintético, en un experimento discontinuo. Se midió la producción de CO<sub>2</sub> constantemente. El crecimiento se monitorizó siguiendo la densidad óptica del cultivo. El precultivo se cultivó sobre 2 % de glucosa.
- 45 La Figura 15 expone la conversión de azúcar y formación de producto de un cultivo mixto de la cepa BIE252 sobre medio sintético, en un experimento discontinuo. Se midió la producción de CO<sub>2</sub> constantemente. El crecimiento se monitorizó siguiendo la densidad óptica del cultivo. El precultivo se cultivó sobre 2 % de glucosa.
- La Figura 16 expone el CO<sub>2</sub> total que se produjo en los experimentos discontinuos.
- La Figura 17 expone el rendimiento de la cepa BIE252 en fibra de maíz hidrolizada al 13,8 % de m.s. Se muestran la tasa de producción de CO<sub>2</sub>, producción de etanol y conversión de azúcar.
- La Figura 18 expone el rendimiento de la cepa BIE252 en hojas y tallos de maíz hidrolizados al 10 % de m.s. Se muestran la tasa de producción de CO<sub>2</sub>, producción de etanol y conversión de azúcar.
- 50 La Figura 19 expone el rendimiento de la cepa BIE252 en paja de trigo hidrolizada al 10 % de m.s. Se muestran la tasa de producción de CO<sub>2</sub>, producción de etanol y conversión de azúcar.
- La Figura 20a expone la conversión de azúcar (tasa) de la cepa BIE252.

La Figura 20b expone la conversión de azúcar (tasa) de una cepa aislada de colonias individuales de la cepa BIE252 después de cultivo durante más de 100 generaciones en medio YEP con 2 % de glucosa.

La Figura 20c expone la conversión de azúcar (tasa) de una cepa aislada de colonias individuales de la cepa BIE252 después de cultivo durante más de 100 generaciones en medio YEP con 2 % de glucosa.

5 La Figura 21 expone el mapa físico del plásmido pPWT148.

La Figura 22 expone el mapa físico del plásmido pPWT152.

10 La Figura 23 expone un gel de CHEF, teñido con bromuro de etidio. Los cromosomas se separaron por su tamaño usando la técnica de CHEF. Las cepas analizadas son BIE104 (célula de levadura no transformada), BIE104A2P1a (transformante primario incapaz de consumir arabinosa, sinónimo de BIE104A2P1), BIE104A2P1c, una cepa derivada de BIE104A2P1 por evolución adaptativa, que es capaz de crecer sobre arabinosa, cepa BIE201, derivada de BIE104A2P1c por evolución adaptativa, que puede crecer sobre arabinosa en condiciones anaerobias y cepa BIE252, que es una cepa de azúcares mixtos. Se observa desplazamiento en los cromosomas (véase el texto). La cepa YNN295 es una cepa marcadora (Bio-Rad), usada como referencia para el tamaño de los cromosomas.

15 **Breve descripción del listado de secuencias**

SEQ ID NO: 1 expone la secuencia del plásmido pPWT018

SEQ ID NO: 2 expone la secuencia del cebador para comprobar la integración de pPWT018

SEQ ID NO: 3 expone la secuencia del cebador para comprobar la integración de pPWT018 (con SEQ ID NO: 2) y para comprobar el número de copias de pPWT018 (con SEQ ID NO: 4)

20 SEQ ID NO: 4 expone la secuencia del cebador para comprobar el número de copias de pPWT018

SEQ ID NO: 5 expone la secuencia del cebador para comprobar la presencia de pPWT018 en el genoma en combinación con SEQ ID NO: 4

SEQ ID NO: 6 expone la secuencia del cebador directo para generar la sonda SIT2

SEQ ID NO: 7 expone la secuencia del cebador inverso para generar la sonda SIT2

25 SEQ ID NO: 8 expone la secuencia del plásmido pPWT080

SEQ ID NO: 9 expone la secuencia del cebador directo para comprobar la correcta integración de pPWT080 en el extremo del locus GRE3 (con SEQ ID NO: 10) y para comprobar el número de copias del plásmido pPWT080 (con SEQ ID NO: 11)

30 SEQ ID NO: 10 expone la secuencia del cebador inverso para comprobar la correcta integración de pPWT080 en el extremo del locus GRE3

SEQ ID NO: 11 expone la secuencia del cebador inverso para comprobar el número de copias del plásmido pPWT080

SEQ ID NO: 12 expone la secuencia del cebador directo para generar una sonda RKI1

SEQ ID NO: 13 expone la secuencia del cebador inverso para generar una sonda RKI1

35 SEQ ID NO: 14 expone la secuencia del plásmido pPWT042

SEQ ID NO: 15 expone la secuencia del cebador para comprobar la integración de pPWT042

SEQ ID NO: 16 expone la secuencia del cebador para comprobar la integración de pPWT042

SEQ ID NO: 17 expone la secuencia del cebador para comprobar la integración de pPWT042

SEQ ID NO: 18 expone la secuencia del cebador para comprobar la pérdida de marcador de pPWT042

40 SEQ ID NO: 19 expone la secuencia del cebador para la amplificación del casete con xylA

SEQ ID NO: 20 expone la secuencia del cebador para la amplificación del casete con xylA

SEQ ID NO: 21 expone la secuencia del cebador para la amplificación del casete con kanMX

SEQ ID NO: 22 expone la secuencia del cebador para la amplificación del casete con kanMX

SEQ ID NO: 23 expone la secuencia del cebador para PCR cuantitativa

SEQ ID NO: 24 expone la secuencia del cebador para PCR cuantitativa.

SEQ ID NO: 25 expone la secuencia del cebador para PCR cuantitativa

SEQ ID NO: 26 expone la secuencia del cebador para PCR cuantitativa.

5 **Descripción detallada de la invención**

En toda la presente memoria descriptiva y las reivindicaciones adjuntas, las palabras "comprender" e "incluir", y variaciones tales como "comprende", "que comprende", "incluye" y "que incluye", deben interpretarse de forma incluyente. Es decir, estas palabras pretenden expresar la posible inclusión de otros elementos o números enteros no específicamente citados, donde lo permita el contexto.

10 Los artículos "un" y "una" se usan en el presente documento para referirse a uno o a más de uno (es decir, a uno o al menos uno) del objeto gramatical del artículo. A modo de ejemplo, "un elemento" puede significar un elemento o más de un elemento.

Las diversas realizaciones de la invención descritas en el presente documento pueden ser combinadas de forma cruzada.

15 **La composición de azúcares**

La composición de azúcares según la invención comprende glucosa, arabinosa y xilosa. Puede usarse cualquier composición de azúcares en la invención que cumpla aquellos criterios. Azúcares opcionales en la composición de azúcares son galactosa y manosa. En una realización preferida, la composición de azúcares es un hidrolizado de uno o más materiales lignocelulósicos. Lignocelulosa en el presente documento incluye hemicelulosa y partes de hemicelulosa de biomasa. Lignocelulosa también incluye fracciones lignocelulósicas de biomasa. Materiales lignocelulósicos adecuados pueden encontrarse en la siguiente lista: iniciadores de huerta, chaparral, desechos de molino, desechos de madera urbanos, residuos municipales, residuos de podas, materiales de poda de bosques, cultivos de arbolados de rotación corta, residuos industriales, paja de trigo, paja de avena, paja de arroz, paja de cebada, paja de centeno, paja de lino, cáscaras de soja, cáscaras de arroz, paja de arroz, pienso de gluten de trigo, cáscaras de avena, caña de azúcar, tallos y hojas de maíz, cañas de maíz, mazorcas de maíz, vainas de maíz, pasto varilla, miscanthus, sorgo dulce, tallos de colza, tallos de soja, pasto de pradera, pasto gama, cola de zorro; pulpa de remolacha dulce, pulpa de frutas cítricas, cáscaras de semillas, residuos animales celulósicos, recortes de césped, algodón, algas marinas, árboles, madera blanda, madera dura, álamo, pino, arbustos, hierbas, trigo, paja de trigo, bagazo de caña de azúcar, maíz, vainas de maíz, mazorcas de maíz, grano de maíz, fibra de granos, productos y subproductos de la molienda en seco o húmedo de granos, residuos sólidos municipales, papel residual, residuos de jardinería, material herbáceo, residuos agrícolas, residuos forestales, residuos sólidos municipales, papel residual, pulpa, residuos de molienda de papel, ramas, arbustos, cañas, maíz, vainas de maíz, un cultivo energético, bosque, una fruta, una flor, un grano, una hierba, un cultivo herbáceo, una hoja, corteza, una espina, un tronco, una raíz, un árbol joven, un matorral, pasto varilla, un árbol, una verdura, piel de fruta, una vid, pulpa de remolacha azucarera, cascarillas de trigo, cáscaras de avena, madera dura o blanda, material de residuos orgánicos generado de un proceso agrícola, residuos de madera de silvicultura, o una combinación de cualesquiera dos o más de los mismos.

Una visión general de algunas composiciones de azúcares adecuadas derivadas de lignocelulosa y la composición de azúcares de sus hidrolizados se da en la Tabla 1. Las lignocelulosas enumeradas incluyen: mazorcas de maíz, fibra de maíz, cáscaras de arroz, cáscaras de melón, pulpa de remolacha azucarera, paja de trigo, bagazo de caña de azúcar, madera, pasto y prensados de aceituna.

Tabla 1: Visión general de las composiciones de azúcares de materiales lignocelulósicos. Gal=galactosa, Xil=xilosa, Ara=arabinosa, Man=manosa, Glu=glutamato, Ram=ramnosa. Se da el porcentaje de galactosa (% de Gal) y la fuente bibliográfica.

Material lignocelulósico	Gal	Xil	Ara	Man	Glu	Ram	Suma	% de Gal.	Bibl.
Mazorca de maíz a	10	286	36		227	11	570	1,7	(1)
Mazorca de maíz b	131	228	160		144		663	19,8	(1)
Cáscaras de arroz a	9	122	24	18	234	10	417	2,2	(1)
Cáscaras de arroz b	8	120	28		209	12	378	2,2	(1)
Cáscaras de melón	6	120	11		208	16	361	1,7	(1)

Material lignocelulósico	Gal	Xil	Ara	Man	Glu	Ram	Suma	% de Gal.	Bibl.
Pulpa de remolacha azucarera	51	17	209	11	211	24	523	9,8	(2)
Paja de trigo Idaho	15	249	36		396		696	2,2	(3)
Fibra de maíz	36	176	113		372		697	5,2	(4)
Bagazo de caña de azúcar	14	180	24	5	391		614	2,3	(5)
Hojas y tallos de maíz	19	209	29		370		626		(6)
Athel (madera)	5	118	7	3	493		625	0,7	(7)
Eucalipto (madera)	22	105	8	3	445		583	3,8	(7)
CWR (pasto)	8	165	33		340		546	1,4	(7)
JTW (pasto)	7	169	28		311		515	1,3	(7)
MSW	4	24	5	20	440		493	0,9	(7)
Vegetación de pasto alpiste	16	117	30	6	209	1	379	4,2	(8)
Semilla de pasto alpiste	13	163	28	6	265	1	476	2,7	(9)
Residuos del prensado de aceitunas	15	111	24	8	329		487	3,1	(9)

Es evidente de la Tabla 1 que en estas lignocelulosas una alta cantidad de azúcar está presente en forma de glucosa, xilosa, arabinosa y galactosa. La conversión de glucosa, xilosa, arabinosa y galactosa en el producto de fermentación es así de gran importancia económica. También está presente manosa en algunos materiales de lignocelulosa, aunque normalmente en cantidades más bajas que los azúcares previamente mencionados. Ventajosamente, por tanto, también se convierte manosa por la célula de azúcares mixtos.

La célula de azúcares mixtos

La célula de azúcares mixtos comprende dos a quince copias de uno o más genes de xilosa isomerasa o dos a quince copias de uno o más genes de xilosa reductasa y xilitol deshidrogenasa, y dos a diez copias de *araA*, *araB* y *araD*, en la que estos genes están integrados en el genoma de la célula.

En una realización, la célula de azúcares mixtos comprende aproximadamente ocho copias u ocho copias de xilosa isomerasa, en la que estos genes están integrados en el genoma de la célula de azúcares mixtos. En otra realización, la célula de azúcares mixtos comprende ocho copias de una xilosa reductasa y xilitol deshidrogenasa, en la que estos genes están integrados en el genoma de la célula de azúcares mixtos.

El número de copias puede ser determinado por el experto por cualquier método conocido. En los ejemplos se describe un método adecuado.

La célula de azúcares mixtos comprende dos a diez copias de genes *araA*, *araB* y *araD*, en la que estos genes están integrados en el genoma de la célula. Es capaz de fermentar glucosa, arabinosa, xilosa y galactosa.

En una realización, la célula de azúcares mixtos comprende dos copias, tres copias, cuatro copias cinco copias, seis copias, siete copias, ocho copias, nueve copias o diez copias, dos a diez, dos a nueve, dos a ocho, dos a siete, dos a seis, dos a cinco, dos a cuatro, tres a diez, tres a nueve, tres a ocho, tres a siete, tres a seis, tres a cinco, cuatro a diez, cuatro a nueve, cuatro a ocho, cuatro a siete o cuatro a seis copias de cada uno de los genes *araA*, *araB* y *araD*, en la que estos genes están integrados en el genoma de la célula. En una realización, la célula de azúcares mixtos comprende aproximadamente cuatro, o cuatro copias de cada uno de los genes *araA*, *araB* y *araD*, en la que estos genes están integrados en el genoma de la célula.

En una realización, el número de copias de gen de xilosa isomerasa o xilosa reductasa y xilitol deshidrogenasa en la célula de azúcares mixtos es ocho o nueve. En una realización, el número de copias de los genes *araA*, *araB* y *araD* en la célula de azúcares mixtos es 3 o 4.

En una realización, la célula es capaz de convertir el 90 % o más de glucosa, xilosa arabinosa, galactosa y manosa disponible en un producto de fermentación. En una realización, la célula es capaz de convertir el 91 % o más, el 92 % o más, el 94 % o más, el 95 % o más, el 96 % o más, el 97 % o más, el 98 % o más o el 100 % de toda la glucosa, xilosa arabinosa, galactosa y manosa disponible en un producto de fermentación.

En una realización, la célula tiene una interrupción o delección del gen GAL80.

En una realización de la invención, la célula de azúcares mixtos es capaz de fermentar uno o más azúcares adicionales, preferentemente azúcar C5 y/o C6, por ejemplo, manosa. En una realización de la invención, la célula de azúcares mixtos comprende uno o más de: un gen *xyIA*, gen *XYL1* y gen *XYL2* y/o gen *XKS1*, para permitir que la célula de azúcares mixtos fermente xilosa; delección del gen de aldosa reductasa (*GRE3*); expresión en exceso de los genes de PPP *TAL1*, *TKL1*, *RPE1* y *RK11* para permitir el aumento del flujo a través de la vía de la pentosa fosfato en la célula.

En una realización, la célula de azúcares mixtos es una célula industrial, más preferentemente una levadura industrial. Una célula industrial y célula de levadura industrial puede definirse del siguiente modo. Los entornos vivos de células (de levadura) en procesos industriales son significativamente diferentes de aquellos en el laboratorio. Células de levadura industriales deben ser capaces de rendir bien en múltiples condiciones medioambientales que pueden variar durante el proceso. Tales variaciones incluyen cambio en las fuentes de nutriente, pH, concentración de etanol, temperatura, concentración de oxígeno, etc., que juntos tienen el potencial de tener un impacto sobre el crecimiento celular y la producción de etanol de *Saccharomyces cerevisiae*. En condiciones industriales adversas, las cepas tolerantes medioambientales deben permitir el crecimiento y la producción robustos. Las cepas de levadura industriales son generalmente más robustas hacia estos cambios en condiciones medioambientales que pueden producirse en las aplicaciones en las que se usan, tales como en la industria de la panadería, industria cervecera, viticultura y la industria del etanol. En una realización, la célula de azúcares mixtos industrial se construye basándose en una célula hospedadora industrial, en la que la construcción se realiza como se describe en lo sucesivo. Ejemplos de levadura industrial (*S. cerevisiae*) son Ethanol Red® (Fermentis), Fermiol® (DSM) y Thermosacc® (Lallemand).

En una realización, la célula de azúcares mixtos es tolerante a inhibidor. La tolerancia a inhibidor es resistencia a inhibir compuestos. La presencia y nivel de compuestos inhibidores en lignocelulosa puede variar ampliamente con la variación de materia prima, proceso de hidrólisis del método de pretratamiento. Ejemplos de categorías de inhibidores son ácidos carboxílicos, furanos y/o compuestos fenólicos. Ejemplos de ácidos carboxílicos son ácido láctico, ácido acético o ácido fórmico. Ejemplos de furano son furfural e hidroximetilfurfural. Ejemplos de compuestos fenólicos son vainillina, ácido siríngico, ácido ferúlico y ácido cumárico. Las cantidades típicas de inhibidores son para ácidos carboxílicos: varios gramos por litro, hasta 20 gramos por litro o más, dependiendo de la materia prima, el pretratamiento y las condiciones de hidrólisis. Para furanos: varios cientos de miligramos por litro hasta varios gramos por litro, dependiendo de la materia prima, el pretratamiento y las condiciones de hidrólisis. Para fenólicos: varias decenas de miligramos por litro, hasta un gramo por litro, dependiendo de la materia prima, el pretratamiento y las condiciones de hidrólisis.

Las cepas de azúcares mixtos según la invención son tolerantes a inhibidor, es decir, pueden resistir a inhibidores comunes al nivel que normalmente tienen con condiciones de pretratamiento e hidrólisis comunes, de manera que las cepas de azúcares mixtos pueden encontrar una amplia aplicación, es decir, tiene alta aplicabilidad para diferente materia prima, diferentes métodos de pretratamiento y diferentes condiciones de hidrólisis.

En una realización, la célula de azúcares mixtos industrial se construye basándose en una célula hospedadora tolerante a inhibidor, en la que la construcción se realiza como se describe en lo sucesivo. Pueden seleccionarse células hospedadoras tolerantes a inhibidor seleccionando cepas para el crecimiento sobre inhibidores que contienen materiales, tal como se ilustra en Kadar et al., Appl. Biochem. Biotechnol. (2007), Vol. 136-140, 847-858, en el que se seleccionó una cepa de *S. cerevisiae* tolerante a inhibidor ATCC 26602.

En una realización, la célula de azúcares mixtos está libre de marcador. Como se usa en el presente documento, el término "marcador" se refiere a un gen que codifica un rasgo o un fenotipo que permite la selección de, o el cribado de, una célula hospedadora que contiene el marcador. Libre de marcador significa que los marcadores están esencialmente ausentes en la célula de azúcares mixtos. El estar libre de marcador es particularmente ventajoso cuando se han usado marcadores de antibiótico en la construcción de la célula de azúcares mixtos y se eliminan a partir de aquí. La eliminación de marcadores puede hacerse usando cualquier técnica del estado de la técnica adecuada, por ejemplo, recombinación intramolecular. Un método adecuado de la eliminación de marcadores se ilustra en los ejemplos.

Una célula de azúcares mixtos puede ser capaz de convertir biomasa de planta, celulosas, hemicelulosas, pectinas, ramnosa, galactosa, fructosa, maltosa, maltodextrinas, ribosa, ribulosa, o almidón, derivados de almidón, sacarosa, lactosa y glicerol, por ejemplo en azúcares fermentables. Por consiguiente, una célula de azúcares mixtos puede expresar una o más enzimas tales como una celulasa (una endocelulasa o una exocelulasa), una hemicelulasa (una endo- o exo-xilanasa o arabinasa) necesaria para la conversión de celulosa en monómeros de glucosa y hemicelulosa en xilosa y monómeros de arabinosa, una pectinasa capaz de convertir pectinas en ácido glucurónico y ácido galacturónico o una amilasa para convertir almidón en monómeros de glucosa.

La célula de azúcares mixtos puede comprender además aquellas actividades enzimáticas requeridas para la conversión de piruvato en un producto de fermentación deseado, tal como etanol, butanol, ácido láctico, ácido 3-

hidroxipropiónico, ácido acrílico, ácido acético, ácido succínico, ácido cítrico, ácido fumárico, ácido málico, ácido itacónico, un aminoácido, 1,3-propanodiol, etileno, glicerol, un antibiótico  $\beta$ -lactámico o una cefalosporina.

5 En una realización, la célula de azúcares mixtos es una célula que es naturalmente capaz de fermentación alcohólica, preferentemente, fermentación alcohólica anaerobia. Una célula de azúcares mixtos tiene preferentemente una alta tolerancia a etanol, una alta tolerancia a pH bajo (es decir, capaz de crecimiento a un pH inferior a aproximadamente 5, aproximadamente 4, aproximadamente 3, o aproximadamente 2,5) y hacia productos orgánicos y/o una alta tolerancia a temperaturas elevadas.

Cualquiera de las características o actividades anteriores de una célula de azúcares mixtos puede estar naturalmente presente en la célula o puede introducirse o modificarse por modificación genética.

#### 10 Construcción de la cepa de azúcares mixtos

Según una realización, los genes pueden introducirse en la célula de azúcares mixtos introduciendo en una célula hospedadora:

a) una agrupación que consiste en los genes *araA*, *araB* y *araD* bajo el control de un promotor constitutivo fuerte

15 b) una agrupación que consiste en los genes de PPP *TAL1*, *TKL1*, *RPE1* y *RK11*, opcionalmente bajo el control de promotor constitutivo fuerte; y delección de un gen de aldosa reductasa;

c) una agrupación que consiste en un gen *xyIA* y un gen *XKS1* bajo el control de promotor constitutivo fuerte;

d) una construcción que comprende un gen *xyIA* bajo el control de un promotor constitutivo fuerte, que tiene la capacidad de integrarse en el genoma en múltiples loci;

20 y evolución adaptativa para producir la célula de azúcares mixtos. La célula anterior puede construirse usando técnicas de expresión recombinante.

#### Expresión recombinante

25 La célula de azúcares mixtos es una célula recombinante. Es decir, una célula de azúcares mixtos comprende, o se transforma con o se modifica genéticamente con una secuencia de nucleótidos que no se produce en la célula en cuestión.

Técnicas para la expresión recombinante de enzimas en una célula, además de para las modificaciones genéticas adicionales de una célula de azúcares mixtos, son muy conocidas para aquellos expertos en la materia. Normalmente, tales técnicas implican la transformación de una célula con construcción de ácidos nucleicos que comprende la secuencia relevante. Tales métodos se conocen, por ejemplo, de manuales estándar, tales como 30 Sambrook y Russel (2001) "Molecular Cloning: A Laboratory Manual (3ª edición), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, o F. Ausubel et al., eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, New York (1987). Métodos para la transformación y modificación genética de células hospedadoras fúngicas se conocen de, por ejemplo, los documentos EP-A-0635 574, WO 98/46772, WO 99/60102, WO 00/37671, WO90/14423, EP-A-0481008, EP-A-0635574 y US 6.265.186.

35 Normalmente, la construcción de ácidos nucleicos puede ser un plásmido, por ejemplo, un plásmido de copia baja o un plásmido de copia alta. La célula según la presente invención puede comprender una única copia o múltiples copias de la secuencia de nucleótidos que codifica una enzima, por ejemplo, por múltiples copias de una construcción de nucleótidos o por el uso de la construcción que tiene múltiples copias de la secuencia de enzima.

40 La construcción de ácidos nucleicos puede mantenerse episómicamente y así comprender una secuencia para la replicación autónoma, tal como una secuencia de secuencia de replicación autosómica. Una construcción episómica adecuada de ácidos nucleicos puede basarse, por ejemplo, en los plásmidos 2 $\mu$  o pKD1 de levadura (Gleer et al., 1991, Biotechnology 9:968-975), o los plásmidos AMA (Fierro et al., 1995, Curr Genet. 29:482-489). Alternativamente, cada construcción de ácidos nucleicos puede integrarse en una o más copias en el genoma de la célula. La integración en el genoma de la célula puede producirse al azar por recombinación no homóloga, pero 45 preferentemente, la construcción de ácidos nucleicos puede integrarse en el genoma de la célula por recombinación homóloga como es muy conocido en la técnica (véanse, por ejemplo, los documentos WO90/14423, EP-A-0481008, EP-A-0635 574 y US 6.265.186).

50 La mayoría de los plásmidos episómicos o 2 $\mu$  son relativamente inestables, perdiéndose en aproximadamente 10<sup>-2</sup> o más células después de cada generación. Incluso en condiciones de crecimiento selectivo, solo del 60 % al 95 % de las células retienen el plásmido episómico. El número de copias de la mayoría de los plásmidos episómicos oscila de 20-100 por célula de hospedadores cir<sup>+</sup>. Sin embargo, los plásmidos no están igualmente distribuidos entre las células, y hay una alta varianza en el número de copias por célula en las poblaciones. Cepas transformadas con plásmidos integrativos son extremadamente estables, incluso en ausencia de presión selectiva. Sin embargo, puede producirse pérdida de plásmidos en aproximadamente 10<sup>-3</sup> a 10<sup>-4</sup> frecuencias por recombinación homóloga entre

ADN repetido en tándem, conduciendo a pérdida por formación de bucle de la secuencia del vector. Preferentemente, el diseño del vector en el caso de integración estable es tal que, ante la pérdida de los genes de marcadores de selección (que también ocurre mediante recombinación homóloga intramolecular), esa pérdida por formación de bucle de la construcción integrada ya no es posible. Preferentemente, los genes están, por lo tanto, integrados de forma estable. La integración estable se define en el presente documento como la integración dentro del genoma, en el que ya no es posible la pérdida por formación de bucle de la construcción integrada. Preferentemente, los marcadores de selección están ausentes. Normalmente, la secuencia codificante de la enzima estará operativamente unida a una o más secuencias de ácidos nucleicos, capaces de proporcionar o ayudar en la transcripción y/o traducción de la secuencia de la enzima.

El término "operativamente unido" se refiere a una yuxtaposición en la que los componentes descritos están en una relación que les permite funcionar en la manera esperada. Por ejemplo, un promotor o potenciador está operativamente unido a una secuencia codificante de modo que dicho promotor o potenciador afecte la transcripción de la secuencia codificante.

Como se usa en el presente documento, el término "promotor" se refiere a un fragmento de ácido nucleico que funciona para controlar la transcripción de uno o más genes, localizado en la dirección 5' con respecto a la dirección de transcripción del sitio de iniciación de la transcripción del gen, y está estructuralmente identificado por la presencia de un sitio de unión para la ARN polimerasa dependiente de ADN, sitios de iniciación de la transcripción y cualquier otra secuencia de ADN conocida por un experto en la materia. Un promotor "constitutivo" es un promotor que está activo bajo la mayoría de las condiciones medioambientales y de desarrollo. Un promotor "inducible" es un promotor que está activo bajo regulación medioambiental o del desarrollo.

El promotor que podría usarse para lograr la expresión de una secuencia de nucleótidos que codifica una enzima según la presente invención puede no ser nativo a la secuencia de nucleótidos que codifica la enzima que va a expresarse, es decir, un promotor que es heterólogo a la secuencia de nucleótidos (secuencia codificante) a la que está operativamente unido. Sin embargo, el promotor puede ser homólogo, es decir, endógeno, a la célula hospedadora.

Los promotores están ampliamente disponibles y son conocidos para el experto. Ejemplos adecuados de tales promotores incluyen, por ejemplo, promotores de genes glucolíticos, tales como los promotores de fosfofructocinasa (PFK), triosa fosfato isomerasa (TPI), gliceraldehído-3-fosfato deshidrogenasa (GPD, TDH3 o GAPDH), piruvato cinasa (PYK), fosfoglicerato cinasa (PGK) de levaduras u hongos filamentosos; más detalles sobre tales promotores de levadura pueden encontrarse en (documento WO 93/03159). Otros promotores útiles son los promotores de genes que codifican proteínas ribosómicas, el promotor del gen de lactasa (LAC4), promotores de alcohol deshidrogenasa (ADH1, ADH4, y similares), y el promotor de enolasa (ENO). Otros promotores, tanto constitutivos como inducibles, y potenciadores o secuencias de activación en la dirección 5' serán conocidos por aquellos expertos en la materia. Los promotores usados en las células hospedadoras de la invención pueden modificarse, si se desea, para afectar sus características de control. Promotores adecuados en este contexto incluyen tanto promotores naturales constitutivos como inducibles, además de promotores modificados por ingeniería, que son muy conocidos por el experto en la materia. Promotores adecuados para células hospedadoras 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*.

En una célula de azúcares mixtos, el extremo 3' de la secuencia de ácidos nucleicos que codifica la enzima está preferentemente unido operativamente a una secuencia terminadora de la transcripción. Preferentemente, la secuencia terminadora es operable en una célula hospedadora de elección, tal como, por ejemplo, las especies de levadura de elección. En cualquier caso, la elección del terminador no es crítica; puede ser, por ejemplo, de cualquier gen de levadura, aunque los terminadores pueden funcionar algunas veces si son de un gen eucariota distinto de levadura. Normalmente, una secuencia de nucleótidos que codifica la enzima comprende un terminador. Preferentemente, tales terminadores se combinan con mutaciones que previenen el decaimiento del ARNm mediado por elementos sin sentido en la célula hospedadora de azúcares mixtos (véase, por ejemplo: Shirley et al., 2002, Genetics 161:1465-1482).

La secuencia de terminación de la transcripción comprende preferentemente además una señal de poliadenilación.

Opcionalmente, puede estar presente un marcador de selección en una construcción de ácido nucleico adecuada para su uso en la invención. Como se usa en el presente documento, el término "marcador" se refiere a un gen que codifica un rasgo o un fenotipo que permite la selección de, o el cribado de, una célula hospedadora que contiene el marcador. El gen marcador puede ser un gen de resistencia a antibióticos por el cual puede usarse el antibiótico apropiado para seleccionar células transformadas de entre células que no están transformadas. Ejemplos de marcadores de resistencia a antibiótico adecuados incluyen, por ejemplo, dihidrofolato reductasa, higromicina-B-fosfotransferasa, 3'-O-fosfotransferasa II (resistencia a kanamicina, neomicina y G418). Los marcadores de resistencia a antibiótico pueden ser más convenientes para la transformación de células hospedadoras poliploides. También pueden usarse marcadores no de resistencia a antibiótico, tales como marcadores auxotróficos (*URA3*, *TRP1*, *LEU2*) o el gen *TPI* de *S. pombe* (descrito por Russell P R, 1985, Gene 40: 125-130). En una realización preferida, las células hospedadoras transformadas con las construcciones de ácido nucleico están libres de genes

5 marcadores. Métodos de construcción de células hospedadoras microbianas recombinantes libres de genes marcadores se desvelan en el documento EP-A-0 635 574 y se basan en el uso de marcadores bidireccionales tales como el gen de *A. nidulans* amdS (acetamidasa) o los genes de levadura *URA3* y *LYS2*. Alternativamente, puede incorporarse un marcador de selección tal como proteína verde fluorescente, lacZ, luciferasa, cloranfenicol acetiltransferasa, beta-glucuronidasa en las construcciones de ácido nucleico de la invención, que permite cribar las células transformadas.

10 Elementos adicionales opcionales que pueden estar presentes en las construcciones de ácido nucleico adecuadas para su uso en la invención incluyen una o más secuencias conductoras, potenciadores, factores de integración, y/o genes indicadores, secuencias intrónicas, centrómeros, telómeros y/o secuencias de unión a matriz (MAR). Las construcciones de ácido nucleico de la invención pueden comprender además una secuencia para replicación autónoma, tal como una secuencia ARS.

15 El proceso de combinación puede así ejecutarse con técnicas de recombinación conocidas. Diversos medios son conocidos por aquellos expertos en la materia para la expresión y expresión en exceso de enzimas en una célula de azúcares mixtos. En particular, una enzima puede expresarse en exceso aumentando el número de copias del gen que codifica la enzima en la célula hospedadora, por ejemplo, mediante la integración de copias adicionales del gen en el genoma de la célula hospedadora, mediante la expresión del gen desde un vector de expresión multicopia episómico o mediante la introducción de un vector de expresión episómico que comprende múltiples copias del gen.

20 Alternativamente, la expresión en exceso de enzimas en las células hospedadoras de la invención puede lograrse usando un promotor que no es nativo a la secuencia que codifica la enzima que va a expresarse en exceso, es decir, un promotor que es heterólogo a la secuencia codificante a la que está operativamente unido. Aunque el promotor es preferentemente heterólogo a la secuencia codificante a la que está operativamente unido, también se prefiere que el promotor sea homólogo, es decir, endógeno a la célula hospedadora. Preferentemente, el promotor heterólogo es capaz de producir un mayor nivel de estado estacionario del transcrito que comprende la secuencia codificante (o es capaz de producir más moléculas de transcrito, es decir, moléculas de ARNm, por unidad de tiempo) que el promotor que es nativo a la secuencia codificante. Promotores adecuados en este contexto incluyen tanto promotores naturales constitutivos como inducibles, además de promotores modificados por ingeniería.

25 En una realización, la célula de azúcares mixtos está libre de marcadores, que significa que no están presentes marcadores auxotróficos o dominantes, en particular marcadores de resistencia a antibióticos, en el genoma o extracromosómicamente.

30 La secuencia codificante usada para la expresión en exceso de las enzimas mencionadas anteriormente puede preferentemente ser homóloga a la célula hospedadora. Sin embargo, pueden usarse secuencias codificantes que son heterólogas al hospedador.

35 La expresión en exceso de una enzima, cuando se refiere a la producción de la enzima en una célula genéticamente modificada, significa que la enzima se produce a un nivel mayor de actividad enzimática específica en comparación con la célula hospedadora sin modificar en condiciones idénticas. Normalmente, esto significa que la proteína enzimáticamente activa (o proteínas en el caso de enzimas con múltiples subunidades) se produce en cantidades mayores, o bien a un nivel de estado estacionario mayor en comparación con la célula hospedadora sin modificar, en condiciones idénticas. Similarmente, esto normalmente significa que el ARNm que codifica la proteína enzimáticamente activa se produce en cantidades mayores, o nuevamente a un nivel de estado estacionario mayor, en comparación con la célula hospedadora sin modificar en condiciones idénticas. Preferentemente, en un hospedador, una enzima que va a expresarse en exceso se expresa en exceso 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 la modificación genética que causa la expresión en exceso. Debe entenderse que estos niveles de expresión en exceso pueden aplicarse al nivel de estado estacionario de la actividad de la enzima, al nivel de estado estacionario de la proteína de la enzima, además de al nivel de estado estacionario del transcrito que codifica la enzima.

#### Adaptación

50 La adaptación es el proceso evolutivo por el cual una población se vuelve más adecuada (adaptada) para su hábitat o hábitats. Este proceso tiene lugar durante de varias a muchas generaciones, y es uno de los fenómenos básicos de la biología.

El término adaptación también se puede referir a una característica que es especialmente importante para la supervivencia de un organismo. Tales adaptaciones se producen en una población variable por las formas mejor adecuadas que se reproducen más satisfactoriamente, mediante selección natural.

55 Los cambios en las condiciones medioambientales alteran el resultado de la selección natural, que afecta los beneficios selectivos de adaptaciones posteriores que mejoran la aptitud de un organismo bajo las nuevas condiciones. En el caso de un cambio medioambiental extremo, la aparición y fijación de adaptaciones beneficiosas puede ser esencial para la supervivencia. Un gran número de factores diferentes, tales como, por ejemplo,

disponibilidad de nutrientes, temperatura, la disponibilidad de oxígeno, etcétera, pueden impulsar la evolución adaptativa.

#### Aptitud

5 Hay una clara relación entre la adaptabilidad (el grado al cual un organismo es capaz de vivir y reproducirse en un grupo dado de hábitats) y la aptitud. La aptitud es una estimación y un factor pronóstico de la tasa de selección natural. Mediante la aplicación de la selección natural, las frecuencias relativas de fenotipos alternativos variarán con el tiempo, si son heredables.

#### Cambios genéticos

10 Cuando la selección natural actúa sobre la variabilidad genética de la población, los cambios genéticos son el mecanismo subyacente. Por este medio, la población se adapta genéticamente a sus circunstancias. Los cambios genéticos pueden producir estructuras visibles, o pueden ajustar la actividad fisiológica del organismo en un modo que se adecue al hábitat cambiado.

15 Puede ocurrir que los hábitats cambien frecuentemente. Por tanto, de esto resulta que el proceso de adaptación nunca está finalmente completo. Con el tiempo, puede ocurrir que el entorno cambie gradualmente, y la especie llegue a adaptarse a su entorno cada vez mejor. Por otra parte, puede ocurrir que los cambios en el entorno ocurran con relativa rapidez, y entonces la especie se vuelva cada menos bien adaptada. La adaptación es un proceso genético, que está ocurriendo todo el tiempo de algún modo, también cuando la población no cambia el hábitat o entorno.

#### La evolución adaptativa

20 Las células de azúcares mixtos pueden someterse en su preparación a evolución adaptativa. Una célula de azúcares mixtos puede adaptarse a la utilización de azúcares mediante selección de mutantes, ya sean espontáneos o inducidos (por ejemplo, mediante radiación o productos químicos), para crecer sobre el azúcar deseado, preferentemente como única fuente de carbono, y más preferentemente en condiciones anaerobias. Puede realizarse selección de mutantes por técnicas que incluyen transferencia en serie de cultivos como, por ejemplo, se describe por Kuyper et al. (2004, FEMS Yeast Res. 4: 655-664) o por cultivo bajo presión selectiva en un cultivo quimiostático. Por ejemplo, en una célula hospedadora preferida, al menos una de las modificaciones genéticas descritas previamente, que incluye las modificaciones obtenidas por selección de mutantes, confiere a la célula hospedadora la capacidad de crecer sobre xilosa como fuente de carbono, preferentemente como única fuente de carbono, y preferentemente en condiciones anaerobias. Cuando se usa XI como gen para convertir xilosa, preferentemente la célula esencialmente no produce xilitol, por ejemplo, el xilitol producido está por debajo del límite de detección o, por ejemplo, es inferior a aproximadamente el 5, aproximadamente el 2, aproximadamente el 1, aproximadamente el 0,5, o aproximadamente el 0,3 % del carbono consumido en base molar.

