

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 648 865**

51 Int. Cl.:

<b>C12P 7/06</b>	(2006.01)
<b>C12P 7/08</b>	(2006.01)
<b>C12P 7/10</b>	(2006.01)
<b>C12N 1/18</b>	(2006.01)
<b>C12N 1/22</b>	(2006.01)
<b>C12N 9/02</b>	(2006.01)
<b>C12N 9/04</b>	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **26.11.2012 PCT/NL2012/050841**
- 87 Fecha y número de publicación internacional: **06.06.2013 WO13081456**
- 96 Fecha de presentación y número de la solicitud europea: **26.11.2012 E 12798012 (6)**
- 97 Fecha y número de publicación de la concesión europea: **27.09.2017 EP 2785849**

54 Título: **Cepas de levadura manipuladas para producir etanol a partir de ácido acético y glicerol**

30 Prioridad:

**30.11.2011 EP 11191333  
30.11.2011 US 201161564932 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**08.01.2018**

73 Titular/es:

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

72 Inventor/es:

**DE BONT, JOHANNES ADRIANUS MARIA;  
TEUNISSEN, ALOYSIUS WILHELMUS  
RUDOLPHUS HUBERTUS;  
KLAASSEN, PAUL;  
HARTMAN, WOUTER WILLEM ANTONIUS y  
VAN BEUSEKOM, SHIMAIRA**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

**ES 2 648 865 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

Cepas de levadura manipuladas para producir etanol a partir de ácido acético y glicerol

### Campo de la invención

5 La presente invención se refiere a la manipulación metabólica en levadura. En particular, la invención se refiere a cepas de levadura que han sido manipuladas para producir etanol a partir de ácido acético y glicerol. Además de ácido acético y glicerol, la cepa de levadura también puede consumir hexosas y pentosas para la producción de etanol. La invención se refiere además a los procesos en los que las cepas manipuladas de la invención producen etanol a partir de ácido acético y glicerol, y a partir de al menos una de una hexosa o una pentosa.

10 El proceso de la invención se define por las reivindicaciones adjuntas 1-13, y las células de levadura de la invención se definen además por las reivindicaciones adjuntas 14 y 15.

### Antecedentes de la invención

15 El bioetanol de segunda generación se produce a partir de, por ejemplo, fracciones lignocelulósicas de biomasa de plantas que se hidroliza en azúcares monoméricos libres, tales como hexosas y pentosas, para la fermentación en etanol. Los hidrolizados lignocelulósicos contienen altas cantidades de ácido acético, que es un potente inhibidor de las capacidades fermentativas de los microorganismos usados para la producción de etanol, tales como levaduras.

20 Sonderegger et al. (2004, Appl. Environ. Microbiol. 70: 2892-2897) desvelan la expresión heteróloga de fosfotransacetilasa y acetaldehído deshidrogenasa en la cepa fermentadora de xilosa *Saccharomyces cerevisiae*. En combinación con la fosfocetolasa nativa, Sonderegger et al. crearon una vía de fosfocetolasa funcional que es capaz de una reoxidación neta de la NADH generada por la expresión heteróloga de una xilosa reductasa y xilitol deshidrogenasa que se usan para la utilización de xilosa en la cepa.

25 Guadalupe et al. (2009, Appl. Environ. Microbiol. doi:10.1128/AEM.01772-09) desvelan una cepa de *Saccharomyces cerevisiae* en la que se elimina la producción del producto secundario glicerol por interrupción de los genes de glicerol 3-fosfato deshidrogenasa dependientes de NAD endógenos (*GPD1* y *GPD2*). La expresión del gen *mhpF* de *E. coli*, que codifica la acetaldehído deshidrogenasa dependiente de NAD acetilante, restauró la capacidad de la cepa interrumpida por GPD para crecer anaeróbicamente por adición de suplementos de ácido acético en el medio.

Yu et al. (2010, Bioresour. Technol. 101(11):4157-61. Publicación electrónica 9 de febrero de 2010) desvelan cepas de *Saccharomyces cerevisiae* manipuladas metabólicamente para la producción mejorada de etanol a partir de glicerol por expresión en exceso simultánea de glicerol deshidrogenasa (*GCY*), dihidroxiacetona cinasa (*DAK*) y la proteína de captación de glicerol (*GUP1*).

30 Lee y Dasilva (2006, Metab Eng. 8(1):58-65) desvelan la levadura *Saccharomyces cerevisiae* manipulada para producir 1,2-propanodiol a partir de glicerol introduciendo, entre otros, la expresión de los genes *mgs* y *gldA* de *Escherichia coli*.

35 Un objeto de la presente invención comprende proporcionar levaduras que son capaces de producir etanol a partir de ácido acético y glicerol (y hexosas y pentosas), además de procesos en los que se usan estas cepas para la producción de etanol.

### Descripción de la invención

#### Definiciones

40 Identidad de secuencia se define en el presente documento como una relación entre dos o más secuencias de aminoácidos (polipéptido o proteína) o dos o más secuencias de ácidos nucleicos (polinucleótido), como se determina por comparación de las secuencias. En la materia, "identidad" también significa el grado de relación de secuencias entre secuencias de aminoácidos o de ácidos nucleicos, según sea el caso, como se ha determinado por la coincidencia de las cadenas de tales secuencias. La "similitud" entre dos secuencias de aminoácidos se determina por comparación de la secuencia de aminoácidos y sus sustitutos de aminoácidos conservados de un polipéptido con la secuencia de un segundo polipéptido. La "identidad" y la "similitud" pueden calcularse fácilmente por métodos conocidos. Los términos "identidad de secuencia" o "similitud de secuencia" significan que dos secuencias de (poli) péptidos o dos secuencias de nucleótidos, cuando se alinean de manera óptima, preferentemente sobre la longitud completa (de al menos la secuencia más corta en la comparación) y se maximiza el número de coincidencias y se minimiza el número de huecos, tal como por los programas ClustalW (1.83), GAP o BESTFIT usando los parámetros por defecto, comparten al menos un porcentaje determinado de identidad de secuencia como se define en cualquier parte en el presente documento. GAP usa el algoritmo de alineamiento global de Needleman y Wunsch para alinear dos secuencias sobre su longitud completa, maximizar el número de coincidencias y minimizar el número de huecos. Generalmente, usan parámetros por defecto de GAP, con una penalización por creación de hueco = 50 (nucleótidos) / 8 (proteínas) y penalización por extensión de hueco = 3 (nucleótidos) / 2 (proteínas). Para nucleótidos, la matriz de puntuación por defecto usada es nws gapdna y para proteínas la matriz de

puntuación por defecto es Blosum62 (Henikoff & Henikoff, 1992, PNAS 89, 915-919). Un programa de alineamientos múltiples preferido para alinear secuencias de proteínas desvelado en el presente documento es (1.83) que usa una matriz blosum y parámetros por defecto (penalización por abertura de hueco: 10; penalización por extensión de hueco: 0,05). Es evidente que cuando se dice que secuencias de ARN son esencialmente similares o tienen un cierto grado de identidad de secuencia con secuencias de ADN, la timina (T) en la secuencia de ADN se considera igual al uracilo (U) en la secuencia de ARN. Pueden determinarse alineamientos de secuencias y puntuaciones para el porcentaje de identidad de secuencia usando programas informáticos, tales como the GCG Wisconsin Package, Versión 10.3, disponible de Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121-3752, EE.UU., o el software de fuente abierta Emboss for Windows (versión actual 2.7.1-07). Alternativamente, el porcentaje de similitud o identidad puede determinarse por búsqueda por comparación con bases de datos tales como FASTA, BLAST, etc.

Métodos preferidos para determinar la identidad están diseñados para obtener la mayor coincidencia posible entre las secuencias probadas. Los métodos de determinación de identidad y similitud están codificados en programas informáticos públicamente disponibles. Métodos con programas informáticos preferidos para determinar la identidad y similitud entre dos secuencias incluyen, por ejemplo, el paquete de programas GCG (Devereux, J., et al., *Nucleic Acids Research* 12 (1): 387 (1984)), BestFit, BLASTP, BLASTN y FASTA (Altschul, S. F. et al., *J. Mol. Biol.* 215:403-410 (1990)). El programa BLASTX está públicamente disponible de NCBI y otras fuentes (BLAST Manual, Altschul, S., et al., NCBI NLM NIH Bethesda, MD 20894; Altschul S., et al., *J. Mol. Biol.* 215: 403-410 (1990)). También puede usarse el muy conocido algoritmo de Smith-Waterman para determinar la identidad.

Parámetros preferidos para la comparación de secuencias de polipéptidos incluyen los siguientes: Algoritmo: Needleman y Wunsch, *J. Mol. Biol.* 48:443-453 (1970); matriz de comparación: BLOSSUM62 de Henikoff y Henikoff, *Proc. Natl. Acad. Sci. USA.* 89:10915-10919 (1992); penalización por hueco: 12; y penalización por longitud de hueco: 4. Un programa útil con estos parámetros está públicamente disponible como el programa "Ogap" de Genetics Computer Group, situado en Madison, WI. Los parámetros anteriormente mencionados son los parámetros por defecto para las comparaciones de aminoácidos (junto con sin penalización para huecos terminales).

Parámetros preferidos para la comparación de ácidos nucleicos incluyen los siguientes: Algoritmo: Needleman y Wunsch, *J. Mol. Biol.* 48:443-453 (1970); matriz de comparación matriz: correspondencias=+10, discordancias=0; penalización por hueco: 50; penalización por longitud de hueco: 3. disponible como el programa Gap de Genetics Computer Group, situado en Madison, Wis. Anteriormente se dan los parámetros por defecto para las comparaciones de ácidos nucleicos.

Opcionalmente, en la determinación del grado de similitud de aminoácidos, el experto también puede tener en cuenta las llamadas sustituciones de aminoácidos "conservativas", como será evidente para el experto. Sustituciones de aminoácidos conservativas se refiere a la intercambiabilidad de restos que tienen cadenas laterales similares. Por ejemplo, un grupo de aminoácidos que tienen cadenas laterales alifáticas es glicina, alanina, valina, leucina e isoleucina; un grupo de aminoácidos que tienen cadenas laterales de hidroxilo alifático es serina y treonina; un grupo de aminoácidos que tienen cadenas laterales que contienen amida es asparagina y glutamina; un grupo de aminoácidos que tienen cadenas laterales aromáticas es fenilalanina, tirosina y triptófano; un grupo de aminoácidos que tienen cadenas laterales básicas es lisina, arginina e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Grupos de sustitución de aminoácidos conservativa preferidos son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina y asparagina-glutamina. Variantes sustitucionales de la secuencia de aminoácidos desveladas en el presente documento son aquellas en las que al menos un resto en las secuencias desveladas ha sido eliminado y se ha insertado un resto diferente en su lugar. Preferentemente, el cambio de aminoácido es conservativo. Sustituciones conservativas preferidas para cada uno de los aminoácidos que existen de forma natural son las siguientes: Ala por ser; Arg por lys; Asn por gln o his; Asp por glu; Cys por ser o ala; Gln por asn; Glu por asp; Gly por pro; His por asn o gln; Ile por leu o val; Leu por ile o val; Lys por arg; gln o glu; Met por leu o ile; Phe por met, leu o tyr; Ser por thr; Thr por ser; Trp por tyr; Tyr por trp o phe; y, Val por ile o leu.

Las secuencias de nucleótidos de la invención también pueden definirse por su capacidad para hibridarse con partes de secuencias de nucleótidos específicas desveladas en el presente documento, respectivamente, en condiciones de hibridación moderadas, o preferentemente rigurosas. Condiciones de hibridación rigurosas se definen en el presente documento como condiciones que permiten que una secuencia de ácidos nucleicos de al menos aproximadamente 25, preferentemente aproximadamente 50 nucleótidos, 75 o 100 y lo más preferentemente de aproximadamente 200 o más nucleótidos, se hibride a una temperatura de aproximadamente 65 °C en una disolución que comprende sal aproximadamente 1 M, preferentemente 6 x SSC o cualquier otra disolución que tiene una fuerza iónica comparable, y lavado a 65 °C en una disolución que comprende sal aproximadamente 0,1 M, o menos, preferentemente 0,2 x SSC o cualquier otra disolución que tiene una fuerza iónica comparable. Preferentemente, la hibridación se realiza durante la noche, es decir, al menos durante 10 horas y preferentemente el lavado se realiza durante al menos una hora con al menos dos cambios de la disolución de lavado. Estas condiciones normalmente permitirán la hibridación específica de secuencias que tienen aproximadamente el 90 % o más identidad de secuencia.

Condiciones moderadas se definen en el presente documento como condiciones que permiten que una secuencia de ácidos nucleicos de al menos 50 nucleótidos, preferentemente de aproximadamente 200 o más nucleótidos, se

hibride a una temperatura de aproximadamente 45 °C en una disolución que comprende sal aproximadamente 1 M, preferentemente 6 x SSC o cualquier otra disolución que tiene una fuerza iónica comparable, y lavado a temperatura ambiente en una disolución que comprende sal aproximadamente 1 M, preferentemente 6 x SSC, o cualquier otra disolución que tiene una fuerza iónica comparable. Preferentemente, la hibridación se realiza durante la noche, es decir, al menos durante 10 horas, y preferentemente el lavado se realiza durante al menos una hora con al menos dos cambios de la disolución de lavado. Estas condiciones normalmente permitirán la hibridación específica de secuencias que tienen hasta el 50 % de identidad de secuencia. El experto en la materia será capaz de modificar estas condiciones de hibridación con el fin de identificar específicamente secuencias que varían en identidad entre el 50 % y el 90 %.

Una "construcción de ácidos nucleicos" o "vector de ácido nucleico" se entiende en el presente documento que significa una molécula de ácidos nucleicos sintética resultante del uso de tecnología de ADN recombinante. El término "construcción de ácidos nucleicos", por tanto, no incluye moléculas de ácidos nucleicos que existen de forma natural, aunque una construcción de ácidos nucleicos puede comprender (partes de) moléculas de ácidos nucleicos que existen de forma natural. Los términos "vector de expresión" o "construcción de expresión" se refieren a secuencias de nucleótidos que son capaces de afectar la expresión de un gen en células hospedadoras u organismos hospedadores compatibles con tales secuencias. Estos vectores de expresión normalmente incluyen al menos secuencias reguladoras de la transcripción adecuadas y opcionalmente señales de terminación de la transcripción de 3'. También puede estar presentes factores necesarios adicionales o útiles en efectuar la expresión, tales como elementos potenciadores de la expresión. El vector de expresión se introducirá en una célula hospedadora adecuada y será capaz de efectuar la expresión de la secuencia codificante en un cultivo celular *in vitro* de la célula hospedadora. El vector de expresión será adecuado para la replicación en célula hospedadora de levadura de la invención.

Como se usa en el presente documento, el término "promotor" o "secuencia reguladora de la transcripción" se refiere a un fragmento de ácido nucleico que funciona controlando la transcripción de una o más secuencias codificantes, y se localiza en la dirección 5' con respecto a la dirección de transcripción del sitio de iniciación de la transcripción de la secuencia codificante, y es estructuralmente identificado por la presencia de un sitio de unión para ARN polimerasa dependiente de ADN, sitios de iniciación de la transcripción y cualquier otra secuencia de ADN, que incluye, pero no se limita a, sitios de unión al factor de transcripción, sitios de unión a la proteína represora y activadora, y cualquier otra secuencia de nucleótidos conocida para un experto en la materia por actuar directamente o indirectamente para regular la cantidad de transcripción del promotor. Un promotor "constitutivo" es un promotor que es activo en la mayoría de los tejidos en la mayoría de las condiciones fisiológicas y de desarrollo. Un promotor "inducible" es un promotor que está regulado desde el punto de vista fisiológico o del desarrollo, por ejemplo, por la aplicación de un inductor químico.

El término "marcador de selección" es un término familiar para un experto habitual en la materia y se usa en el presente documento para describir cualquier entidad genética que, cuando se expresa, puede usarse para seleccionar una célula o células que contienen el marcador de selección. El término "indicador" puede usarse indistintamente con marcador, aunque principalmente se usa para referirse a marcadores visibles, tales como proteína verde fluorescente (GFP). Los marcadores de selección pueden ser dominantes o recesivos o bidireccionales.