La evolución adaptativa también se describe, por ejemplo, en Wisselink H.W. et al., Applied and Environmental Microbiology Aug. 2007, p. 4881-4891

35 En una realización de la evolución adaptativa se aplica un régimen que consiste en cultivo discontinuo repetido con ciclos repetidos de crecimiento consecutivo en diferentes medios, por ejemplo, tres medios con diferentes composiciones (glucosa, xilosa y arabinosa; xilosa y arabinosa. Véase Wisselink et al. (2009) Applied and Environmental Microbiology, Feb. 2009, p. 907-914.

#### Transformación de levaduras y estabilidad genética

40 La ingeniería genética, es decir, la transformación de células de levadura con ADN recombinante, fue posible por primera vez en 1978 [Beggs, 1978; Hinnen et al., 1978]. La tecnología de ADN recombinante en levadura se ha establecido desde entonces. Están disponibles una multitud de diferentes construcciones de vector. Generalmente, estos vectores plasmídicos, denominados vectores lanzadera, contienen material genético derivado de vectores de *E. coli* que consiste en un origen de replicación y un marcador de selección (frecuentemente el gen de  $\beta$ -lactamasa, *ampR*), que les permite propagarse en *E. coli* antes de la transformación en células de levadura. Además, los vectores lanzadera contienen un marcador de selección para selección en levadura. Los marcadores pueden ser genes que codifican enzimas para la síntesis de un aminoácido o nucleótido particular, de manera que las células que llevan la delección genómica correspondiente (o mutación) se complementan para auxotrofia o autotrofia. Alternativamente, estos vectores contienen marcadores de resistencia dominantes heterólogos, que proveen a las células de levadura recombinantes (es decir, las células que han incorporado el ADN y expresan el gen marcador) resistencia hacia ciertos antibióticos, como g418 (Geneticin), higromicina B o fleomicina. Además, estos vectores pueden contener una secuencia de sitios de restricción (combinados) (sitio de clonación múltiple o MCS) que permitirá clonar ADN foráneo en estos sitios, aunque también existen métodos alternativos.

55 Tradicionalmente, pueden distinguirse cuatro tipos de vectores lanzadera por la ausencia o presencia de elementos genéticos adicionales:

- Plásmidos integrativos (YIp) que se integran por recombinación homologa dentro del genoma del hospedador en el locus del gen marcador u otro gen, cuando éste se abre por restricción y se usa el ADN linealizado para la transformación de las células de levadura. Esto generalmente resulta en presencia de una copia del ADN foráneo insertado en este sitio particular en el genoma.
- 5 • Plásmidos episómicos (YE<sub>p</sub>) que llevan parte de la secuencia de ADN del plásmido 2 μ necesaria para la replicación autónoma en células de levadura. Se propagan múltiples copias del plásmido transformado en la célula de levadura y se mantienen como episomas.
- Plásmidos de replicación autónoma (YR<sub>p</sub>) que llevan un origen de replicación para levaduras (ARS, secuencia replicada autónomamente) que permite que los plásmidos transformados se propaguen varios cientos de veces.
- 10 • Plásmidos CEN (YC<sub>p</sub>) que llevan, además de una secuencia ARS, una secuencia centromérica (derivada de uno de los cromosomas nucleares) que normalmente garantiza la segregación mitótica estable y normalmente reduce el número de copias de plásmido autoreplicado a solo una.

15 Estos plásmidos se introducen en las células de levadura por transformación. La transformación de células de levadura puede lograrse por varias técnicas diferentes, tales como permeabilización de las células con acetato de litio (Ito et al., 1983) y métodos de electroporación.

En la aplicación comercial de microorganismos recombinantes, la inestabilidad del plásmido es el problema más importante. La inestabilidad es la tendencia de las células transformadas a perder sus propiedades modificadas por ingeniería debido a cambios o a pérdida de plásmidos. Este problema se trata en detalle por Zhang et al (Plasmid stability in recombinant *Saccharomyces cerevisiae*. Biotechnology Advances, Vol. 14, N.º 4, pp. 401-435, 1996). Las cepas transformadas con plásmidos integrativos son extremadamente estables, incluso en ausencia de presión selectiva (Sherman, F. <http://dbb.urmc.rochester.edu/labs/sherman/f/yeast/9.html> y referencias en la misma).

El ADN heterólogo normalmente se introduce en el organismo en forma de plásmidos extracromosómicos (YE<sub>p</sub>, YC<sub>p</sub> y YR<sub>p</sub>). Desafortunadamente, se ha encontrado, tanto con bacterias como con levaduras, que no pueden ser retenidas las nuevas características, especialmente si la presión de selección no se aplica continuamente. Esto es debido a la inestabilidad segregacional del plásmido híbrido cuando las células recombinantes crecen durante un largo periodo de tiempo. Esto conduce a heterogeneidad de la población y variabilidad clonal, y eventualmente a una población celular en la que la mayoría de las células han perdido las propiedades que se introdujeron por la transformación. Si se usan vectores con marcadores auxotróficos, el cultivo en medios ricos frecuentemente conduce a una pérdida rápida del vector, debido a que el vector solo se retiene en medio mínimo. La alternativa, el uso de marcadores de resistencia a antibióticos dominantes, frecuentemente no es compatible con los procesos de producción. El uso de antibióticos puede ser no deseable desde el punto de vista del registro (la posibilidad de que cantidades traza del antibiótico terminen en el producto final) o por motivos económicos (costes del uso de antibióticos a escala industrial).

35 La pérdida de vectores conduce a problemas en situaciones de producción a gran escala. Existen métodos alternativos para la introducción de ADN para levaduras, tales como el uso de plásmidos integrativos (YIp). El ADN se integra en el genoma hospedador mediante recombinación, produciendo alta estabilidad (Caunt, P. Stability of recombinant plasmids in yeast. Journal of Biotechnology 9(1988) 173 -192). Los presentes inventores han encontrado que es una buena alternativa un método de integración que usa los transposones del hospedador.

#### 40 Transposones

Según la invención, se integran dos a quince genes de xilosa isomerasa o genes de xilosa reductasa y xilitol deshidrogenasa en el genoma de la célula de azúcares mixtos. En una realización, la introducción inicial (es decir, antes de la evolución adaptativa) de múltiples copias, puede ejecutarse de cualquier modo conocido en la técnica que conduzca a la introducción de los genes. En una realización, esto puede llevarse a cabo usando un vector con partes homólogas a secuencias repetitivas (transposones), de la célula hospedadora. Cuando la célula hospedadora es una célula de levadura, secuencias repetidas adecuadas son las repeticiones terminales largas (LTR) del elemento Ty, conocido como secuencia delta.

Los elementos Ty se clasifican en dos subfamilias bastante similares denominadas Ty1 y Ty2. Estos elementos son de aproximadamente 6 kilobases (kb) de longitud y están unidos mediante repeticiones terminales largas (LTR), secuencias de aproximadamente 335 pares de bases (Boeke JD et al., The *Saccharomyces cerevisiae* Genome Contains Functional and Nonfunctional Copies of Transposon Ty1. Molecular and Cellular Biology, Apr. 1988, p. 1432-1442 Vol. 8, N.º 4). En la cepa totalmente secuenciada de *S. cerevisiae*, S288c, los transposones más abundantes son Ty1 (31 copias) y Ty2 (13 copias) (Gabriel A, Dapprich J, Kunkel M, Gresham D, Pratt SC, et al. (2006) Global mapping of transposon location. PLoS Genet 2(12): e212.doi:10.1371/journal.pgen.0020212). Estos transposones consisten en dos marcos de lectura abiertos solapados (ORF), cada uno de los cuales codifica varias proteínas. Las regiones codificantes están flanqueadas por las LTR casi idénticas anteriormente mencionadas. Otros elementos Ty, pero menos abundantes y más distintos, en *S. cerevisiae* comprenden Ty3, Ty4 y Ty5. Para cada

familia de elementos Ty de longitud completa hay un orden de magnitud más de elementos LTR solos dispersos a lo largo del genoma. Se cree que éstos surgen de la recombinación LTR-LTR de elementos de longitud completa, con una escisión por formación de bucle de las regiones internas que codifican proteínas.

5 Se ha usado el mecanismo de retrotransposición del retrotransposón Ty para integrar múltiples copias a lo largo del genoma (Boeke et al., 1988; Jacobs et al., 1988). Las repeticiones terminales largas (LTR) del elemento Ty, conocido como secuencias delta, también son buenas dianas para la integración por recombinación homóloga ya que existen en aproximadamente 150-200 copias que están o bien asociadas a Ty o a sitios solos (Boeke, 1989; Kingsman and Kingsman, 1988). (Parekh R.N. (1996). An Integrating Vector for Tunable, High Copy, Stable Integration into the Dispersed Ty DELTA Sites of *Saccharomyces cerevisiae*. Biotechnol. Prog. 1996, 12, 16-21). Por adaptación evolutiva, puede cambiar el número de copias.

#### La célula hospedadora

15 La célula hospedadora puede ser cualquier célula hospedadora adecuada para la producción de un producto útil. Una célula hospedadora puede ser cualquier célula adecuada, tal como una célula procariota, tal como una bacteria, o una célula eucariota. Normalmente, la célula será una célula eucariota, por ejemplo, una levadura o un hongo filamentoso.

Las levaduras se definen en la presente como microorganismos eucariotas e incluyen todas las especies de la subdivisión Eumycotina (Alexopoulos, C. J., 1962, en: Introductory Mycology, John Wiley & Sons, Inc., Nueva York) que crecen predominantemente en forma unicelular.

20 Las levaduras pueden crecer ya sea por gemación de un tallo unicelular o pueden crecer por fisión del organismo. Una levadura preferida como célula de azúcares mixtos puede pertenecer a los géneros *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces* o *Yarrowia*. Preferentemente, la levadura es una capaz de fermentación anaerobia, más preferentemente una capaz de fermentación alcohólica anaerobia.

25 Los hongos filamentosos se definen en el presente documento como microorganismos eucariotas que incluyen todas las formas filamentosas de la subdivisión Eumycotina. Estos hongos se caracterizan por un micelio vegetativo compuesto de quitina, celulosa y otros polisacáridos complejos. Los hongos filamentosos que son adecuados para su uso como una célula de la presente invención son morfológicamente, fisiológicamente y genéticamente distintos de las levaduras. Pueden usarse ventajosamente células de hongos filamentosos debido a que la mayoría de los hongos no requieren condiciones estériles para la propagación y son insensibles a las infecciones por bacteriófagos.

30 El crecimiento vegetativo por hongos filamentosos es por elongación de hifas y el metabolismo del carbono de la mayoría de los hongos filamentosos es aerobio estricto. Hongos filamentosos preferidos como célula hospedadora pueden pertenecer al género *Aspergillus*, *Trichoderma*, *Humicola*, *Acremonium*, *Fusarium* o *Penicillium*. Más preferentemente, la célula de hongo filamentoso puede ser una célula de *Aspergillus niger*, *Aspergillus oryzae*, un *Penicillium chrysogenum*, o *Rhizopus oryzae*.

35 En una realización, la célula hospedadora puede ser levadura.

Preferentemente, el hospedadora es un hospedador industrial, más preferentemente una levadura industrial. Un hospedador industrial y una célula de levadura industrial pueden definirse como sigue. Los entornos vivos de las células de levadura en procesos industriales son significativamente diferentes de los del laboratorio. Las células de levadura industriales deben ser capaces de rendir bien bajo múltiples condiciones medioambientales que pueden variar durante el proceso. Tales variaciones incluyen cambios en las fuentes de nutrientes, pH, concentración de etanol, temperatura, concentración de oxígeno, etc., que juntos tienen un posible impacto sobre el crecimiento celular y la producción de etanol de *Saccharomyces cerevisiae*. En condiciones industriales adversas, las cepas tolerantes al entorno deben permitir un crecimiento y producción robustos. Las cepas de levaduras industriales son generalmente más robustas a estos cambios en las condiciones medioambientales que pueden producirse en las aplicaciones en las que se usan, tales como en la industria de la panadería, industria cervecera, viticultura y la industria del etanol. Ejemplos de levaduras industriales (*S. cerevisiae*) son Ethanol Red® (Fermentis) Fermiol® (DSM) y Thermosacc® (Lallemand).

45 En una realización, el hospedador es tolerante a inhibidor. Pueden seleccionarse células hospedadoras tolerantes a inhibidor mediante el cribado de cepas para crecimiento sobre materiales que contienen inhibidores, tal como se ilustra en Kadar et al., Appl. Biochem. Biotechnol. (2007), Vol. 136-140, 847-858, en el que se seleccionó una cepa tolerante a inhibidor de *S. cerevisiae* ATCC 26602.

#### Genes *araA*, *araB* y *araD*

55 Una célula de azúcares mixtos es capaz de usar arabinosa. Una célula de azúcares mixtos es, por tanto, capaz de convertir L-arabinosa y L-ribulosa y/o xilulosa 5-fosfato y/o en un producto de fermentación deseado, por ejemplo, uno de aquellos mencionados en el presente documento.

Pueden producirse organismos, por ejemplo cepas de *S. cerevisiae*, capaces de producir etanol a partir de L-arabinosa modificando una célula por introducción de los genes *araA* (L-arabinosa isomerasa), *araB* (L-ribulocinasa) y *araD* (L-ribulosa-5-P4-epimerasa) a partir de una fuente adecuada. Tales genes pueden introducirse en una célula de azúcares mixtos con el objetivo de que sea capaz de usar arabinosa. Un enfoque tal se da se describe en el documento WO2003/095627. Pueden usarse genes *araA*, *araB* y *araD* de *Lactobacillus plantanum* y se desvelan en el documento WO2008/041840. Pueden usarse el gen *araA* de *Bacillus subtilis* y los genes *araB* y *araD* de *Escherichia coli* y se desvelan en el documento EP1499708. En otra realización, los genes *araA*, *araB* y *araD* pueden derivar de al menos uno del género *Clavibacter*, *Arthrobacter* y/o *Gramella*, en particular uno de *Clavibacter michiganensis*, *Arthrobacter aurescens* y/o *Gramella forsetii*, como se desvela en el documento WO 2009011591.

#### 10 Genes de PPP

Una célula de azúcares mixtos puede comprender una o más modificaciones genéticas que aumentan el flujo de la vía de la pentosa fosfato. En particular, la(s) modificación (modificaciones) genética(s) puede(n) conducir a un aumento del flujo a través de la parte no oxidativa de la vía de la pentosa fosfato. Una modificación genética que produce un aumento del flujo de la parte no oxidativa de la vía de la pentosa fosfato se entiende en el presente documento que significa una modificación que aumenta el flujo 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 aumento de flujo. El flujo de la parte no oxidativa de la vía de la pentosa fosfato puede medirse cultivando el hospedador modificado sobre xilosa como única fuente de carbono, determinando la tasa de consumo específico de xilosa y restando la tasa específica de producción de xilitol de la tasa específica de consumo de xilosa, si es que se produce xilitol. Sin embargo, el flujo de la parte no oxidativa de la vía de la pentosa fosfato es proporcional a la tasa de crecimiento sobre xilosa como única fuente de carbono, preferentemente con la tasa de crecimiento anaeróbico sobre xilosa como única fuente de carbono. Hay una relación lineal entre la tasa de crecimiento sobre xilosa como única fuente de carbono ( $\mu_{m\acute{a}x}$ ) y el flujo de la parte no oxidativa de la vía de la pentosa fosfato. La tasa específica de consumo de xilosa ( $Q_s$ ) es igual a la tasa de crecimiento ( $\mu$ ) dividida entre el rendimiento de biomasa en azúcar ( $Y_{xs}$ ) debido a que el rendimiento de biomasa en azúcar es constante (en un grupo de condiciones dadas: anaerobias, medio de crecimiento, pH, antecedentes genéticos de la cepa, etc.; es decir,  $Q_s = \mu / Y_{xs}$ ). Por lo tanto, el aumento de flujo de la parte no oxidativa de la vía de la pentosa fosfato puede deducirse a partir del aumento de la tasa de crecimiento máxima en estas condiciones a pesar del transporte (la captación es limitante).

Pueden introducirse una o más modificaciones genéticas que aumentan el flujo de la vía de la pentosa fosfato en la célula hospedadora de diversas formas. Éstas incluyen, por ejemplo, conseguir mayores niveles de actividad de estado estacionario de xilulosa cinasa y/o una o más de las enzimas de la parte no oxidativa de la vía de la pentosa fosfato y/o un nivel en estado estacionario reducido de la actividad inespecífica de aldosa reductasa. Estos cambios en los niveles de actividad de estado estacionario pueden estar afectados por la selección de mutantes (espontáneos o inducidos por productos químicos o radiación) y/o por tecnología de ADN recombinante, por ejemplo, por expresión en exceso o inactivación, respectivamente, de genes que codifican para las enzimas o factores que regulan estos genes.

En una célula hospedadora preferida, la modificación genética comprende la expresión en exceso de al menos una enzima de la (parte no oxidativa) de la vía de la pentosa fosfato. Preferentemente, la enzima se selecciona del grupo que consiste en las enzimas que codifican ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa, transcetolasa y transaldolasa. Pueden expresarse en exceso diversas combinaciones de las enzimas de la (parte no oxidativa) vía de la pentosa fosfato. Por ejemplo, las enzimas que se expresan en exceso pueden ser al menos las enzimas ribulosa-5-fosfato isomerasa y ribulosa-5-fosfato epimerasa; o al menos las enzimas ribulosa-5-fosfato isomerasa y transcetolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato epimerasa y transaldolasa; o al menos las enzimas transcetolasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato epimerasa, transcetolasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa, transcetolasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa, y transaldolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa y transcetolasa. En una realización de la invención, cada una de las enzimas ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa, transcetolasa y transaldolasa se expresan en exceso en la célula hospedadora. Se prefiere más una célula hospedadora en la que la modificación genética comprenda al menos la expresión en exceso de ambas enzimas transcetolasa y transaldolasa de modo que una célula hospedadora ya sea capaz de crecimiento anaerobio sobre xilosa. En realidad, en algunas condiciones, las células hospedadoras que expresan en exceso solo la transcetolasa y la transaldolasa ya tienen la misma tasa de crecimiento anaerobio sobre xilosa que las células hospedadoras que expresan en exceso las cuatro enzimas, es decir, la ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato epimerasa, transcetolasa y transaldolasa. Además, se prefieren las células hospedadoras que expresan en exceso ambas enzimas ribulosa-5-fosfato isomerasa y ribulosa-5-fosfato epimerasa con respecto a las células hospedadoras que expresan en exceso solo la isomerasa o solo la epimerasa ya que la expresión en exceso de solo una de estas enzimas puede producir desequilibrios metabólicos.

La enzima "ribulosa 5-fosfato epimerasa" (EC 5.1.3.1) se define en el presente documento como una enzima que cataliza la epimerización de D-xilulosa 5-fosfato a D-ribulosa 5-fosfato y viceversa. La enzima también se conoce como fosforribulosa epimerasa; eritrosa-4-fosfato isomerasa; fosfocetopentosa 3-epimerasa; xilulosa fosfato 3-epimerasa; fosfocetopentosa epimerasa; ribulosa 5-fosfato 3-epimerasa; D-ribulosa fosfato-3-epimerasa; D-ribulosa 5-fosfato epimerasa; D-ribulosa-5-P 3-epimerasa; D-xilulosa-5-fosfato 3-epimerasa; pentosa-5-fosfato 3-epimerasa; o D-ribulosa-5-fosfato 3-epimerasa. Una ribulosa 5-fosfato epimerasa también puede definirse además por su secuencia de aminoácidos. Asimismo, una ribulosa 5-fosfato epimerasa puede definirse por una secuencia de nucleótidos que codifica la enzima, además de por una secuencia de nucleótidos de referencia que codifica una ribulosa 5-fosfato epimerasa. La secuencia de nucleótidos que codifica ribulosa 5-fosfato epimerasa se designa en el presente documento *RPE1*.

La enzima "ribulosa 5-fosfato isomerasa" (EC 5.3.1.6) se define en el presente documento como una enzima que cataliza la isomerización directa de D-ribosa 5-fosfato en D-ribulosa 5-fosfato y viceversa. La enzima también se conoce como fosfopentosisomerasa; fosforribosisomerasa; ribosa fosfato isomerasa; 5-fosforribosa isomerasa; D-ribosa 5-fosfato isomerasa; D-ribosa-5-fosfato cetol-isomerasa; o D-ribosa-5-fosfato aldosa-cetosa-isomerasa. Una ribulosa 5-fosfato isomerasa también puede definirse por su secuencia de aminoácidos. Asimismo, una ribulosa 5-fosfato isomerasa puede definirse por una secuencia de nucleótidos que codifica la enzima, además de por una secuencia de nucleótidos que se hibrida con una secuencia de nucleótidos de referencia que codifica una ribulosa 5-fosfato isomerasa. La secuencia de nucleótidos que codifica ribulosa 5-fosfato isomerasa se designa en el presente documento *RK11*.

La enzima "transcetolasa" (EC 2.2.1.1) se define en el presente documento como una enzima que cataliza la reacción: D-ribosa 5-fosfato + D-xilulosa 5-fosfato  $\leftrightarrow$  sedoheptulosa 7-fosfato + D-gliceraldehído 3-fosfato y viceversa. La enzima también se conoce como glicolaldehídotransferasa o sedoheptulosa-7-fosfato:D-gliceraldehído-3-fosfato glicolaldehídotransferasa. Una transcetolasa también puede definirse por su aminoácido. Asimismo, una transcetolasa puede definirse por una secuencia de nucleótidos que codifica la enzima, además de por una secuencia de nucleótidos que se hibrida con una secuencia de nucleótidos de referencia que codifica una transcetolasa. La secuencia de nucleótidos que codifica transcetolasa se designa en el presente documento *TKL1*.

La enzima "transaldolasa" (EC 2.2.1.2) se define en el presente documento como una enzima que cataliza la reacción: sedoheptulosa 7-fosfato + D-gliceraldehído 3-fosfato  $\leftrightarrow$  D-eritrosa 4-fosfato + D-fructosa 6-fosfato y viceversa. La enzima también se conoce como dihidroxiacetona transferasa; dihidroxiacetona sintasa; formaldehído transcetolasa; o sedoheptulosa-7-fosfato:D-gliceraldehído-3-fosfato glicerina transferasa. Una transaldolasa también puede definirse por su secuencia de aminoácidos. Asimismo, una transaldolasa puede definirse por una secuencia de nucleótidos que codifica la enzima, además de por una secuencia de nucleótidos que se hibrida con una secuencia de nucleótidos de referencia que codifica una transaldolasa. La secuencia de nucleótidos que codifica transcetolasa se designa en el presente documento *TAL1*.

### 35 Genes de xilosa isomerasa o xilosa reductasa

Según la invención, se introducen dos a quince copias de uno o más genes de xilosa isomerasa y/o uno o más genes de xilosa reductasa y xilitol deshidrogenasa en el genoma de la célula hospedadora. La presencia de estos dos a quince elementos genéticos confiere a la célula la capacidad de convertir xilosa por isomerización o reducción.

En una realización, las dos a quince copias de uno o más genes de xilosa isomerasa se introducen en el genoma de la célula hospedadora.

Una "xilosa isomerasa" (EC 5.3.1.5) se define en el presente documento como una enzima que cataliza la isomerización directa de D-xilosa en D-xilulosa y/o viceversa. La enzima también se conoce como una D-xilosa cetoisomerasa. Una xilosa isomerasa en el presente documento también puede ser capaz de catalizar la conversión de D-glucosa y D-fructosa (y, por consiguiente, puede, por tanto, denominarse una glucosa isomerasa). Una xilosa isomerasa en el presente documento puede requerir un catión bivalente, tal como magnesio, manganeso o cobalto como cofactor.

Por consiguiente, una célula de azúcares mixtos tal es capaz de isomerizar xilosa en xilulosa. La capacidad de isomerizar xilosa en xilulosa se confiere a la célula hospedadora por la transformación de la célula hospedadora con una construcción de ácido nucleico que comprende una secuencia de nucleótidos que codifica una xilosa isomerasa definida. Una célula de azúcares mixtos isomeriza xilosa en xilulosa por la isomerización directa de xilosa en xilulosa.

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

El gen de xilosa isomerasa puede tener diversos orígenes, tales como, por ejemplo, de *Pyromyces* sp. como se desvela en el documento WO2006/009434. Otros orígenes adecuados son *Bacteroides*, en particular *Bacteroides uniformis* como se describe en el documento PCT/EP2009/52623, *Bacillus*, en particular *Bacillus stearothermophilus* como se describe en el documento PCT/EP2009/052625.

En otra realización, las dos a quince copias de uno o más genes de xilosa reductasa y xilitol deshidrogenasa se introducen en el genoma de la célula hospedadora. En esta realización, la conversión de xilosa se realiza en una conversión de dos etapas de xilosa en xilulosa a través de un producto intermedio de xilitol como se cataliza por xilosa reductasa y xilitol deshidrogenasa, respectivamente. En una realización de la misma, pueden expresarse en exceso xilosa reductasa (XR), xilitol deshidrogenasa (XDH) y xilocinasa (XK), y opcionalmente se regulan por incremento uno o más de genes que codifican enzimas productoras de NADPH y se regulan por incremento uno o más de los genes que codifican enzimas consumidoras de NADH, como se desvela en el documento WO 2004085627.

#### Gen XKS1

Una célula de azúcares mixtos puede comprender una o más modificaciones genéticas que aumentan la actividad específica de xilulosa cinasa. Preferentemente, la modificación o modificaciones genéticas producen la expresión en exceso de una xilulosa cinasa, por ejemplo mediante expresión en exceso de una secuencia de nucleótidos que codifica una xilulosa cinasa. El gen que codifica la xilulosa cinasa puede ser endógeno a la célula hospedadora o puede ser una xilulosa cinasa que es heteróloga a la célula hospedadora. Una secuencia de nucleótidos usada para la expresión en exceso de xilulosa cinasa en la célula hospedadora es una secuencia de nucleótidos que codifica un polipéptido con actividad de xilulosa cinasa.

La enzima "xilulosa cinasa" (EC 2.7.1.17) se define en el presente documento como una enzima que cataliza la reacción  $ATP + D\text{-xilulosa} = ADP + D\text{-xilulosa 5-fosfato}$ . La enzima también se conoce como una xilocinasa fosforilante, D-xilocinasa o ATP :D-xilulosa 5-fosfotransferasa. Una xilulosa cinasa de la invención también puede definirse por su secuencia de aminoácidos. Asimismo, una xilulosa cinasa puede definirse por una secuencia de nucleótidos que codifica la enzima, además de por una secuencia de nucleótidos que se hibrida con una secuencia de nucleótidos de referencia que codifica una xilulosa cinasa.

En una célula de azúcares mixtos, puede combinarse una modificación o modificaciones genéticas que aumenta(n) la actividad específica de xilulosa cinasa con cualquiera de las modificaciones que aumentan el flujo de la vía de la pentosa fosfato como se describió anteriormente. Esto no es, sin embargo, esencial.

Así, una célula hospedadora puede comprender solo una modificación o modificaciones genéticas que aumentan la actividad específica de xilulosa cinasa. Los diversos medios disponibles en la materia para conseguir y analizar la expresión en exceso de una xilulosa cinasa en las células hospedadoras de la invención son los mismos que se describieron anteriormente para enzimas de la vía de la pentosa fosfato. Preferentemente, en las células hospedadoras de la invención, una xilulosa cinasa que va a expresarse en exceso se expresa en exceso 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 la(s) modificación (modificaciones) genética(s) que causan la expresión en exceso. Debe entenderse que estos niveles de expresión en exceso pueden aplicarse al nivel de estado estacionario de la actividad de la enzima, al nivel de estado estacionario de la proteína de la enzima, además de al nivel de estado estacionario del transcrito que codifica la enzima.

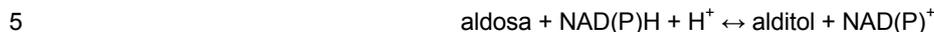
#### Delección del gen de aldosa reductasa (GRE3)

En una realización, donde XI se usa como gen para convertir xilosa, puede ser ventajoso reducir la actividad de aldosa reductasa. Una célula de azúcares mixtos puede, por tanto, comprender una o más modificaciones genéticas que reduzcan la actividad inespecífica de aldosa reductasa en la célula hospedadora. Preferentemente, la actividad inespecífica de aldosa reductasa se reduce en la célula hospedadora mediante una o más modificaciones genéticas que reducen la expresión o inactivan un gen que codifica una aldosa reductasa inespecífica. Preferentemente, la(s) modificación (modificaciones) genética(s) reducen o inactiva la expresión de cada copia endógena de un gen que codifica una aldosa reductasa inespecífica en la célula hospedadora (denominada en el presente documento delección de GRE3). Las células de azúcares mixtos pueden comprender múltiples copias de genes que codifican aldosas reductasas inespecíficas como resultado de di-, poli-o aneuploidía, y/o la célula hospedadora puede contener varias (iso)enzimas diferentes con actividad de aldosa reductasa que se diferencian en la secuencia de aminoácidos y que son cada una codificadas por un gen diferente. También en dichos casos se reduce o inactiva preferentemente la expresión de cada gen que codifica una aldosa reductasa inespecífica. Preferentemente, el gen se inactiva mediante delección de al menos parte del gen o mediante interrupción del gen, por lo que en este contexto el término gen también incluye cualquier secuencia no codificante en la dirección 5' o 3' de la secuencia codificante, cuya delección (parcial) o inactivación produce una reducción de la expresión de la actividad inespecífica de aldosa reductasa en la célula hospedadora.

Una secuencia de nucleótidos que codifica una aldosa reductasa cuya actividad va a reducirse en la célula hospedadora es una secuencia de nucleótidos que codifica un polipéptido con actividad de aldosa reductasa.

Así, una célula hospedadora que comprende solo una modificación o modificaciones genéticas que reduzca(n) la actividad inespecífica de aldosa reductasa en la célula hospedadora está especialmente incluida en la invención.

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



La enzima tiene una amplia especificidad y también se conoce como aldosa reductasa; poliol deshidrogenasa (NADP<sup>+</sup>); alditol:NADP oxidorreductasa; alditol:NADP<sup>+</sup> 1-oxidorreductasa; NADPH-aldopentosa reductasa; o NADPH-aldosa reductasa.

10 Un ejemplo particular de una aldosa reductasa inespecífica tal que es endógena a *S. cerevisiae* y que está codificada por el gen GRE3 (Traff et al., 2001, Appl. Environ. Microbiol. 67: 5668-74). Así, una aldosa reductasa de la invención también puede definirse por su secuencia de aminoácidos. Asimismo, una aldosa reductasa puede definirse por la secuencia de nucleótidos que codifica la enzima, además de por una secuencia de nucleótidos que se hibrida con una secuencia de nucleótidos de referencia que codifica una aldosa reductasa.

#### Identidad de secuencia

15 Se dice que secuencias de aminoácidos o de nucleótidos son homólogas cuando presentan un cierto nivel de similitud. Dos secuencias que son homólogas indican un origen evolutivo común. Si dos secuencias homologas están estrechamente relacionadas o más distantemente relacionadas, se indica por el "porcentaje de identidad" o "porcentaje de similitud", que es alto o bajo, respectivamente. Aunque se cuestiona, la indicación de "porcentaje de identidad" o "porcentaje de similitud", "nivel de homología" o "porcentaje de homología" se usan frecuentemente  
20 indistintamente.

Los términos "homología", "porcentaje de homología", "porcentaje de identidad" o "porcentaje de similitud" se usan indistintamente en el presente documento. Con el fin de la presente invención, se define aquí que con el fin de determinar el porcentaje de identidad de dos secuencias de aminoácidos o de dos secuencias de ácidos nucleicos, se alinean las secuencias completas para fines de comparación óptima. Con el fin de optimizar el alineamiento entre  
25 las dos secuencia, pueden introducirse huecos en cualquiera de las dos secuencias que se comparan. Tal alineamiento se lleva a cabo a lo largo de la longitud completa de las secuencias que se están comparando. Alternativamente, el alineamiento puede llevarse a cabo a lo largo de una longitud más corta, por ejemplo a lo largo de aproximadamente 20, aproximadamente 50, aproximadamente 100 o más ácidos nucleicos/bases o aminoácidos. La identidad es el porcentaje de coincidencias idénticas entre las dos secuencias a lo largo de la región alineada informada.  
30

Puede llevarse a cabo una comparación de secuencias y la determinación del porcentaje de identidad entre dos secuencias usando un algoritmo matemático. El experto conocerá el hecho de que están disponibles varios programas informáticos diferentes para alinear dos secuencias y determinar la homología entre dos secuencias (Kruskal, J. B. (1983) An overview of squence comparison In D. Sankoff y J. B. Kruskal, (ed.), Time warps, string  
35 edits and macromolecules: the theory and practice of sequence comparison, pp. 1-44 Addison Wesley). Puede determinarse el porcentaje de identidad entre dos secuencias de aminoácidos usando el algoritmo de Needleman y Wunsch para el alineamiento de dos secuencias. (Needleman, S. B. y Wunsch, C. D. (1970) J. Mol. Biol. 48, 443-453). El algoritmo alinea secuencias de aminoácidos, además de secuencias de nucleótidos. El algoritmo de Needleman-Wunsch se ha implementado en el programa informático NEEDLE. Con el fin de la presente invención se usó el programa NEEDLE del paquete EMBOSS (versión 2.8.0 o superior, EMBOSS: The European Molecular  
40 Biology Open Software Suite (2000) Rice, P. Longden, I. and Bleasby, A. Trends in Genetics 16, (6) pp 276-277, <http://emboss.bioinformatics.nl/>). Para secuencias de proteínas, se usa EBLOSUM62 para la matriz de sustitución. Para secuencias de nucleótidos, se usó EDNAFULL. Pueden especificarse otras matrices. Los parámetros óptimos usados para el alineamiento de secuencias de aminoácidos son una penalidad por abertura de hueco de 10 y una penalidad por extensión de hueco de 0,5. El experto apreciará que todos estos parámetros diferentes darán resultados ligeramente diferentes, pero que el porcentaje de identidad global de dos secuencias no se altera significativamente si se usan diferentes algoritmos.  
45

#### *Definición de homología global*

La homología o identidad es el porcentaje de coincidencias idénticas entre las dos secuencias completas a lo largo de la región alineada total que incluye cualquier hueco o extensión. La homología o identidad entre las dos secuencias alineadas se calcula como sigue: Número de posiciones correspondientes en el alineamiento que muestran un aminoácido idéntico en ambas secuencias dividido entre la longitud total del alineamiento incluyendo los huecos. La identidad definida como en el presente documento puede obtenerse de NEEDLE y se etiqueta en la salida del programa como "IDENTITY".  
50

#### *Definición de identidad más larga*

La homología o identidad entre las dos secuencias alineadas se calcula como sigue: Número de posiciones correspondientes en el alineamiento que muestran un aminoácido idéntico en ambas secuencias dividido entre la

longitud total del alineamiento después de la resta del número total de huecos en el alineamiento. La identidad definida como en el presente documento puede obtenerse de NEEDLE usando la opción NOBRIEF y se etiqueta en la salida del programa como "identidad más larga". Para los fines de la invención, el nivel de identidad (homología) entre dos secuencias (aminoácidos o nucleótidos) se calcula según la definición de "identidad más larga" como puede llevarse a cabo usando el programa NEEDLE.

Las secuencias de proteínas usadas en la presente invención pueden usarse además como una "secuencia de consulta" para realizar una búsqueda por comparación con bases de datos de secuencias, por ejemplo para identificar otros miembros de la familia o secuencias relacionadas. Tales búsquedas pueden realizarse usando los programas BLAST. El software para realizar análisis por BLAST está públicamente disponible a través del Centro Nacional para Información Biotecnológica (<http://www.ncbi.nlm.nih.gov>). BLASTP se usa para secuencias de aminoácidos y BLASTN para secuencias de nucleótidos. El programa BLAST usa como parámetros por defecto:

- Coste por hueco abierto: por defecto = 5 para nucleótidos/ 11 para proteínas
- Coste por extensión de hueco: por defecto = 2 para nucleótidos/ 1 para proteínas
- Penalización por no coincidencia de nucleótido: por defecto = -3
- Recompensa por coincidencia de nucleótido: por defecto = 1
- Valor esperado: por defecto = 10
- Tamaño de palabra: por defecto = 11 para nucleótidos/ 28 para megablast/ 3 para proteínas

Además, el grado de identidad local (homología) entre la consulta de secuencia de aminoácidos o la consulta de secuencia de nucleótidos y las secuencias homologas recuperadas se determina mediante el programa BLAST. Sin embargo, solo se comparan aquellos segmentos de secuencia que dan una coincidencia por encima de un cierto umbral. Por consiguiente, el programa calcula la identidad solo para estos segmentos coincidentes. Por tanto, la identidad calculada de esta forma se denomina identidad local.

#### Producción de bioproductos

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 bioproducción de etanol han continuado usando las levaduras del género *Saccharomyces* como productor de etanol. Esto es debido a las muchas características atractivas de las especies de *Saccharomyces* para los procesos industriales, es decir, una alta capacidad de crecimiento anaerobio con tolerancia a ácido, etanol y osmolaridad, y, por supuesto, su alta capacidad fermentativa alcohólica. Las especies de levadura preferidas como células hospedadoras incluyen *S. cerevisiae*, *S. bulderi*, *S. barnetti*, *S. exiguus*, *S. uvarum*, *S. diastaticus*, *K. lactis*, *K. marxianus* o *K. fragilis*.

Una célula de azúcares mixtos puede ser una célula adecuada para la producción de etanol. Una célula de azúcares mixtos puede, sin embargo, ser adecuada para la producción de productos de fermentación distintos de etanol

Tales productos de fermentación no etanólicos incluyen en principio cualquier producto químico a granel o refinado que pueda ser producido mediante un microorganismo eucariota tal como una levadura o un hongo filamentoso.

Una célula de azúcares mixtos que puede usarse para la producción de productos de fermentación no etanólicos es una célula hospedadora que contiene una modificación genética que produjo una actividad de alcohol deshidrogenasa reducida.

En una realización, la célula de azúcares mixtos puede usarse en un proceso en el que los azúcares que se originan a partir de lignocelulosa se convierten en etanol.

#### Lignocelulosa

La lignocelulosa, que puede considerarse como una materia prima potencialmente renovable, generalmente comprende los polisacáridos celulosa (glucanos) y hemicelulosas (xilanos, heteroxilanos y xiloglucanos). Además, algo de la hemicelulosa puede estar presente como glucomananos, por ejemplo en materias primas derivadas de madera. La hidrólisis enzimática de estos polisacáridos en azúcares solubles, que incluyen tanto monómeros como multímeros, por ejemplo glucosa, celobiosa, xilosa, arabinosa, galactosa, fructosa, manosa, ramnosa, ribosa, ácido galacturónico, ácido glucurónico y otras hexosas y pentosas, se produce bajo la acción de diferentes enzimas que actúan en concierto.

Además, pectinas y otras sustancias pécticas tales como arabinanos pueden constituir una proporción considerable de la masa seca de paredes celulares típicas de tejidos vegetales distintos de madera (aproximadamente entre un cuarto y la mitad de la masa seca puede ser pectinas).

Pretratamiento

5 Antes del tratamiento enzimático, el material lignocelulósico puede ser pretratado. El pretratamiento puede comprender exponer el material lignocelulósico a un ácido, una base, un disolvente, calor, un peróxido, ozono, ruptura mecánica, trituración, molienda o despresurización rápida, o una combinación de dos o más cualquiera de los mismos. Este pretratamiento químico se combina frecuentemente con pretratamiento con calor, por ejemplo entre 150-220 °C durante 1 a 30 minutos.

Hidrólisis enzimática

10 El material pretratado comúnmente se somete a hidrólisis enzimática para liberar los azúcares que pueden ser fermentados según la invención. Esto se puede llevar a cabo con métodos convencionales, por ejemplo poniendo en contacto con celulasas, por ejemplo celobiohidrolasa(s), endoglucanasa(s), beta-glucosidasa(s) y opcionalmente otras enzimas. La conversión con las celulasas puede llevarse a cabo a temperatura ambiente o a temperaturas más altas, a un tiempo de reacción para liberar cantidades suficientes de azúcar(es). El resultado de la hidrólisis enzimática es un producto de hidrólisis que comprende azúcares C5/C6, designados en el presente documento como la composición de azúcares.

Fermentación

15 El proceso de fermentación puede ser un proceso de fermentación aerobio o anaerobio. Un proceso de fermentación anaerobio se define en el presente documento como un proceso de fermentación realizado en ausencia de oxígeno o en el que sustancialmente no se consume oxígeno, preferentemente se consume menos de aproximadamente 5, aproximadamente 2,5 o aproximadamente 1 mmol/l/h, más preferentemente 0 mmol/l/h (es decir, el consumo de oxígeno no es detectable), y en el que las moléculas orgánicas sirven tanto de donante de electrones como de aceptores de electrones. En ausencia de oxígeno, el NADH producido en la glicolisis y formación de biomasa no puede ser oxidado mediante la fosforilación oxidativa. Para resolver este problema, muchos microorganismos usan piruvato o uno de sus derivados como aceptor de electrones y de hidrógeno, regenerando así NAD<sup>+</sup>.

20

25 Así, en un proceso de fermentación anaerobio preferido se usa 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, un aminoácido, 1,3-propano-diol, etileno, glicerol, un antibiótico β-lactámico y una cefalosporina.

30 El proceso de fermentación se realiza preferentemente a una temperatura que es óptima para la célula. Así, para la mayoría de las células hospedadoras de levadura o fúngicas, el proceso de fermentación se realiza a una temperatura que es inferior a aproximadamente 42 °C, preferentemente inferior a aproximadamente 38 °C. Para células hospedadoras de levadura o de hongos filamentosos, el proceso de fermentación se realiza preferentemente 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.

35 El rendimiento de etanol sobre xilosa y/o glucosa en el proceso es preferentemente de al menos aproximadamente el 50, aproximadamente el 60, aproximadamente el 70, aproximadamente el 80, aproximadamente el 90, aproximadamente el 95 o aproximadamente el 98 %. El rendimiento de etanol se define en el presente documento como un porcentaje del rendimiento teórico máximo.

La invención también se refiere a un proceso de producción de un producto de fermentación.

40 El proceso de fermentación según la presente invención puede realizarse en condiciones aerobias y anaerobias. En una realización, el proceso se lleva a cabo bajo condiciones microaerófilas o de oxígeno limitado.

45 Un proceso de fermentación anaerobio se define en el presente documento como un proceso de fermentación realizado en ausencia de oxígeno o en el que sustancialmente no se consume oxígeno, preferentemente menos de aproximadamente 5, aproximadamente 2,5 o aproximadamente 1 mmol/l/h, y en el que las moléculas orgánicas sirven de donante de electrones como de aceptores de electrones.

50 Un proceso de fermentación de oxígeno limitado es un proceso en el que el consumo de oxígeno está limitado por la transferencia de oxígeno del gas al líquido. El grado de limitación de oxígeno se determina por la cantidad y composición del flujo de gas entrante, además de las propiedades de mezcla/transferencia de masa reales del equipo de fermentación usado. Preferentemente, en un proceso en condiciones de oxígeno limitado, la tasa de consumo de oxígeno es de al menos aproximadamente 5,5, más preferentemente al menos aproximadamente 6, tal como al menos 7 mmol/l/h. Un proceso de la invención puede comprender la recuperación del producto de fermentación.

55 En un proceso preferido, la célula fermenta tanto la xilosa como la glucosa, preferentemente simultáneamente, en cuyo caso preferentemente se usa una célula que no es sensible a la represión por glucosa para prevenir el crecimiento diaúxico. Además de una fuente de xilosa (y glucosa) como fuente de carbono, el medio de fermentación

comprenderá además el componente apropiado requerido para el crecimiento de la célula. Composiciones de medio de fermentación para el crecimiento de microorganismos tales como levaduras son muy conocidas en la técnica.

- 5 Los procesos de fermentación pueden llevarse a cabo en modo discontinuo, de lotes alimentados o continuo. También puede aplicarse un proceso de hidrólisis y fermentación (SHF) separado o un proceso de sacarificación y fermentación (SSF) simultáneo. También es posible una combinación de estos modos de procesos de fermentación para una productividad óptima. Estos procesos se describen más adelante en más detalle.

#### Modo SSF

- 10 Para el modo de sacarificación y fermentación simultánea (SSF), el tiempo de reacción para la etapa de licuefacción/hidrólisis o presacarificación depende del tiempo para lograr un rendimiento deseado, es decir, el rendimiento de conversión de celulosa a glucosa. Tal rendimiento es preferentemente tan alto como sea posible, preferentemente 60 % o más, 65 % o más, 70 % o más, 75 % o más 80 % o más, 85 % o más, 90 % o más, 95 % o más, 96 % o más, 97 % o más, 98 % o más, 99 % o más, incluso 99,5 % o más o 99,9 % o más.

- 15 Según la invención, se logran concentraciones de azúcar muy altas en el modo SHF y concentraciones de producto muy altas (por ejemplo, etanol) en el modo SSF. En la operación de SHF, la concentración de glucosa es 25 g/l o más, 30 g/l o más, 35 g/l o más, 40 g/l o más, 45 g/l o más, 50 g/l o más, 55 g/l o más, 60 g/l o más, 65 g/l o más, 70 g/l o más, 75 g/l o más, 80 g/l o más, 85 g/l o más, 90 g/l o más, 95 g/l o más, 100 g/l o más, 110 g/l o más, 120 g/l o más o puede ser, por ejemplo, 25 g/l-250 g/l, 30 g/l-200 g/l, 40 g/l-200 g/l, 50 g/l-200 g/l, 60 g/l-200 g/l, 70 g/l-200 g/l, 80 g/l-200 g/l, 90 g/l, 80 g/l-200 g/l.

#### Concentración de producto en modo SSF

- 20 En la operación de SSF, la concentración de producto (g/l) depende de la cantidad de glucosa producida, pero esto no es visible debido a que los azúcares se convierten en producto en SSF, y las concentraciones de productos pueden relacionarse con la concentración de glucosa subyacente por multiplicación con el rendimiento máximo teórico (Yps máximo en g de producto por gramo de glucosa)

- 25 El rendimiento máximo teórico (Yps máximo en g de producto por gramo de glucosa) de un producto de fermentación puede derivar de textos de bioquímica. Para etanol, 1 mol de glucosa (180 gramos) da según la vía normal de fermentación por glicolisis en levadura 2 moles de etanol (=  $2 \times 46 = 92$  g de etanol. El rendimiento máximo teórico de etanol en glucosa es, por tanto,  $92/180 = 0,511$  g de etanol/g de glucosa.

Para butanol (MW 74 g/mol) o isobutanol, el rendimiento teórico máximo es 1 mol de butanol por mol de glucosa. Así, el Yps máximo para (iso-)butanol =  $74/180 = 0,411$  g de (iso-)butanol/g de glucosa.

- 30 Para el ácido láctico, el rendimiento de fermentación para la fermentación homoláctica es 2 moles de ácido láctico (MW = 90 g/mol) por mol de glucosa. Según esta estequiometría, el Yps máximo = 1 g de ácido láctico/g de glucosa.

Para otros productos de fermentación puede hacerse un cálculo similar.

#### Modo SSF

- 35 En la operación SSF, la concentración de producto es 25 g \* Yps g/l/l o más, 30 \* Yps g/l o más, 35 g \* Yps g/l o más, 40 \* Yps g/l o más, 45 \* Yps g/l o más, 50 \* Yps g/l o más, 55 \* Yps g/l o más, 60 \* Yps g/l o más, 65 \* Yps g/l o más, 70 \* Yps g/l o más, 75 \* Yps g/l o más, 80 \* Yps g/l o más, 85 \* Yps g/l o más, 90 \* Yps g/l o más, 95 \* Yps g/l o más, 100 \* Yps g/l o más, 110 \* Yps g/l o más, 120 g/l \* Yps o más o puede ser, por ejemplo, 25 \* Yps g/l-250 \* Yps g/l, 30 \* Yps g/l-200 \* Yps g/l, 40 \* Yps g/l-200 \* Yps g/l, 50 \* Yps g/l-200 \* Yps g/l, 60 \* Yps g/l-200 \* Yps g/l, 70 \* Yps g/l-200 \* Yps g/l, 80 \* Yps g/l-200 \* Yps g/l, 90 \* Yps g/l, 80 \* Yps g/l-200 \* Yps g/l

- 40 Por consiguiente, la invención proporciona un método para la preparación de un producto de fermentación, método que comprende:

- a. degradar lignocelulosa usando un método como se describe en el presente documento; y
- b. fermentar el material resultante,

preparando así un producto de fermentación.

- 45 Producto de fermentación

- El producto de fermentación de la invención puede ser cualquier producto útil. En una realización, es un producto seleccionado del grupo que consiste en etanol, n-butanol, isobutanol, ácido láctico, ácido 3-hidroxiisovalérico, ácido acrílico, ácido acético, ácido succínico, ácido fumárico, ácido málico, ácido itacónico, ácido maleico, ácido cítrico, ácido adípico, un aminoácido, tal como lisina, metionina, triptófano, treonina y ácido aspártico, 1,3-propano-diol, etileno, glicerol, un antibiótico  $\beta$ -lactámico y una cefalosporina, vitaminas, productos farmacéuticos, suplementos para piensos para animales, especialidades químicas, materias primas químicas, plásticos, disolventes,

combustibles, que incluyen biocombustibles y biogás o polímeros orgánicos, y una enzima industrial, tal como una proteasa, una celulasa, una amilasa, una glucanasa, una lactasa, una lipasa, una liasa, una oxidoreductasa, una transferasa o una xilanas. Por ejemplo, los productos de fermentación pueden ser producidos por células según la invención, siguiendo métodos de preparación de células del estado de la técnica y procesos de fermentación, cuyos ejemplos, sin embargo, no deben interpretarse como limitantes en el presente documento. Puede producirse n-butanol mediante células como se describe en los documentos WO2008121701 o WO2008086124; ácido láctico como se describe en los documentos US2011053231 o US2010137551; ácido 3-hidroxipropiónico como se describe en el documento WO2010010291; ácido acrílico como se describe en el documento WO2009153047; ácido acético como se describe en..., ácido succínico como se describe en..., las células del mutante deficiente en fumarasa que fermentan glucosa acumularon ácido fumárico extracelular (\*\*Nota todavía por añadir..\_

#### Recuperación del producto de fermentación

Para la recuperación del producto de fermentación se usan tecnologías existentes. Para diferentes productos de fermentación son apropiados diferentes procesos de recuperación. Los métodos existentes de recuperación de etanol de mezclas acuosas comúnmente usan técnicas de fraccionamiento y adsorción. Por ejemplo, todavía puede usarse una cerveza para procesar un producto fermentado, que contiene etanol en una mezcla acuosa, para producir una mezcla enriquecida que contiene etanol que luego se somete a fraccionamiento (por ejemplo, destilación fraccionada u otras técnicas similares). A continuación, las fracciones que contienen las concentraciones de etanol más altas pueden pasarse a través de un adsorbente para eliminar la mayoría, sino toda el agua remanente del etanol.

Los siguientes ejemplos ilustran la invención:

#### **EJEMPLOS**

A menos que se indique lo contrario, los métodos descritos aquí son técnicas bioquímicas convencionales. Ejemplos de textos adecuados de metodología general incluyen Sambrook *et al.*, Molecular Cloning, a Laboratory Manual (1989), y Ausubel *et al.*, Current Protocols in Molecular Biology (1995), John Wiley & Sons, Inc.

#### Composición del medio

Experimentos de crecimiento: Se cultivan cepas de *Saccharomyces cerevisiae* en un medio que tiene la siguiente composición: 0,67 % (p/v) de base nitrogenada o medio sintético para levaduras (Verduyn *et al.*, Yeast 8:501-517, 1992) y glucosa, arabinosa, galactosa o xilosa, o una combinación de estos sustratos, en concentraciones variables (véanse los ejemplos para detalles específicos; concentraciones en % en peso con respecto al volumen (p/v)). Para las placas con agar, el medio se complementa con 2 % (p/v) de agar bacteriológico.

#### Producción de etanol

Se prepararon precultivos inoculando 25 ml de medio de Verduyn (Verduyn *et al.*, Yeast 8:501-517, 1992) complementado con 2 % de glucosa en un matraz oscilante 100 ml, con un cultivo madre congelado o una colonia individual de una placa con agar. Después de la incubación a 30 °C en un agitador orbital (280 rpm) durante aproximadamente 24 horas, se recogió este cultivo y se usó para la determinación de experimentos de desprendimiento de CO<sub>2</sub> y de producción de etanol.

Se realizaron cultivos para la producción de etanol a 30 °C en 100 ml de medio modelo sintético (medio de Verduyn (Verduyn *et al.*, Yeast 8:501-517, 1992), con 5 % de glucosa, 5 % de xilosa, 3,5 % de arabinosa y 1 % de galactosa), en BAM (monitor de la actividad biológica, Los Países Bajos). El pH del medio se ajustó a 4,2 con NaOH/H<sub>2</sub>SO<sub>4</sub> 2 M antes de la esterilización. Se complementó el medio sintético para el cultivo anaerobio con 0,01 g l<sup>-1</sup> de ergosterol y 0,42 g l<sup>-1</sup> de Tween 80 disuelto en etanol (Andreasen y Stier. J. Cell Physiol. 41:23-36, 1953; y Andreasen y Stier. J. Cell Physiol. 43:271-281, 1954). El medio se inoculó a una DO<sub>600</sub> inicial de aproximadamente 2. Los cultivos se agitaron con un agitador magnético. Las condiciones anaerobias se desarrollaron rápidamente durante la fermentación, ya que los cultivos no fueron aireados. Se monitorizó constantemente la producción de CO<sub>2</sub>. La conversión de azúcar y la formación de productos (etanol, glicerol) por RMN. El crecimiento se monitorizó siguiendo la densidad óptica del cultivo a 600 nm en un espectrofotómetro LKB Ultrospec K.

#### Transformación de *S. cerevisiae*

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

#### PCR de colonias

Se recogió una cepa aislada de colonias individuales con un palillo de plástico y se resuspendió en 50 µl de agua MilliQ. La muestra se incubó durante 10 minutos a 99 °C. Se usaron 5 µl de la muestra incubada como molde para la reacción de PCR, usando ADN polimerasa Phusion® (Finnzymes) según las instrucciones proporcionadas por el proveedor.

Condiciones de la 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	

Aislamiento de ADN cromosómico

5 Se cultivaron células de levadura en medio YEP que contenía 2 % de glucosa en un agitador rotatorio (durante la noche, a 30 °C y 280 rpm en un agitador orbital). Se transfirieron 1,5 ml de estos cultivos a un tubo Eppendorf y se centrifugaron durante 1 minuto a velocidad máxima. El sobrenadante se decantó y el sedimento se resuspendió en 200 µl de YCPS (0,1 % de SB3-14 (Sigma Aldrich, Los Países Bajos) en Tris.HCl 10 mM a pH 7,5; EDTA 1 mM) y 1 µl de RNasa (20 mg/ml de RNasa A de páncreas bovino, Sigma, Los Países Bajos). La suspensión de células se incubó durante 10 minutos a 65 °C. La suspensión se centrifugó en una centrífuga Eppendorf durante 1 minuto a 7000 rpm. El sobrenadante se desechó. El sedimento se disolvió cuidadosamente en 200 µl de CLS (EDTA 25 mM, 10 2 % de SDS) y 1 µl de RNasa A. Después de la incubación a 65 °C durante 10 minutos, la suspensión se enfrió sobre hielo. Después de la adición de 70 µl de PPS (acetato de amonio 10 M), las disoluciones se mezclaron minuciosamente en una mezcladora con vórtex. Después de la centrifugación (5 minutos en centrífuga Eppendorf a velocidad máxima), el sobrenadante se mezcló con 200 µl de isopropanol helado. El ADN precipitó fácilmente y se sedimentó por centrifugación (5 minutos, velocidad máxima). El sedimento se lavó con 400 µl de 70 % de etanol helado. El sedimento se secó a temperatura ambiente y se disolvió en 50 µl de TE (Tris.HCl 10 mM a pH 7,5, EDTA 1 mM).

Transformación de células de levadura por electroporación

Se cultivaron células de levadura inoculando 25 ml de medio YEP que contenía 2 % de glucosa con una colonia de levadura individual. El matraz se incubó durante la noche a 30 °C.

20 Se determinó la densidad óptica a 600 nm y se transfirió la cantidad necesaria para obtener una densidad óptica de 0,2 a 100 ml de medio YEP con 2 % de glucosa. Las células se cultivaron durante 4 a 5 horas, con el fin de alcanzar una densidad óptica de aproximadamente 1,2 a 1,3, que se corresponde con 2 a 3 generaciones. Se recogieron las células por centrifugación y se suspendieron en 28 ml de TE (Tris-HCl 10 mM, EDTA 1 mM, pH 7,5). Se añadieron 3 ml de una disolución 1 M de LiAC (establecida a pH 7,5 con HAc concentrado). Las células se agitaron suavemente en una estufa de incubación rotatoria (150 rpm, 30 °C) durante 45 minutos. Después de la adición de 25 500 µl de una disolución 1 M de DTT (ditiotreitól), las células se incubaron una vez más bajo estas condiciones, durante 15 minutos. El volumen se enrasó a 100 ml con agua MilliQ helada estéril. Las células se recogieron por centrifugación.

30 Se desechó el sobrenadante y las células sedimentadas se lavaron con 50 ml de agua MilliQ helada estéril, y se recogieron por centrifugación. Se hizo un tratamiento de lavado posterior con 30 ml de una disolución 1 M helada de sorbitol. Después de la centrifugación, el sobrenadante se desechó y el sedimento de células se resuspendió en 4 ml de una disolución 1 M helada de sorbitol. Después de la centrifugación, el sobrenadante se desechó y el sedimento de células se resuspendió en 300 µl de una disolución 1 M helada de sorbitol.

35 Para cada transformación, se transfieren 40 µl de la suspensión de células a tubos Eppendorf helados. Se añaden el ADN transformante y 5 µg de ADN de esperma de salmón (como ADN vehículo), juntos en un volumen máximo de 20 µl. El ADN debe disolverse en TE. Se golpea cuidadosamente el tubo Eppendorf con el fin de mezclar el contenido suavemente. Transferir el contenido a una cubeta de electroporación previamente enfriada (sobre hielo) con un hueco de 0,2 cm. Aplicar un pulso (usando, por ejemplo, un dispositivo de electroporación BioRad) a 1,5 kV, 200 Ohm y 25 µF. La duración del pulso debería ser de aproximadamente 5 ms.

40 Transferir las células inmediatamente a 200 µl de sorbitol 1 M. Añadir 4 ml de YEP con 2 % de glucosa e incubar a 30 °C durante 1 hora. Recoger las células por centrifugación, desechar el sobrenadante y resuspender el sedimento en 4 ml de sorbitol 1 M. Recoger las células por centrifugación, desechar el sobrenadante y resuspender el sedimento en 300 µl de sorbitol 1 M. Diluir las células según convenga y extenderlas sobre placas selectivas o usarlas en medios selectivos.

45

Prueba de aplicación de la levadura sobre hidrolizados reales

Se hidrolizaron enzimáticamente muestras pretratadas con ácido diluido de hojas y tallos de maíz y paja de trigo usando una preparación de celulasa de amplio espectro experimental a 60 °C durante 3 días (72 horas). El pH al comienzo de la hidrólisis fue 5,0. El contenido de materia seca al comienzo de la hidrólisis fue 10 % p/p.

- 5 Las condiciones para la hidrólisis de las muestras de fibra de maíz pretratadas fueron esencialmente las mismas, excepto que la temperatura de hidrólisis fue 50 °C y el contenido de materia seca al comienzo de la hidrólisis fue del 13,8 %.

Después de la hidrólisis (72 h), se dejó que las muestras se enfriaran a temperatura ambiente. El pH se ajustó a 5,5 usando 10 % de NaOH. Posteriormente, se añadió 1 mililitro de 200 gramos por litro de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> y 1 mililitro de 100 gramos por litro de KH<sub>2</sub>PO<sub>4</sub>. Finalmente, se añadieron muestras de levadura correspondientes a un contenido de materia seca de levadura de 1 gramo de levadura por kilogramo de hidrolizado. Se siguió el desprendimiento de CO<sub>2</sub> con el tiempo usando el AFM (monitor de fermentación de alcohol; HaloteC Instruments BV, Veenendaal, Los Países Bajos). Los experimentos se realizaron al menos por triplicado, durante 72 horas a 33 °C. Uno de estos se muestrea a intervalos regulares con el fin de ser capaz de analizar la formación de etanol y concentraciones de azúcar residual. Estos datos pueden usarse para calcular los rendimientos de fermentación. No se muestrea el caldo de los otros dos experimentos. En su lugar, al final de la fermentación, el caldo se destila usando una unidad de destilación Buchi K-355 al 45 % de vapor durante 15 minutos. El alcohol producido se determina usando un medidor de densidad Anton Paar DMA 5000 (Anton Paar Benelux BVBA, Dongen, Los Países Bajos).

PCR cuantitativa

20 Se realizaron reacciones de PCR cuantitativa para determinar el número de copias de los genes que estaban presentes en el genoma, especialmente en el caso de los genes de xilosa isomerasa. Para este fin, se usó el sistema Bio-Rad iCycler iQ de Bio-Rad (Bio-Rad Laboratories, Hercules, CA, EE.UU.). Se usó iQ SYBR Green Supermix (Bio-Rad). Los experimentos se configuraron como se sugería en el manual del proveedor. Como cebadores para la detección del gen de xilosa isomerasa se usaron los cebadores de SEQ ID NO: 23 y SEQ ID NO: 24.

Las condiciones de PCR fueron del siguiente modo:

incubación de 3 minutos a 95 °C  
 40 ciclos: 10 segundos a 95 °C  
                   45 segundos a 58 °C  
 30                   45 segundos a 72 °C  
 incubación de 1 minuto a 95 °C  
 incubación de 1 minuto a 65 °C

La curva de fusión se determina empezando a medir la fluorescencia a 65 °C durante 10 segundos. La temperatura se aumenta cada 10 segundos con 0,5 °C, hasta que se alcanza una temperatura de 95 °C. A partir de las lecturas, puede calcularse y/o estimarse el número de copias del gen de interés. El método tiene sus limitaciones con respecto a una determinación precisa del número de copias por encima de un cierto umbral, como será apreciado por aquellos expertos en la materia.

Ejemplo 1Construcción de la cepa BIE104A2P140 1.1 Construcción de un vector de expresión que contiene los genes para la vía de arabinosa

Se construyó el plásmido pPWT018, como se expone en la Figura 2, del siguiente modo: Se digirió el vector pPWT006 (Figura 1, que consiste en un locus *SIT2* (Gottlin-Ninfa y Kaback (1986) Molecular and Cell Biology, vol. 6, N.º 6, 2185-2197) y los marcadores que permiten la selección de transformantes en el antibiótico G418 y la capacidad para cultivar en acetamida con las enzimas de restricción *Bsi*WI y *Mlu*I. El marcador kanMX, que confiere resistencia a G418, se aisló de p427TEF (Dualsystems Biotech) y un fragmento que contiene el marcador *amdS* se ha descrito en la bibliografía (Swinkels, B.W., Noordermeer, A.C.M. y Renniers, A.C.H.M (1995) The use of the *amdS* cDNA of *Aspergillus nidulans* as a dominant, bidirectional selection marker for yeast transformation. Yeast, volumen 11, número 1995A, página S579; y documento US 6051431). Se sintetizaron los genes que codifican arabinosa isomerasa (*araA*), L-ribulocinasa (*araB*) y L-ribulosa-5-fosfato-4-epimerasa (*araD*) de *Lactobacillus plantarum*, como se desvela en la solicitud de patente WO2008/041840, por BaseClear (Leiden, Los Países Bajos). Se sintetizó un fragmento grande que alojaba los tres genes de arabinosa mencionados anteriormente, bajo el control de (o unidos operativamente a) promotores fuertes de *S. cerevisiae*, es decir, 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*. Éste estaba rodeado por las enzimas de restricción únicas *Acc65I* y *MluI*. La clonación de este fragmento en pPWT006 digerido con *MluI* y *BsiWI* produjo el plásmido pPWT018 (Figura 2). La secuencia del plásmido pPWT018 expone en SEQ ID NO: 1.

### 1.2 Transformación de levadura

5 Se transformó CEN.PK113-7D (*MATa URA3 HIS3 LEU2 TRP1 MAL2-8 SUC2*) con el plásmido pPWT018, que había sido previamente linealizado con *SfiI* (New England Biolabs), según las instrucciones del proveedor. Se diseñó un sitio *SfiI* sintético en el flanco 5' del sitio *SIT2* (véase la Figura 2). Se sembraron mezclas de transformación en YPD-agar (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 contenía 100 µg de G418 (Sigma Aldrich) por ml. Después de dos a cuatro días, aparecieron colonias sobre las placas, mientras que el control negativo (es decir, sin adición de ADN en el experimento de transformación) produjo placas YPD/G418 vacías. La integración del plásmido pPWT018 se dirige al locus *SIT2*. Se caracterizaron los transformantes usando PCR y técnicas de transferencia Southern.

15 Se realizaron reacciones de PCR, que son indicativas de la correcta integración de una copia de plásmido pPWT018, con los cebadores indicados por SEQ ID NO: 2 y SEQ ID NO: 3, y SEQ ID NO: 3 y SEQ ID NO: 4. Con los pares de cebadores de SEQ ID NO: 2 y SEQ ID NO: 3, se comprobó la correcta integración en el locus *SIT2*. Si el plásmido pPWT018 se integró en múltiples copias (integración de cabeza a cola), el par de cebadores de SEQ ID NO: 3 y SEQ ID NO: 4 dará un producto de PCR. Si el último producto de PCR está ausente, esto es indicativo de una integración de una copia de pPWT018. Una cepa en la que se integró una copia del plásmido pPWT018 en el locus *SIT2* se designó BIE104R2.

### 20 1.3 Rescate de los marcadores

Con el fin de ser capaces de transformar la cepa de levadura con otras construcciones usando los mismos marcadores de selección, es necesario eliminar los marcadores de selección. El diseño del plásmido pPWT018 fue tal que tras la integración de pPWT018 en el cromosoma, las secuencias homólogas están en estrecha proximidad entre sí. Este diseño permite que los marcadores de selección sean perdidos por recombinación intramolecular espontánea de estas regiones homólogas.

Tras el crecimiento vegetativo, tendrá lugar la recombinación intramolecular, aunque a baja frecuencia. La frecuencia de esta recombinación depende de la extensión de la homología y el locus en el genoma (resultados no publicados). Tras la transferencia secuencial de una subfracción del cultivo a medio fresco, se acumularán recombinantes intramoleculares con el tiempo.

30 Para este fin, se cultivó la cepa BIE104R2 en medio YPD (por litro: 10 gramos de extracto de levadura, 20 gramos por litro de peptona, 20 gramos por litro de dextrosa), a partir de una cepa aislada de colonias individuales. Se usaron 25 µl de un cultivo 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 se diluyeron a una concentración de aproximadamente 5000 por ml. Se sembraron 100 µl de la suspensión de células sobre medio de carbono básico para levadura (Difco) que contenía KPi 30 mM (pH 6,8), 0,1 % de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, fluoroacetamida 40 mM (Amersham) y 1,8 % de agar (Difco). Células idénticas a las células de la cepa BIE104R2, es decir, sin recombinación intracelular, todavía contienen el gen *amdS*. Para aquellas células, la fluoroacetamida es tóxica. Estas células no serán capaces de crecer y no formarán colonias sobre un medio que contiene fluoroacetamida. Sin embargo, si ha ocurrido recombinación intramolecular, variantes de BIE104R2 que han perdido los marcadores de selección serán capaces de crecer en el medio de fluoroacetamida, ya que son incapaces de convertir fluoroacetamida en compuestos inhibidores del crecimiento. Aquellas células formarán colonias sobre este medio de agar.

45 Las colonias resistentes a fluoroacetamida obtenidas se sometieron a un análisis de PCR usando cebadores de SEQ ID NO: 2 y SEQ ID NO: 3, y SEQ ID NO: 4 y SEQ ID NO: 5. Los cebadores de SEQ ID NO: 2 y SEQ ID NO: 3 darán una banda si la recombinación de los marcadores de selección ha tenido lugar como era previsto. Como resultado, el casete con los genes *araA*, *araB* y *araD* bajo el control de los promotores fuertes de levadura ha sido integrado en el locus *SIT2* del genoma de la cepa hospedadora. En ese caso, una reacción de PCR usando cebadores de SEQ ID NO: 4 y SEQ ID NO: 5 no debe producir un producto de PCR, ya que el cebador 4 sensibiliza en una región que debe ser perdida debido a la recombinación. Si se obtiene una banda con los últimos cebadores, esto es indicativo de la presencia del plásmido pPWT018 completo en el genoma, por lo que no ha tenido lugar recombinación.

50 Si los cebadores de SEQ ID NO: 2 y SEQ ID NO: 3 no producen un producto de PCR, la recombinación ha tenido lugar, pero de tal forma que el plásmido pPWT018 completo se ha eliminado del genoma por recombinación. No solo se han perdido los marcadores de selección, sino también los genes de arabinosa. En realidad, se ha recuperado levadura no mutada.

55 Las cepas aisladas que mostraron resultados de PCR según la integración de una copia de pPWT018 se sometieron a análisis de transferencia Southern. El ADN cromosómico de las cepas CEN.PK113-7D y los recombinantes correctos se digirieron con *EcoRI* y *HindIII* (digestión doble). Se preparó una sonda de *SIT2* con cebadores de

SEQ ID NO: 6 y SEQ ID NO: 7 usando ADN cromosómico de la cepa CEN.PK113-7D como molde. El resultado del experimento de hibridación se muestra en la Figura 3.

5 En la cepa no mutada, se observa una banda de 2,35 kb, que es según el tamaño esperado del gen no mutado. Tras la integración y pérdida parcial por recombinación del plásmido pPWT018, era de esperar una banda de 1,06 kb. De hecho, esta banda se observa, como se muestra en la Figura 3 (carril 2).

Una de las cepas que mostró el patrón correcto de bandas en la transferencia Southern (como puede deducirse de la Figura 3) es la cepa designada BIE104A2.

#### 1.4 Introducción de cuatro genes expresados constitutivamente de la vía de la pentosa fosfato no oxidativa

10 Se transformó *Saccharomyces cerevisiae* BIE104A2, que expresa los genes *araA*, *araB* y *araD* constitutivamente, con el plásmido pPWT080 (Figura 4). La secuencia del plásmido pPWT080 se expone en SEQ ID NO: 8. El procedimiento para la transformación y selección, después de seleccionar un transformante de integración de una copia, es el mismo que se ha descrito anteriormente en las Secciones 1.1, 1.2 y 1.3. En resumen, se transformó BIE104A2 con pPWT080 digerido con *Sfi*I. Se sembraron mezclas de transformación en YPD-agar (por litro: 10  
15 gramos de extracto de levadura, 20 gramos por litro de peptona, 20 gramos por litro de dextrosa, 20 gramos de agar) que contenía 100 µg de G418 (Sigma Aldrich) por ml.

Después de dos a cuatro días, aparecieron colonias en las placas, mientras que el control negativo (es decir, sin adición de ADN en el experimento de transformación) produjo placas YPD/G418 vacías.

20 La integración del plásmido pPWT080 se dirige al locus GRE3. Se caracterizaron transformantes usando PCR y técnicas de transferencia Southern. Se comprobó la correcta integración del plásmido pPWT080 en el locus GRE3 por PCR usando los pares de cebadores de SEQ ID NO: 9 y SEQ ID NO: 10, y el par de cebadores de SEQ ID NO: 9 y SEQ ID NO: 11 se usó para detectar integración de una copia o múltiples copias del plásmido pPWT080. Para el análisis de Southern, se preparó una sonda por PCR usando SEQ ID NO: 12 y SEQ ID NO: 13, amplificando una parte del gen *RK11* de *S. cerevisiae*. A continuación del gen *RK11* nativo, se obtuvo una señal adicional resultante de la integración del plásmido pPWT080 (datos no mostrados)

25 Un transformante que mostro la correcta integración de una copia del plásmido pPWT080, según el patrón de hibridación esperado, se designó BIE104A2F1. Con el fin de eliminar los marcadores de selección introducidos por la integración del plásmido pPWT080, se cultivó la cepa BIE104A2F1 en medio YPD, a partir de una cepa aislada de colonia. Se usaron 25 µl de un cultivo durante la noche para inocular medio YPD nuevo. Después de cinco  
30 transferencias en serie, se determinó la densidad óptica del cultivo y las células se diluyeron a una concentración de aproximadamente 5000 por ml. Se sembraron 100 µl de la suspensión de células en medio de carbono básico para levadura (Difco) que contenía KPi 30 mM (pH 6,8), 0,1 % de (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, fluoroacetamida 40 mM (Amersham) y 1,8 % de agar (Difco). Colonias resistentes a fluoroacetamida se sometieron a un análisis de PCR usando los  
35 cebadores de SEQ ID NO: 9 y SEQ ID NO: 10. En caso de perfiles de PCR correctos, se realizó análisis de transferencia Southern con el fin de verificar la correcta integración, otra vez usando la sonda del gen *RK11*. Una de las cepas que mostró el patrón correcto de bandas en la transferencia Southern es la cepa designada BIE104A2P1.

## **Ejemplo 2**

### **Evolución adaptativa en matraz oscilante que conduce a BIE104A2P1c y BIE201.**

#### 2.1 Evolución adaptativa (aerobia)

40 Se usó una cepa aislada de colonias individuales de la cepa BIE104A2P1 para inocular medio YNB (Difco) complementado con 2 % de galactosa. El precultivo se incubó durante aproximadamente 24 horas a 30 °C y 280 rpm en un agitador orbital. Se recogieron las células y se inocularon en medio YNB que contenía 1 % de galactosa y 1 % de arabinosa a una DO<sup>600</sup> inicial de 0,2 (Figura 5). Las células se cultivaron a 30 °C y 280 rpm en un agitador orbital. Se monitorizó regularmente la densidad óptica a 600 nm.

45 Cuando la densidad óptica alcanzó un valor de 5, se transfirió una alícuota del cultivo a medio YNB fresco que contenía el mismo medio. La cantidad de células añadidas fue tal que la DO<sup>600</sup> inicial del cultivo fue 0,2. Después de alcanzar una DO<sup>600</sup> de 5 otra vez, se transfirió una alícuota del cultivo a medio YNB que contenía 2 % de arabinosa como única fuente de carbono (evento indicado por (1) en la Figura 5).

50 Tras la transferencia a YNB con 2 % de arabinosa como única fuente de carbono, pudo observarse el crecimiento después de aproximadamente dos semanas. Cuando la densidad óptica a 600 nm alcanzó un valor al menos de 1, se transfirieron células a un matraz oscilante con medio YNB fresco complementado con 2 % de arabinosa a una DO<sup>600</sup> inicial de 0,2 (Figura 5, día 28).

Se repitió tres veces la transferencia secuencial, como se expone en la Figura 5. La cepa resultante que fue capaz de crecer rápido en arabinosa se designó BIE104A2P1c.