Como se usa en el presente documento, el término "operativamente unido" se refiere a un enlace de elementos de polinucleótido en una relación funcional. Un ácido nucleico está "operativamente unido" cuando se pone en una relación funcional con otra secuencia de ácidos nucleicos. Por ejemplo, una secuencia reguladora de la transcripción está operativamente unida a una secuencia codificante si afecta la transcripción de la secuencia codificante. Operativamente unido significa que las secuencias de ADN que se unen normalmente son contiguas y, donde sea necesario unir dos regiones codificantes de proteína, contiguas y en marco de lectura.

Los términos "proteína" o "polipéptido" se usan indistintamente y se refieren a moléculas que consisten en una cadena de aminoácidos, sin referencia a un modo específico de acción, tamaño, estructura tridimensional u origen.

"Hongos" (singular hongo) se entienden en el presente documento como microorganismos eucariotas heterotróficos que digieren su alimento externamente, absorbiendo moléculas nutrientes en sus células. Los hongos son un reino separado de organismos eucariotas e incluyen levaduras, mohos y setas. Los términos hongos, hongo y fúngico, como se usan en el presente documento, incluyen así expresamente levaduras, además de hongos filamentosos.

El término "gen" significa un fragmento de ADN que comprende una región (región transcrita), que se transcribe en una molécula de ARN (por ejemplo, un ARNm) en una célula, operativamente unida a regiones reguladoras adecuadas (por ejemplo, un promotor). Un gen normalmente comprenderá varios fragmentos operativamente unidos, tales como un promotor, una secuencia conductora de 5', una región codificante y una secuencia no traducida de 3' (extremo 3') que comprende un sitio de poliadenilación. "Expresión de un gen" se refiere al proceso en el que una región de ADN que está operativamente unida a regiones reguladoras apropiadas, particularmente un promotor, se transcribe en un ARN, que es biológicamente activo, es decir, que es capaz de ser traducido en una proteína o péptido biológicamente activo.

El término "homólogo", cuando se usa para indicar la relación entre una molécula de ácido nucleico (recombinante) dada o de polipéptido y un organismo hospedador dado o célula hospedadora, se entiende que significa que en la naturaleza el ácido nucleico o molécula de polipéptido se produce por una célula hospedadora u organismos de la misma especie, preferentemente de la misma variedad o cepa. Si es homóloga a una célula hospedadora, una secuencia de ácidos nucleicos que codifica un polipéptido normalmente se unirá operativamente (pero no necesariamente) a otra secuencia promotora (heteróloga) y, si es aplicable, otra secuencia señal secretora (heteróloga) y/o secuencia terminadora distinta de en su entorno natural. Se entiende que las secuencias reguladoras, secuencias señal, secuencias terminadoras, etc., también pueden ser homólogas a la célula hospedadora. En este contexto, el uso de solo elementos de secuencia "homólogos" permite la construcción de organismos genéticamente modificados (OGM) "auto-clonados" (la auto-clonación se define en el presente documento como en la Directiva europea 98/81/CE, Anexo II). Cuando se usa para indicar la vinculación de dos secuencias de ácidos nucleicos, el término "homólogo" significa que una secuencia de ácidos nucleicos monocatenaria puede hibridarse con una secuencia de ácidos nucleicos monocatenaria complementaria. El grado de hibridación puede depender de varios factores, que incluyen la cantidad de identidad entre las secuencias y las condiciones de hibridación, tales como la temperatura y concentración de sales como se trata después.

Los términos "heterólogo" y "exógeno", cuando se usan con respecto a un ácido nucleico (ADN o ARN) o proteína, se refieren a un ácido nucleico o proteína que no se produce naturalmente como parte del organismo, célula, genoma o secuencia de ADN o ARN en la que está presente, o que se encuentra en una célula o localización o localizaciones en el genoma o secuencia de ADN o ARN que se diferencian de aquella en la que se encuentra en la naturaleza. Los ácidos nucleicos heterólogos y exógenos o proteínas no son endógenos a la célula en la que se introducen, pero han sido obtenidos de otra célula o producidos sintéticamente o recombinantemente. Generalmente, aunque no necesariamente, tales ácidos nucleicos codifican proteínas, es decir, proteínas exógenas, que normalmente no son producidas por la célula en la que el ADN se transcribe o expresa. Similarmente, el ARN exógeno codifica proteínas que normalmente no se expresan en la célula en la que está presente el ARN exógeno. Los ácidos nucleicos heterólogos/exógenos y proteínas también pueden denominarse ácidos nucleicos o proteínas extraños. Cualquier ácido nucleico o proteína que un experto en la materia reconocería como extraño a la célula en la que se expresa está englobado en el presente documento por el término ácido nucleico heterólogo o exógeno o proteína. Los términos heterólogo y exógeno también se aplican a combinaciones no naturales de secuencias de ácidos nucleicos o aminoácidos, es decir, combinaciones donde al menos dos de las secuencias combinadas son extrañas la una con respecto a la otra.

Se entiende en el presente documento que la "actividad específica" de una enzima significa la cantidad de actividad de una enzima particular por la cantidad de proteína de célula hospedadora total, normalmente expresada en unidades de actividad enzimática por mg de proteína de célula hospedadora total. En el contexto de la presente invención, la actividad específica de una enzima particular puede aumentar o disminuir en comparación con la actividad específica de esa enzima en una célula hospedadora no mutada (por lo demás idéntica).

"Condiciones anaerobias" o un proceso de fermentación anaerobia se define en el presente documento como condiciones o un proceso de fermentación realizado en ausencia de oxígeno o en el que sustancialmente no se consume oxígeno, preferentemente menos de 5, 2,5 o 1 mmol/l/h, más preferentemente se consumen 0 mmol/l/h (es decir, el consumo de oxígeno no es detectable), y en el que las moléculas orgánicas sirven de tanto donante de electrones como aceptores de electrones.

#### Descripción detallada de la invención

La expresión de una acetaldéhidó deshidrogenasa exógena en levadura permite que la levadura convierta ácido acético, que puede estar presente en altas cantidades en hidrolizados lignocelulósicos, en etanol. La reducción dependiente de NADH del ácido acético a etanol ha sido propuesta como una sustitución para la formación de glicerol como un disipador de redox en cultivos anaerobios cultivados en glucosa de *S. cerevisiae*, proporcionando así una base estequiométrica para la eliminación de la producción de glicerol (como subproducto) durante la producción industrial del etanol y, por consiguiente, un rendimiento de etanol más alto (Guadalupe et al., arriba). Sin embargo, la estequiometría de estas reacciones es tal que la reducción de una molécula de ácido acético a etanol requeriría que no se produjeran dos moléculas de glicerol. Los presentes inventores han encontrado, sin embargo, que en la práctica la cantidad de ácido acético que normalmente está presente en hidrolizados lignocelulósicos industriales es tal que la cantidad de NADH requerida para que se reduzca a etanol supera la cantidad de NADH que estaría disponible para prevenir la producción de glicerol en levaduras cultivadas en condiciones anaerobias. Los presentes inventores han encontrado ahora sorprendentemente que cantidades mucho más altas de ácido acético pueden reducirse a etanol por consumo simultáneo de glicerol por la levadura, en vez de previniendo su producción.

Se generan grandes cantidades de glicerol como subproducto en la producción de biodiésel a partir de reacciones de transesterificación usando aceites vegetales o grasas animales y un alcohol. Por tanto, se predice que la disponibilidad del glicerol bruto aumentará durante los siguientes años como resultado del crecimiento en la producción de biodiésel en el mundo. Por consiguiente, estarán disponibles grandes cantidades de glicerol a bajo coste. La presente invención proporciona medios y métodos de valorización del glicerol, obtenido, por ejemplo, como subproducto de la producción de biodiésel, convirtiéndolo en etanol que puede usarse como biocombustible. Al mismo tiempo, la presente invención trata el problema de altas cantidades de ácido acético que están presentes en

hidrolizados lignocelulósicos y que inhiben las capacidades fermentativas de las levaduras para producir etanol a partir de estos hidrolizados. Una ventaja adicional de la presente invención es que dejando intacta la vía de respuesta de glicerol de alta osmolaridad en las células de levadura de la invención (a diferencia de las cepas en las que (todos) los genes de glicerolfosfato deshidrogenasa se inactivan como se describe por Guadalupe et al., arriba), se obtienen cepas de levadura más robustas que son más capaces de manipular el choque osmótico que puede producirse bajo condiciones de fermentaciones industriales.