## 2.2 Evolución adaptativa (anaerobia)

Después de la adaptación al crecimiento en arabinosa en condiciones aerobias, se inoculó una colonia individual de la cepa BIE104A2P1c en medio YNB complementado con 2 % de glucosa. El precultivo se incubó durante aproximadamente 24 horas a 30 °C y 280 rpm en un agitador orbital. Se recogieron las células y se inocularon en medio YNB que contenía 2 % de arabinosa, con una densidad óptica inicial  $DO^{600}$  de 0,2. Los matraces se cerraron con tapones impermeables, asegurando condiciones de crecimiento anaerobias después de que el oxígeno se agotara del medio y el espacio de cabeza. Después de alcanzar una  $DO^{600}$  mínima de 3, se transfirió una alícuota del cultivo a medio YNB nuevo que contenía 2 % de arabinosa (Figura 6), cada vez a un valor de  $DO^{600}$  inicial de 0,2. Después de varias transferencias la cepa resultante se designó BIE104A2P1d (=BIE201).

### 10 **Ejemplo 3**

#### **Construcción de la cepa BIE201X9**

##### 3.1 Transformación de la cepa BIE201

Se transformó la cepa BIE201 con el plásmido pPWT042. El mapa físico del plásmido pPWT042 se expone en la Figura 7. El plásmido pPWT042 (SEQ ID NO: 14) contenía una inserción de 4630 pb que contenía un gen *xyIA* optimizado en el par de codones de *Bacteroides uniformis* bajo el control del promotor TP11 y el gen *XKS1* de *S. cerevisiae* bajo el control del promotor TDH1. Antes de la transformación de BIE201, se linealizó pPWT042 usando la enzima de restricción Sfil, según las instrucciones proporcionadas por el proveedor. Se sembraron las mezclas de transformación en YPD-agar (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 colonias sobre las placas, mientras que el control negativo (es decir, sin adición de ADN en el experimento de transformación) produjo placas YPD/G418 vacías.

Tras la digestión del plásmido pPWT042 con Sfil, su integración se dirige al locus SIT4 (Gottlin-Ninfa y Kaback (1986) Molecular and Cellular Biology Vol. 6, N.º 6, 2185-2197) en el genoma. Se caracterizaron transformantes usando PCR y técnicas de transferencia Southern. Se realizaron reacciones de PCR, usando ADN polimerasa Phusion® (Finnzymes), que son indicativas de la correcta integración de una copia de plásmido pPWT042, con los cebadores indicados por SEQ ID NO: 15 y 16 (integración correcta de pPWT042 en el locus SIT4) y SEQ ID NO: 16 y 17 (la banda solo aparece en caso de integración de múltiples copias). Una cepa con una copia de plásmido pPWT042 integrado en el genoma se designó BIE201Y9.

##### 3.2 Rescate de marcadores

Con el fin de eliminar los marcadores de selección introducidos por la integración del plásmido pPWT042, se realizó el siguiente procedimiento. El diseño del plásmido pPWT042 es tal que, tras la integración de pPWT042 en el cromosoma, secuencias homólogas están en estrecha proximidad entre sí. Este diseño permite que los marcadores de selección sean perdidos por recombinación intramolecular espontánea de estas regiones homólogas. Tras el crecimiento vegetativo, tendrá lugar la recombinación intramolecular, aunque a baja frecuencia. La frecuencia de esta recombinación depende de la extensión de la homología y el locus en el genoma (resultados no publicados). Tras la transferencia secuencial de una subfracción del cultivo a medio fresco, se acumularán recombinantes intramoleculares con el tiempo.

Para este fin, se cultivó la cepa BIE201Y9 en YPD-2 % de glucosa, a partir de una cepa aislada de colonias individuales. Se usaron 25 µl de un cultivo durante la noche para inocular medio YPD-2 % de glucosa fresco. Después de cinco transferencias en serie, se determinó la densidad óptica del cultivo y las células se diluyeron a una concentración de aproximadamente 5000 por ml. Se sembraron 100 µl de la suspensión de células sobre medio de carbono básico para levadura (Difco) que contenía KPi 30 mM (pH 6,8), 0,1 % de  $(NH_4)_2SO_4$ , fluoroacetamida 40 mM (Amersham) y 1,8 % de agar (Difco). Células idénticas a las células de la cepa BIE201Y9, es decir, sin recombinación intracelular, todavía contienen el gen *amdS*. Para aquellas células, la fluoroacetamida es tóxica. Estas células no serán capaces de crecer y no formarán colonias sobre un medio que contiene fluoroacetamida. Sin embargo, si ha ocurrido recombinación intramolecular, variantes de BIE201Y9 que han perdido los marcadores de selección serán capaces de crecer en el medio de fluoroacetamida, ya que son incapaces de convertir fluoroacetamida en compuestos inhibidores del crecimiento. Aquellas células formarán colonias sobre este medio de agar.

Las colonias resistentes a fluoroacetamida así obtenidas se sometieron a análisis de PCR usando cebadores de SEQ ID NO: 15 y SEQ ID NO: 16, y SEQ ID NO: 17 y SEQ ID NO: 18. Los cebadores de SEQ ID NO: 15 y SEQ ID NO: 16 darán una banda si la recombinación de los marcadores de selección ha tenido lugar como era previsto. Como resultado, los genes *xyIA* y *XKS1* han sido integrados en el locus *SIT4*. En ese caso, una reacción de PCR usando cebadores de SEQ ID NO: 17 y SEQ ID NO: 18 no debe producir un producto de PCR, ya que el cebador 17 sensibiliza en una región que debe ser perdida debido a la recombinación. Si se obtiene una banda con estos cebadores, esto es indicativo de la presencia del plásmido pPWT042 completo en el genoma, por lo que no ha tenido lugar recombinación.

Si los cebadores de SEQ ID NO: 15 y SEQ ID NO: 16 no producen un producto de PCR, la recombinación ha tenido lugar, pero de tal forma que el plásmido pPWT042 completo se ha eliminado del genoma por recombinación. No solo se han perdido los marcadores de selección, sino también el gen *xyIA* y *XKS1*. En realidad, se ha recuperado levadura no mutada.

- 5 Las cepas aisladas que presentaron los resultados de PCR esperados se sometieron a análisis de transferencia Southern (véase arriba). Una de las cepas que mostró el patrón correcto de bandas en la transferencia Southern es la cepa designada BIE201X9.

#### **Ejemplo 4**

##### **Construcción y selección de la cepa BIE252**

###### 10 **4.1 Amplificación del casete con *xyIA***

Con el fin de introducir copias adicionales del gen *xyIA* en el genoma, se realizó una reacción de PCR usando ADN polimerasa Phusion® (Finnzymes) con el plásmido pPWT042 como molde y los oligonucleótidos con SEQ ID NO: 19 y SEQ ID NO: 20 como cebadores. Con estos cebadores, se amplifica el casete con *xyIA*, que comprende el promotor *TPI1*, el gen *xyIA* optimizado en el par de codones y el terminador *PMA1*. El primer diseño es tal que los flancos del fragmento de PCR son homólogos a la secuencia consenso de las delta-secuencias del transposón de levadura Ty-1. Estas secuencias pueden obtenerse de NCBI (<http://www.ncbi.nlm.nih.gov/>) y alinearse usando un paquete de software que permite hacer esto como, por ejemplo, Clone Manager 9 Professional Edition (Scientific & Educational Software, Cary, EE.UU.).

15 El casete con *xyIA* no contiene un marcador de selección con el que pueda seleccionarse la integración en el genoma. Con el fin de estimar la frecuencia de transformación, se hizo una segunda transformación de control con el marcador *kanMX*. Para este fin, se amplificó el casete con *kanMX* del plásmido p427TEF (Dualsystems Biotech) en una reacción de PCR usando los cebadores correspondientes a SEQ ID NO: 21 y SEQ ID NO: 22.

###### 20 **4.2 Transformación de BIE201X9**

Se transformó BIE201X9 según el protocolo de electroporación (como se ha descrito anteriormente) con los fragmentos que comprenden ya sea 30 µg del casete con *xyIA* (designado X18-Ty1) o 10 µg del casete con *kanMX*. La mezcla de transformación con *kanMX* se sembró en YPD-agar (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 contenía 100 µg de G418 (Sigma Aldrich) por ml. Después de dos a cuatro días, aparecieron colonias sobre las placas, mientras que el control negativo (es decir, sin adición de ADN en el experimento de transformación) produjo placas YPD/G418 vacías. La frecuencia de transformación pareció ser superior a 600 colonias por µg de casete con *kanMX*.

Se usó la mezcla de transformación de X18-Ty1 para inocular un matraz oscilante que contenía 100 ml de medio de Verduyn, complementado con 2 % de xilosa. Como control, se usó el control negativo de la transformación (es decir, sin adición de ADN en el experimento de transformación). Los matraces oscilantes se incubaron a 30 °C y 280 rpm en un agitador orbital. El crecimiento fue seguido de medición de la densidad óptica a 600 nm regularmente.

35 El resultado de la curva de crecimiento se representa en la Figura 8. Como puede apreciarse, entre los días 20 y 25, la densidad óptica del matraz oscilante con X18-Ty1 aumentó espectacularmente, mientras que el crecimiento en el control negativo estuvo todavía ausente. En el día 25, se inoculó un matraz que contenía medio de Verduyn fresco complementado con 2 % de xilosa del cultivo X18-Ty1 a una densidad óptica inicial a 600 nm de 0,15. A partir de la Figura 8 es evidente que tras la re-inoculación el cultivo empezó a crecer sobre xilosa inmediatamente y rápidamente. Como es probable que el cultivo consista en una mezcla de subcultivos, consistiendo así en células con diferencias en el número de copias del casete X18-Ty1 y en la tasa de crecimiento sobre xilosa, se diluyeron células en agua MilliQ y se sembraron en placas de YPD-agar con el fin de conseguir cepas aisladas de colonias individuales. Las cepas aisladas de colonias individuales se probaron para su capacidad para utilizar diferentes fuentes de carbono.

###### 40 **4.3 Selección de la cepa BIE252**

Con el fin de seleccionar una cepa que hubiera adquirido crecimiento sobre xilosa como única fuente de carbono sin perder su capacidad para utilizar los otros azúcares importantes (glucosa, arabinosa y galactosa), se volvieron a cultivar en línea diez cepas aisladas de colonias individuales del cultivo de evolución adaptativa sobre YPD-agar. Posteriormente, se hizo un precultivo en medio YPD complementado con 2 % de glucosa. Los diez cultivos se incubaron durante la noche a 30 °C y 280 rpm. Se usaron alícuotas de cada cultivo para inocular medio de Verduyn fresco complementado con ya sea 2 % de glucosa, o 2 % de xilosa, o 2 % de arabinosa o 2 % de galactosa, a una densidad óptica inicial de 0,15. Como controles, se incluyeron cepas BIE201, BIE201X9 y la población mixta (de la que se recuperaron las diez cepas aisladas de colonias individuales) en el experimento. Se cultivaron células a 30 °C y 280 rpm en un agitador orbital. El crecimiento se evaluó basándose en las medidas de densidad óptica a 600 nm. Los resultados de la densidad óptica a 600 nm después de 19 horas de incubación se presentan en la Figura 9.

Los resultados de la Figura 9 muestran que tanto el cultivo mixto como las diez cepas aisladas de colonias individuales presentan una densidad óptica final más alta a 600 nm en xilosa. Aunque estas cepas crecieron sobre arabinosa como única fuente de carbono, la densidad óptica final a 600 nm fue menor en comparación con las cepas parentales, BIE201 y BIE201X9, 19 horas después de la inoculación. Después del crecimiento prolongado, durante varios días, se seleccionó la colonia 3 debido a que había alcanzado la densidad óptica más alta en comparación con las otras cepas aisladas de colonias individuales (datos no mostrados). El crecimiento en xilosa no fue diferente en todas las colonias probadas.

El cultivo de la colonia 3 se diluyó y se sembró en YPD-agar. Se probaron once cepas aisladas de colonias individuales para su capacidad para crecer en medio de Verduyn, complementado con ya sea 2 % de arabinosa o 2 % de xilosa. Para este fin, se hizo un precultivo en Verduyn con 2 % de glucosa. Las células se cultivaron durante la noche a 30 °C y 280 rpm en un agitador orbital. Se transfirió una alícuota a medio de Verduyn, complementado con ya sea 2 % de arabinosa o 2 % de xilosa, a una densidad óptica inicial de 0,15. El cultivo de colonia 3 se incluyó como control en esta prueba. Las células se cultivaron durante la noche a 30 °C y 280 rpm en un agitador orbital. El crecimiento se evaluó basándose en las mediciones de densidad óptica a 600 nm. Los resultados de la densidad óptica a 600 nm después de 4 días de incubación se presentan en la Figura 10.

Los resultados muestran que la colonia 3,7 ha alcanzado una densidad óptica similar al cultivo del control, la colonia 3, en xilosa. Además, su capacidad para crecer en arabinosa como única fuente de carbono ha mejorado ampliamente en comparación con el cultivo de la colonia 3. Las otras cepas aisladas de colonias individuales también muestran un crecimiento mejorado en arabinosa, aparentemente a costa de la capacidad para crecer en xilosa. El cultivo de la cepa aislada de colonias individuales 3,7 se designó la cepa BIE252. Se determinaron SNP y la amplificación en BIE252 siguiendo los mismos procedimientos que se describen en la solicitud de patente en tramitación junto con la presente PCT, presentada el 19 de abril de 2011, que reivindica prioridad del documento EP10160647.3. Esta prueba mostró que BIE252 tiene los SNP: G1363T en el gen *SSY1*, A512T en el gen *YJR154w*, A1186G en el gen *CEP3* y A436C en el gen *GAL80*. Éstos son los mismos que en BIE201. También la amplificación en el cromosoma VII que existe en BIE201 está presente en BIE252. Esto se muestra en la Figura 23.

### **Ejemplo 5**

#### **Prueba de rendimiento en BAM**

Con el fin de probar el rendimiento de la cepa BIE252 seleccionada, la cepa se inoculó en medio de Verduyn, complementado con 2 % de glucosa. Como controles, se incluyeron las cepas BIE201 y BIE104P1Y9mc. La primera es una cepa capaz de fermentar glucosa, galactosa y arabinosa (véase arriba), y la última es una cepa capaz de fermentar glucosa, galactosa y xilosa.

Después de la incubación durante la noche a 30 °C y 280 rpm en un agitador rotatorio, se recogieron las células por centrifugación y se realizaron cultivos para la producción de CO<sub>2</sub> a 33 °C en 100 ml de medio de Verduyn complementado con 5 % de glucosa, 5 % de xilosa, 3,5 % de arabinosa y 1 % de galactosa en BAM (monitor de la actividad biológica). En un experimento, el experimento discontinuo, las células se añadieron a 100 ml de medio de Verduyn complementado con los azúcares. En un segundo experimento, el experimento de lotes alimentados, los 100 ml de medio de Verduyn complementado con los azúcares se añadieron a las células a una tasa de 3 ml por minuto. La producción de CO<sub>2</sub> se monitorizó constantemente a intervalos, y se tomaron muestras para el análisis (densidad óptica a 600 nm, etanol, azúcares residuales). Las cepas BIE201 y BIE104P1Y9mc se mezclaron en una relación 1:1, basada en la densidad óptica, justo antes del experimento de BAM.

Los resultados del experimento de BAM se muestran en las Figuras 11, 12, 13, 14, 15 y 16.

El rendimiento de las cepas mixtas (BIE201 y BIE104P1Y9mc) en el modo de lotes alimentados (Figura 11) muestra que a pesar de que la xilosa y arabinosa se acumulan a un cierto grado con el tiempo, ya antes del final de la alimentación (a aproximadamente 33 h) la tasa de consumo es más rápida que la tasa de alimentación. Después de 72 h, toda la arabinosa se ha convertido, mientras que la xilosa no se ha consumido completamente. La glucosa y galactosa están siempre en o cerca del nivel de detección, que indica que estos azúcares son preferidos con respecto a las pentosas arabinosa y xilosa.

El rendimiento de la cepa BIE252, como se prueba en el modo de lotes alimentados en el BAM (véase Figura 12), muestra esencialmente las mismas características que el cultivo de cepas mixtas. Sin embargo, se acumulan xilosa y arabinosa a concentraciones más altas, pero la posterior conversión en etanol tiene lugar a una tasa más alta, que produce el mismo título de etanol y producción de CO<sub>2</sub> acumulada idéntica después de 120 h (Figura 13).

En el modo discontinuo, el rendimiento de las cepas mixtas (BIE201 y BIE104P1Y9mc) en el modo discontinuo muestra que la glucosa es fácilmente consumida después del comienzo del experimento de fermentación (Figura 14). Posteriormente, se cofermentan la galactosa, arabinosa y xilosa. Después de aproximadamente 30 h, se consumió la galactosa. Después de aproximadamente 72 horas, también se consumió completamente la arabinosa. Ambas contribuyeron significativamente a la formación de CO<sub>2</sub> y etanol. El consumo de xilosa, y así de etanol y la producción de CO<sub>2</sub>, se ralentizaron significativamente después del agotamiento de la arabinosa. La cepa BIE252 rindió mejor en ese respecto (Figura 15). Inmediatamente después del agotamiento de glucosa en el plazo de 10

5 horas, la galactosa, arabinosa y xilosa se cofermentaron rápidamente y eficientemente. Incluso después del agotamiento de galactosa, ambas pentosas arabinosa y xilosa se co-fermentaron completamente, contribuyendo ambas a la producción de CO<sub>2</sub> y formación de etanol. La fermentación de todos los azúcares estuvo completa después de aproximadamente 72 horas, en caso de la cepa BIE252. Esto produjo una producción de CO<sub>2</sub> acumulada más alta de la cepa BIE252 en comparación con el cultivo de cepas mixtas (BIE201 y BIE104P1Y9mc) como se expone en la Figura 16.

**Ejemplo 6**

Prueba de rendimiento en hidrolizados reales

10 Se realizó la prueba de rendimiento en hidrolizados reales usando la cepa BIE252 que se cultivó durante la noche en matraces oscilantes que contenían medio YEP complementado con 2 % de glucosa. Se recogieron las células por centrifugación y se resuspendieron a una concentración de 50 gramos de materia seca por litro.

Las materias primas que se probaron consistieron en lotes de fibra de maíz, tallos y hojas de maíz y paja de trigo. Se realizaron la hidrólisis y fermentación como se describe en la sección de materiales y métodos.

15 Los resultados se presentan en las Figuras 17 (fibra de maíz), 18 (tallos y hojas de maíz) y 19 (paja de trigo). En caso de tallos y hojas de maíz y paja de trigo, materias primas con una cantidad relativamente baja de galactosa y arabinosa, pero que principalmente consisten en glucosa y xilosa, todos los azúcares se convirtieron en 72 horas en CO<sub>2</sub> y etanol. En caso de fibra de maíz (Figura 17), queda una baja cantidad residual de arabinosa, galactosa y xilosa.

20 En las tablas a continuación se calculó el rendimiento de la fermentación, basándose en los azúcares liberados al final de la hidrólisis y la cantidad de etanol que se produjo al final de la fermentación.

Tabla 2: Azúcar total liberado (g/l), etanol producido (g/l) y rendimiento de etanol ( $\frac{g_{\text{etanol}}}{g_{\text{azúcar}}}$ ) de fermentación en BAM de BIE252, para diferente materia prima lignocelulósica

Materia prima lignocelulósica	Azúcar total* (g/l)	EtOH producido (g/l)	Rendimiento de EtOH ( $\frac{g_{\text{etanol}}}{g_{\text{azúcar}}}$ )
Tallos y hojas de maíz	45,4	22,7	0,50
Paja de trigo	42,7	20,7	0,48
Fibra de maíz	58,1	27,4	0,47
*(azúcar monomérica liberada al comienzo de la fermentación)			

25 Basándose en la cantidad de etanol producida al final de la fermentación, como se ha determinado por la medición del medidor de densidad Anton Paar DMA 5000, y la cantidad de materia prima pretratada que se estaba usando, se calcularon los rendimientos de la hidrólisis global y fermentación, por duplicado. Estas cifras se explican en la Tabla 3.

Tabla 3: Rendimiento global de la hidrólisis y fermentación (galones de etanol por tonelada de materia seca) de fermentación en BAM de BIE252, para diferente materia prima lignocelulósica

Materia prima lignocelulósica	Rendimiento global de la hidrólisis y fermentación (galones de etanol por tonelada de materia seca) 1ª fermentación	Rendimiento global de la hidrólisis y fermentación (galones de etanol por tonelada de materia seca) 2ª fermentación
Tallos y hojas de maíz	81	81
Paja de trigo	72	73
Fibra de maíz	67	69

30

**Ejemplo 7****Prueba de estabilidad de la cepa BIE252****7.1 Estabilidad de la cepa BIE252**

5 Con el fin de probar la estabilidad de la cepa BIE252, se usó una cepa aislada de colonias individuales para inocular 25 ml de YEP con 2 % de glucosa. La densidad óptica del cultivo se midió a 600 nm. El cultivo se incubó durante la noche a 30 °C y 280 rpm en un agitador rotatorio.

10 Después de la incubación durante la noche, se determinó la densidad óptica. Basándose en los valores de  $DO^{600}$  antes y después de la incubación, se calculó el número de generaciones hechas durante las incubaciones. Se usaron 25  $\mu$ l del cultivo durante la noche para inocular un matraz que contenía 25 ml de medio fresco. El cultivo se incubó otra vez bajo las mismas condiciones que se han descrito anteriormente. Este procedimiento se repitió hasta que se obtuvo un cultivo en el que las células estuvieron al menos cien generaciones separadas de la cepa aislada de colonias individuales inicial. Se eligió medio YEP complementado con 2 % de glucosa, debido a que bajo estas condiciones no se aplica presión de selección para mantener los genes introducidos en la cepa BIE252, necesaria para la conversión de arabinosa y xilosa.

15 Se diluyó el cultivo celular obtenido después de al menos 100 generaciones y se sembró en YPD-agar. Se inocularon dos cepas aisladas de colonias individuales en medio Verduyn con 2 % de glucosa. Como control, se incluyó la cepa parental BIE252. Los cultivos se incubaron durante la noche a 30 °C y 280 rpm en un agitador rotatorio. Se usó una alícuota del cultivo durante la noche para inocular un matraz que contenía 25 ml de medio de Verduyn complementado con 1 % de glucosa, 1 % de xilosa, 0,7 % de arabinosa, 0,3 % de galactosa y 0,1 % de manosa a una densidad óptica inicial de 0,15. A intervalos regulares, se tomaron muestras para el análisis: mediciones de densidad óptica a 600 nm, y para la determinación de azúcares residuales por RMN. Los resultados, como se representa en las Figuras 20a (cultivo de BIE252 antes de la prueba de estabilidad), 20b (cultivo de la colonia de BIE252 después de más de 100 generaciones) y 20c (cultivo de la segunda colonia de BIE252 después de más de 100 generaciones), muestran que el comportamiento de los tres cultivos de cepas es prácticamente idéntico en las condiciones probadas. En otras palabras, el rendimiento en este medio de crecimiento no ha cambiado; la cepa es todavía capaz de utilizar cinco azúcares diferentes con las mismas tasas después de cultivar durante más de 100 generaciones, que indica que la cepa es genéticamente estable.

25 Además, se hizo un experimento de PCR cuantitativa (Q-PCR) para evaluar el número de copias de BIE252 antes y después del cultivo de 100 generaciones en medio YEP complementado con 2 % de glucosa. El análisis de Q-PCR se realizó usando el sistema Bio-Rad iCycler iQ de Bio-Rad (Bio-Rad Laboratories, Hercules, CA, EE.UU.). Se usó iQ SYBR Green Supermix (Bio-Rad). Los experimentos se configuraron como se sugirió en el manual del proveedor.

30 Se evaluó la estabilidad de la cepa BIE252 determinando el número de copias del gen *xyIA* que codifica xilosa isomerasa. Como gen de una sola copia de referencia, se eligió el gen *ACT1*.

Los cebadores para la detección de los genes *xyIA* y *ACT1* son:

35            cebador directo para *xyIA*        SEQ ID NO 23  
                   cebador inverso para *xyIA*        SEQ ID NO 24  
                   cebador directo para *ACT1*        SEQ ID NO 25  
                   cebador inverso para *ACT1*        SEQ ID NO 26

40 Los resultados indicaron que el número de copias de BIE252 es aproximadamente 8 copias del gen *xyIA*, con respecto al gen *ACT1*. El número de copias se determinó antes y después del cultivo de 100 generaciones en medio YEP complementado con 2 % de glucosa, y pareció que era esencialmente el mismo, teniendo en cuenta las limitaciones del análisis de PCR cuantitativa.

45 Se aisló el ADN cromosómico de varias colonias, antes y después del cultivo de 100 generaciones en medio YEP complementado con 2 % de glucosa. Se corta el ADN cromosómico por XbaI. Se realiza un análisis de transferencia Southern usando el producto de PCR del par de cebadores SEQ ID NO: 23 y SEQ ID NO: 24 como sonda. El autorradiograma resultante está mostrando que el patrón de bandas no ha cambiado, que indica que la cepa es genéticamente estable después de 100 generaciones de crecimiento bajo condiciones no selectivas.

**Experimento comparativo A****7.2 Estabilidad de BIE201X9 transformada con vectores YEp y YCp**

50 Se transformó la cepa BIE201X9 con los vectores pPWT148 (un plásmido YCp) y pPWT152 (un plásmido YEp). El mapa físico del plásmido pPWT148 se expone en la Figura 21. El mapa físico del plásmido pPWT152 se expone en la Figura 22. Se sembraron mezclas de transformación en YPD-agar (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 contenía 100 µg de G418 (Sigma Aldrich) por ml. Después de dos a cuatro días, aparecieron colonias sobre las placas, mientras que el control negativo (es decir, sin adición de ADN en el experimento de transformación) produjo placas YPD/G418 vacías.

5 Se usaron tres colonias de cada transformación como una mezcla para probar la estabilidad. Para este fin, se usaron las colonias mixtas para inocular medio YEP, complementado con 2 % de glucosa y 100 µg de G418/ml. Las células se incubaron durante la noche a 30 °C y 280 rpm en un agitador rotatorio. La densidad óptica se midió a 600 nm. El cultivo durante la noche se usó para inocular medio YEP nuevo, complementado con 2 % de glucosa, y el mismo medio complementado con 100 µg de G418/ml, a una densidad óptica inicial a 600 nm de 0,02. Basándose en lo que se informa en la bibliografía (véase arriba), la esperanza es que las células que se cultivaron en presencia de G418 (es decir, se aplica presión selectiva), la mayoría de las células retendrá el plásmido dentro de la célula. Sin embargo, en ausencia de G418, se espera que las células pierdan su plásmido rápidamente.

10 Se incubaron los cultivos durante la noche a 30 °C y 280 rpm en un agitador rotatorio. La densidad óptica se midió a 600 nm. Basándose en las mediciones de densidad óptica, se calculó que el cultivo había hecho 8 generaciones en ambos medios (es decir, con y sin G418). Se hicieron diluciones y se sembraron en placas de YPD con y sin 100 µg de G418/ml. Las placas se incubaron a 30 °C durante aproximadamente dos días, o más, según se necesite para producir colonias visibles y contables.

15 Además, se usaron 25 µl del cultivo durante la noche para inocular 25 ml de medio YEP fresco, complementado con 2 % de glucosa, y el mismo medio complementado con 100 µg de G418/ml. Los cultivos se incubaron durante la noche a 30 °C y 280 rpm en un agitador rotatorio. La densidad óptica se midió a 600 nm. Basándose en las mediciones de densidad óptica, se calculó que el cultivo había hecho 8 generaciones adicionales en ambos medios (es decir, con y sin G418). Se hicieron diluciones y se sembraron en placas de YPD con y sin 100 µg de G418/ml. Las placas se incubaron a 30 °C durante aproximadamente dos días, o más, según se necesite para producir colonias visibles y contables.

20 Basándose en las colonias que se contaron en YDP-agar con y sin G418, la estabilidad podría calcularse después de 8 y 16 generaciones en presencia o ausencia de G418. Los resultados se resumen en la Tabla 4.

Tabla 4: Estabilidad de las cepas del experimento comparativo A (% de clones resistentes a G418)

	BIE201X9-pPWT148	BIE201X9-pPWT152
% de clones resistentes a G418 después de 8 generaciones de crecimiento en YEP con 2 % de glucosa	4	12
% de clones resistentes a G418 después de 16 generaciones de crecimiento en YEP con 2 % de glucosa	1	5
% de clones resistentes a G418 después de 8 generaciones de crecimiento en YEP con 2 % de glucosa y 100 µg de G418/ml	9	33
% de clones resistentes a G418 después de 16 generaciones de crecimiento en YEP con 2 % de glucosa y 100 µg de G418/ml	28	40

30 A partir de la tabla, puede observarse que tanto los plásmidos YCp como YEp (es decir, pPWT148 y pPWT152, respectivamente) no son establemente mantenidos en las células de levadura transformadas, aunque la presión de selección está siendo aplicada durante el cultivo, en este caso mediante la adición de G418 al medio de cultivo. Debe observarse, sin embargo, que la pérdida de plásmido es mucho más severa en ausencia de presión de selección.

Este experimento en combinación con el Ejemplo 7 proporciona el conocimiento de que la integración cromosómica de los genes de xilosa isomerasa en el genoma es necesaria para obtener transformantes de levadura estables.

### 35 **Ejemplo 8**

#### **La delección de GAL80 conduce a una mejor conversión de arabinosa**

En el Ejemplo 4 se mostró que el SNP identificado en el gen *GAL80* tiene un efecto aditivo positivo sobre el crecimiento en arabinosa, si también está presente la amplificación de una parte del cromosoma VII.

40 *GAL80* codifica un represor transcripcional implicado en la regulación transcripcional en respuesta a galactosa (Timson DJ, et al. (2002) *Biochem J* 363(Pt 3):515-20). Conjuntamente con Gal4p y Gal3p, Gal80p regula coordinadamente la expresión de genes que contienen un sitio de activación GAL en la dirección 5' en su promotor (UAS-GAL), que incluye los genes metabólicos de *GAL GAL1*, *GAL10*, *GAL2* y *GAL7* (revisado en Lohr D, et al.

(1995) FASEB J 9(9):777-87). Células nulas para gal80 expresan constitutivamente genes *GAL*, incluso en medios no inductores (Torchia TE, et al. (1984) Mol Cell Biol 4(8):1521-7).

La hipótesis es que el SNP que se identificó en el gen *GAL80* influye en la interacción entre Gal80p, Gal3p y Gal4p. De ahí que la expresión de los genes metabólicos de galactosa, que incluyen galactosa permeasa que codifica *GAL2*, se cambie también en comparación con una célula de levadura con un alelo *GAL80* no mutado. Gal2p (galactosa permeasa) es el principal transportador de azúcar para arabinosa (Kou et al (1970) J Bacteriol. 103(3):671-678; Becker y Boles (2003) Appl Environ Microbiol. 69(7): 4144-4150).

Aparentemente, el SNP en el gen *GAL80* tiene un efecto positivo sobre la capacidad de convertir L-arabinosa. Con el fin de investigar si el fenotipo de crecimiento de arabinosa podría mejorarse adicionalmente, se deletó la secuencia codificante del gen *GAL80* en su totalidad, usando una estrategia de sustitución génica mediada por PCR.

### 8.1 Alteración del gen *GAL80*

Se usaron cebadores de SEQ ID NO 58 y 59 (los cebadores directos e inversos, respectivamente) para la amplificación del marcador kanMX del plásmido p427-TEF (Dualsystems Biotech, Schlieren, Suiza). Los flancos de los cebadores son homólogos a la región 5' y la región 3 del gen *GAL80*. Tras la recombinación homóloga, el ORF del gen *GAL80* se sustituirá por el marcador kanMX, similar a como se describe por Wach (Wach et al (1994) Yeast 10, 1793-1808). El fragmento obtenido se designa el fragmento *GAL80::kanMX*.

Se hizo una transformación de levadura de la cepa BIE252 con el fragmento *GAL80::kanMX* purificado según el protocolo descrito por Gietz y Woods (2002), Methods in Enzymology 350: 87-96). La construcción de la cepa BIE252 se ha descrito en el documento EP10160622.6. La cepa BIE252 es una cepa fermentadora de xilosa y arabinosa de *S. cerevisiae*, que es un derivado de BIE201. La cepa BIE252 también contenía el SNP de *GAL80*.

Se sembraron las células transformadas en YEPD-agar que contenía 100 µg/ml de G418 para la selección. Las placas se incubaron a 30 °C hasta que las colonias fueron visibles. Se incluyó el plásmido p427-TEF como control positivo y dio muchas colonias. Se incluyó MilliQ (es decir, sin ADN) como control negativo y no dio colonias. El fragmento *GAL80::kanMX* dio muchas colonias. Se probaron dos colonias independientes por transferencia Southern con el fin de verificar la correcta integración (datos no mostrados). Una colonia con la delección correcta del gen *GAL80* se designó BIE252Δ*GAL80*.

### 8.2 Efecto de la sustitución del gen *GAL80* sobre el rendimiento en BAM

Se realizó un experimento de BAM (monitor de la actividad biológica; Halotec BV, Veenendaal, Los Países Bajos). Se usaron cepas aisladas de colonias individuales de la cepa BIE252 y cepa BIE252Δ*GAL80* (un transformante en el que el ORF del gen *GAL80* se sustituyó correctamente por el marcador kanMX) para inocular medio de Verduyn (Verduyn et al., Yeast 8:501-517, 1992) complementado con 2 % de glucosa. Los precultivos se incubaron durante aproximadamente 24 horas a 30 °C y 280 rpm. Se recogieron las células y se inocularon en un medio modelo sintético (medio de Verduyn complementado con 5 % de glucosa, 5 % de xilosa, 3,5 % de arabinosa, 1 % de galactosa y 0,5 % manosa, pH 4,2) a una densidad celular de aproximadamente 1 gramo de peso seco por kg de medio. Se monitorizó constantemente la producción de CO<sub>2</sub>. Se analizó la conversión de azúcar y la formación de producto por RMN. Los datos representan la cantidad residual de azúcares en los momentos de tiempo indicados (glucosa, arabinosa, galactosa, manosa y xilosa en gramos por litro) y la formación de (sub)productos (etanol, glicerol). Se monitorizó el crecimiento siguiendo la densidad óptica del cultivo a 600nm. El experimento se realizó durante aproximadamente 72 horas.

Los gráficos se presentan en la Figura 23 (BIE252) y 24 (BIE252Δ*GAL80*).

Los experimentos muestran claramente que la cepa BIE252 de referencia convirtió glucosa y manosa rápidamente. Después del agotamiento de glucosa (aproximadamente 10 horas), comenzó la conversión de xilosa y arabinosa. Algo de galactosa ya estaba siendo fermentado aproximadamente en el momento de tiempo de 10 horas, que podría ser debido al SNP de *GAL80* en esta cepa, que permitiría la utilización simultánea (parcial) de glucosa y galactosa. Al final del experimento, aproximadamente 72 horas, se habían convertido casi todos los azúcares (aproximadamente el 99 %). Se obtuvo un rendimiento de etanol de 0,37 gramos de etanol por gramo de azúcar.

La cepa BIE252Δ*GAL80* presenta capacidad de conversión de azúcar más rápida que la cepa BIE252. También en el caso de esta cepa, se convierten manosa y glucosa en las primeras horas de la fermentación. Sin embargo, a diferencia de la cepa BIE252, en este transformante hay algo de co-consumo de glucosa, galactosa y manosa con arabinosa y especialmente xilosa. En general, el consumo de azúcar es más rápido, conduciendo a un uso más completo de todos los azúcares disponibles. Esto también es evidente del desprendimiento de CO<sub>2</sub> con el tiempo. En el caso de BIE252, se observa un primer pico, que es básicamente el CO<sub>2</sub> formado a partir de glucosa y manosa. Después de alcanzar un mínimo de justo por encima de 10 ml/h (Figura 20), se observa un segundo pico más plano. En caso de BIE252Δ*GAL80*, sin embargo (Figura 21), el segundo pico aparece como una cola del primer pico, debido a un co-uso intensificado de glucosa, xilosa, arabinosa, manosa y galactosa, como es evidente del análisis de azúcares por RMN. En la cepa parental BIE252, el uso de los diferentes azúcares es más secuencial. De ahí que el

rendimiento de la cepa BIE252ΔGAL80 sea más alto al final del experimento (72 h): 0,40 gramos de etanol por gramo de azúcar.

En conclusión, la delección del ORF del gen *GAL80* produjo un rendimiento mejorado adicional, como se probó en la cepa BIE252.

5 **LISTADO DE SECUENCIAS**

<110> DSM IP Assets B.V.

<120> Célula de azúcares mixtos

10 <130> 27758-WO-PCT

<160> 26

<170> PatentIn versión 3.5

15

<210> 1

<211> 18215

<212> ADN

<213> Secuencia artificial

20

<220>

<223> Plásmido sintético

<400> 1

ES 2 651 076 T3

ggccaagatg gccgatctgc atttttcata ataatcctcg gtactttcta caagatcaat	60
taaattccaa tcaaaaatcg tcttttgcaa gattttgaag tcacagtact tttcattttc	120
aatgtcaaca ggcgccatt tgtattgtct tcctttaact ttttcgcct tttcattaaa	180
aatgtactca ttagatgcaa ttatactgaa tggatatttt tgaaaaatat cttgtgttgc	240
attcaaaact tcatcgccga aaaagaaaca tacagggata tcttgtactc ttattatttc	300
tctaacttgt gttttgaagt ttttcaattc ctctttcgtt agcaaatctg atttagcaat	360
aaccgggatt aaattcactc tcttcgctaa ttttttcatt gttacgacgt ctaaagtatc	420
aattccctta tttgaaggtc tcagaaagta caaacaacaa tggactctat tatcaacat	480
ttttgtccta tcaggttgtt cttcttgaa aatgtacgat cttatttctt catcaatata	540
gtttctagac tgcagcccg gatccgtcga caagcttgtg gagaggtgac ttcatgaacc	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
ttctgctggt taaataagta ttcaactact attgaagccg actcatgaag ccggttacgg	900
acaaaaccgg gcaaatttcg 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
taaaggattt tcgcgaggc ggccttcaat catcgaatac tacgtcttaa tatgatgtac	1200
tgtggttcat attttcaagt agtgtagta aatttgtata cgttcatgta agtgtgtatc	1260
ttgagtgtct gtatgggccc ataaacgtaa gcgagacttc caaatggagc aaacgagaag	1320

ES 2 651 076 T3

agatctttaa agtattatag aagagctggg caggaactat tatgacgtaa agccttgacc 1380  
 ataataaaga cgattctttg tccctctata caaacatctt gcaaagatac caaatatfff 1440  
 caaatcctac tcaataaaaa attaatgaat aaattagtgt gtgtgcatta tatatatata 1500  
 aaattaagaa ttagactaaa taaagtgttt ctaaaaaaat attaaagttg aaatgtgcgt 1560  
 gttgtgaatt gtgctctatt agaataatta tgacttgtgt gcgtttcata ttttaaaata 1620  
 ggaaataacc aagaagaaa 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 cgtagggtgc tgggtgaaca gtttattcct ggcacccact 1920  
 aaatataatg gagcccgctt tttagctgg catccagaaa aaaaaagaat cccagcacca 1980  
 aaatattggt ttcttcacca accatcagtt cataggtcca ttctcttagc gcaactacag 2040  
 agaacagggg cacaaacagg caaaaaacgg gcacaacctc aatggagtga tgcaacctgc 2100  
 ctggagtaaa tgatgacaca aggcaattga cccacgcag tatctatctc attttcttac 2160  
 accttctatt accttctgct ctctctgatt tggaaaaagc tgaaaaaaa ggttgaacc 2220  
 agttccctga aattattccc ctacttgact aataagtata taaagacggt aggtattgat 2280  
 tgtaattctg taaatctatt tcttaactt cttaaattct acttttatag ttagtctttt 2340  
 ttttagtttt aaaacaccaa gaacttagtt tcgaataaac acacataaac aaacaaatg 2400  
 ttatcagtac ctgattatga gttttggttt gttaccggtt cacaacacct ttatggtgaa 2460  
 gaacaattga agtctgttgc taaggatgcg caagatattg cggataaatt gaatgcaagc 2520  
 ggcaagttac cttataaagt agtctttaag gatggtatga cgacggctga aagtatcacc 2580  
 aactttatga aagaagttaa ttacaatgat aaggtagccg gtgttattac ttggatgcac 2640  
 acattctcac cagctaagaa ctggattcgt ggaactgaac tgttacaaaa accattatta 2700  
 cacttagcaa cgcaatattt gaataatatt ccatatgcag acattgactt tgattacatg 2760  
 aaccttaacc aaagtgccca tggcgaccgc gagtatgcct acattaacgc ccggttgca 2820  
 aaacataata agattgttta cggctattgg ggcgatgaag atgtgcaaga gcagattgca 2880  
 cgttgggaag acgtcgccgt agcgtacaat gagagcttta aagttaaggt tgctcgcttt 2940  
 ggcgacacaa tgcgtaatgt ggcggttact gaaggtgaca aggttgaggc tcaaatgaa 3000  
 atgggctgga cagttgacta ttatggtatc ggtgacttag ttgaagagat caataaggtt 3060  
 toggatgctg atgttgataa ggaatacgcct gacttgaggt ctoggtatga aatggtccaa 3120  
 ggtgataacg atgcggacac gtataaacat tcagttcggg ttcaattggc acaatatctg 3180  
 ggtattaagc ggttcttaga aagaggcggg tacacagcct ttaccacgaa ctttgaagat 3240  
 ctttggggga tggagcaatt acctggtcta gcttcacaat tattaattcg tgatgggtat 3300

ES 2 651 076 T3

ggttttgggtg ctgaaggtga ctggaagacg gctgcttttag gacgggttat gaagattatg 3360  
 tctcacaaca agcaaaccgc ctttatggaa gactacacgt tagacttgcg tcatgggtcat 3420  
 gaagcgatct taggttcaca catgttggaa gttgatccgt ctatcgcaag tgataaacca 3480  
 cgggtcgaag ttcattccatt ggatattggg ggtaaagatg atcctgctcg cctagtattt 3540  
 actggttcag aaggtgaagc aattgatgtc accggtgccc atttccgtga tgggttcaag 3600  
 atgattagct acgcggtaga tgcgaataag ccagaagccc aaacacctaa tttaccagtt 3660  
 gctaagcaat tatggacccc aaagatgggc ttaaagaaaag gtgcactaga atggatgcaa 3720  
 gctgggtggg gtcaccacac gatgctgtcc ttctcgttaa ctgaagaaca aatggaagac 3780  
 tatgcaacca tggttggcat gactaaggca ttcttaaagt aagtgaattt actttaaatc 3840  
 ttgcatttaa ataaattttc tttttatagc tttatgactt agtttcaatt tataactat 3900  
 tttaatgaca ttttcgattc attgattgaa agctttgtgt tttttcttga tgcgctattg 3960  
 cattgttctt gtctttttcg ccacatgtaa tatctgtagt agatacctga tacattgtgg 4020  
 atgctgagtg aaattttagt taataatgga ggcgctctta ataatttttg ggatattggc 4080  
 tttttttttt aaagtttaca aatgaatttt ttcgccagg atcgtacgcc gcggaaccgc 4140  
 cagatattca ttacttgacg caaaagcgtt tgaataatg acgaaaaaga aggaagaaaa 4200  
 aaaaagaaaa ataccgcttc taggcgggtt atctactgat ccgagcttcc actaggatag 4260  
 cacccaaaaca cctgcatatt tggacgacct ttacttacac caccaaaaac cactttcgcc 4320  
 totcccgcc ctgataacgt ccaactaattg agcgattacc tgagcgggtcc tcttttgttt 4380  
 gcagcatgag acttgcatc tgc aaatcgt aagtagcaac gtctcaaggt caaaactgta 4440  
 tggaaacctt gtcacctcac ttaattctag ctagcctacc ctgcaagtca agaggctccc 4500  
 gtgattccta gccacctcaa ggtatgcctc tccccggaaa ctgtggcctt ttctggcaca 4560  
 catgatctcc acgatttcaa catataaata gcttttgata atggcaatat taatcaaatt 4620  
 tattttactt ctttcttga acatctctct tgtaatccct ttttcttct agctattttt 4680  
 cataaaaaac caagcaactg cttatcaaca cacaaacct aaatcaaat gaatttagtt 4740  
 gaaacagccc aagcgattaa aactggcaaa gtttcttttag gaattgagct tggctcaact 4800  
 cgaattaaag ccgttttgat cacggacgat ttaatacga ttgcttcggg aagttacgtt 4860  
 tgggaaaaac aatttgttga tggacttgg acttacgcac ttgaagatgt ctggaccgga 4920  
 attcaacaaa gttatacga attagcagca gatgtccgca gtaaataca catgagtttg 4980  
 aagcatatca atgctattgg cattagtgcc atgatgcacg gatacctagc atttgatcaa 5040  
 caagcgaat tattagttcc gtttcggact tggcgtaata acattacggg gcaagcagca 5100  
 gatgaattga ccgaattatt tgatttcaac attccacaac ggtggagtat cgcacactta 5160  
 taccaggcaa tcttaataa tgaagcgcac gttaaacagg tggacttcat aacaacgctg 5220

ES 2 651 076 T3

gctggetatg taacctggaa attgtcgggt gagaaagttc taggaatcgg tgatgcgtct 5280  
ggcgttttcc caattgatga aacgactgac acatacaatc agacgatggt aaccaagttt 5340  
agccaacttg acaaagttaa accgtattca tgggatatcc ggcataatctt accgcggggt 5400  
ttaccagcgg gagccattgc tggaaagtta acgctgccc gggcgagctt acttgatcag 5460  
agcggcacgc tgcacgctgg cagtgttatt gcaccgccag aaggggatgc tggaacagga 5520  
atggtcggta cgaacagcgt ccgtaaacgc acgggtaaca tctcgggtggg aacctcagca 5580  
ttttcogatga acgttctaga taaaccattg tctaaagtct atcgcgatat tgatattggt 5640  
atgacgccag atgggtcacc agttgcaatg gtgcatgtta ataattgttc atcagatatt 5700  
aatgcgtggg caacgatttt tcatgagttt gcagcccggg tgggaatgga attgaaaccg 5760  
gatcgattat atgaaacggt attcctggaa tcaactcgcg ctgatgcgga tgctggaggg 5820  
ttggctaatt atagttatca atccggtgag aatattacta agattcaagc tggtcggccg 5880  
ctatttgtac ggacacccaa cagtaaattt agttaccga actttatggt gactcaatta 5940  
tatgcggcgt tgcacccct ccaacttggg atggatattc ttgttaacga agaacatggt 6000  
caaacggacg ttatgattgc acagggtgga ttgttccgaa cgcgggtaat tggccaacaa 6060  
gtattggcca acgcactgaa cattccgatt actgtaatga gtactgctgg tgaaggcggc 6120  
coatggggga tggcagtgtt agccaacttt gcttgcggc aaactgcaat gaacctagaa 6180  
gatttcttag atcaagaagt ctttaaagag ccagaaagta tgacgttgag tccagaaccg 6240  
gaacgggtgg ccgatatcg tgaatttatt caacgttatc aagctggctt accagtgaa 6300  
gcagcggctg ggcaagcaat caaatattag agcttttgat taagccttct agtccaaaa 6360  
acacgttttt ttgtcattta tttcattttc ttagaatagt ttagtttatt cattttatag 6420  
tcacgaatgt tttatgattc tatatagggt tgcaaacaaag catttttcat tttatgttaa 6480  
aacaatttca ggtttacctt ttattctgct tgtggtgacg cgggtatccg cccgctcttt 6540  
tggtcaccca tgtatttaat tgcataaata attcttaaaa gtggagctag tctatttcta 6600  
tttacatacc tctcatttct catttcctcc actagtagag aattttgcca tcggacatgc 6660  
taccttacgc ttatatctct cattggaata tcgttttctg attaaaacac ggaagtaaga 6720  
acttaattcg tttttcgttg aactatggtg tgccagcgtg acattaaaaa agagtgtaca 6780  
aggccaogtt ctgtcaccgt cagaaaaata tgtcaatgag gcaagaaccg ggatggtaac 6840  
aaaaatcacg atctgggtgg gtgtgggtgt attggattat aggaagccac gcgctcaacc 6900  
tgggaattaca ggaagctggt aattttttg gtttgcaatc atcaccatct gcacgttgtt 6960  
ataatgtccc gtgtctatat atatocattg acggtattct atttttttgc tattgaaatg 7020  
agcgtttttt gttactacaa ttggttttac agacggaatt ttccctattt gtttcgccc 7080  
atttttcctt ttctcattgt tctcatatct taaaaaggtc ctttcttcat aatcaatgct 7140  
ttcttttact taatatttta cttgcattca gtgaatttta atacatattc ctctagtctt 7200

ES 2 651 076 T3

gcaaaatcga tttagaatca agataccagc ctaaaaatgc tagaagcatt aaaacaagaa 7260  
gtttatgagg ctaacatgca gttccaaaag ctgggcctgg ttacttttac ctggggcaat 7320  
gtctcgggca ttgaccggga aaaaggccta ttcgtgatca agccatctgg tgttgattat 7380  
ggtgaattaa aaccaagcga tttagtcggt gttaacttac aggtgaagt ggttgaaggt 7440  
aaactaaatc cgtctagtga tacgccgact catacgggtg tatataacgc ttttcctaata 7500  
attggcggaa ttgtccatac tcattcgcca tgggcagttg cctatgcagc tgctcaaata 7560  
gatgtgccag ctatgaacac gaccatgct gatacgttct atggtgacgt gccggccgcg 7620  
gatgcgctga ctaaggaaga aattgaagca gattatgaag gcaacacggg taaaaccatt 7680  
gtgaagacgt tccaagaacg gggcctcgat tatgaagctg taccagcctc attagtcagc 7740  
cagcacggcc catttgcttg gggaccaacg ccagctaaag ccgtttaca tgctaaagt 7800  
ttggaagtgg ttgccgaaga agattatcat actgcgcaat tgaccctgac aagtagcgaa 7860  
ttaccacaat atttattaga taagcattat ttacgtaagc atggtgcaag tgcctattat 7920  
ggtcaaaata atgcgcattc taaggatcat gcagttcgca agtaaacaaa tcgctcttaa 7980  
atatatacct aaagaacatt aaagctatat tataagcaaa gatacgtaaa ttttgcttat 8040  
attattatac acatatcata tttctatatt ttaagattt ggttatataa tgtacgtaat 8100  
gcaaaggaaa taaattttat acattattga acagcgtcca agtaactaca ttatgtgcac 8160  
taatagttaa gcgtcgtgaa gactttattg tgtcgcgaaa agtaaaaatt ttaaaaatta 8220  
gagcaccttg aacttgcgaa aaaggttctc atcaactgtt taaaaacgag tgtcttctgt 8280  
gtttcagttc agggcttttc ggaggatgtg aatcgacggc gtactgtcct tgggaacttt 8340  
gtctacgtat tttcacttcc tcagcgaatc cagagactat cttgggaaat tcgacaggac 8400  
agtctgttga caaccgactc ctttttgact tcataataaa aattcaatga cgcaaaagga 8460  
atthtaggtt tttattattt atthattat ttctgttaat tgatcctttt ctttccacta 8520  
ccaacaacaa aaaagggggg aaaaagatgt ataactaaa agacactaat ctgctcttga 8580  
tatacttatt atgtaatgga ataactcata taaatgtaaa atagaacttc aaattaatat 8640  
tataatgata gtcgaggtca gacacactta taatacatta agtaaaagaaa aaaaaatgtc 8700  
tgtcatcgag gtctcttttg tgtcgtctaac aaaacatcac taaatacgaa gacactttgc 8760  
atgggaagga tgcagcaaat ggcaaaactaa cgggccattg attggtttac ctcttctatt 8820  
tgtattacga ccagaaagaa cgaatggttt tcataatga ggtaggaaac gacctaaata 8880  
taatgtagca tagataaaat ctttgtactg tatggttgca atgccttctt gattagtatc 8940  
gaatttcctg aataattttg ttaatctcat tagccaaact aacgcctcaa cgaatttatc 9000  
aaactttagt tcttttctg ttocatttct gttataaac tcagcatatt ggtcaaatgt 9060  
tttctcgcta acttcaaaag gtattagata tcctagttct tgaagtgagt tatgaaattc 9120

ES 2 651 076 T3

gcttacagaa atggtgagcg atccgttgat atcattgtcc acataaaactt ttctccaact 9180  
tttcaactctt ttgtataggg cgatgaatc tgccctggtg acagtgccaa acctggaagc 9240  
accaaataaa tttatcagcg catctactga tgatatacaa aaatgggagt tgcctgctgtt 9300  
ttgtagtaag ttctgtagtt cctcagctgt cagtcggttt ttgcccttta catcatggtt 9360  
atgaaatagc tgtgtggcca cttgcatgtc tcgtacatct tctctgctat cgaacgaagc 9420  
aggtgcaact ttcttcaaga gttgtgcagc cactgcttga ttgtgaatta ggggaggagg 9480  
agaggaagct atccgttgag cggaaagtgt caagttgtta taatggggtg gcgctggagg 9540  
tataggcctg cctgctggtt tctgtgcgat aacattatat ctaggatcca cagggtgtttt 9600  
cgtatgtctt ggagaataac tttggggaga accataggag tggtgaccgt tttctgctct 9660  
gtttttgtta tattgagttt gtaaggaat tggagctgag tggactctag tgttgggagt 9720  
ttgtgcttga gtaaccggta ccacggctcc tcgctgcaga cctgcgagca gggaaaacgct 9780  
cccctcacag tcgcttgaa ttgtcccccac gccgcgcccc thtagagaaa tataaaaggt 9840  
taggatttgc cactgaggtt cttctttcat ataactcctt ttaaaatctt gctaggatac 9900  
agttctcaca tcacatccga acataaacia ccatgggtaa ggaaaagact caagtttoga 9960  
ggccgcgatt aaattccaac atggatgctg atttatatgg gtataaatgg gctcgcgata 10020  
atgtcgggca atcaggtgag acaatctatc gattgtatgg gaagcccgat gcgccagagt 10080  
tgtttctgaa acatggcaaa ggtagcgttg ccaatgatgt tacagatgag atggtcagac 10140  
taaactggct gacggaattt atgcctcttc cgaccatcaa gcattttatc cgtactcctg 10200  
atgatgcatg gttactcacc actgcgatcc ccggcaaaac agcattccag gtattagaag 10260  
aatatcctga ttcaggtgaa aatattgttg atgogctggc agtgttcctg ccgccggttc 10320  
attecattcc tgtttgaat tgcctttta acagcgatcg cgtatttctg ctgcctcagg 10380  
cgcaatcacg aatgaataac ggtttggttg atgcgagtga tttgatgac gagcgtaatg 10440  
gctggcctgt tgaacaagtc tggaaagaaa tgcataagct tttgccattc tcaccggatt 10500  
cagtcgtcac tcatggtgat ttctcacttg ataaccttat ttttgacgag gggaaattaa 10560  
taggttgtat tgatgttga cgagtcggaa tcgcagaccg ataaccaggat cttgccatcc 10620  
tatggaactg cctcgggtgag tttctcctt cattacagaa acggcttttt caaaaatatg 10680  
gtattgataa tcctgatatg aataaattgc agtttcattt gatgctcgat gagttttctt 10740  
aatcagctact gacaataaaa agattcttgt tttcaagaac ttgtcatttg tatagttttt 10800  
ttatattgta gttgttctat tttaatcaaa tgttagcgtg atttatattt tttttcgctt 10860  
cgacatcatc tgcccagatg cgaagttaag tgcgcagaaa gtaatatcat gcgtcaatcg 10920  
tatgtgaatg ctggtcgcta tactgctgtc gattogatac taacgccgcc atccagggta 10980  
ccatcctttt gttgtttccg ggtgtacaat atggacttcc tcttttctgg caaccaaac 11040  
catacatcgg gattcctata ataccttctg tggctcctt aacatgtagg tggcggaggg 11100

ES 2 651 076 T3

gagatataca atagaacaga taccagacaa gacataatgg gctaaacaag actacaccaa 11160  
ttacactgcc tcattgatgg tggtagataa cgaactaata ctgtagccct agacttgata 11220  
gccatcatca tatcgaagtt tcactaccct ttttcattt gccatctatt gaagtaataa 11280  
taggcgcatg caacttcttt tctttttttt tcttttctct ctccccggtt gttgtctcac 11340  
catatccgca atgacaaaaa aatgatgga agacactaaa ggaaaaaatt aacgacaaa 11400  
acagcaccaa cagatgtcgt tgttccagag ctgatgaggg gtatcttoga acacacgaaa 11460  
ctttttcctt ccttcattca cgcacactac tctctaataa gcaacgggat acggccttcc 11520  
ttccagttac ttgaatttga aataaaaaaa gtttgcgct ttgctatcaa gtataaatag 11580  
acctgcaatt attaactctt tgtttcctcg tcattgttct cgttcccttt cttccttgtt 11640  
tctttttctg cacaatattt caagctatac caagcataca atcaactatc tcatatacaa 11700  
tgcctcaatc ctgggaagaa ctggccgctg ataagcgcgc ccgctcgcga aaaaccatcc 11760  
ctgatgaatg gaaagtccag acgctgcctg cgaagacag cgttattgat ttcccaaaga 11820  
aatcggggat cctttcagag gccgaactga agatcacaga ggctccgct gcagatcttg 11880  
tgtccaagct ggcggccgga gaggtagcct cggtggaagt tacgctagca ttctgtaaac 11940  
gggcagcaat cgcacagcag ttaacaaaact gcgcccacga gttcttccct gacgcccctc 12000  
tcgcgaggc aagggaaact gatgaatact acgcaaagca caagagacc gttggtccac 12060  
tccatggcct ccccatctct ctcaaagacc agcttcgagt caagggctac gaaacatcaa 12120  
tgggctacat ctcatggcta aacaagtacg acgaagggga ctcggttctg acaaccatgc 12180  
tccgcaaagc cggtagcctc ttctacgtca agacctctgt cccgcagacc ctgatggtct 12240  
gcgagacagt caacaacatc atcgggcgca ccgtcaacc acgcaacaag aactggctgt 12300  
gcggcgccag ttctgggtgt gagggtgcga tcgttgggat tcgtgggtgc gtcacgggtg 12360  
taggaacgga tatcgggtgc tcgattcgag tgcggccgc gttcaacttc ctgtacggtc 12420  
taaggccgag tcattggcgg ctgccgtatg caaagatggc gaacagcatg gagggtcagg 12480  
agacggtgca cagcgttgtc gggccgatta cgcactctgt tgaggacctc cgcctcttca 12540  
ccaaatccgt cctcggtcag gagccatgga aatacgaact caaggtcatc cccatgccct 12600  
ggcggccagtc cgagtcggac attattgcct ccaagatcaa gaacggcggg ctcaatatcg 12660  
gctactacaa ctctgacggc aatgtccttc cacaccctcc tctctgcgc ggcgtggaaa 12720  
ccaccgtgc cgcactcgcc aaagccggtc acaccgtgac cccgtggacg ccatacaagc 12780  
acgatttcgg ccacgatctc atctcccata tctacgcggc tgacggcagc gccgacgtaa 12840  
tgcgcgatat cagtgcctcc ggcgagccgg cgattccaaa tatcaaagac ctactgaacc 12900  
cgaacatcaa agctgttaac atgaacgagc tctgggacac gcatctccag aagtggaatt 12960  
accagatgga gtacctgag aatggcggg aggctgaaga aaaggccggg aaggaaactgg 13020

ES 2 651 076 T3

acgccatcat cgcgcggatt acgcctaccg ctgcggtagc gcatgaccag ttccgggtact 13080  
 atgggtatgc ctctgtgatc aacctgctgg atttcacgag cgtggttggt ccggttacct 13140  
 ttgcggataa gaacatcgat aagaagaatg agagtttcaa ggcggttagt gagcttgatg 13200  
 ccctcgtgca ggaagagtat gatccggagg cgtaccatgg ggcaccggtt gcagtgcagg 13260  
 ttatcggacg gagactcagt gaagagagga cgttggcgat tgcagaggaa gtggggaagt 13320  
 tgctgggaaa tgtggtgact ccataggtcg agaatttata cttagataag tatgtactta 13380  
 caggtatatt tctatgagat actgatgtat acatgcatga taatatttaa acggttatta 13440  
 gtgccgattg tcttgtgcga taatgacggt cctatcaaag caatacactt accacctatt 13500  
 acatgggcca agaaaatatt ttccaacttg 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 cgtaagcgc 13800  
 acaacagggt attagtacag agaaagcatt ctgtggtggt gcccggact ttcttttgcg 13860  
 acataggtaa atcgaatacc atcatactat cttttccaat gactccctaa agaaagactc 13920  
 ttcttcgatg ttgtatacgt tggagcatag ggcaagaatt gtggcttgag atctagatta 13980  
 cgtggaagaa aggtagtaaa agtagtagta taagtagtaa aaagaggtaa aaagagaaaa 14040  
 ccggctacat actagagaag cacgtacaca aaaactcata ggcacttcat catacgcag 14100  
 tttcttgatg cattataata gtgtattaga tattttcaga aatatgcata gaacctcttc 14160  
 ttgcctttac tttttataca tagaacattg gcagatttac ttacactact ttgtttctac 14220  
 gccatttctt ttgttttcaa cacttagaca agttggtgag aaccggacta ctaaaaagca 14280  
 atgttcccac tgaaaatcat gtacctgcag gataataacc ccctaattct gcatcgatcc 14340  
 agtatgtttt tttttctcta ctcattttta cctgaagata gagcttctaa aacaaaaaaa 14400  
 atcagcggatt acatgcata tgtgtgttct agaattggcg atcaccagat cgccattaca 14460  
 atgtatgcag gcaaatattt ctcagaatga aaaatagaga aaaggaaaog 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 gggaatagta tataaccttc cgttctgta ataatcaaa tctttcatct 14820  
 agttctotta agatttcaat attttgcttt cttgaagaaa gaatctactc tctccccca 14880  
 ttgcactgc aaagctagct tggcactggc cgtogtttta caacgctcgtg actgggaaaa 14940  
 ccctggcctt acccaactta atgccttgc agcacatccc cctttcgcca gctggcgtaa 15000

ES 2 651 076 T3

tagcgaagag gccgcaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg 15060  
ggaaattgta aacgttaata ttttgttaaa attcgcgta aatttttggt aaatcagctc 15120  
attttttaac caataggccg aaatcggcaa aatcccttat aaatcaaaag aatagaccga 15180  
gatagggttg agtggtgttc cagtttgtaa caagagtcca ctattaaaga acgtggactc 15240  
caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg aaccatcacc 15300  
ctaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc ctaaagggag 15360  
ccccgattt agagcttgac ggggaaagcc ggcgaacgtg gcgagaaag aaggaagaa 15420  
agcgaagga gcgggcgcta gggcgctggc aagtgtagcg gtcacgctgc gcgtaaccac 15480  
cacaccgcc gcgcttaatg ccgcgctaca gggcgctca ggtggcactt ttcggggaaa 15540  
tgtgcgcgga acccctatth gtttattttt ctaaatacat tcaaataatgt atccgctcat 15600  
gagacaataa ccctgataaa tgcttcaata atattgaaaa aggaagagta tgagtattca 15660  
acatttccgt gtcgccctta ttcccttttt tgcggcattt tgccttcctg tttttgctca 15720  
cccagaaacg ctggtgaaag taaaagatgc tgaagatcag ttgggtgcac gagtgggtta 15780  
catcgaactg gatctcaaca gcggtgaagat ccttgagagt tttcgccccg aagaacgtht 15840  
tccaatgatg agcactthta aagttctgct atgtggcgcg gtattatccc gtattgacgc 15900  
cgggcaagac caactcggtc gccgcataca ctattctcag aatgacttgg ttgagtactc 15960  
accagtcaca gaaaagcatc ttacggatgg catgacagta agagaattat gcagtgtgc 16020  
cataaccatg agtgataaca ctgcccga aettactctg acaacgatcg gaggaccgaa 16080  
ggagctaacc gcttttttgc acaacatggg ggatcatgta actcgccttg atcgttggga 16140  
accggagctg aatgaagcca taccaaaacga 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 ccgtatcgta gttatctaca cgacggggag 16440  
tcaggcaact atggatgaac gaaatagaca gatcgctgag ataggtgcct cactgattaa 16500  
gcattggtaa ctgtcagacc aagtttactc atatatactt tagattgatt taaaacttca 16560  
tttttaattt aaaaggatct aggtgaagat cttttttgat aatctcatga ccaaaatccc 16620  
ttaacgtgag ttttcgttcc actgagcgtc agaccccgtg gaaaagatca aaggatcttc 16680  
ttgagatcct tttttctg gcgtaatctg ctgcttgcaa acaaaaaaac caccgctacc 16740  
agcggtggtt tgtttgccgg atcaagagct accacctctt tttccgaag taactggctt 16800  
cagcagagcg cagataccaa atactgtcct tctagtgtag ccgtagttag gccaccactt 16860  
caagaactct gtagcacgcg ctacatacct cgctctgcta atcctgttac cagtggctgc 16920

ES 2 651 076 T3

tgccagtggc gataagtcgt gtcttaccgg gttggactca agacgatagt taccggataa 16980  
 ggcgcagcgg tcgggctgaa cgggggggttc gtgcacacag cccagcttgg agcgaacgac 17040  
 ctacaccgaa ctgagatacc tacagcgtga gcattgagaa agcggccacgc ttcccgaagg 17100  
 gagaaaggcg gacaggtatc cggtaagcgg cagggctcga acaggagagc gcacgagggga 17160  
 gcttccaggg ggaaacgcct ggtatcttta tagtcctgtc gggtttcgcc acctctgact 17220  
 tgagcgtcga tttttgtgat gctcgtcagg gggggcggagc ctatggaaaa acgccagcaa 17280  
 cgcggccttt ttacggttcc tggccttttg ctggcctttt gctcacatgt tctttcctgc 17340  
 gttatcccct gattctgtgg ataaccgtat taccgccttt gagtgagctg ataccgctcg 17400  
 ccgcagccga acgaccgagc gcagcagtc agtgagcag gaagcgggaag agcgcoccaat 17460  
 acgcaaaccg cctctccccg cgcgttggcc gattcattaa tgcagctggc acgacaggtt 17520  
 tcccgactgg aaagcgggca gtgagcgoaa cgcaattaat gtgagttagc tcactcatta 17580  
 ggcaccccag gctttacact ttatgcttcc ggctcgtatg ttgtgtggaa ttgtgagcgg 17640  
 ataacaattt cacacaggaa acagctatga catgattacg aatttaatac gactcacaat 17700  
 aggaattag cttgcgcgaa attattggct tttttttttt ttttaattaat actacctttt 17760  
 gatgtgaacg tttactaaag tagcactatc tgtggaatgg ctggttgaac tttttccgat 17820  
 taacagcttg tattccaagt cctgacattc cagttgtaag ttttccaact tgtgattcaa 17880  
 ttgttcaatc tcttggttaa aattctcttg ttccatgaat aggctctttt tccagtctcg 17940  
 aaatittgaa atttctctgt tggacagctc gttgaatttt ttcttagctt ctaattgtct 18000  
 agttataaat tcaggatccc attctgtagc caccttatcc atgaccgttt tattaattat 18060  
 ttcatagcac ttgtaatttt tgagtttgtt ttccctgatt tcatcgaagt tcatttcttc 18120  
 ctccaaaaat ttcctttggt cttccgttat gtcaaacact ttcgttgta agcaatctct 18180  
 ggcctttaat agcctagttc ttagcatttc agatc 18215

<210> 2

<211> 23

5

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador sintético

10

<400> 2

tgatcttga gaaagtaccg agg 23

<210> 3

15

<211> 24

<212> ADN  
<213> Secuencia artificial

<220>

5 <223> Cebador sintético

<400> 3  
ggaaacagct atgacatgat tacg 24

10 <210> 4  
<211> 24  
<212> ADN  
<213> Secuencia artificial

15 <220>  
<223> Cebador sintético

<400> 4  
ggaaacagct atgacatgat tacg 24

20 <210> 5  
<211> 25  
<212> ADN  
<213> Secuencia artificial

25 <220>  
<223> Cebador sintético

<400> 5  
30 cttgttctt ccggtatgac aacac 25

<210> 6  
<211> 23  
<212> ADN  
35 <213> Secuencia artificial

<220>

# ES 2 651 076 T3

<223> Cebador sintético

<400> 6

ttccaagaag aacaacctga tag 23

5

<210> 7

<211> 21

<212> ADN

<213> Secuencia artificial

10

<220>

<223> Cebador sintético

<400> 7

15 tgatgtgaac gtttactaaa g 21

<210> 8

<211> 16176

<212> ADN

20 <213> Secuencia artificial

<220>

<223> Plásmido sintético

25 <400> 8

tcgcgcggtt cggtgatgac ggtgaaaacc tcttgacaca tgcagctccc ggagacggtc 60

acagcttgtc tgtaagcggg tgccggggagc agacaagccc gtcagggcgc gtcagcgggt 120

ES 2 651 076 T3

gttgccgggt gtcggggctg gcttaactat gcggcatcag agcagattgt actgagagtg	180
caccatattc ggtgtgaaat accgcacaga tgcgtaagga gaaaataccg catcaggcgc	240
cattcgcctt tcaggctgcg caactgttgg gaagggcgat cggtgccggc ctcttcgcta	300
ttacgccagc tggcgaaag gggatgtgct gcaaggcgat taagttgggt aacgccaggg	360
ttttcccagt cacgaogttg taaaacgacg gccagtaagc ttgcatgcct gcaggtcagc	420
gcggcccgcat attttttgta actgtaattt cactcatgca caagaaaaa aaaactggat	480
taaaagggag cccaaggaaa actcctcagc atatatttag aagtctcctc agcatatagt	540
tgtttgtttt ctttacacat tcactgttta ataaaacttt tataatattt cattatcgga	600
actctagatt ctatacttgt ttcccaattg ggcgatcgg gccttgctgg tagtaaactg	660
atacgtcata aaagggaaaa gccacatgcg gaagaatttt atggaaaaa aaaaaactc	720
gaagttacta cttctagggg gcctatcaag taaattactc ctggtacact gaagtatata	780
agggatatag aagcaaatag ttgtcagtc aatccttcaa gacgattggg aaaatactgt	840
aggtaccgga gacctaaacta catagtgttt aaagattacg gatatttaac ttacttagaa	900
taatgccatt tttttgagtt ataataatcc tacgtagtg tgagcgggat ttaaactgtg	960
aggacctaa tacattcaga cacttctgcg gtatcacct acttattccc ttogagatta	1020
tatctaggaa cccatcaggt tgggtgaaga ttacccttc taagactttt cagcttctc	1080
tattgatgtt acacctggac acccctttc tggcatccag tttttaatct tcagtggcat	1140
gtgagattct ccgaaattaa ttaaagcaat cacacaattc tctcggatc cacctcgggt	1200
gaaactgaca ggtggtttgt tacgcatgct aatgcaaagg agcctatata cctttggctc	1260
ggctgctgta acagggaaata taaagggcag cataatttag gagtttagtg aacttgcaac	1320
atttactatt ttcccttctt acgtaaataat ttttctttt aattctaaat caatctttt	1380
caattttttg tttgtattct tttcttgctt aatctataa ctacaaaaa cacatacata	1440
aactaaaaat gtctgaacca gctcaaaaga aacaaaagg tgcatacaac tctctagaac	1500
aattgaaagc ctccggcact gtcgttgtg ccgacactgg tgatttcggc tctattgcca	1560
agtttcaacc tcaagactcc acaactaacc catcattgat cttggctgot gccaaagcaac	1620
caacttacgc caagttgatc gatgttgccg tgaatacgg taagaagcat ggtaagacca	1680
cogaagaaca agtcgaaaat gctgtggaca gattgttagt cgaattcggg aaggagatct	1740
taaagattgt tccaggcaga gtctccaccg aagttgatgc tagattgtct tttgacactc	1800
aagctaccat tgaaaaggct agacatatca ttaaattgtt tgaacaagaa ggtgtctcca	1860
aggaaagagt ccttattaa attgcttcca cttgggaagg tattcaagct gccaaagaat	1920
tggaagaaaa ggacggtatc cactgtaatt tgactctatt attctcctc gttcaagcag	1980
ttgcctgtgc cgaggcccaa gttactttga tttcccatt tgttgtaga attctagact	2040
ggtacaaatc cagcactggt aaagattaca agggtgaagc cgaccaggt gttatttccg	2100

ES 2 651 076 T3

tcaagaaaat ctacaactac tacaagaagt acggttacia gactattggt atgggtgctt 2160  
ctttcagaag cactgacgaa atcaaaaact tggctgggtg tgactatcta acaatttctc 2220  
cagctttatt ggacaagtg atgaacagta ctgaaccttt cccaagagt ttggaccctg 2280  
tctccgctaa gaaggaagcc ggcgacaaga tttcttacat cagcgcgaa tctaaattca 2340  
gattegactt gaatgaagac gctatggcca ctgaaaaatt gtccgaaggt atcagaaaat 2400  
tctctgccga tattgttact ctattcgact tgattgaaaa gaaagttacc gcttaaggaa 2460  
gtatctcggg aatattaatt taggccatgt ccttatgcac gtttcttttg atacttacgg 2520  
gtacatgtac acaagtatat ctatatatat aaattaatga aaatccccta tttatatata 2580  
tgactttaac gagacagaac agttttttat tttttatcct atttgatgaa tgatacagtt 2640  
tcttattcac gtggtatacc cacaccaaat ccaatagcaa taccggccat cacaatcact 2700  
gtttcggcag cccctaagat cagacaaaac atccggaacc accttaaatc aacgtcccat 2760  
atgaatcctt gcagcaaagc cgctcgtacc ggagatatac aatagaacag ataccagaca 2820  
agacataatg ggctaaacaa gactacacca attacactgc ctcatgatg gtggtacata 2880  
acgaactaat actgtagccc tagacttgat agccatcatc atatogaagt ttcactacct 2940  
tttttccatt tgccatctat tgaagtaata ataggcgcac gcaacttctt ttctttttt 3000  
ttcttttctc tctccccgtg tgttgctca ccatatccgc aatgacaaaa aatgatgga 3060  
agacactaaa ggaaaaaatt aacgacaaag acagcaccaa cagatgctgt tgttccagag 3120  
ctgatgaggg gtatctcgaa gcacacgaaa ctttttccct ccttcattca cgcacactac 3180  
tctctaatag gcaacggtat acggccttcc ttccagttac ttgaatttga aataaaaaaa 3240  
agtttgctgt cttgctatca agtataaata gacctgcaat tattaatctt ttgtttcctc 3300  
gtcattgttc tcgttccctt tcttccctgt ttctttttct gcacaatatt tcaagctata 3360  
ccaagcatac aatcaactat ctcatataca atgactcaat tcactgacat tgataagcta 3420  
gccgtctcca ccataagaat tttggctgtg gacaccgtat ccaaggccaa ctcaggtcac 3480  
ccaggtgctc cattgggtat ggcaccagct gcacacgttc tatggagtca aatgcgcatg 3540  
aacccaacca acccagactg gatcaacaga gatagatttg tcttgtctaa cggtcacgcg 3600  
gtcgctttgt tgtattctat gctacatttg actggttacg atctgtctat tgaagacttg 3660  
aaacagttca gacagttggg ttccagaaca ccaggctatc ctgaatttga gttgccaggt 3720  
gttgaagtta ctaccggtcc attaggtcaa ggtatctcca acgctggttg tatggccatg 3780  
gctcaagcta acctggctgc cacttacaac aagccgggct ttacctgtc tgacaactac 3840  
acctatgttt tcttgggtga cggttgtttg caagaaggta tttcttcaga agcttctctc 3900  
ttggctggtc atttgaaatt gggtaacttg attgccatct acgatgacaa caagatcact 3960  
atcgatgggt ctaccagtat ctcatcgat gaagatggtg ctaagagata cgaagcctac 4020