En un primer aspecto, la invención se refiere a una célula de levadura según la reivindicación 14 adjunta, que comprende un gen exógeno que codifica una enzima con la capacidad para reducir acetilCoA en acetaldehído, gen que confiere a la célula la capacidad para convertir ácido acético en etanol. Una enzima con la capacidad para reducir acetilCoA en acetaldehído se entiende en el presente documento como una enzima que cataliza la reacción (ACDH; EC 1.2.1.10):



Así, la enzima cataliza la conversión de acetilCoA en acetaldehído (y viceversa) y también se denomina una acetaldehído deshidrogenasa (dependiente de NAD acetilante) o una acetil-CoA reductasa. La enzima puede ser una enzima bifuncional que cataliza además la conversión de acetaldehído en etanol (y viceversa; véase más adelante). Por comodidad, los presentes inventores deben referirse en el presente documento a una enzima que tiene al menos la capacidad de reducir acetilCoA en ya sea acetaldehído o etanol como "acetaldehído deshidrogenasa". Se entiende además en el presente documento que la célula hospedadora de levadura tiene actividades de acetil-CoA sintetasa y alcohol deshidrogenasa endógenas que permiten que la célula, que está provista de actividad de acetaldehído deshidrogenasa, complete la conversión de ácido acético en etanol.

El gen exógeno puede codificar una enzima monofuncional que tiene solo actividad de acetaldehído deshidrogenasa (es decir, una enzima solo que tiene la capacidad de reducir acetilCoA en acetaldehído) tal como, por ejemplo, la acetaldehído deshidrogenasa codificada por el gen *mhpF* de *E. coli*. Un gen exógeno adecuado que codifica una enzima con actividad de acetaldehído deshidrogenasa comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos el 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 % de identidad de secuencia de aminoácidos con SEQ ID NO: 1. Ejemplos adecuados de procariontes que comprenden enzimas monofuncionales con actividad de acetaldehído deshidrogenasa se proporcionan en la Tabla 1. Las secuencias de aminoácidos de estas enzimas monofuncionales están disponibles en bases de datos públicas y pueden ser usadas por el experto para diseñar secuencias de nucleótidos de codón optimizado que codifican la enzima monofuncional correspondiente (véase, por ejemplo, SEQ ID NO: 2).

Tabla 1: Enzimas con actividad de acetaldehído deshidrogenasa relacionadas con *mhpF* de *E. coli*

Organismo	Identidad de aminoácidos (%)
<i>Escherichia coli</i> str. K12 substr. MG1655	100 %
<i>Shigella sonnei</i>	100 %
<i>Escherichia coli</i> IAI39	99 %
<i>Citrobacter youngae</i> ATCC 29220	93 %
<i>Citrobacter</i> sp. 30_2	92 %
<i>Klebsiella pneumoniae</i> 342)	87 %
<i>Klebsiella variicola</i>	87 %
<i>Pseudomonas putida</i>	81 %
<i>Ralstonia eutropha</i> JMP134	82 %
<i>Burkholderia</i> sp. H160	81 %
<i>Azotobacter vinelandii</i> DJ	79 %
<i>Ralstonia metallidurans</i> CH34	70 %
<i>Xanthobacter autotrophicus</i> Py2	67 %
<i>Burkholderia cenocepacia</i> J2315	68 %
<i>Frankia</i> sp. EAN1pec	67 %

Organismo	Identidad de aminoácidos (%)
<i>Polaromonas</i> sp. JS666	68 %
<i>Burkholderia phytofirmans</i> PsJN	70 %
<i>Rhodococcus opacus</i> B4	64 %

Preferentemente, la célula hospedadora comprende un gen exógeno que codifica una enzima bifuncional con actividad de acetaldehído deshidrogenasa y alcohol deshidrogenasa, gen que confiere a la célula la capacidad para convertir el ácido acético en etanol. La ventaja de uso de una enzima bifuncional con actividades de acetaldehído deshidrogenasa y alcohol deshidrogenasa en lugar de separar enzimas para cada una de las actividades de acetaldehído deshidrogenasa y alcohol deshidrogenasa, es que permite que la directa canalización del producto intermedio entre enzimas que catalizan reacciones consecutivas en una vía ofrezca la posibilidad de un medio eficiente, exclusivo y protegido de administración de metabolitos. La canalización de sustratos disminuye así el tiempo de tránsito de los productos intermedios, previene la pérdida de productos intermedios por difusión, protege los productos intermedios lábiles del disolvente, y anticipa la entrada de productos intermedios en las vías metabólicas de competición. La enzima bifuncional, por tanto, permite una conversión más eficiente de ácido acético en etanol en comparación con las enzimas acetaldehído deshidrogenasa y alcohol deshidrogenasa separadas. Una ventaja adicional de uso de la enzima bifuncional es que también puede usarse en células hospedadoras que tienen poca o ninguna actividad de alcohol deshidrogenasa en la condición usada, tal como, por ejemplo, condiciones anaerobias y/o condiciones de represión de catabolitos.

Enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y alcohol deshidrogenasa son procariotas y protozoarios conocidos en la técnica, que incluyen, por ejemplo, las enzimas bifuncionales codificadas por los genes *adhE* de *Escherichia coli* y *ADH2* de *Entamoeba histolytica* (véase, por ejemplo, Bruchaus and Tannich, 1994, J. Biochem. 303: 743-748; Burdette and Zeikus, 1994, J. Biochem. 302: 163-170; Koo et al., 2005, Biotechnol. Lett. 27: 505-510; Yong et al., 1996, Proc Natl Acad Sci USA, 93: 6464-6469). Las enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y alcohol deshidrogenasa son proteínas más grandes que consisten en aproximadamente 900 aminoácidos y son bifuncionales por que presentan tanto actividad de acetaldehído deshidrogenasa (ACDH; EC 1.2.1.10) como de alcohol deshidrogenasa (ADH; EC 1.1.1.1). *adhE* de *E. coli* y *ADH2* de *Entamoeba histolytica* muestran el 45 % de identidad de aminoácidos. Por tanto, en una realización de la invención, un gen exógeno adecuado que codifica una enzima bifuncional con actividad de acetaldehído deshidrogenasa y alcohol deshidrogenasa comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos el 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 % de identidad de secuencia de aminoácidos con al menos una de SEQ ID NO: 3 y 5. Ejemplos adecuados de procariotas que comprenden enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y alcohol deshidrogenasa se proporcionan en las Tablas 2 y 3. Las secuencias de aminoácidos de estas enzimas bifuncionales están disponibles en bases de datos públicas y pueden ser usadas por el experto para diseñar secuencias de nucleótidos de codón optimizado que codifican la enzima bifuncional correspondiente (véase, por ejemplo, SEQ ID NO: 4 o 6).

Tabla 2: Enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y alcohol deshidrogenasa relacionadas con *adhE* de *E. coli*

Organismo	Identidad de aminoácidos (%)
<i>Escherichia coli</i> O157:H7 str. Sakai	100 %
<i>Shigella sonnei</i>	100 %
<i>Shigella dysenteriae</i> 1012	99 %
<i>Klebsiella pneumoniae</i> 342	97 %
<i>Enterobacter</i> sp. 638	94 %
<i>Yersinia pestis</i> biovar Microtus str. 91001	90 %
<i>Serratia proteamaculans</i> 568	90 %
<i>Pectobacterium carotovorum</i> WPP14	90 %
<i>Sodalis glossinidius</i> str. 'morsitans'	87 %
<i>Erwinia tasmaniensis</i> Et1/99	86 %

Organismo	Identidad de aminoácidos (%)
<i>Aeromonas hydrophila</i> ATCC 7966	81 %
<i>Vibrio vulnificus</i> YJ016]	76 %

Tabla 3: Enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y alcohol deshidrogenasa relacionadas con ADH2 de *Entamoeba histolytica*

Organismo	Identidad de aminoácidos (%)
<i>Entamoeba histolytica</i> HM-1:IMSS	99 %
<i>Entamoeba dispar</i> SAW760	98 %
<i>Mollicutes bacterium</i> D7	65 %
<i>Fusobacterium mortiferum</i> ATCC 9817	64 %
<i>Actinobacillus succinogenes</i> 130Z	63 %
<i>Pasteurella multocida</i> Pm70	62 %
<i>Mannheimia succiniciproducens</i> MBEL55E	61 %
<i>Streptococcus</i> sp. 2_1_36FAA]	61 %

5 El gen exógeno que codifica la enzima bifuncional que tiene actividades de acetaldehído deshidrogenasa y alcohol deshidrogenasa, una enzima que tiene actividad de acetaldehído deshidrogenasa, preferentemente es una construcción de expresión que comprende una secuencia de nucleótidos que codifica la enzima operativamente unida a regiones/secuencias reguladoras de la expresión adecuadas para garantizar la expresión de la enzima tras la transformación de la construcción de expresión en la célula hospedadora de levadura de la invención. Así, el gen o construcción de expresión comprenderá al menos un promotor que es funcional en la célula hospedadora operativamente unido a la secuencia codificante. El gen o construcción puede comprender además una secuencia conductora de 5' en la dirección 5' de la región codificante y una secuencia no traducida de 3' (extremo 3') que comprende un sitio de poliadenilación y un sitio de terminación de la transcripción en la dirección 3' de la secuencia codificante. Métodos de preparación o construcción de las células de levadura de la invención se desvelan además en el presente documento. Para este fin, se usan técnicas de biología genética y molecular estándar que son generalmente conocidas en la técnica y han sido descritas, por ejemplo, por Sambrook y Russell (2001, "Molecular cloning: A laboratory manual" (3ª edición), Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press) y Ausubel et al. (1987, eds., "Current protocols in molecular biology", Green Publishing and Wiley Interscience, Nueva York). Además, la construcción de cepas de levadura hospedadoras mutadas se lleva a cabo por cruces genéticos, esporulación de los diploides resultantes, disección de tétradas de las esporas haploides que contienen los marcadores auxotróficos deseados, y purificación de colonias de tales levaduras hospedadoras haploides en el medio de selección apropiado. Todos estos métodos son métodos genéticos de levadura estándar conocidos para aquellos en la materia. Véase, por ejemplo, Sherman et al., Methods Yeast Genetics, Cold Spring Harbor Laboratory, NY (1978) y Guthrie et al. (Eds.) Guide To Yeast Genetics and Molecular Biology Vol. 194, Academic Press, San Diego (1991).

Promotores adecuados para la expresión de la secuencia de nucleótidos que codifica la enzima que tienen actividad de acetaldehído deshidrogenasa y opcionalmente alcohol deshidrogenasa (además de otras enzimas, véase más adelante) incluyen promotores que son preferentemente insensibles a la represión de catabolitos (glucosa), que son activos en condiciones anaerobias y/o que preferentemente no requieren xilosa o arabinosa para la inducción. Los promotores que tienen estas características 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 (*PPK*), triosa fosfato isomerasa (*TPI*), gliceraldehído-3-fosfato deshidrogenasa (*GDP*, *TDH3* o *GAPDH*), piruvato cinasa (*PYK*), de fosfoglicerato cinasa (*PGK*), del promotor de glucosa-6-fosfato isomerasa (*PGI1*) de levaduras. Más detalles sobre tales promotores de levadura pueden encontrarse en (documento WO 93/03159). Otros promotores útiles son promotores del gen que codifica proteína ribosómica (*TEF1*), el promotor del gen de lactasa (*LAC4*), promotores de alcohol deshidrogenasa (*ADH1*, *ADH4*, y similares), el promotor de enolasa (*ENO*) y el promotor del transportador de hexosa (glucosa) (*HXT7*). Alternativamente, la secuencia de nucleótidos que codifica la enzima que tiene actividad de acetaldehído deshidrogenasa y opcionalmente alcohol deshidrogenasa se expresa en exceso en condiciones anaerobias usando un promotor anóxico tal como, por ejemplo, el promotor *ANB1* de *S. cerevisiae* (SEQ ID NO: 19). Otros promotores, tanto

constitutivos como inducibles, y potenciadores o secuencias activadoras en la dirección 5', serán conocidos para aquellos expertos en la materia. Preferentemente, el promotor que está operativamente unido a la secuencia de nucleótidos como se ha definido anteriormente es homólogo a la célula hospedadora. Secuencias terminadoras adecuadas son obtenibles, por ejemplo, del gen de citocromo c1 (*CYC1*) o un gen de alcohol deshidrogenasa (por ejemplo, *ADH1*).

Para aumentar la probabilidad de que la enzima que tiene actividades de acetaldehído deshidrogenasa y opcionalmente alcohol deshidrogenasa se exprese a niveles suficientes y en forma activa en las células hospedadoras de levadura transformadas de la invención, la secuencia de nucleótidos que codifica estas enzimas, además de otras enzimas (véase más adelante), se adapta preferentemente para optimizar su uso de codones al de la célula hospedadora en cuestión. La capacidad de adaptación de una secuencia de nucleótidos que codifica una enzima al uso de codones de una célula hospedadora puede expresarse como el índice de adaptación de codones (IAC). El índice de adaptación de codones se define en el presente documento como una medición de la capacidad de adaptación relativa del uso de codones de un gen hacia el uso de codones de genes altamente expresados en una célula hospedadora u organismo particular. La capacidad de adaptación relativa ( $w$ ) de cada codón es la relación del uso de cada codón, con la del codón más abundante para el mismo aminoácido. El índice IAC se define como la media geométrica de estos valores de capacidad de adaptación relativa. Se excluyen codones y codones de terminación no sinónimos (dependientes del código genético). Los valores de IAC oscilan de 0 a 1, indicando valores más altos una proporción más alta de los codones más abundantes (véanse Sharp and Li, 1987, *Nucleic Acids Research* 15: 1281-1295; véase también: Jansen et al., 2003, *Nucleic Acids Res.* 31(8):2242-51). Una secuencia de nucleótidos adaptada preferentemente tiene un IAC de al menos 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8 o 0,9. Las preferidas son las secuencias que han sido optimizadas en codones para la expresión en la célula hospedadora fúngica en cuestión tal como, por ejemplo, células de *S. cerevisiae*.

La secuencia de nucleótidos codifica una enzima que tiene actividades de acetaldehído deshidrogenasa y opcionalmente alcohol deshidrogenasa que se expresa preferentemente en forma activa en la célula hospedadora transformada. Así, la expresión de la secuencia de nucleótidos en la célula hospedadora produce una acetaldehído deshidrogenasa con una actividad específica de al menos 0,005, 0,010, 0,020, 0,050 o 0,10  $\mu\text{mol min}^{-1}$  (mg de proteína)<sup>-1</sup>, determinada como la velocidad dependiente de acetyl-CoA de la reducción de NADH en extractos de células de la célula hospedadora transformada a 30 °C como se describe en los ejemplos en el presente documento.

La célula hospedadora que va a transformarse con una construcción de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica una enzima con acetaldehído deshidrogenasa y opcionalmente alcohol deshidrogenasa es una célula hospedadora de levadura. Preferentemente, la célula hospedadora es una célula cultivada. La célula hospedadora de levadura de la invención, preferentemente es hospedador capaz de transporte activo o pasivo de pentosa (xilosa y preferentemente también arabinosa) en la célula. La célula hospedadora contiene preferentemente glicólisis activa. La célula hospedadora puede contener además preferentemente una vía de pentosa fosfato endógena y puede contener actividad de xilulosa cinasa endógena de manera que la xilulosa isomerizada a partir de xilosa puede ser metabolizada a piruvato. El hospedador contiene además preferentemente enzimas para la conversión de una pentosa (preferentemente mediante piruvato) en un producto de fermentación deseado tal como etanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, 1,3-propano-diol, butanoles (1-butanol, 2-butanol, isobutanol) y productos derivados de isoprenoide. Una célula hospedadora particularmente preferida es una célula de levadura que es naturalmente capaz de fermentación alcohólica, preferentemente, fermentación alcohólica anaerobia. La célula hospedadora de levadura adicional preferentemente tiene una alta tolerancia al etanol, una alta tolerancia a pH bajo (es decir, capaz de crecimiento a un pH inferior a 5, 4, o 3) y hacia ácidos orgánicos como ácido láctico, ácido acético o ácido fórmico y productos de degradación de azúcar tales como furfural e hidroximetilfurfural, y una alta tolerancia a temperaturas elevadas. Cualquiera de estas características o actividades de la célula hospedadora de levadura pueden estar naturalmente presentes en la célula hospedadora o pueden introducirse o modificarse por modificación genética, preferentemente por auto-clonación o por los métodos descritos a continuación.