ES 2 651 076 T3

ggttggaag tttgtacgt agaaaatggt aacgaagatc tagccggtat tgccaaggct 4080  
 attgctcaag ctaagttatc caaggacaaa ccaactttga tcaaaatgac cacaaccatt 4140  
 ggttacggtt ccttgcatgc cggctctcac tctgtgcacg gtgccccatt gaaagcagat 4200  
 gatgttaaac aactaaagag caaattcggc ttcaaccag acaagtcctt tgttgttcca 4260  
 caagaagttt acgaccacta caaaagaca attttaaagc cagggtgctga agccaacaac 4320  
 aagtgaaca agttgttcag cgaataccaa aagaaattcc cagaattagg tgctgaattg 4380  
 gctagaagat tgagcggcca actaccgca aattgggaat ctaagttgcc aacttacacc 4440  
 gccaaaggact ctgccgtggc cactagaaaa ttatcagaaa ctggttctga ggatgtttac 4500  
 aatcaattgc cagagttgat tgggtggtct gccgattaa caccttctaa cttgaccaga 4560  
 tggaaggaag cccttgactt ccaacctcct tcttccggtt caggtaacta ctctggtaga 4620  
 tacattaggt acggtattag agaacacgct atgggtgcca taatgaacgg tatttcagct 4680  
 ttcggtgcca actacaaacc atacggtggt actttcttga acttcgttc ttatgctgct 4740  
 ggtgccgta gattgtccgc tttgtctggc caccagtta tttgggttc tacacatgac 4800  
 tctatcggtg tcggtgaaga tggccaaca catcaacctt ttgaaacttt agcacacttc 4860  
 agatccctac caaacattca agtttgaga ccagctgatg gtaacgaagt ttctgccc 4920  
 tacaagaact ctttagaatc caagcactt ccaagtatca ttgotttgc cagacaaaac 4980  
 ttgccacaat tggaaggtag ctctattgaa agcgttctc aggggtgta cgtactacaa 5040  
 gatgttgcta acccagatat tatttttagt gctactggtt ccgaagtgc tttgagtgtt 5100  
 gaagctgcta agactttggc cgcaaagaac atcaaggctc gtgttgttc tctaccagat 5160  
 ttcttcaact ttgacaaaac acccctagaa tacagactat cagtcttacc agacaacggt 5220  
 ccaatcatgt ctgttgaagt tttggctacc acatgttggg gcaaatacgc tcatcaatcc 5280  
 ttcggtattg acagatttgg tgcctccggt aaggcaccag aagtcttcaa gttcttcggt 5340  
 ttcaccccag aagggtgttc tgaaagagct caaaagacca ttgcattcta taagggtgac 5400  
 aagctaattt ctctttgaa aaaagcttct taaattctga tcgtagatca tcagatttga 5460  
 tatgatatta tttgtgaaaa aatgaaataa aactttatac aacttaata caactttttt 5520  
 tataaacgat taagcaaaaa aatagtttca aacttttaac aatattcaa acaactcagtc 5580  
 ctttctcttc ttatattata ggtgtacgta ttatagaaaa atttcaatga ttactttttc 5640  
 tttctttttc cttgtaccag cacatggccg agcttgaatg ttaaaccctt cgagagaatc 5700  
 acaccattca agtataaagc caataaagaa tategtacca gagaattttg ccatcggaca 5760  
 tgctacctta cgcttatatc tctcattgga atatcgtttt ctgattaaaa cacggaagta 5820  
 agaacttaat tcgtttttcg ttgaactatg ttgtgccagc gtaacattaa aaaagagtgt 5880  
 acaaggccac gttctgtcac cgtcagaaaa atatgtcaat gaggcaagaa ccgggatggt 5940  
 aacaaaaatc acgatctggg tgggtgtggg tgtattggat tataggaagc cacgcgctca 6000

ES 2 651 076 T3

acctggaatt acaggaagct ggtaatTTTT tgggtttgca atcatcacca tctgcacgtt 6060  
 gttataatgt cccgtgtcta tatatatcca ttgacgggat tctatTTTT tgctattgaa 6120  
 atgagcgttt tttgttacta caattggttt tacagacgga atttcccta tttgtttcgt 6180  
 cccatTTTT cttttctcat tgttctcata tettaaaaaa gtcctttctt cataatcaat 6240  
 gttttctttt acttaatatt ttacttgcac tcagtgaatt ttaatacata ttctctagc 6300  
 cttgcaaaaat cgatttagaa tcaagatacc agcctaaaaa tggcacaacc aattatagct 6360  
 cccagtatcc 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  
 aatgggtcg acgattttgc taaatgtggt gctgaccaat ttacgttcca ctacgaggcc 6660  
 acacaagacc ctttgcattt agttaagttg attaagtcta agggcatcaa agctgcatgc 6720  
 gccatcaaac ctggtacttc tgttgacggt ttatttgaac tagctcctca tttggatatg 6780  
 gctcttgta tgactgtgga acctgggttt ggaggccaaa aatcatgga agacatgatg 6840  
 ccaaaagtgg aaactttgag agccaagttc cccatttga atatccaagt cgatggtggt 6900  
 ttgggcaagg agaccatccc gaaagccgcc aaagccggtg ccaacgttat tgtcgtgga 6960  
 accagtgttt tcaactgcagc tgaccgcac gatgttatct cctcatgaa agaagaagtc 7020  
 tgaaggaat tgcgttctag agatttgeta gattagttgt acatatgcg catttcttat 7080  
 atttatactc tctatactat acgatatggt attttttct cgttttgatc tcctaataata 7140  
 cataaacga gccattccta ctatacaaga tacgtaagt cctaactcat gggaaaaatg 7200  
 ggccgccag ggtggtgcct tgcctgtttt cgatgatcaa tccctgggat gcagtatcgt 7260  
 caatgacact ccataaggct tccttaacca aagtcaaaga actcttcttt tcattctctt 7320  
 tcactttctt accgccatct agatcaatat ccatttctga ccccgggaa ccgccagata 7380  
 ttacttactt gacgcaaaag cgtttgaat aatgacgaaa aagaaggaag aaaaaaaaaag 7440  
 aaaaataacc cttctaggcg gttatctac tgatccgagc ttccactag atagcaccca 7500  
 aacacctgca tatttggacg accttactt acaccacca aaaccacttt cgcctctccc 7560  
 gccctgata acgtccacta attgagcgat tacctgagcg gtcctctttt gtttgcagca 7620  
 tgagacttgc atactgcaa tcgtaagtag caacgtctca aggtcaaac tgatggaaa 7680  
 cttgtcacc tcaactaatt ctagctagcc taccctgcaa gtcgaagagt ctccgtgatt 7740  
 cctagccacc tcaaggtatg cctctccccg gaaactgtgg ccttttctgg cacacatgat 7800  
 ctccacgatt tcaacatata aatagctttt gataatggca atattaatca aatttatttt 7860  
 acttctttct tgtaacatct ctottgtaat ccottattcc ttctagctat ttttcataaa 7920

ES 2 651 076 T3

aaaccaagca actgcttata aacacacaaa cactaaatca aaatggctgc cgggtgccca 7980  
aaaattgatg cgttagaatc tttgggcaat cctttggagg atgccaagag agctgcagca 8040  
tacagagcag ttgatgaaaa tttaaaattt gatgatcaca aaattattgg aattggtagt 8100  
ggtagcacag tggtttatgt tgccgaaaga attggacaat atttgcataa ccctaaattt 8160  
tatgaagtag cgtctaaatt catttgcatt ccaacaggat tccaatcaag aaacttgatt 8220  
ttggataaca agttgcaatt aggtccatt gaacagatc ctcgcattga tatagcgttt 8280  
gacggtgctg atgaagtga tgagaattta caattaatta aagggtggtg tgcttgcata 8340  
tttcaagaaa aattggttag tactagtgc taaaacctca ttgtcgttcg tgattcaaga 8400  
aaaaagtcac caaaacattt aggtaagaac tggaggcaag gtgttcccat tgaaaatgta 8460  
ccttcctcat acgtgagggt caagaatgat ctattagaac aattgcatgc tgaaaaagtt 8520  
gacatcagac aaggagggtc tgctaaagca ggtcctgttg taactgaca taataacttc 8580  
attatcgatg cggatttcgg tgaaatttcc gatccaagaa aattgcatag agaaatcaaa 8640  
ctgttagtgg gcgtggtgga aacaggttta ttcacgaca acgcttcaaa agcctacttc 8700  
ggtaattctg acggtagtgt tgaagttacc gaaaagtgag cagatcaaag gcaagacag 8760  
aaaccgtagt aaaggttgac ttttcacaac agtgtctcca tttttatat tgtattatta 8820  
aagctattta gttatttga tactgtttt tttccagaag ttttctttt agtaaagtac 8880  
aatccagtaa aatgaagga tgaacaatcg gtgtatgcag attcaacacc aataaatgca 8940  
atgtttatct ctttggaaac tttgtgttgt tcgaaatcca ggataatcct tcaacaagac 9000  
cctgtccgga taaggcgtta ctaccgatga cacaccaagc tcgagtaacg gagcaagaat 9060  
tgaagatat ttctgcacta aatgccaaca tcagatttaa tgatccatgg acctggttgg 9120  
atggtaaatt ccccactttt gcctgatcca gccagtaaaa tccatactca acgacgatat 9180  
gaacaaattt ccctcattcc gatgctgtat atgtgtataa atttttacat gctcttctgt 9240  
ttagacacag aacagcttta aataaaatgt tggatatact ttttctgctt gtggtgtcat 9300  
ccacgctttt aattcatctc ttgtatggtt gacaatttgg ctatttttta acagaacca 9360  
acggtaattg aaattaaaag ggaacagat gggggcgatg agtgagtgat actaaaatag 9420  
acaccaagag agcaaagcgg tcccagcggc cgcgaattcg gcgtaatcat ggtcatagct 9480  
gtttcctgtg tgaaattggt atccgctcac aattccacac aacatacag cgggaagcat 9540  
aaagtgtaaa gcctggggtg cctaattgag gagctaactc acattaattg cgttgcgctc 9600  
actgccgctc ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg 9660  
cgcggggaga ggcggtttgc gtattggcg ctcttccgct tcctcgctca ctgactcgct 9720  
gcgctcgctc gttcggctgc ggcgagcgg atcagctcac tcaaagcgg taatacggtt 9780  
atccacagaa tcaggggata acgcaggaaa gaacatgta gcaaaaggcc agcaaaaggc 9840  
caggaaccgt aaaaaggccg cgttgcctgc gtttttccat aggtccgcc ccctgacga 9900

ES 2 651 076 T3

gcatcacaaa aatcgacgct caagtcagag gtggcgaaaac ccgacaggac tataaagata 9960  
 ccaggcggtt cccctcgaa gctccctcgt gcgctctcct gttccgacce tgccgcttac 10020  
 cggatacctg tccgccttcc tcccttcggg aagcgtggcg ctttctcaat gctcacgctg 10080  
 taggtatctc agttcgggtg aggtcgttcg ctccaagctg ggetgtgtgc acgaaccccc 10140  
 cgttcagccc gaccgctcgc cttatccgg taactatcgt cttgagtcca acccggttaag 10200  
 acacgactta tcgccactgg cagcagccac tggtaacagg attagcagag cgaggatagt 10260  
 aggcggtgct acagagttct tgaagtgggt gcctaactac ggctacacta gaaggacagt 10320  
 atttggatcc tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtatctcttg 10380  
 atccggcaaa caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac 10440  
 gcgcagaaaa aaaggatctc aagaagatcc tttgatcttt tctacggggg ctgacgctca 10500  
 gtggaacgaa aactcacgtt aagggtttt ggtcatgaga ttatcaaaaa ggatcttcac 10560  
 ctagatcctt ttaaattaa aatgaagttt taaatcaatc taaagtatat atgagtaaac 10620  
 ttggtctgac agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt 10680  
 tcgttcaccc atagttgcct gactccccgt cgtgtagata actacgatac gggagggcct 10740  
 accatctggc cccagtgctg caatgatacc gcgagacca cgctcaccg ctccagattt 10800  
 atcagcaata aaccagccag ccggaagggc cgagcgcaga agtggctctg caactttatc 10860  
 cgctccatc cagtctatta attggttcgg ggaagctaga gtaagtagtt cgccagttaa 10920  
 tagtttgccg aacgttgttg ccattgctac aggcacgtg gtgtcacgct cgtcgtttgg 10980  
 tatggcttca ttcagctccg gttcccacg atcaaggcga gttacatgat ccccatggtt 11040  
 gtgcaaaaaa gcggttagct ccttcggctc tccgatcgtt gtcagaagta agttggccgc 11100  
 agtggtatca ctcatgggta tggcagcact gcataattct cttactgtca tgccatccgt 11160  
 aagatgcttt tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg 11220  
 gcgaccgagt tgctcttgcc cggcgtcaat accgggataat accgcgccac atagcagaac 11280  
 tttaaaagtg ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc 11340  
 gctgttgaga tccagttcga tgaaccac tctgacccc aactgatctt cagcatcttt 11400  
 tactttcacc agcgtttctg ggtgagcaaa aacaggaagg caaatgccg caaaaaggg 11460  
 aataaggcgc 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

ES 2 651 076 T3

tggatcgggt ctgcgattcc gactcgtcca acatcaatac aacctattaa tttccctcgg 11880  
 tcaaaaataa ggttatcaag tgagaaatca ccatgagtga cgactgaatc cggtgagaat 11940  
 ggcaaaagct tatgcatttc tttccagact tgttcaacag gccagccatt acgctcgtca 12000  
 tcaaaatcac tcgcatcaac caaaccgtta ttcattcgtg attgcgccctg agcgagacga 12060  
 aatcgcgat cgctgttaaa aggacaatta caaacaggaa tcgaatgcaa ccggcgcagg 12120  
 aacctgccca gcgcatcaac aatattttca cctgaatcag gatattcttc taatacctgg 12180  
 aatgctgttt tgccggggat cgcagtgtg agtaacctag catcatcagg agtacggata 12240  
 aaatgcttga tggtcggaag aggcataaat tccgtcagcc agtttagtct gaccatctca 12300  
 tctgtaacat cattggcaac gctacctttg ccatgtttca gaaacaactc tggcgcacg 12360  
 ggcttcccat acaatcgata gattgtcgca cctgattgcc cgacattatc gcgagcccat 12420  
 ttatacccat ataaatcagc atccatgttg gaatttaatc gcggcctcga aacgtgagtc 12480  
 ttttccttac ccatggttgt ttatgttcgg atgtgatgtg agaactgtat cctagcaaga 12540  
 ttttaaaag aagtatatga aagaagaacc tcagtggcaa atcctaacct tttatatttc 12600  
 tctacagggg cgcggcgttg ggacaattca acgcgactgt gacgcgttct agaacacaca 12660  
 atatgcatgt aatcgtgat ttttttgtt ttagaagctc tatcttcagg taaaaatgag 12720  
 tagagaaaaa aaaacatact ggatcgatgc agaattaggg ggttattatc ctgcaggtag 12780  
 atgattttca gtgggaacat tgctttttag tagtccggtt ctcaacaact tgtctaagt 12840  
 ttgaaaacaa aagaaatgac gtagaaacaa agtagtataa gtaaactctc caatgttcta 12900  
 tgtataaaaa gtaaaggcaa gaagaggttc tatgcatatt tctgaaaata tctaatacac 12960  
 tattataatg catcaagaaa ctgtcgtatg atgaagtgcc tatgagtttt tgtgtacgtg 13020  
 cttctctagt atgtagccgg tttctctttt ttacctcttt ttactactta tactactact 13080  
 ttactacct ttcttcacag taatctagat ctcaagccac aattccttgc ctatgctcca 13140  
 acgtatacaa catcgaagaa gagtctttct ttagggagtc attggaaaag atagtatgat 13200  
 ggtattcgat ttacctatgt cgcaaaagaa agtccggggc aacaccacag aatgctttct 13260  
 ctgtactaat aacctgttgt gcgcttaacg gtctaactgt taatcagcgg tggttaaatt 13320  
 tttgtaaatc taatgttcca tgattttctt tcttcaaaaag gaacatgtag cgaaaatctt 13380  
 ttttttactt tgatacactg caattgtttc tgagcatgct gaaattttct cgatgttttt 13440  
 tttttttatt ggcatccaag taattaatcc ttatgctacg aaaaagttgt agaatgaat 13500  
 catgcataat ctaacggata tcatcatata ctctgtgcta atattctaaa caagttcgaa 13560  
 aatattttct tggcccatgt aataggtggg aagtgtattg ctttgatagg aacgtcatta 13620  
 tcgcacaaga caatcggcac taataaccgt ttaaatatta tcatgcatgt atacatcagt 13680  
 atctcataga aatatacctg taagtacata cttatctaag tataaattct cgacctatgg 13740  
 agtcaccaca tttccagca acttcccac ttcctctgca atcgccaacg tcctctcttc 13800

ES 2 651 076 T3

actgagtctc cgtccgataa cctgcactgc aaccggtgcc ccatggtacg cctccggatc 13860  
atactcttcc tgcacgaggg catcaagctc actaaccgcc ttgaaactct cattcttctt 13920  
atcgatgttc ttatccgcaa aggtaaccgg aacaaccacg ctcgtgaaat ccagcaggtt 13980  
gatcacagag gcatacccat agtaccggaa ctgggtcatgc cgtaccgcag cggtaggcgt 14040  
aatcggcgcg atgatggcgt ccagttcctt cccggccttt tcttcagcct cccgccattt 14100  
ctcaaggtag tccatctggt aattccactt ctggagatgc gtgtccaga gctcgttcat 14160  
gttaacagct ttgatgttcg ggttcagtag gtctttgata tttggaatcg cggctcggcc 14220  
ggatgcactg atatcgcgca ttacgtcggc getgcctca gccgcgtaga tatgggagat 14280  
gagatcgtgg ccgaaatcgt gcttgatggc cgtccacggg gtcacgggtg gaccggcttt 14340  
ggcgagtgcg gcgacggtag tttccacgcc gcgcaggata ggaggggtg gaaggacatt 14400  
gccgtcgaag ttgtagtagc cgatattgag cccgccgttc ttgatcttgg aggcaataat 14460  
gtccgactcg gactggcgcc agggcatggg gatgacctg gactcgtatt tccatggctc 14520  
ctgaccgagg acggatttgg tgaagaggcg gaggtcctca acagagtgcg taatcggccc 14580  
gacaacgctg tgcaccgtct cctgaccctc catgctgttc gccatctttg catacggcag 14640  
ccgccatga ctggcctta gaccgtacag gaagttgaac gcggccggca ctcgaatcga 14700  
gccaccgata tccgttctta caccgatgac gccaccacga atcccaacga tcgcaccctc 14760  
accaccagaa ctgccgccgc acgaccagt tctgttgcgt gggttgacg tgcgccgat 14820  
gatgttggtg actgtctcgc agaccatcag ggtctgcggg acagaggtct tgacgtagaa 14880  
gacggcaccg gctttgcgga gcatggttgt cagaaccgag tccccttctg cgtacttgtt 14940  
tagccatgag atgtagccca ttgatgttcc gtagcccttg actcgaagct ggtctttgag 15000  
agagatgggg aggccatgga gtggaccaac ggtctctctg tgctttgcgt agtattcatc 15060  
gagttccctt gcctgcgcga gagcggcgtc agggaagaac tcgtgggcgc agtttgtaa 15120  
ctgctgggcg attgctgcc gtttacagaa tctagcgtta acttcaccg aggtcaactc 15180  
tccggccgcc agcttgaca caagatctgc agcggaggcc tctgtgatct tcagttcggc 15240  
ctctgaaagg atccccgatt tctttgggaa atcaataacg ctgtcttccg caggcagcgt 15300  
ctggactttc cattcatcag ggatggtttt tgcgagggcg gcgcgcttat cagcggccag 15360  
ttcttcccag gattgaggca ttgtatatga gatagttgat tgtatgcttg gtatagcttg 15420  
aaatattgtg cagaaaaaga aacaaggag aaagggaacg agaacaatga cgaggaaaca 15480  
aaagattaat aattgcaggt ctatttatac ttgatagcaa agcggcaaac tttttttatt 15540  
tcaaattcaa gtaactggaa ggaaggccgt ataccgttgc tcattagaga gtagtgtgcg 15600  
tgaatgaagg aaggaaaaag ttctgtgtgt tcaagatac ccctcatcag ctctggaaca 15660  
acgacatctg ttggtgctgt ctttgcgtt aatttttcc tttagtgtct tccatcattt 15720

ES 2 651 076 T3

```

tttttgcac tgccgatatg gtgagacaac aacgggggag agagaaaaga aaaaaaaga 15780
aaagaagttg catgcgccta ttattacttc aatagatggc aaatggaaaa agggtagtga 15840
aacttcgata tgatgatggc tatcaagtct agggctacag tattagtctg ttatgtacca 15900
ccatcaatga ggcagtgtaa tttgtgtagt cttgttttagc ccattatgtc ttgtctggta 15960
tctgttctat tgtatatctc ccctccgcca cctacatggt agggagacca acgaaggat 16020
tataggaatc ccgatgtatg ggtttggttg ccagaaaaga ggaagtccat attgtacacc 16080
cggaacaac aaaaggatgg gcccatgacg tctaagaaac cattattatc atgacattaa 16140
cctataaaaa taggcgtatc acgagccct ttcgctc 16176

```

<210> 9

<211> 22

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Cebador sintético

10

<400> 9

acgccagggt ttcccagtc ac 22

<210> 10

15 <211> 22

<212> ADN

<213> Secuencia artificial

<220>

20 <223> Cebador sintético

<400> 10

ccagcaccct aagccgacta gg 22

25 <210> 11

<211> 22

<212> ADN

<213> Secuencia artificial

<220>  
 <223> Cebador sintético  
  
 <400> 11  
 5 caccaacctg atgggtcct ag 22  
  
 <210> 12  
 <211> 21  
 <212> ADN  
 10 <213> Secuencia artificial  
  
 <220>  
 <223> Cebador sintético  
  
 <400> 12  
 15 acggtgctga tgaagtggat g 21  
  
 <210> 13  
 <211> 21  
 20 <212> ADN  
 <213> Secuencia artificial  
  
 <220>  
 <223> Cebador sintético  
 25  
 <400> 13  
 accacgccca ctaacagttt g 21  
  
 <210> 14  
 30 <211> 16580  
 <212> ADN  
 <213> Secuencia artificial  
  
 <220>  
 35 <223> Plásmido sintético  
  
 <400> 14

ES 2 651 076 T3

taaaagaaa cattctctag ggattacgag gtaaagatac attttcaagg cttattcgat	60
tctgtgaact cagttggaat attaagggac aggttgtttc cttgcaccca gagaagcaat	120
atcgttgagc atgttcgaca ttgcgtatcc ttggatgaaa gacgtggaaa attcaagcag	180
ttatgtttca ctccgatgcc gtacattccg aaactatttt cattgacata ttgtaatcat	240
ataactgacc agtgttcgcc ggtgcccaact tctaattgcat taatgcgtga tctaaccctg	300
gaaaatcctt tgataaaata cactttaaaa agtggcgcac attctattag taatccttct	360
ccactcattc ctgataacc c tggaaggttg ttatcgagca aaagcgagga aactacagag	420
ttgtgtttg acctgaactc attcttagaa ggtaattcat acgcgagaga tacagaatgt	480
tacaacaagag gaattgaagc cttttccaa cttcaatcta tccaaggcag cggtacatca	540
agtagaatga ctatgacacc cgacttgatt gaaaaatggt ttccaggtga tcggccatct	600
tggccgatca ttctgacgtt ggtggagggt gggcgcctga ctgtgagaca gaagagaact	660
tgtcaaattt aacgctgca tggattttag cagaggcaat caaatttgggt gttaaattca	720
aacctggtgc aatacatgat ttcgctacca aacacacttc gattggatct ttattcgcag	780
acacacatga ttaccttagt ttcaactcac caaagaaatg ttccctacta ggagtgagtg	840
ataatgagga tggagcccga gaggataaat ctggcagaaa tgagagaatg gaagattgtc	900
taaaaaatat aaaagagact agattgagct tgaaagatga aaaagaaaa gtgaaggatg	960
cttttactct taaatgtgga catgcaaata aatttatgag attggtgtgg tgggtattgg	1020
aactgctccc cattggaata cgaatggaaa ataaagaagg aaagtggcaa aattttcata	1080
cacctaacct cggaagatcg tcgacaagct tgtggagagg tgacttcatg aaccaagtgt	1140
ctgtcgatat acaacaaaaa ggaaccattt tcatcttgat ggacaacatg tgcatcaaaa	1200
accttatcgt aaagagttct tggacccttg gatggagtgt aaacctgat ttaaacagc	1260
aaataataaa aatcgatagc gacaaaaact gtcaatttca atattcttta tatttggtga	1320

ES 2 651 076 T3

ctgcttagat attttgagaa aattcagcgg aaacagcgtg atgagtgagt taagttctgc	1380
tgtttaaata agtattcaac tactattgaa gccgactcat gaagccggtt acggacaaaa	1440
ccgggcaaat ttccgccggtc ccggaatttt cgtttccgca ataaaagaac cgctcatcat	1500
catagcgcca gggtagtata ctatagaagg tcagactaaa ctgagtcatc tagagtaatg	1560
acgccttagt agcttttaca tcttcataag aaaaggaaac ttgtagaatg gcctggcgat	1620
ttgtttgctt tcttgtgatg aagaaatttc gatgcgatta accggcaaaa tcagtaaagg	1680
tatttcgctg aggcggcctt caatcatcga atactacgtc ttaatgatg gtactgtggt	1740
tcatattttc aagtagtggt agtaaatttg tatacgttca tgtaagtgtg tatcttgagt	1800
gtctgtatgg gcgcataaac gtaagcgaga cttccaaatg gagcaaacga gaagagatct	1860
ttaaagtatt atagaagagc tgggcaggaa ctattatgac gtaaagcctt gaccataata	1920
aagacgattc tttgtccctc tatacaaca tcttgcaaag ataccaaata ttttcaaatc	1980
ctactcaata aaaaattaat gaataaatta gtgtgtgtgc attatatata ttaaaaatta	2040
agaattagac taaataaagt gtttctaaaa aatattaaa gttgaaatgt gcgtgtgtg	2100
aattgtgctc tattagaata attatgactt gtgtgcgttt catattttaa aataggaaat	2160
aaccaagaaa gaaaaagtac catccagaga aaccaattat atcaaatcaa ataaaaaac	2220
cagcttcggg gtgtgtgtgt gtgtgaagct aagagttgat gccatttaat ctaaaaattt	2280
taaggtgtgt gtgtggataa aatattagaa tgacaattcc ccggaattgc gtacgcccgg	2340
gtcgcgaccg cggcttaaca gtacatgttg acaatggctt cgtacaattc ttgcttacca	2400
gaagtttgct ttggttcacc gttagccttg gcgtaagcaa ccaaattctc caaagtcaac	2460
ttaccatctt caaattcctt acccttacca gagtcgaaag aagcgtatct gtcagccaac	2520
atcttctgtt aaggagattc ttccaataat tttagcagcag attccaaagc tctggccatg	2580
acatccatac cggcaatgtg agcgatgaag atatcttcca agtcagtaga gtttcttctg	2640
gtcttagcat cgaagtgtgt accaccgtta ccgaaaccac cgtttctgat gatttgcac	2700
atagcttgag tcaattcaaa gttgtcgatt gggaaattgg cgggtgtcca accgttttgg	2760
tagtcacctc tgtagcacc aatggaacct aacataccgt tgtogacagc aacagccaat	2820
togtgttoga aagtgtgacc ggccaaagta gcgtggttga cttogatgtt gaccttgaag	2880
tcttgtcca agttgtgagc cttcaagaaa ccgatgacag tttcgggtgc aacatcgtat	2940
tggtgcttgg ttggttccat tggctttggt tcaatcaaga aagtaccctt gaaacctctg	3000
gcacgagcgt agtcacgagc aatggtcaac atttgagcca agtgttcctt ttctctcttt	3060
tggtcagtg tcaacaagga catgtaacct tctctaccac ccagaaaaac gtagttgaa	3120
ccacctaat caatggtagc atcgatggcg ttcttgattt ggatggcagc tctggcaaca	3180
acatcgaaat ctgggttggg agcggcaccg ttcattgtatc tggcatgacc aaagacgtta	3240
gcagtacccc ataatagctt gataccagtt tcagcttgc tttgcttagc gtaagcaaca	3300

ES 2 651 076 T3

atttccttca agttagcttc gtattcttcg atggtttcag cttcttcaca caagtcaaca 3360  
 tcgtggaaac agtagtattc aatacccatc ttttgcataa attcgaaacc agcgtccatc 3420  
 ttgttcttag cagcttggaac cttgtcagct tcaccgttcc atgggaattt cttgttacca 3480  
 ccaccgaatt ggtcaccacc ttcagcacac aaggtatgcc accaagccat agcgaacttt 3540  
 aaccattcag acatcttctt acccatgata accttgtcag catcgtagta tctgaaggcc 3600  
 attgggttct tggattcctt accttcgaat ttgatcttac caatacctgg gaagtattcc 3660  
 ttggtagcca tttttagttt atgtatgtgt ttttactagt tatagattta agcaagaaaa 3720  
 gaatacaaac aaaaaattga aaaagattga tttagaatta aaaagaaaaa tatttacgta 3780  
 agaagggaaa atagtaaatg ttgcaagttc actaaactcc taaattatgc tgcctttat 3840  
 attcctgtt acagcagccg agccaaaggc atataggctc ctttgcatta gcatgcgtaa 3900  
 caaacacct gtcagtttca accgaggtgg tatccgagag aattgtgtga ttgctttaat 3960  
 taatttcgga gaatctcaca tgccactgaa gattaaaaac tggatgccag aaaaggggtg 4020  
 tccaggtgta acatcaatag aggaagctga aaagtcttag aacgggtaat cttccaccaa 4080  
 cctgatgggt tcttagatat aatctcgaag ggaataagta gggatgatac gcagaagtgt 4140  
 ctgaatgtat taaggctctc acagttttaa tcccgctcac actaacgtag gattattata 4200  
 actcaaaaa atggcattat tctaagtaag ttaaataatcc gtaatcttta aacactatgt 4260  
 agttaggtct cgggccccag cgcagtagg gttgttgagc ttagtaaaaa tgtgcgcacc 4320  
 acaagcctac atgactccac gtcacatgaa accacaccgt ggggccttgt tgcgctagga 4380  
 ataggatatg cgacgaagac gcttctgctt agtaaccaca ccacatttc agggggctga 4440  
 tctgcttctc tctttactg tcacgagcgg ccataatcg cgcttttttt ttaaaaggcg 4500  
 cgagacagca aacaggaagc tcgggtttca accttcggag tggctgcaga tctggagact 4560  
 ggatctttac aatacagtaa ggcaagccac catctgcttc ttaggtgcat gcgacggtat 4620  
 ccacgtgcag aacaacatag tctgaagaag ggggggagga gcatgttcat tctctgtagc 4680  
 agtaagagct tggtgataat gacccaaact ggagtctcga aatcatataa atagacaata 4740  
 tattttcaca caatgagatt tgtagtacag tctattctc tctcttgcac aaataagaaa 4800  
 ttcaccaaga acttggtttg atatttcacc aacacacaca aaaaacagta cttcactaaa 4860  
 ttacacaca aaacaaaatg ttgtgttcag taattcagag acagacaaga gaggtttcca 4920  
 acacaatgtc tttagactca tactatcttg ggtttgatct ttcgaccaa caactgaaat 4980  
 gtctcgccat taaccaggac ctaaaaattg tccattcaga aacagtggaa tttgaaaagg 5040  
 atcttcgca ttatcacaca aagaaggggtg tctatataca cggcgacact atcgaatgtc 5100  
 ccgtagccat gtggtttagag gctctagatc tggttctctc gaaatctgc gaggtctaat 5160  
 ttccattgaa caaagttatg gcgctctcag ggtcctgcca gcagcacggg tctgtctact 5220

ES 2 651 076 T3

ggtcctccca agccgaatct ctgtagagc aattgaataa gaaaccggaa aaagatttat 5280  
 tgcactacgt gagctctgta gcatttgcaa ggcaaaccgc cccaattgg caagaccaca 5340  
 gtactgcaaa gcaatgtcaa gagtttgaag agtgcatagg tgggcctgaa aaaatggctc 5400  
 aattaacagg gtccagagcc cattttagat ttactggtcc tcaaattctg aaaattgcac 5460  
 aattagaacc agaagcttac gaaaaaaca agaccatttc ttagtgtct aatTTTTTga 5520  
 cttctatctt agtgggcat cttgttgaat tagaggaggc agatgcctgt ggtatgaacc 5580  
 tttatgatat acgtgaaaga aaattcagtg atgagctact acatctaatt gatagtctt 5640  
 ctaaggataa aactatcaga caaaaattaa tgagagcacc catgaaaaat ttgatagcgg 5700  
 gtaccatctg taaatatttt attgagaagt acggtttcaa tacaactgc aaggctctc 5760  
 ccatgactgg ggataattta gccactatat gttctttacc cctgcggaag aatgacgttc 5820  
 tcgtttccct aggaacaagt actacagttc ttctggctac cgataagtat caccctctc 5880  
 cgaactatca tcttttcatt catccaactc tgccaaacca ttatatgggt atgatttgtt 5940  
 attgtaatgg ttctttggca agggagagga taagagacga gttaaacaaa gaacgggaaa 6000  
 ataattatga gaagactaac gattggactc tttttaatca agctgtgcta gatgactcag 6060  
 aaagtagtga aatgaatta ggtgtatatt ttctctggg ggagatcgtt cctagcgtaa 6120  
 aagccataaa caaaagggtt atcttcaatc caaaaacggg tatgattgaa agagaggtgg 6180  
 ccaagttcaa agacaagag caccgatcca aaaatattgt agaatcacag gctttaagtt 6240  
 gcagggtaa aatatctccc ctgctttcgg attcaaaccg aagctcacia cagagactga 6300  
 acgaagatac aatcgtgaag tttgattacg atgaatctcc gctgcgggac tacctaata 6360  
 aaaggccaga aaggactttt tttgtagtg gggcttctaa aaacgatgct attgtgaaga 6420  
 agtttgctca agtcattggt gctacaaagg gtaattttag gctagaaaca ccaactcat 6480  
 gtgcccttg tggttgttat aaggccatgt ggtcattggt atatgactct aataaaattg 6540  
 cagttccttt tgataaattt ctgaatgaca attttccatg gcatgtaatg gaaagcatat 6600  
 ccgatgtgga taatgaaaat tgggatcgtc ataattccaa gattgtcccc ttaagcgaac 6660  
 tggaaaagac tctcatctaa aatatgtttg aataatttat catgcctga caagtacaca 6720  
 caaacacaga cacataatat acatacatat atatatatca ccgttattat gcgtgcacat 6780  
 gacaatgcc ttgtatgttt cgtatactgt agcaagtagt catcattttg ttccccgttc 6840  
 ggaaaatgac aaaaagtaaa atcaataaat gaagagtaaa aaacaattta tgaaagggtg 6900  
 agcgaccagc aacgagagag acaaatcaaa ttagcgcttt ccagtgagaa tataagagag 6960  
 cattgaaaga gctaggttat acgctggaag atctcgttat gtaccggaat atgtcagttt 7020  
 acattggtca gtctattgga gaattaagt ttgatogtagg tatagaccgg acaatatgcc 7080  
 ggaatatgta aggcaattgt tccaagattt ggaaggattt gatttaaaaa gtaataaagt 7140  
 ttcaataaaa tatgataagc aagataatag caacgggagt gaaatcaatg ggggctttt 7200

ES 2 651 076 T3

tgataatgag gaagggcagg aactccacat gggtaaaaa gcaagttatt ttgcaacgac	7260
atacaattca agattatttg acagtaaata ctcccaatta aaaaagaaat tcatggactg	7320
ggatagtaat tcttggacag atattccaga tgatttaaaa atatacctac agcaagatga	7380
atcgctttag cattaaaaa accccttcgg tacgtaatat aaaaaatttt ataggtaata	7440
tacatatata aaaatacttc aatcattttt acaatcttgt atactttata caacatgtga	7500
aatcttctgc ttctggacat caatattcaa atacaggcca atcttaggta aaacatttgg	7560
agaaaagaag gataaggcag gacgaggaa gataaatagt ttcgttaatt ataaatacat	7620
gcagataaat aaaggaatat caaatattat gaatagaaaa agaagatggt gagacaaaa	7680
agtagtaata aataggcca aatcttcttt atttcccctt tcttttctta tccttttgtt	7740
ttctccatat tgtataagaa tatattctta ggaaaatcaa caggaatac agtatagtga	7800
ttttcgttcc tttttgagcg taatcccttc gagactgtga tgttgattat tttgttgtg	7860
atttcaaaat tcttaggta gttgtatagt tcccgttcat aacataatgg atagtaaag	7920
aaaaatcaaa ataagggtga acaaataga caataaagat gtagttttcg aggacgaaaa	7980
acaaacctaa ccaacaatga ccttatcacc atcgaattca taagcaggaa tttctaagtt	8040
taagggggca ggtcccttcc tgattctacc ggaatatca taatgtgaac catggcaagg	8100
acagaaccaa ccacaaaat caccggcttc accaattgga acacaaccta agtgagtaca	8160
aatacccagc ataattaacc attgagggtc ttgactctg tcagcatcgg tctgtgggtc	8220
cttcaaagcg gacatatcca cactgttggc ttctgaatt tcatgaggag ttctgtgtct	8280
aatgaacaca ggcttacctt gccatttgac aaccacgttt ttaccaatg ggatagccgc	8340
taaattaact tcaactttag ccatagccaa aacatcggca gtacgggtca tagatgaaat	8400
aaaggtttct acggttgatt tggcacctgc agatgacaaa agaccatag caccgacct	8460
aaagtaagca taagaacggc ctttatcagc atcgttattt tcctttaaaa cgtcatcaaa	8520
atttggggtc ctgtacgtgg atttgctagc cagcaaagat tgagaaatca ggtaccacgg	8580
ctcctcgctg cagacctgcg agcagggaaa cgtcccctc acagtcgctg tgaattgtcc	8640
ccacgccgcg cccctgtaga gaaatataaa aggttaggat ttgccactga ggttcttctt	8700
tcatatactt ccttttaaaa tcttgctagg atacagttct cacatcacat ccgaacataa	8760
acaacctagg gtaaggaaaa gactcacgtt tcgaggccgc gattaaattc caacatggat	8820
gctgatttat atgggtataa atgggctcgc gataatgtcg ggcaatcagc tgcgacaatc	8880
tatcgattgt atgggaagcc cgatgcgcca gagttgttcc tgaacatgg caaaggtagc	8940
gttgccaatg atgttacaga tgagatggtc agactaaact ggctgacgga atttatgcct	9000
cttccgacca tcaagcattt tatccgtact cctgatgatg catggttact caccactgcg	9060
atccccggca aaacagcatt ccaggatta gaagaatc ctgattcagc tgaaaatatt	9120