Levaduras se definen en el presente documento como microorganismos eucariotas e incluyen todas las especies de la subdivisión Eumycotina (Yeasts: characteristics and identification, J.A. Barnett, R.W. Payne, D. Yarrow, 2000, 3<sup>a</sup> ed., Cambridge University Press, Cambridge RU; y, The yeasts, a taxonomic study, C.P. Kurtzman and J.W. Fell (eds) 1998, 4<sup>a</sup> ed., Elsevier Science Publ. B.V., Ámsterdam, Los Países Bajos) que crecen predominantemente en forma unicelular. Las levaduras pueden o bien crecer por gemación de un talo unicelular o bien crecer por fisión del organismo. Células de levadura preferidas para su uso en la presente invención pertenecen a los géneros *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces* y *Yarrowia*. Preferentemente, la levadura es capaz de fermentación anaerobia, más preferentemente fermentación alcohólica anaerobia. A lo largo de los años se han hecho sugerencias para la introducción de diversos organismos para la producción de bioetanol a partir de azúcares de cultivo. En la práctica, sin embargo, todos los procesos de producción de bioetanol importantes 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 procesos industriales, es decir, una alta tolerancia a ácidos, etanol y osmotolerancia, capacidad de crecimiento anaerobio, y, por supuesto, su alta capacidad fermentativa alcohólica. Especies de levadura preferidas como células hospedadoras incluyen *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis*, *K. marxianus* y *Schizosaccharomyces pombe*. La célula hospedadora de levadura de la invención comprende además una modificación genética que

introduce actividad de glicerol deshidrogenasa unida a NAD<sup>+</sup> en la célula. La glicerol deshidrogenasa codificada por el gen *GCY1* de levadura endógeno parece ser específica para el cofactor NADP<sup>+</sup> (EC 1.1.1.72), a diferencia de NAD<sup>+</sup> (EC 1.1.1.6). Levaduras tales como *S. cerevisiae* parecen carecer de actividad de glicerol deshidrogenasa dependiente de NAD<sup>+</sup> (EC 1.1.1.6) (véase, por ejemplo, la vía de KEGG 00561). Una glicerol deshidrogenasa unida a NAD<sup>+</sup> se entiende en el presente documento como una enzima que cataliza la reacción química (EC 1.1.1.6):



Otros nombres en uso común incluyen glicerina deshidrogenasa y glicerol:NAD<sup>+</sup> 2-oxidoreductasa. La modificación genética que introduce actividad de glicerol deshidrogenasa unida a NAD<sup>+</sup> en la célula hospedadora es la expresión de una glicerol deshidrogenasa unida a NAD<sup>+</sup> que es heteróloga a la célula hospedadora de levadura. La secuencia de nucleótidos para la expresión de una glicerol deshidrogenasa heteróloga en las células de levadura de la invención es una secuencia que codifica una glicerol deshidrogenasa bacteriana que usa NAD<sup>+</sup> como cofactor (EC 1.1.1.6). Un ejemplo adecuado de una glicerol deshidrogenasa unida a NAD<sup>+</sup> bacteriana para la expresión en una célula hospedadora de levadura de la invención es, por ejemplo, el gen *gldA* de *E. coli* descrito por Truniger y Boos (1994, J Bacteriol. 176(6):1796-1800), cuya expresión en levadura ya se ha informado (Lee y Dasilva, 2006, Metab Eng. 8(1):58-65). Preferentemente, la secuencia de nucleótidos que codifica una glicerol deshidrogenasa heteróloga comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos el 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 % de identidad de secuencia de aminoácidos con SEQ ID NO: 7 o una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tienen una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 7. En una realización preferida, una secuencia de nucleótidos optimizada en codones (véase anteriormente) que codifica la glicerol deshidrogenasa heteróloga se expresa en exceso, tal como, por ejemplo, una secuencia de nucleótidos optimizada en codones que codifica la secuencia de aminoácidos de la glicerol deshidrogenasa de SEQ ID NO: 7. Una secuencia de nucleótidos optimizada en codones tal se proporciona, por ejemplo, en SEQ ID NO: 21 (posiciones 10 - 1113; IAC = 0,976).

Para la expresión en exceso de la secuencia de nucleótidos que codifica la glicerol deshidrogenasa, la secuencia de nucleótidos (que va a expresarse en exceso) se pone en una construcción de expresión en la que está operativamente unida a regiones/secuencias reguladoras de la expresión adecuadas para garantizar la expresión en exceso de la enzima glicerol deshidrogenasa tras la transformación de la construcción de expresión en la célula hospedadora de levadura de la invención (véase anteriormente). Promotores adecuados para la expresión (en exceso) de la secuencia de nucleótidos que codifica la enzima que tiene actividad de glicerol deshidrogenasa incluyen promotores que son preferentemente insensibles a la represión de catabolitos (glucosa), que son activos en condiciones anaerobias y/o que preferentemente no requieren xilosa o arabinosa para la inducción. Ejemplos de tales promotores se dan anteriormente. La expresión de la secuencia de nucleótidos en la célula hospedadora produce una actividad de glicerol deshidrogenasa unida a NAD<sup>+</sup> específica de al menos 0,2, 0,5, 1,0, 2,0, o 5,0 U min<sup>-1</sup> (mg de proteína)<sup>-1</sup>, determinada en extractos de células de las células hospedadoras transformadas a 30 °C como se describe en los ejemplos en el presente documento. La célula hospedadora de levadura de la invención comprende además una modificación genética que aumenta la actividad específica de dihidroxiacetona cinasa en la célula. Datos de transcriptomas han mostrado que la dihidroxiacetona cinasa *DAK1* endógena ya se expresa a altos niveles en *S. cerevisiae*. Sin embargo, para velocidades de conversión óptimas, la célula hospedadora de levadura de la invención comprende así una modificación genética que aumenta la actividad específica de dihidroxiacetona cinasa en la célula. Una dihidroxiacetona cinasa se entiende en el presente documento como una enzima que cataliza la reacción química (EC 2.7.1.29):



Otros nombres en uso común incluyen glicerona cinasa, ATP:glicerona fosfotransferasa y acetol cinasa (fosforilante). Se entiende que la glicerona y dihidroxiacetona son la misma molécula. Preferentemente, la modificación genética produce la expresión en exceso de una dihidroxiacetona cinasa, por ejemplo por expresión en exceso de una secuencia de nucleótidos que codifica una dihidroxiacetona cinasa. La secuencia de nucleótidos que codifica la dihidroxiacetona cinasa puede ser endógena a la célula o puede ser una dihidroxiacetona cinasa que es heteróloga a la célula. Secuencias de nucleótidos que pueden usarse para la expresión en exceso de dihidroxiacetona cinasa en las células de la invención son, por ejemplo, los genes de dihidroxiacetona cinasa de *S. cerevisiae* (*DAK1*) y (*DAK2*) como se describe, por ejemplo, por Molin et al. (2003, J. Biol. Chem. 278:1415-1423). Preferentemente, la secuencia de nucleótidos que codifica la dihidroxiacetona cinasa comprende una secuencia de aminoácidos con al menos el 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 % de identidad de secuencia de aminoácidos con al menos una de SEQ ID NO: 8 y 9. En una realización preferida, una secuencia de nucleótidos optimizada en codones (véase anteriormente) que codifica la dihidroxiacetona cinasa se expresa en exceso, tal como, por ejemplo, una secuencia de nucleótidos optimizada en codones que codifica la dihidroxiacetona cinasa de SEQ ID NO: 8 o una secuencia de nucleótidos optimizada en codones que codifica la dihidroxiacetona cinasa de SEQ ID NO: 9. Una secuencia de nucleótidos preferida para la expresión en exceso de una dihidroxiacetona cinasa es una secuencia de nucleótidos que codifica una dihidroxiacetona cinasa que comprende una secuencia de aminoácidos con al menos el 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 % de identidad de secuencia de aminoácidos con al menos una de SEQ ID NO: 8 (*S. cerevisiae* (*DAK1*)) o que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 8.

Secuencias de nucleótidos que pueden usarse para la expresión en exceso de una dihidroxiacetona cinasa heteróloga en las células de la invención son, por ejemplo, secuencias que codifican dihidroxiacetona cinasas bacterianas tales como el gen *dhaK* de *Citrobacter freundii* descrito, por ejemplo, por Daniel et al. (1995, J. Bacteriol. 177:4392-4401). Preferentemente, la secuencia de nucleótidos que codifica una dihidroxiacetona cinasa heteróloga comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos el 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 % de identidad de secuencia de aminoácidos con SEQ ID NO: 25 o una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tienen una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 25. En una realización preferida, una secuencia de nucleótidos optimizada en codones (véase anteriormente) que codifica la dihidroxiacetona cinasa heteróloga se expresa en exceso, tal como por ejemplo una secuencia de nucleótidos optimizada en codones que codifica la secuencia de aminoácidos de la dihidroxiacetona cinasa de SEQ ID NO: 25. Una secuencia de nucleótidos optimizada en codones tal se proporciona, por ejemplo, en SEQ ID NO: 26 (posiciones 10 - 1668).

Para la expresión en exceso de la secuencia de nucleótidos que codifica la dihidroxiacetona cinasa, la secuencia de nucleótidos (que va a expresarse en exceso) se pone en una construcción de expresión en la que está operativamente unida a regiones/secuencias reguladoras de la expresión adecuadas para garantizar la expresión en exceso de la enzima dihidroxiacetona cinasa tras la transformación de la construcción de expresión en la célula hospedadora de la invención (véase anteriormente). Promotores adecuados para la expresión (en exceso) de la secuencia de nucleótidos que codifica la enzima que tiene actividad de dihidroxiacetona cinasa incluyen promotores que son preferentemente insensibles a la represión de catabolitos (glucosa), que son activos en condiciones anaerobias y/o que preferentemente no requieren xilosa o arabinosa para la inducción. Ejemplos de tales promotores se dan anteriormente. En las células de la invención, una dihidroxiacetona cinasa que va a expresarse en exceso se expresa preferentemente en exceso al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que causa la expresión en exceso. Preferentemente, la dihidroxiacetona cinasa se expresa en exceso en condiciones anaerobias por al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 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 en estado estacionario de la actividad de la enzima (actividad específica en la célula), el nivel en estado estacionario de la proteína de la enzima, además de al nivel en estado estacionario del transcrito que codifica la enzima en la célula. La expresión en exceso de la secuencia de nucleótidos en la célula hospedadora produce una actividad de dihidroxiacetona cinasa específica de al menos 0,002, 0,005, 0,01, 0,02 o 0,05 U min<sup>-1</sup> (mg de proteína)<sup>-1</sup>, determinada en extractos de células de las células hospedadoras transformadas a 30 °C como se describe en los ejemplos en el presente documento.

En una realización adicional, la célula hospedadora de levadura de la invención comprende además una modificación genética que aumenta el transporte de glicerol en la célula. Preferentemente, la modificación genética que aumenta el transporte de glicerol en la célula preferentemente es una modificación genética que produce la expresión en exceso de una secuencia de nucleótidos que codifica al menos uno de una proteína de captación de glicerol y un canal de glicerol.

Una proteína de captación de glicerol se entiende en el presente documento como una proteína que atraviesa múltiples membranas que pertenece a la incluida en la superfamilia de O-aciltransferasas unidas a membrana (MBOAT) que incluye, por ejemplo, las proteínas de captación de glicerol de *S. cerevisiae* codificadas por los genes *GUP1* y *GUP2*. Preferentemente, la modificación genética produce la expresión en exceso de una proteína de captación de glicerol, por ejemplo por expresión en exceso de una secuencia de nucleótidos que codifica una proteína de captación de glicerol. La secuencia de nucleótidos que codifica la proteína de captación de glicerol puede ser endógena a la célula o puede ser una proteína de captación de glicerol que es heteróloga a la célula. Secuencias de nucleótidos que pueden usarse para la expresión en exceso de proteína de captación de glicerol en las células de la invención son, por ejemplo, los genes de la proteína de captación de glicerol de *S. cerevisiae* (*GUP1*) y (*GUP2*) y ortólogos de los mismos como se describe, por ejemplo, por Neves et al. (2004, FEMS Yeast Res. 5:51-62). Preferentemente, la secuencia de nucleótidos que codifica la proteína de captación de glicerol comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos el 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 % de identidad de secuencia de aminoácidos con al menos una de SEQ ID NO: 10 (Gup1p) y 11 (Gup2p). En una realización preferida, una secuencia de nucleótidos optimizada en codones (véase anteriormente) que codifica la proteína de captación de glicerol se expresa en exceso, tal como, por ejemplo, una secuencia de nucleótidos optimizada en codones que codifica la proteína de captación de glicerol SEQ ID NO: 10 o una secuencia de nucleótidos optimizada en codones que codifica la proteína de captación de glicerol de SEQ ID NO: 11. Aunque la naturaleza exacta de la influencia de *GUP1* sobre el transporte de glicerol todavía no está clara, Yu et al. (2010, arriba) han mostrado que la expresión en exceso de *GUP1* en *S. cerevisiae* mejora la producción de etanol en células cultivadas en glicerol. Una secuencia de nucleótidos preferida para la expresión en exceso de una proteína de captación de glicerol es, por tanto, una secuencia de nucleótidos que codifica una proteína de captación de glicerol que es capaz de rescatar el fenotipo asociado al estrés por sales de un mutante *gup1Δ* de *S. cerevisiae* por complementación como se describe por Neves et al. (2004, arriba). Tales ortólogos de complementación de *GUP1* de *S. cerevisiae* incluyen secuencias de nucleótidos que codifican secuencias de aminoácidos que tienen al menos el 60, 68, 72, 75, 80, 85, 90, 95, 98, 99 % de identidad con la secuencia de aminoácidos de SEQ ID NO: 10 y pueden obtenerse a partir de especies de levadura que pertenecen a los géneros

de *Saccharomyces*, *Zygosaccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Hansenula*, *Kloeckera*, *Schwanniomyces* y *Yarrowia*.

Un canal de glicerol se entiende en el presente documento como un miembro de la familia MIP de proteínas de canales revisado por Reizer et al. (1993, CRC Crit. Rev. Biochem. Mol. Biol., 28: 235-257), cuyas proteínas de canales comprenden un dominio transmembranario de 250 - 280 aminoácidos que consiste en seis dominios que atraviesan la membrana y tienen al menos el 30, 35, 40, 45, 50, 60, 70, 80, 90, 95, 98 o 99 % de identidad de aminoácidos, o al menos el 55, 60, 65, 70, 80, 90, 95, 98 o 99 % de similitud de aminoácidos con la secuencia de aminoácidos entre los aminoácidos 250 y 530 de SEQ ID NO: 12, la acuagliceroporina de FPS1 de *S. cerevisiae*. Secuencias de nucleótidos que pueden usarse para la expresión en exceso de un canal de glicerol en las células de la invención incluyen secuencias de nucleótidos que codifican el gen FPS1 de acuagliceroporina de levadura de, por ejemplo, *S. cerevisiae* (Van Aelst et al., 1991, EMBO J. 10:2095-2104) y ortólogos del mismo de otras levaduras que incluyen *Kluyveromyces lactis*, *Kluyveromyces marxianus* y *Zygosaccharomyces rouxii* como se describe, por ejemplo, por Neves et al. (2004, arriba). Sin embargo, el uso de canales de glicerol bacteriano o de planta no se excluye, por ejemplo, Luyten et al. (1995, EMBO J. 14:1360-1371) han mostrado que el facilitador de glicerol de *E. coli*, que tiene solo el 30 % de identidad de secuencia con la secuencia de aminoácidos entre los aminoácidos 250 y 530 de la acuagliceroporina de FPS1 de *S. cerevisiae*, puede complementar la captación de glicerol en un mutante *fps1Δ* de *S. cerevisiae*. La secuencia de nucleótidos que codifica el canal de glicerol puede ser endógena a la célula o puede ser un canal de glicerol que es heterólogo a la célula. En una realización preferida, una secuencia de nucleótidos optimizada en codones (véase anteriormente) que codifica el canal de glicerol se expresa en exceso, tal como, por ejemplo, una secuencia de nucleótidos optimizada en codones que codifica la acuagliceroporina de SEQ ID NO: 12.