ES 2 651 076 T3

gttgatgccc tggcagtgtt cctgcgccgg ttgcattcga ttctgtttg taattgtcct 9180  
 tttaacagcg atcgcgtatt tcgtctcgtt caggcgcaat cacgaatgaa taacggtttg 9240  
 gttgatgcca gtgattttga tgaocgagct aatggctggc ctggtgaaca agtctggaaa 9300  
 gaaatgcata agcttttgcc attctcaccg gattcagtcg tcaactcatgg tgattttctca 9360  
 ctgataaacc ttatttttga cgaggggaaa ttaataggtt gtattgatgt tggacgagtc 9420  
 ggaatcgcag accgatacca ggatcttgcc atcctatgga actgcctcgg tgagttttct 9480  
 ccttcattac agaaaaggct ttttcaaaa tatggatttg ataactcctga tatgaataaa 9540  
 ttgcagtttc atttgatgct cgatgagttt ttctaatacag tactgacaat aaaaagattc 9600  
 ttgttttcaa gaacttgta tttgtatagt ttttttatat thtagttgtt ctattttaat 9660  
 caaatgttag cgtgatttat attttttttc gcctcgacat catctgccc gatgcgaagt 9720  
 taagtgcgca gaaagtaata tcatgcgtca atcgtatgtg aatgctggtc gctatactgc 9780  
 tgtcgattcg ataactaacc cgccatccag ggtaccatcc tttgtttgtt tccgggtgta 9840  
 caatatggac ttctctttt ctggcaacca aaccataca tccggattcc tataatacct 9900  
 tcgttggctc ccctaaccatg tagtgggcgg aggggagata tacaatagaa cagataccag 9960  
 acaagacata atgggctaaa caagactaca caaattacac tgctcattg atggtggtac 10020  
 ataacgaact aatactgtag ccctagactt gatagccatc atcatatcga agtttcaacta 10080  
 ccctttttcc atttgccatc tattgaagta ataataggcg catgcaactt cttttctttt 10140  
 tttttctttt ctctctcccc cgttgtttgc tcaccatata cgcaatgaca aaaaaaatga 10200  
 tggaagacac taaaggaaaa aattaacgac aaagacagca ccaacagatg tcgttgttcc 10260  
 agagctgatg aggggtatct tcgaacacac gaaacttttt ccttccttca ttcacgcaca 10320  
 ctactctcta atgagcaacg gtatacggcc ttctctccag ttacttgaat ttgaaataaa 10380  
 aaaagtttgc cgctttgcta tcaagtataa atagacctgc aattattaat cttttgtttc 10440  
 ctcgtcattg ttctcgttcc ctttcttctt tgtttctttt tctgcacaat atttcaagct 10500  
 ataccaagca tacaatcaac tatctcatat acaatgcctc aatcctggga agaactggcc 10560  
 gctgataagc ggcgccgctt cgcaaaaacc atcctgatg aatggaaagt ccagacgctg 10620  
 cctgcggaag acagcgttat tgatttcca aagaaatcgg ggatcctttc agaggccgaa 10680  
 ctgaagatca cagaggcctc cgctgcagat cttgtgtcca agctggcggc cggagagttg 10740  
 acctcgggtg aagttacgct agcattctgt aaacgggcag caatcgcaca gcagttaaca 10800  
 aactgcgcc acgagttctt ccctgacgcc gctctcgcgc aggcaaggga actcgtgaa 10860  
 tactacgcaa agcacaagag acccgttggc ccaactccatg gcctcccat ctctctcaaa 10920  
 gaccagcttc gagtcaaggc ctacgaaaca tcaatgggct acatctcatg gctaaacaag 10980  
 tacgacgaag gggactcgtt tctgacaacc atgctccgca aagccggtgc cgtcttctac 11040  
 gtcaagacct ctgtcccga gacctgatg gtctgcgaga cagtcaaca catcatcggg 11100

ES 2 651 076 T3

cgcaccgtca acccacgcaa caagaactgg tcgtgcggcg gcagttctgg tggtaggggt 11160  
 gcgatcgttg ggattcgtgg tggcgtcacc ggtgtaggaa cggatatacg tggetcgatt 11220  
 cgagtgccgg ccgcgttcaa ctctctgtac ggtctaagac cgagtcattg gcggtgccc 11280  
 tatgcaaaga tggcgaacag catggagggt caggagacgg tgcacagcgt tgtcgggccc 11340  
 attacgcaat ctgttgagga cctccgcctc ttacacaaat ccgtcctcgg tcaggagcca 11400  
 tggaaatacg actccaaggt catccccatg ccctggcgcc agtccgagtc ggacattatt 11460  
 gcctccaaga tcaagaacgg cgggctcaat atcggctact acaacttcga cggcaatgct 11520  
 ctccacacc ctctatcct gcgcggcgtg gaaaccaccg tcgccgcaat cgcctaacgc 11580  
 ggtcacaccg tgaccccggt gacgccatac aagcacgatt tcggccacga tctcatctcc 11640  
 catatctacg cggctgacgg cagcgcggac gtaatgcgcg atatcagtc atccggcgag 11700  
 ccggcgattc caaatatcaa agacctactg aaccgaaca tcaaagctgt taacatgaac 11760  
 gagctctggg acacgcatct ccagaagtgg aattaccaga tggagtacct tgagaaatgg 11820  
 cgggagcgtg aagaaaaggc cgggaaggaa ctggacgcca tcatcgccc gattacgcct 11880  
 accgctgcgg tacggcatga ccagttccgg tactatgggt atgcctctgt gatcaacctg 11940  
 ctggatttca cgagcgtggt tgttccggtt acctttgcgg ataagaacat cgataagaag 12000  
 aatgagagtt tcaaggcggg tagtgagctt gatgccctcg tgcaggaaga gtatgatccg 12060  
 gaggcgtacc atggggcacc ggttgcagtg caggttatcg gacgggagact cagtgaagag 12120  
 aggacgttgg cgattgcaga ggaagtgggg aagtgtctgg gaaatgtggt gactccatag 12180  
 gtcgagaatt tatacttaga taagtatgta cttacaggta tatttctatg agatactgat 12240  
 gtatacatgc atgataatat ttaaaccggtt attagtgccg attgtcttgt gcgataatga 12300  
 cgttctctac aaagcaatac acttaccacc tattacatgg gccaaagaaa tattttcgaa 12360  
 ctgttttaga atattagcac agagtatatg atgatatacg ttagattatg catgattcat 12420  
 tcctacaact ttttcgtagc ataaggatta attacttggg tgccaataaa aaaaaaaaaa 12480  
 atcgagaaaa tttcagcatg ctcagaaaca attgcagtg atcaaagtaa aaaaaagatt 12540  
 ttcgctacat gttccttttg aagaaagaaa atcatggaac attagattta caaaaattta 12600  
 accaccgctg attaacgatt agaccgttaa gcgcacaaca ggttattagt acagagaaag 12660  
 cattctgttg tgttgccccg gactttcttt tgcgacatag gtaaatacga taccatcata 12720  
 ctatcttttc caatgactcc ctaaagaaag actcttcttc gatgttgat acgttgagc 12780  
 ataggccaag aattgtggct tgagatctag attacgtgga agaaaggtag taaaagtagt 12840  
 agtataagta gtaaaaagag gtaaaaagag aaaaccggct acatactaga gaagcacgta 12900  
 cacaaaaact cataggcaat tcatcatacg acagtttctt gatgcattat aatagtgtat 12960  
 tagatatttt cagaaatatg catagaacct cttcttgcct ttacttttta tacatagaac 13020

ES 2 651 076 T3

attggcagat ttacttacac tactttgttt ctaoqccatt tcttttgttt tcaaacetta 13080  
 gacaagttgt tgagaaccgg actactaaaa agcaatgttc ccaactgaaa tcatgtacct 13140  
 gcaggataat aaccccctaa ttctgcatcg atccagtatg tttttttttc totactcatt 13200  
 tttacctgaa gatagagcct ctaaaacaaa aaaaatcagc gattacatgc atattgtgtg 13260  
 ttctagaatt gcgcatcacc agatcgccat tacaatgtat gcaggcaaat atttctcaga 13320  
 atgaaaaata gagaaaagga aacgaaaatt ctgtaagatg ccttcgaaga gatttctoga 13380  
 tatgcaaggc gtgcatcagg gtgatccaaa ggaactcgag agagagggcg aaaggcaatt 13440  
 taatgcattg cttctccatt gacttctagt tgagcggata agttcggaaa tgtaagtcac 13500  
 agctaatagac aatccactt taggtttcga ggcactattht aggcataaag acgagtgggg 13560  
 aaataacaaa cgctcaaaca tattagcata taccttcaaa aatgggaat agtatataac 13620  
 cttccgggtc gttaataaat caaatctttc atctagttct cttagattt caatattttg 13680  
 ctttcttgaa gaaagaatct actctcctcc cccattcgca ctgcaaaagt agcttgccac 13740  
 tggccgtcgt tttacaacgt cgtgactggg aaaaccctgg ccttacccaa cttaatcgcc 13800  
 ttgcagcaca tccccctttc gccagctggc gtaatagcga agaggcccgcc accgatcgcc 13860  
 cttcccaaca gttgocgagc ctgaatggcg aatgggaaat tgtaaacgtt aatattttgt 13920  
 taaaattcgc gttaaatttt tgttaaatca gctcattttt taaccaatag gcogaaatcg 13980  
 gcaaaatccc ttataaatca aaagaataga ccgagatagg gttgagtgtt gttccagttt 14040  
 ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga aaaaccgtct 14100  
 atcagggcga tggcccacta cgtgaaccat caccctaac aagttttttg gggtcgaggt 14160  
 gccgtaaagc actaaatcgg aaccctaaag ggagcccccg atttagagct tgacggggaa 14220  
 agccggcgaa cgtggcgaga aaggaaggga agaaagcga aggagcgggc gctagggcgc 14280  
 tggcaagtgt agcggtcacg ctgcgcgtaa ccaccacacc cgccgcgctt aatgcgcgcg 14340  
 tacagggcgc gtcaggtggc acttttcggg gaaatgtgcg cggaaccctt atttgtttat 14400  
 ttttctaaat acattcaaat atgtatccgc tcatgagaca ataaccctga taaatgcttc 14460  
 aataatattg aaaaaggaag agtatgagta ttcaacattt ccgtgtcgcc cttattccct 14520  
 tttttgoggc attttgctt cctgtttttg ctcaccacga aacgctgggt aaagtaaaag 14580  
 atgctgaaga tcagttgggt gcacgagtgg gttacatcga actggatctc aacagcggta 14640  
 agatcctga gagttttcgc cccgaagaac gttttccaat gatgagcact tttaaagttc 14700  
 tgctatgtgg cgcggtatta tcccgtattg acgccgggca agaccaactc ggtcgcgcga 14760  
 tacactattc tcagaatgac ttggttgagt actcaccagt cacagaaaag catcttacgg 14820  
 atggcatgac agtaagagaa ttatgcagtg ctgccataac catgagtgat aacactgcgg 14880  
 ccaacttact tctgacaacg atcggaggac cgaaggagct aaccgctttt ttgcacaaca 14940  
 tgggggatca tgtaactcgc cttgatcgtt gggaaaccgga gctgaatgaa gccataccaa 15000

ES 2 651 076 T3

acgacgagcg tgacaccacg atgcctgtag caatggcaac aacgttgcgc aaactattaa 15060  
 ctggcgaact acttagtcta gcttcccggc aacaattaat agactggatg gaggcggata 15120  
 aagttgcagg accacttctg cgctcggccc ttccggctgg ctggtttatt gctgataaat 15180  
 ctggagccgg tgagcgtggg tctcgcggta tcattgcagc actggggcca gatggttaagc 15240  
 cctcccgat cgtagttatc tacacgacgg ggagtcaggc aactatggat gaacgaaata 15300  
 gacagatcgc tgagataggt gcctcactga ttaagcattg gtaactgtca gaccaagttt 15360  
 actcatatat actttagatt gatttaaaac ttcattttta atttaaaagg atctaggtga 15420  
 agatcctttt tgataatctc atgaccaaaa tccttaacg tgagttttcg ttccactgag 15480  
 cgtcagaccc cgtagaaaag atcaaaggat cttcttgaga tccttttttt ctgcgcgtaa 15540  
 tctgctgctt gcaaacaaaa aaaccaccgc taccagcggg ggtttgttt cccgatcaag 15600  
 agtaccacc tctttttccg aaggttaactg gcttcagcag agcgcagata ccaaatactg 15660  
 tccttctagt gtagccgtag ttagccacc acttcaagaa ctctgtagca ccgcctacat 15720  
 acctcgctct gctaactctg ttaccagtgg ctgctgccag tggcgataag tcgtgtctta 15780  
 ccgggttga ctcaagacga tagttaccgg ataaggcgca gcggtcgggc tgaacggggg 15840  
 gttcgtgcac acagcccagc ttggagcgaa cgacctacac cgaactgaga tacctacagc 15900  
 gtgagcattg agaaagcgc acgcttccc aagggagaaa ggcggacagg tatccggtaa 15960  
 gcggcagggg cggaacagga gagcgcacga gggagcttc aggggaaac gcctggtatc 16020  
 tttatagtcc tgtcgggtt cgccacctct gacttgagcg tcgattttt tgatgctcgt 16080  
 cagggggggc gagcctatgg aaaaacgcca gcaacgcggc ctttttacgg ttctggcct 16140  
 tttgctggcc ttttgctcac atgttctttc ctgcggtatc ccctgattct gtggataacc 16200  
 gtattaccgc ctttgagtga gctgataccg ctgcgcgag ccgaacgacc gagcgcagcg 16260  
 agtcagtgag cgaggaagcg gaagagcgc caatacgcaa accgcctctc cccgcgcgtt 16320  
 ggccgattca ttaatgcagc tggcacgaca ggtttccga ctggaaagcg ggcagtgagc 16380  
 gcaacgcaat taatgtgagt tagctcactc attagccacc ccaggcttta cactttatgc 16440  
 ttccggctcg tatgttgtgt ggaattgtga gcggataaca atttcacaca ggaaacagct 16500  
 atgacatgat tacgaattta atacgactca caatagggaa ttagcttgcg cgaaattatt 16560  
 ggctttttt ttttttaaat 16580

<210> 15

<211> 24

5

<212> ADN

<213> Secuencia artificial

<220>

<223> Cebador sintético

10

# ES 2 651 076 T3

<400> 15  
ccaaggcagc ggtacatcaa gtag 24

5 <210> 16  
<211> 23  
<212> ADN  
<213> Secuencia artificial

10 <220>  
<223> Cebador sintético

<400> 16  
tgcacatggt gtccatcaag atg 23

15 <210> 17  
<211> 24  
<212> ADN  
<213> Secuencia artificial

20 <220>  
<223> Cebador sintético

<400> 17  
ggaaacagct atgacatgat tacg 24

25 <210> 18  
<211> 23  
<212> ADN  
<213> Secuencia artificial

30 <220>  
<223> cebador sintético

<400> 18  
gtagcgaaat catgtattgc acc 23

35 <210> 19

ES 2 651 076 T3

<211> 120

<212> ADN

<213> Secuencia artificial

5 <220>

<223> cebador sintético

<400> 19

10           tttctcatgg tagcgctgt gcttcgggta cttctaagga agtccacaca aatcaagatc           60  
              cgttagacgt ttcagcttcc aaaacagaag aatgtgagac gctggggccc gagacctaac           120

<210> 20

<211> 120

<212> ADN

15 <213> Secuencia artificial

<220>

<223> cebador sintético

20 <400> 20

              ggaggtggta ctgaagcagg ttgaggagag gcatgatggg ggttctctgg aacagctgat           60  
              gaagcaggtg ttgttgtctg ttgagagtta gccttagtgt tgtggagagg tgacttcatg           120

<210> 21

25 <211> 120

<212> ADN

<213> Secuencia artificial

<220>

30 <223> cebador sintético

<400> 21

              tttctcatgg tagcgctgt gcttcgggta cttctaagga agtccacaca aatcaagatc           60  
              cgttagacgt ttcagcttcc aaaacagaag aatgtgagag ctcccctcac agacgcggtg           120

ES 2 651 076 T3

<210> 22  
<211> 120  
<212> ADN  
<213> Secuencia artificial

5

<220>  
<223> cebador sintético

<400> 22

10

gaggtggtac tgaagcaggt tgaggagagg catgatgggg gttctctgga acagctgatg 60  
aagcaggtgt tgttgtctgt tgagagttag ccttagtgca aatgacaagt tcttgaaaac 120

<210> 23  
<211> 22  
<212> ADN  
<213> Secuencia artificial

15

<220>  
<223> cebador sintético

20

<400> 23  
caccgtagc cttggcgtaa gc 22

<210> 24  
<211> 21  
<212> ADN  
<213> Secuencia artificial

25

<220>  
<223> cebador sintético

30

<400> 24  
cacttgcga cacgaattgg c 21

35

<210> 25  
<211> 21

<212> ADN

<213> ADN artificial

<220>

5

<223> Cebador sintético

<400> 25

gttacgtcgc ctggacttc g 21

10

<210> 26

<211> 21

<212> ADN

<213> ADN artificial

15

<220>

<223> Cebador sintético

<400> 26

cggcaatacc tgggaacatg g 21

20

**REIVINDICACIONES**

- 5 1. Célula adecuada para la producción de uno o más productos de fermentación a partir de una composición de azúcares que comprende glucosa, galactosa, xilosa, arabinosa y manosa, en la que la célula comprende dos a quince copias de uno o más genes de xilosa isomerasa o dos a quince copias de uno o más genes de xilosa reductasa y xilitol deshidrogenasa, y dos a diez copias de cada uno de *araA*, *araB* y *araD*, en la que estos genes están integrados en el genoma de la célula.
2. Célula según la reivindicación 1, en la que la célula es capaz de convertir el 90 % o más de glucosa, xilosa arabinosa, galactosa y manosa disponible, en un producto de fermentación.
3. Célula según la reivindicación 1, en la que la célula tiene una interrupción o delección del gen GAL80.
- 10 4. Célula según cualquiera de las reivindicaciones 1-3, en la que la célula es del género *Saccharomyces*, opcionalmente de la especie *Saccharomyces cerevisiae*.
5. Célula según cualquiera de las reivindicaciones 1-4, en la que la célula comprende expresados en exceso los genes de PPP *TAL1*, *TKL1*, *RPE1* y *RK11*.
6. Célula según cualquiera de las reivindicaciones 1-5, en la que la célula comprende un gen *XKS1*.
- 15 7. Célula según cualquiera de la reivindicación 1-6, en la que un gen de aldosa reductasa está delecionado.
8. Célula según cualquiera de las reivindicaciones 1-7, en la que todos los genes exógenos a la célula se integran en el genoma de la célula.
9. Célula según cualquiera de las reivindicaciones 1-8, en la que los genes han sido introducidos en la célula por introducción en una célula hospedadora de:
  - 20 a) una agrupación que consiste en los genes *araA*, *araB* y *araD* bajo el control de un promotor constitutivo fuerte
  - b) una agrupación que consiste en los genes de PPP *TAL1*, *TKL1*, *RPE1* y *RK11*, opcionalmente bajo el control de un promotor constitutivo fuerte; y delección de un gen de aldosa reductasa;
  - c) una agrupación que consiste en un gen *xyIA* y un gen *XKS1* bajo el control de promotor constitutivo fuerte;
  - 25 d) una construcción que comprende un gen *xyIA* bajo el control de un promotor constitutivo fuerte, que tiene la capacidad de integrarse en el genoma en múltiples loci;

y evolución adaptativa de la construcción de azúcares mixtos para producir la célula de azúcares mixtos.
10. Célula según la reivindicación 9, en la que la célula es una célula resistente a inhibidor.
11. Célula según la reivindicación 9 o 10, en la que la célula es una célula industrial.
- 30 12. Proceso para la producción de uno o más productos de fermentación a partir de una composición de azúcares que comprende glucosa, galactosa, arabinosa y xilosa, en el que la composición de azúcares se fermenta con una célula según cualquiera de las reivindicaciones 1-11.
13. Proceso según la reivindicación 12, en el que la composición de azúcares se produce a partir de material lignocelulósico por:
  - 35 a) pretratamiento de uno o más materiales lignocelulósicos para producir material lignocelulósico pretratado;
  - b) tratamiento enzimático del material lignocelulósico pretratado para producir la composición de azúcares.
14. Proceso según las reivindicaciones 12 o 13, en el que la fermentación se realiza anaeróbicamente.
15. Proceso según cualquiera de las reivindicaciones 12-14, en el que el producto de fermentación es etanol.

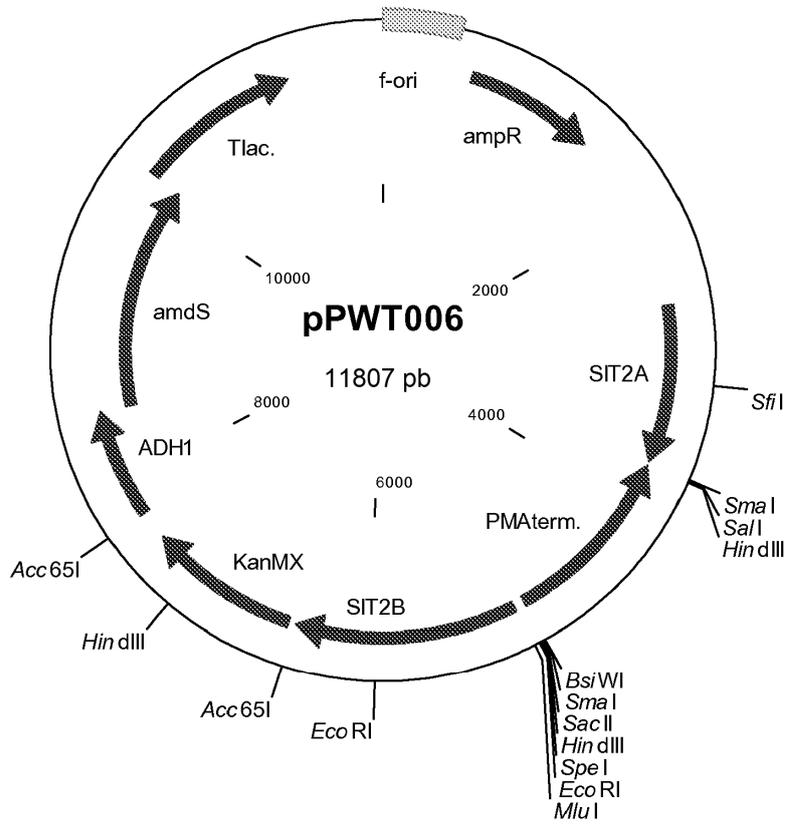


Fig. 1

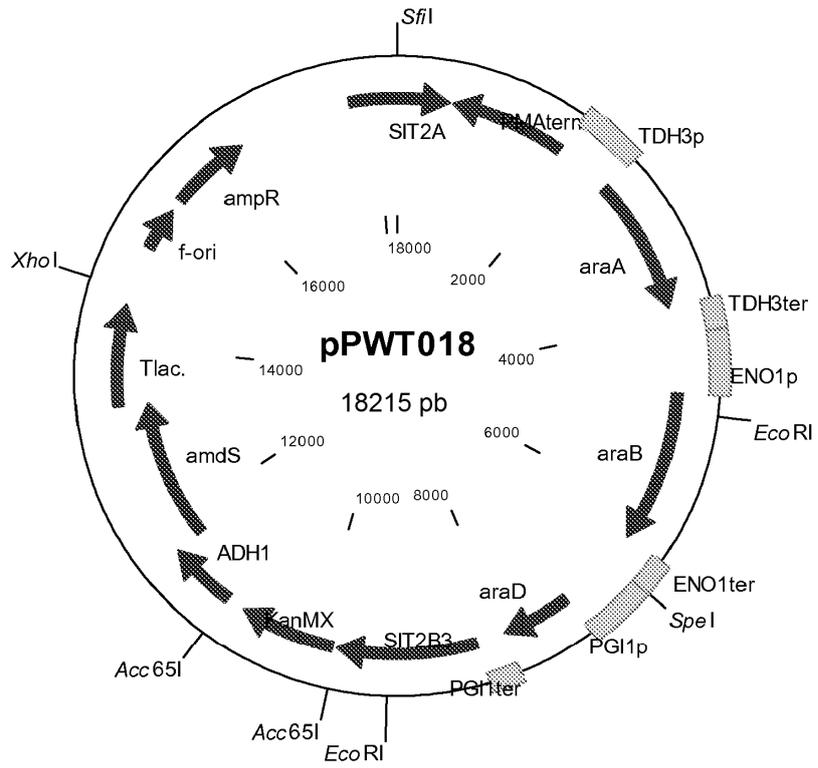


Fig. 2

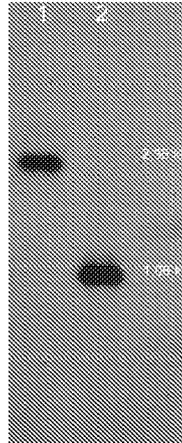


Fig. 3

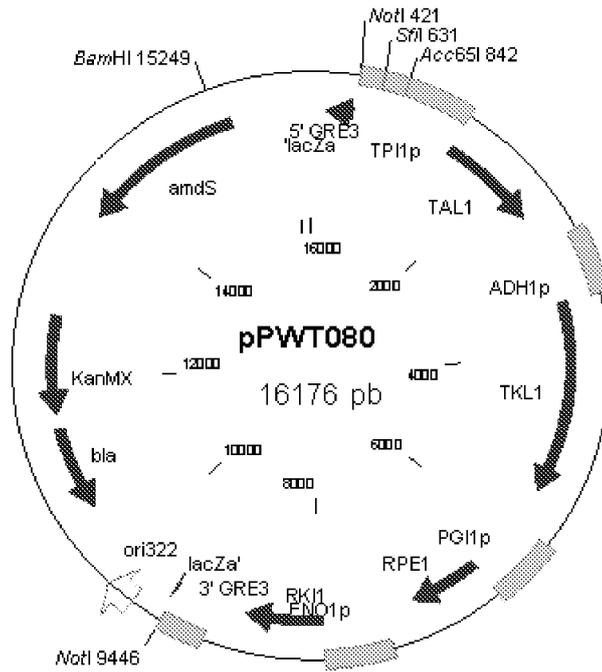


Fig. 4

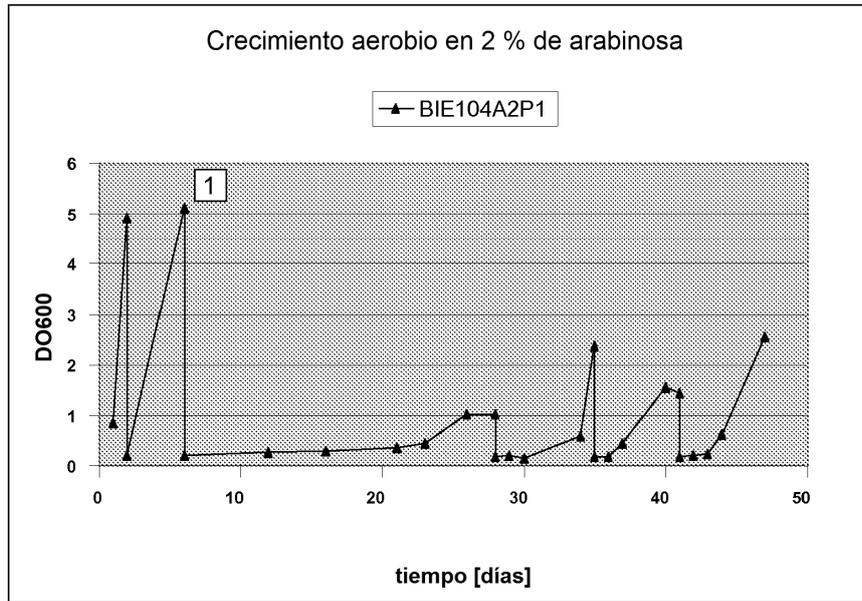


Fig. 5

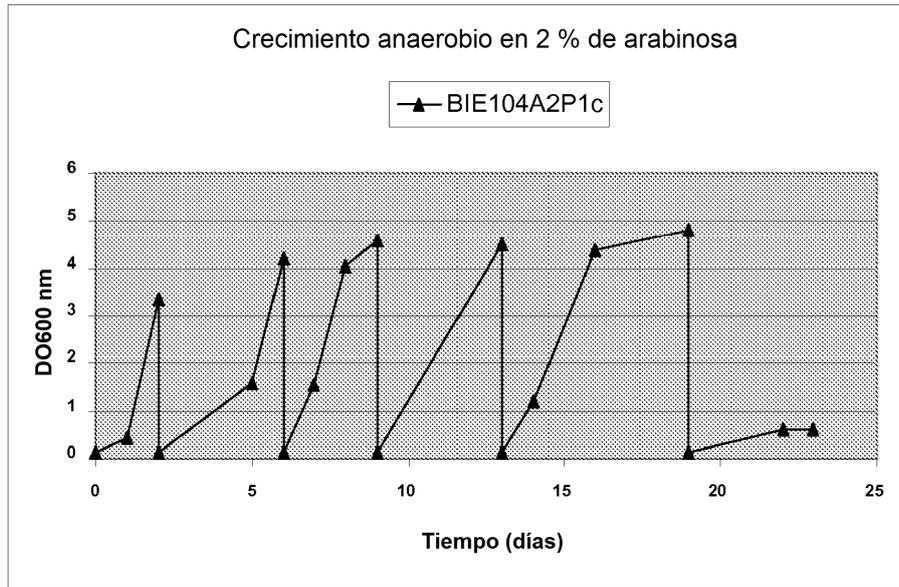


Fig. 6

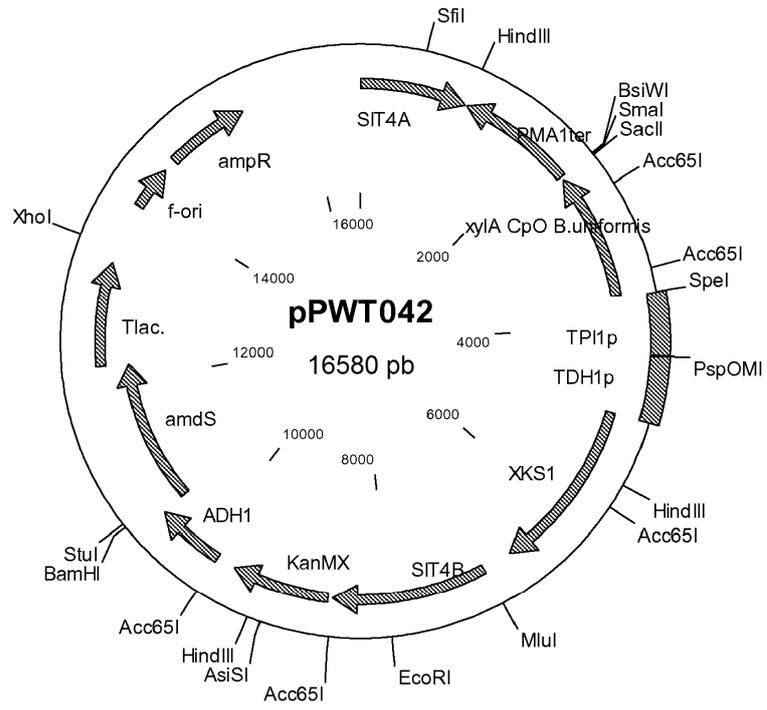


Fig. 7

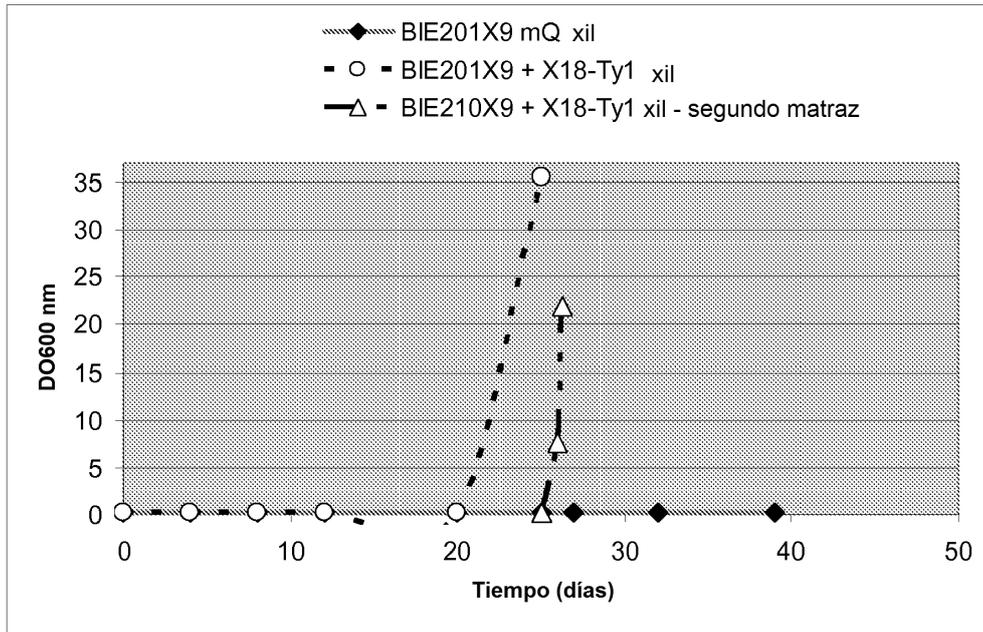


Fig. 8

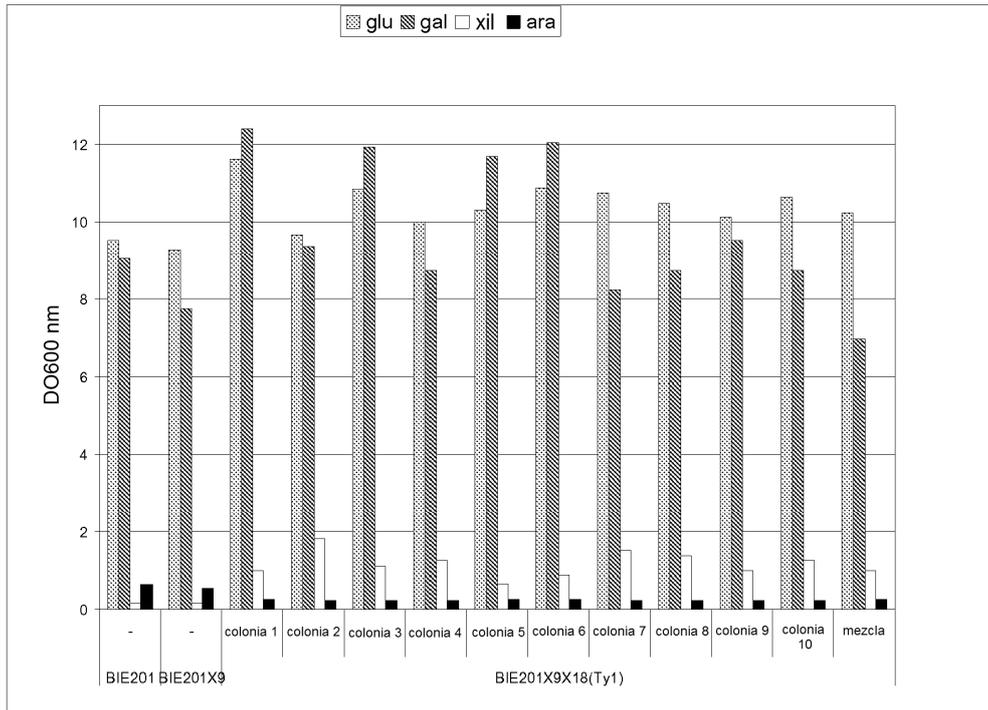


Fig. 9

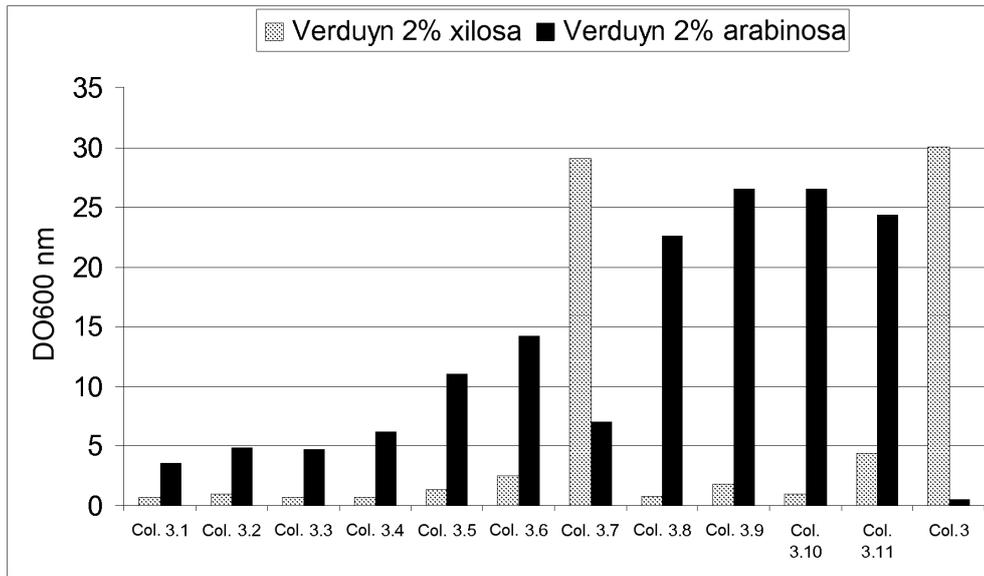


Fig. 10

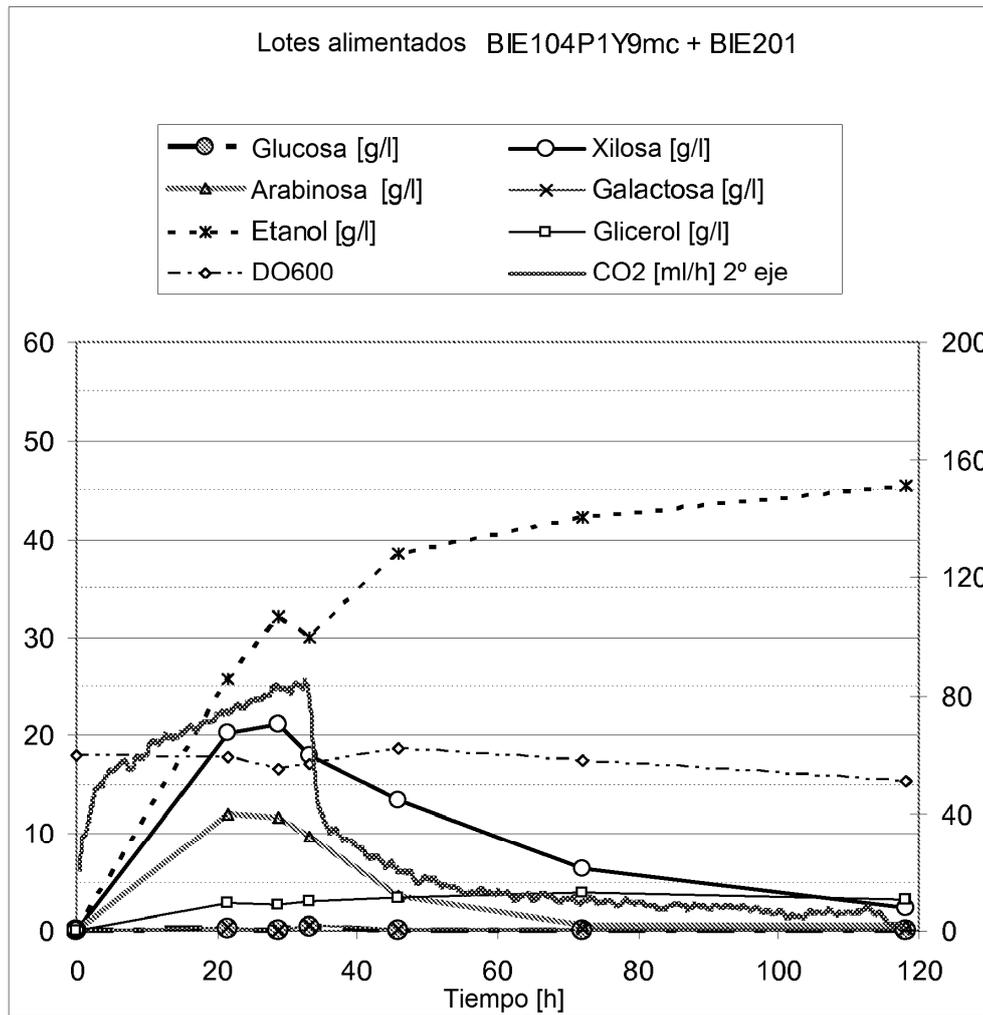


Fig. 11

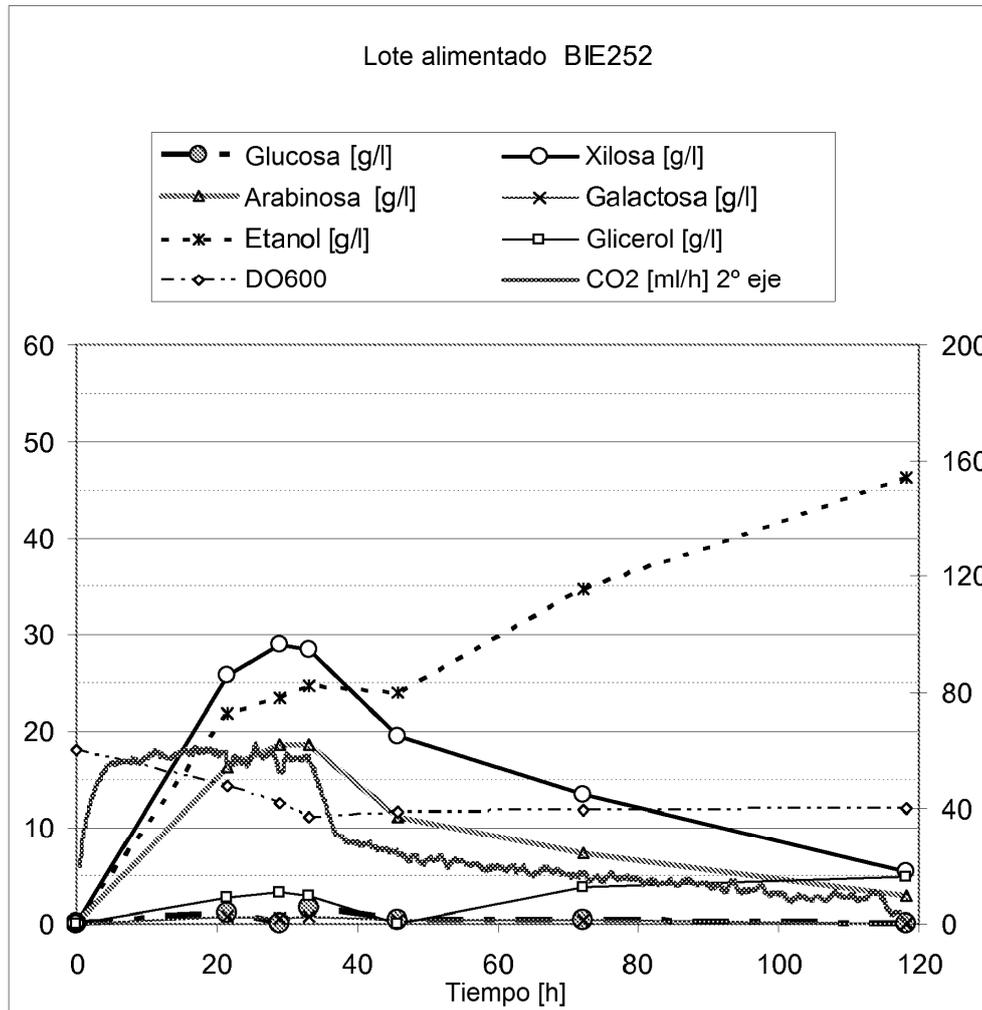


Fig. 12

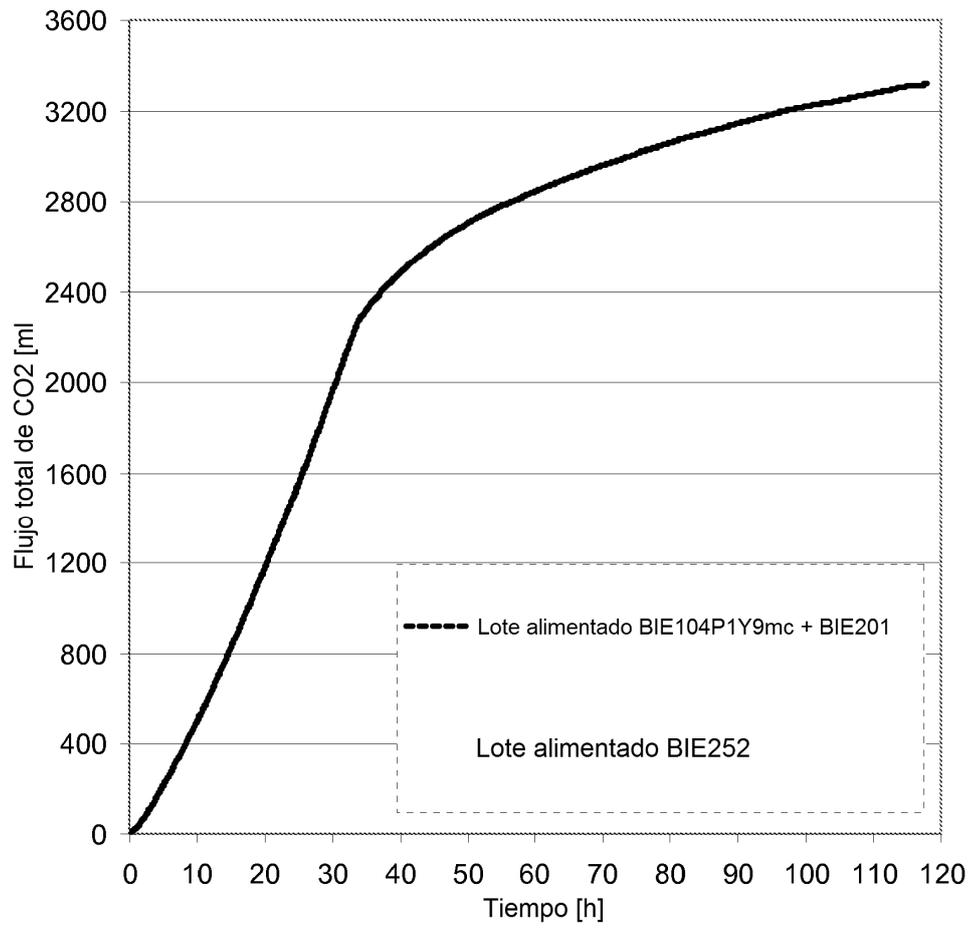


Fig. 13

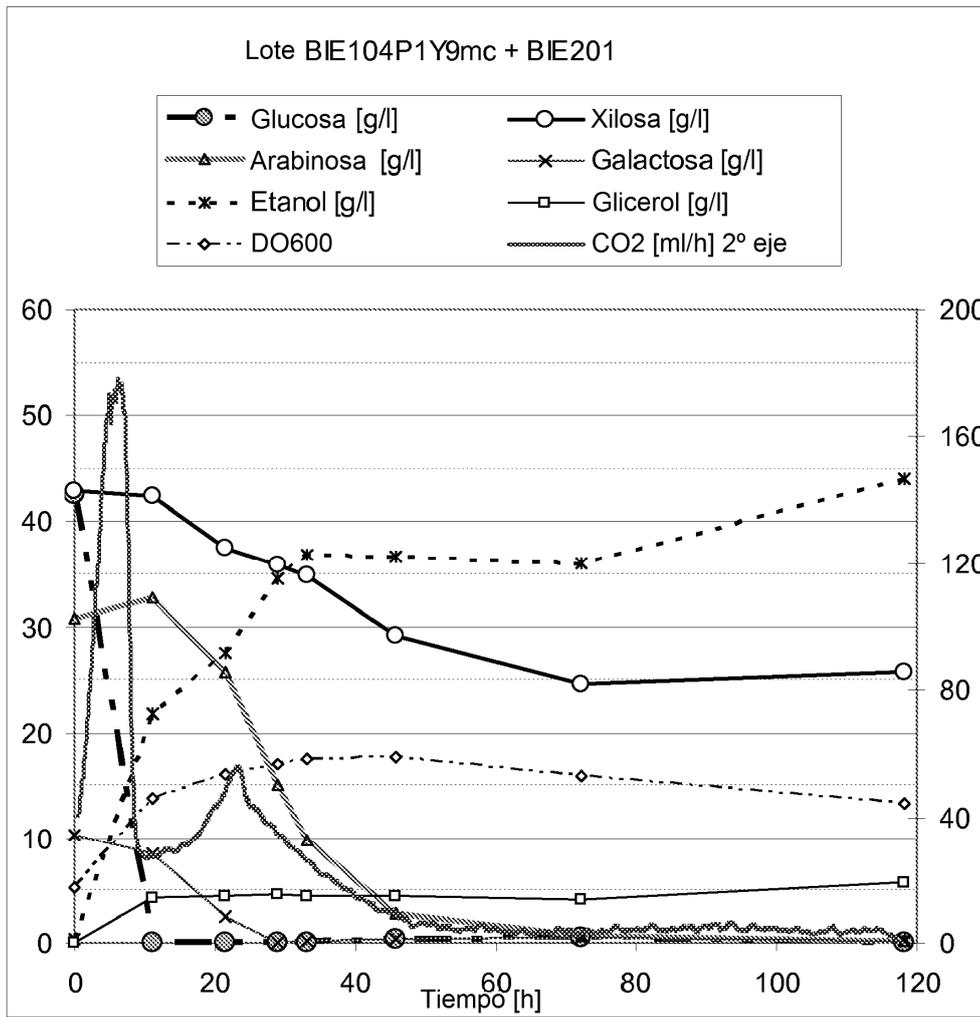


Fig. 14

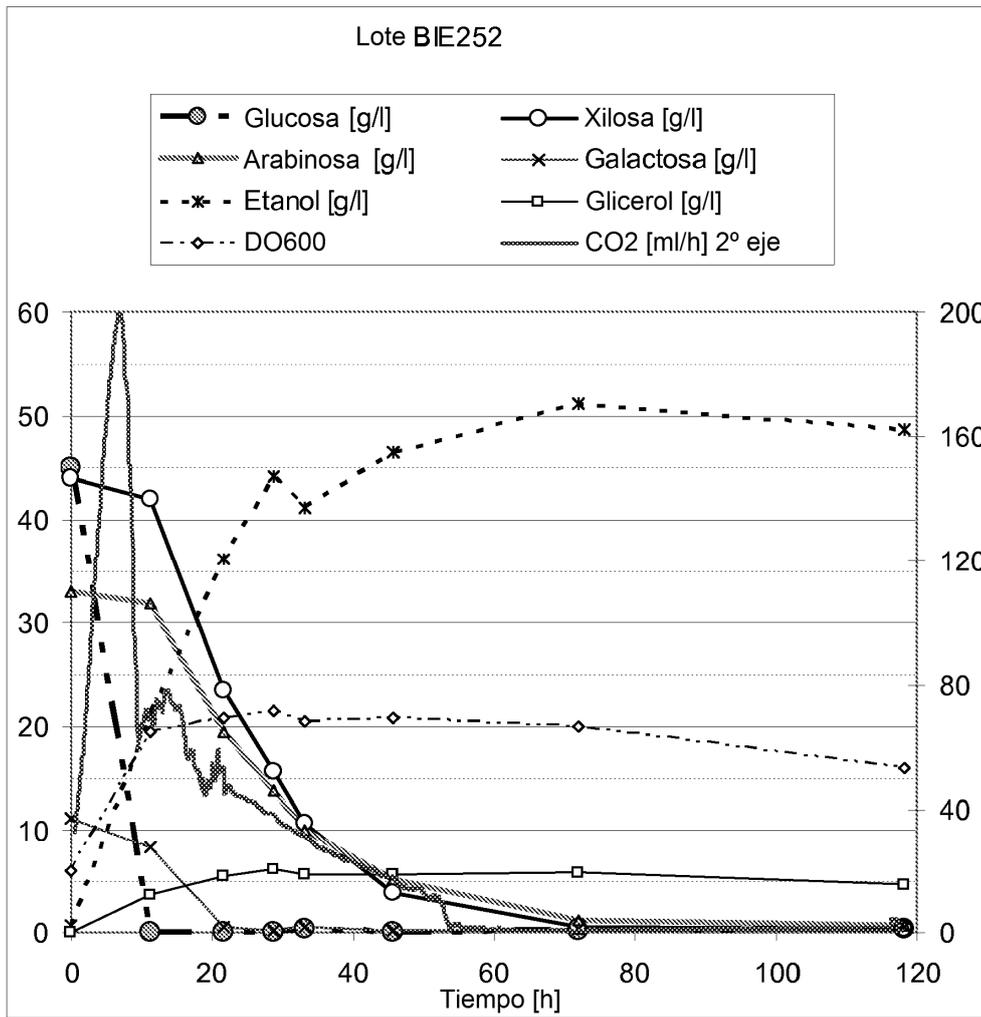


Fig. 15

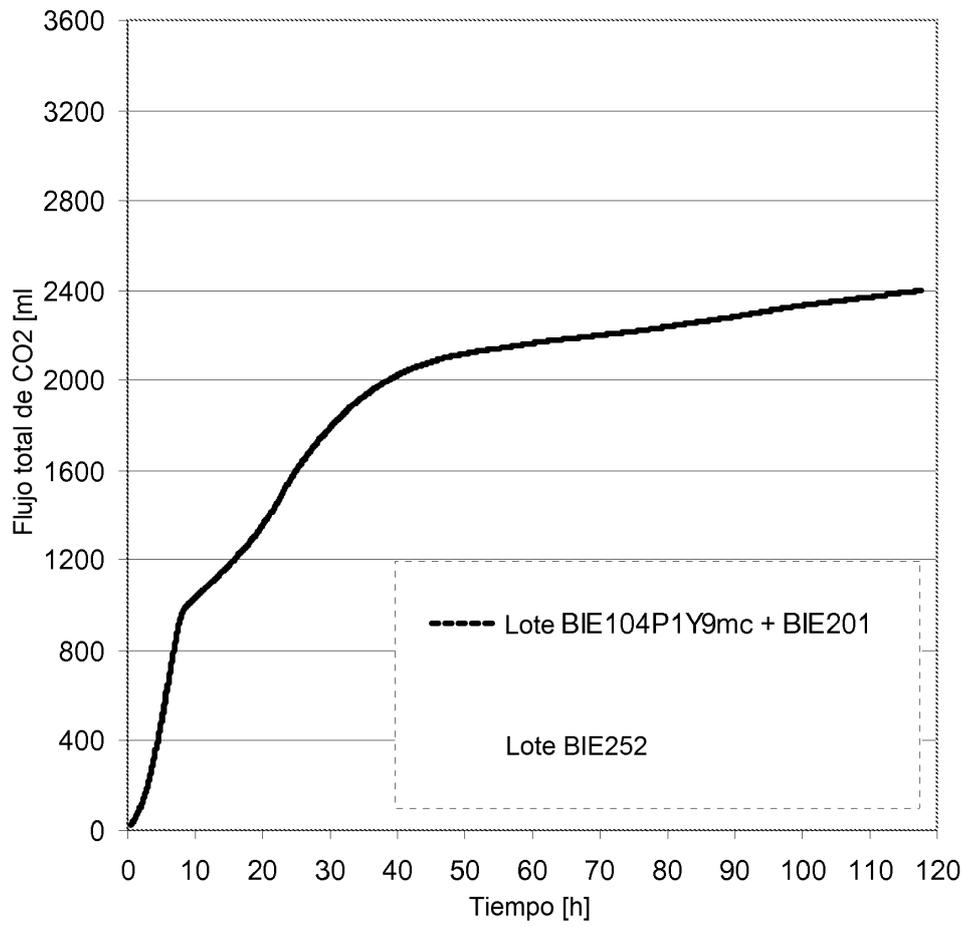


Fig. 16

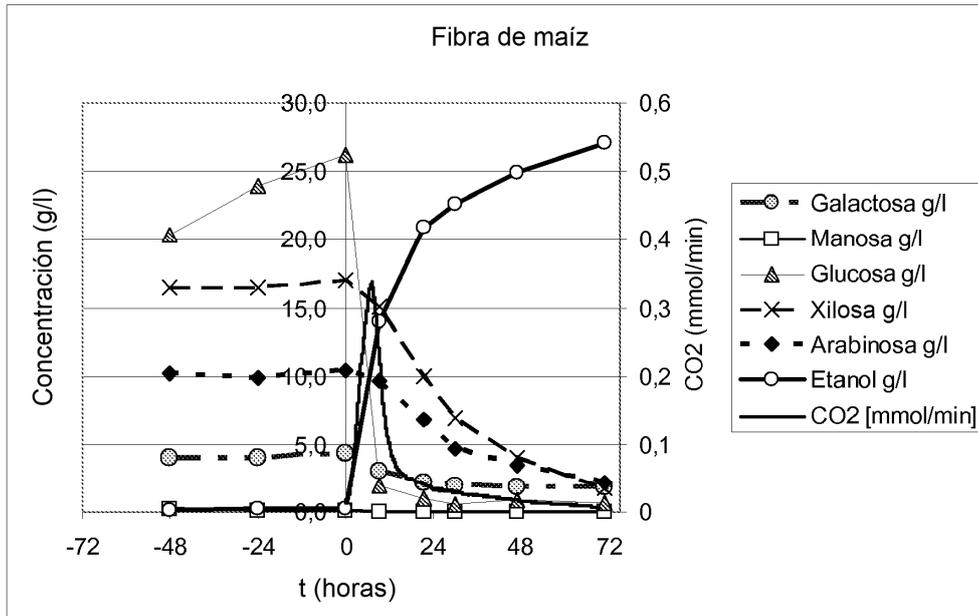


Fig. 17

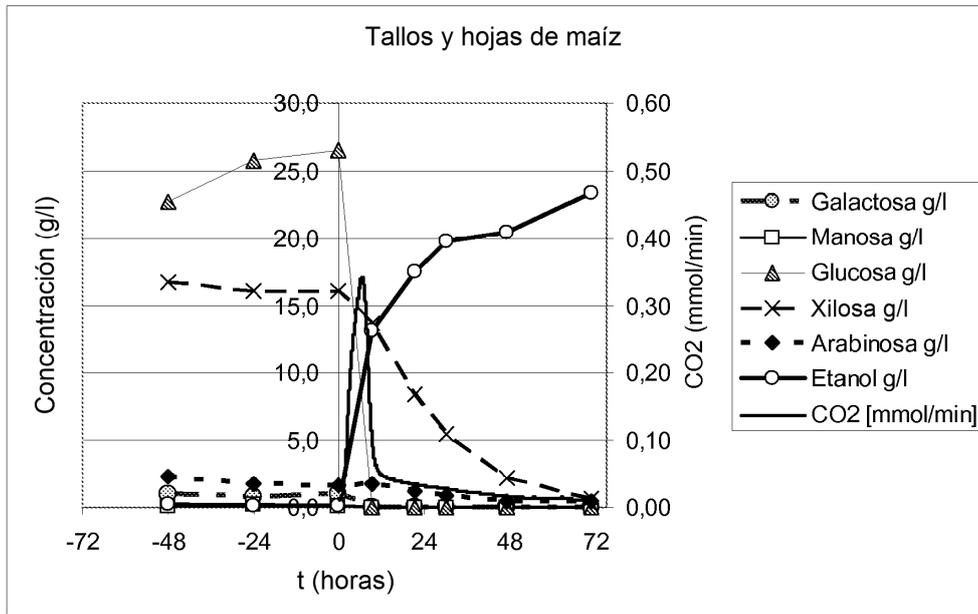


Fig. 18

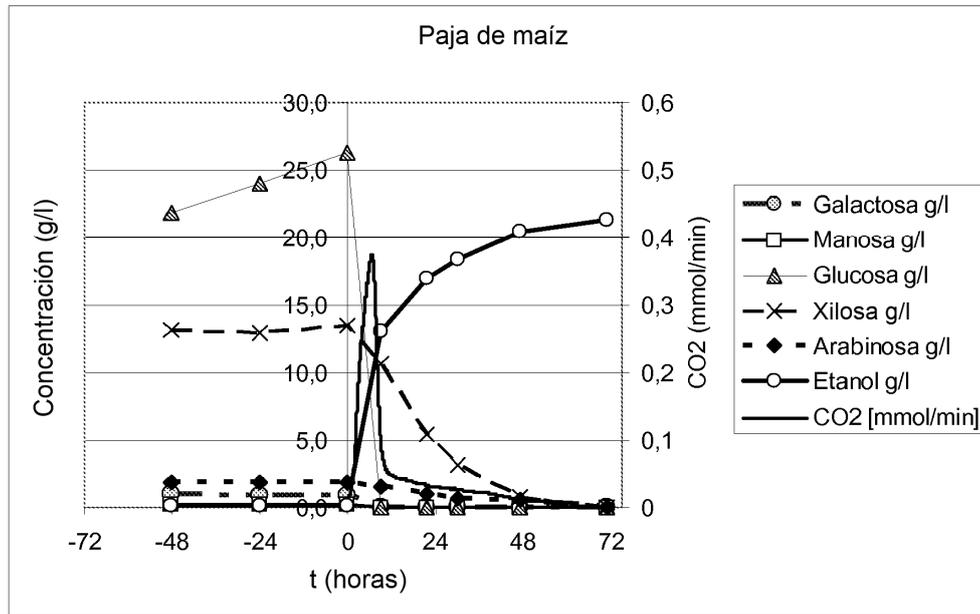


Fig. 19

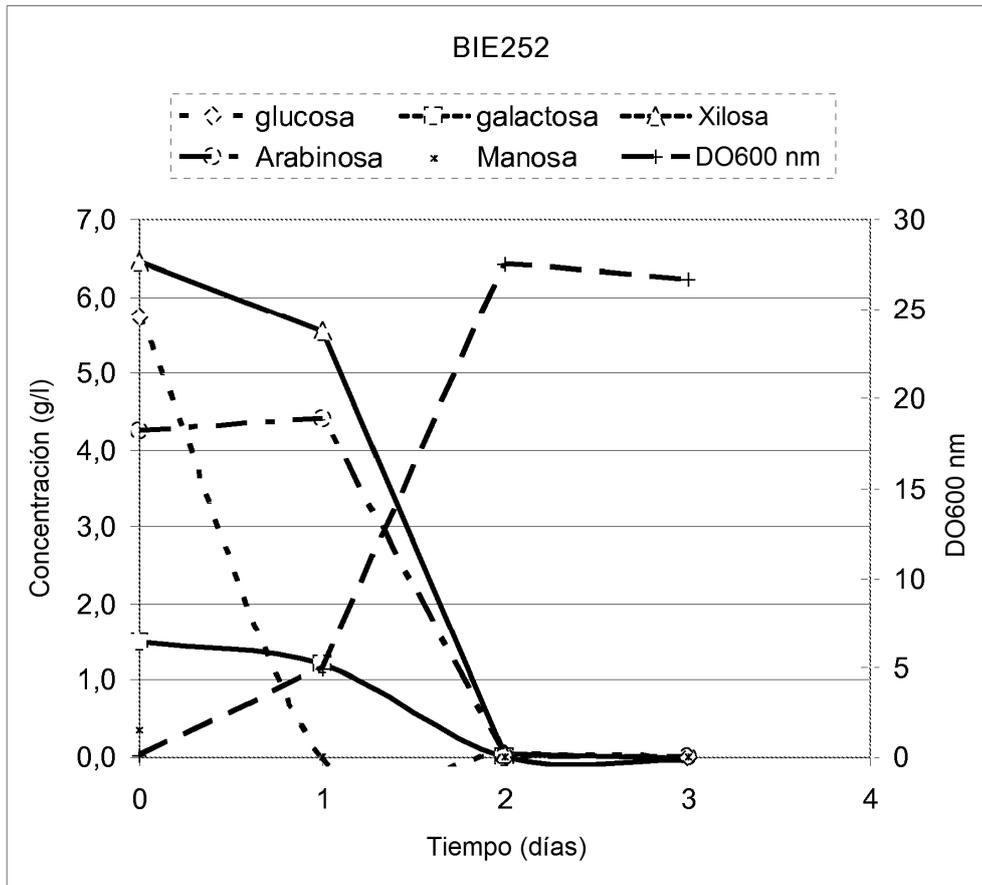


Fig. 20a

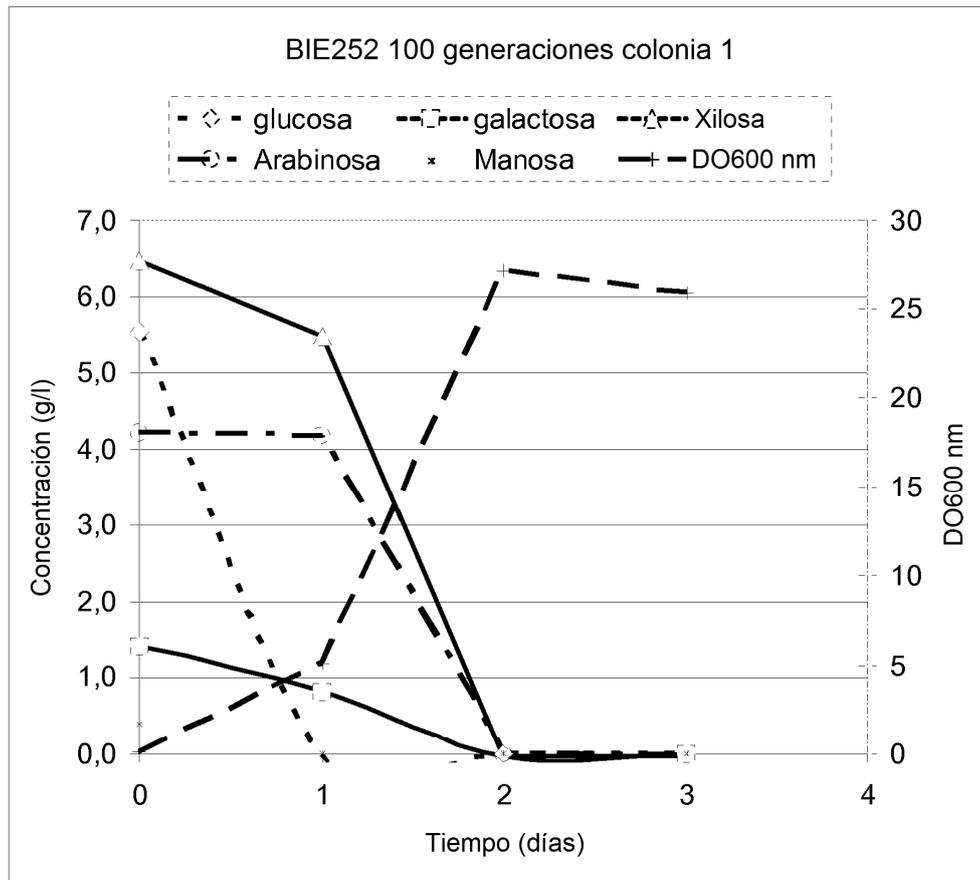


Fig. 20b

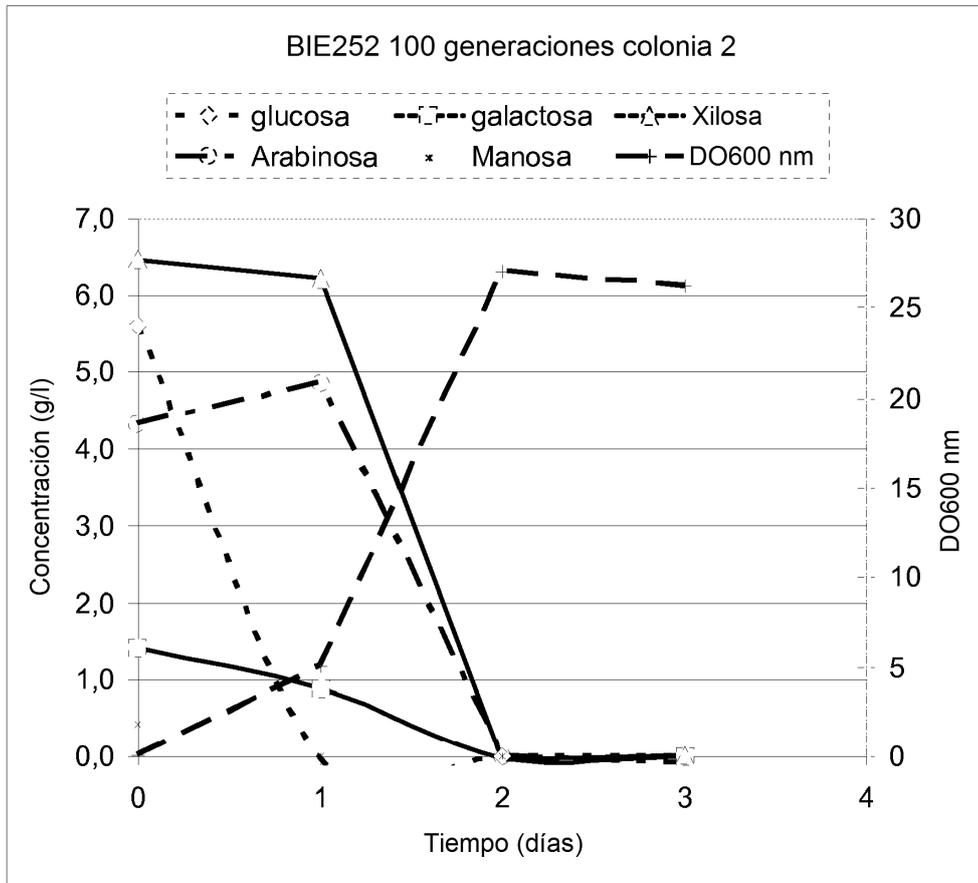


Fig. 20c

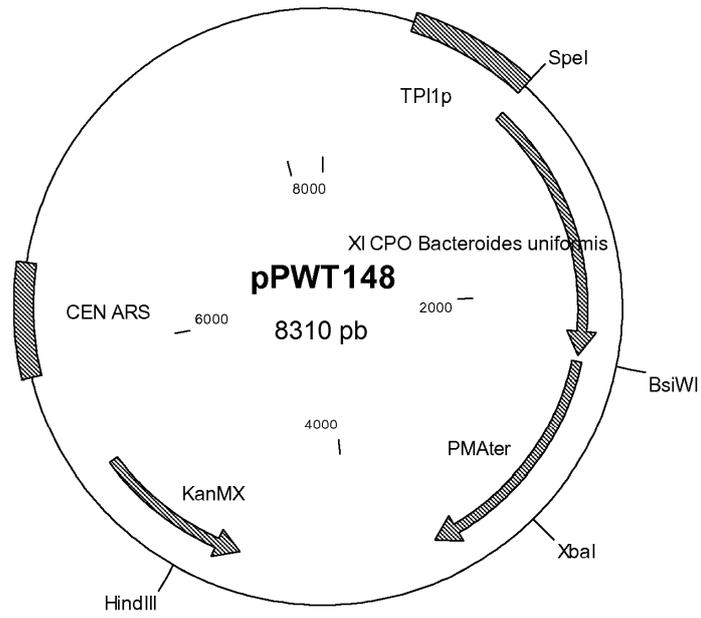


Fig. 21

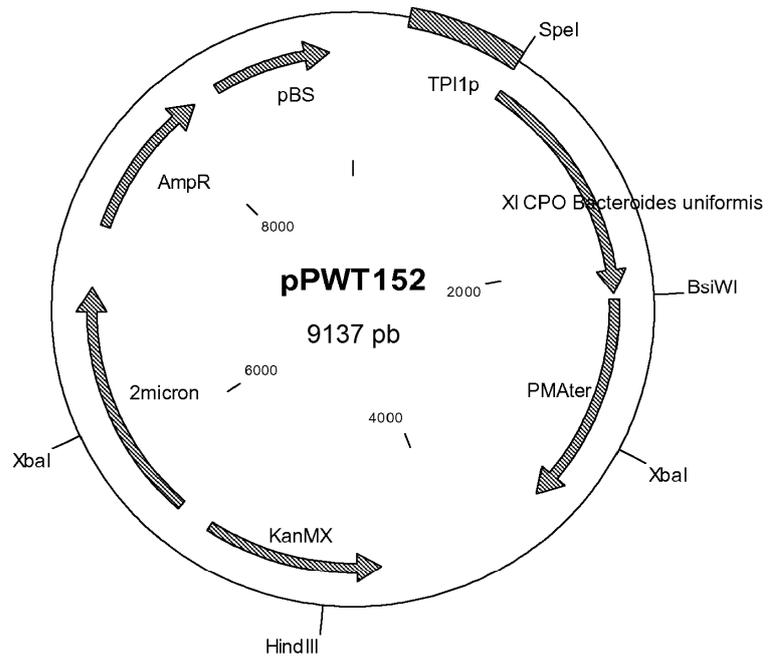


Fig. 22

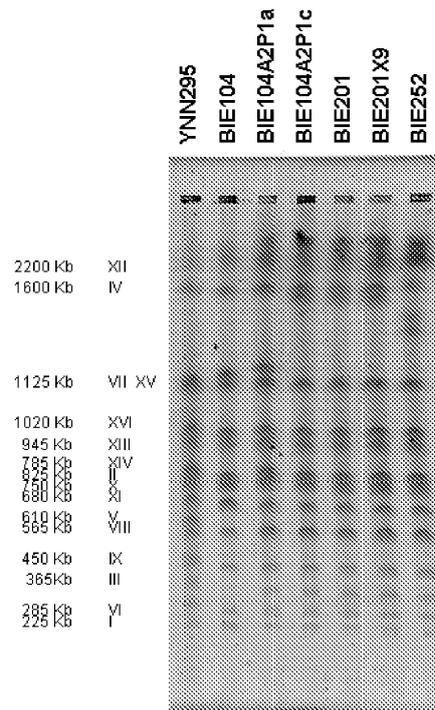


Fig. 23