Para la expresión en exceso de la secuencia de nucleótidos que codifica la proteína de captación de glicerol y/o la proteína de canales de glicerol, la secuencia de nucleótidos (que va a expresarse en exceso) se pone en una construcción de expresión en la que está operativamente unida a regiones/secuencias reguladoras de la expresión adecuadas para garantizar la expresión en exceso de la proteína de captación de glicerol y/o la proteína de canales de glicerol tras la transformación de la construcción de expresión en la célula hospedadora de la invención (véase anteriormente). Promotores adecuados para la expresión (en exceso) de la secuencia de nucleótidos que codifica la proteína de captación de glicerol y/o la proteína de canales de glicerol incluyen promotores que son preferentemente insensibles a la represión de catabolitos (glucosa), que son activos en condiciones anaerobias y/o que preferentemente no requieren xilosa o arabinosa para la inducción. Ejemplos de tales promotores se dan anteriormente. En las células de la invención, una proteína de captación de glicerol y/o una proteína de canales de glicerol que va a expresarse en exceso se expresan preferentemente en exceso al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que causa la expresión en exceso. Preferentemente, la proteína de captación de glicerol y/o la proteína de canales de glicerol se expresan en exceso en condiciones anaerobias al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 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 en estado estacionario de la actividad de la enzima (actividad específica en la célula), el nivel en estado estacionario de la proteína de la enzima, además de al nivel en estado estacionario del transcrito que codifica la enzima en la célula.

En una realización preferida de la célula hospedadora de la invención, la expresión de la proteína de canales de glicerol como se ha definido anteriormente se reduce o inactiva. Una modificación genética que reduce o que inactiva la expresión de la proteína de canales de glicerol puede ser útil para reducir o prevenir el transporte de glicerol fuera de la célula. Preferentemente, la reducción o inactivación de la expresión de la proteína de canales de glicerol se combina con la expresión en exceso de la secuencia de nucleótidos que codifica la proteína de captación de glicerol como se ha definido anteriormente.

En una realización adicional, la célula hospedadora de la invención comprende además una modificación genética que aumenta la actividad de acetil-CoA sintetasa específica en la célula, preferentemente en condiciones anaerobias ya que esta actividad es limitante de la velocidad en estas condiciones. Acetil-CoA sintetasa o acetato-CoA ligasa (EC 6.2.1.1) se entiende en el presente documento como una enzima que cataliza la formación de un nuevo enlace químico entre el acetato y la coenzima A (CoA). Preferentemente, la modificación genética produce la expresión en exceso de una acetil-CoA sintetasa, por ejemplo, por expresión en exceso de una secuencia de nucleótidos que codifica una acetil-CoA sintetasa. La secuencia de nucleótidos que codifica la acetil-CoA sintetasa puede ser endógena a la célula o puede ser una acetil-CoA sintetasa que es heteróloga a la célula. Secuencias de nucleótidos que pueden usarse para la expresión en exceso de acetil-CoA sintetasa en las células de la invención son, por ejemplo, los genes de acetil-CoA sintetasa de *S. cerevisiae* (*ACS1* y *ACS2*) como se describen, por ejemplo, por de Jong-Gubbels et al. (1998, FEMS Microbiol Lett. 165: 15-20). Preferentemente, la secuencia de nucleótidos que codifica la acetil-CoA sintetasa comprende una secuencia de aminoácidos con al menos el 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99 % de identidad de secuencia de aminoácidos con al menos una de SEQ ID NO: 13 y 14.

En una realización, la secuencia de nucleótidos que se expresa en exceso codifica una acetil-CoA sintetasa con una alta afinidad por acetato. El uso de una acetil-CoA sintetasa con una alta afinidad por acetato se prefiere para condiciones en las que hay una concentración relativamente baja de ácido acético en el medio de cultivo, por ejemplo, no más de 2 g de ácido acético/l de medio de cultivo. Una acetil-CoA sintetasa con una alta afinidad por

acetato se define en el presente documento como una acetil-CoA sintetasa con una afinidad más alta por acetato que la acetil-CoA sintetasa codificada por ACS2 de *S. cerevisiae*. Preferentemente, una acetil-CoA sintetasa con una alta afinidad por acetato tiene una Km por acetato de no más de 10, 5, 2, 1, 0,5, 0,2 o 0,1 mM, tal como, por ejemplo, la acetil-CoA sintetasa codificada por el gen ACS1 de *S. cerevisiae*. Más preferentemente, una secuencia de nucleótidos optimizada en codones (véase anteriormente) que codifica la secuencia de aminoácidos de SEQ ID NO: 13 se expresa en exceso.

En otra realización, la secuencia de nucleótidos que se expresa en exceso codifica una acetil-CoA sintetasa con una velocidad máxima ( $v_{m\acute{a}x}$ ) alta. El uso de una acetil-CoA sintetasa con una velocidad máxima alta se prefiere para la condición en la que hay una concentración relativamente alta de ácido acético en el medio de cultivo, por ejemplo, al menos 2, 3, 4 o 5 g de ácido acético/l de medio de cultivo. Una acetil-CoA sintetasa con una velocidad máxima alta se define en el presente documento como una acetil-CoA sintetasa con una velocidad máxima más alta que la acetil-CoA sintetasa codificada por ACS1 de *S. cerevisiae*. Preferentemente, la acetil-CoA sintetasa con una velocidad máxima alta es la acetil-CoA sintetasa codificada por el gen ACS2 de *S. cerevisiae*. Más preferentemente, una secuencia de nucleótidos optimizada en codones (véase anteriormente) que codifica la secuencia de aminoácidos de SEQ ID NO: 14 se expresa en exceso.

Para la expresión en exceso de la secuencia de nucleótidos que codifica la acetil-CoA sintetasa (que va a expresarse en exceso) se pone en una construcción de expresión en la que está operativamente unida a regiones/secuencias reguladoras de la expresión adecuadas para garantizar la expresión en exceso de la enzima acetil-CoA sintetasa tras la transformación de la construcción de expresión en la célula hospedadora de la invención (véase anteriormente). Promotores adecuados para la expresión (en exceso) de la secuencia de nucleótidos que codifica la enzima que tiene actividad de acetil-CoA sintetasa incluyen promotores que son preferentemente insensibles a la represión de catabolitos (glucosa), que son activos en condiciones anaerobias y/o que preferentemente no requieren xilosa o arabinosa para la inducción. Ejemplos de tales promotores se dan anteriormente. En las células de la invención, una acetil-CoA sintetasa que va a expresarse en exceso se expresa en exceso al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que causa la expresión en exceso. Preferentemente, la acetil-CoA sintetasa se expresa en exceso en condiciones anaerobias al menos un factor 2, 5, 10, 20, 50, o 100 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 en estado estacionario de la actividad de la enzima (actividad específica), el nivel en estado estacionario de la proteína de la enzima, además de al nivel en estado estacionario del transcrito que codifica la enzima.

En una realización adicional, la célula hospedadora de la invención comprende además una modificación genética que reduce la actividad de glicerol 3-fosfato deshidrogenasa dependiente de  $NAD^+$  específica en la célula. La glicerol 3-fosfato deshidrogenasa o glicerolfosfato deshidrogenasa (EC 1.1.1.8) cataliza la reducción de dihidroxiacetona fosfato a sn-glicerol 3-fosfato mientras que oxida  $NADH$  a  $NAD^+$ . En las células de la invención, la actividad de glicerolfosfato deshidrogenasa específica se reduce preferentemente al menos un factor 0,8, 0,5, 0,3, 0,1, 0,05 o 0,01 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que causa la expresión en exceso, preferentemente en condiciones anaerobias.

Preferentemente, la actividad de glicerolfosfato deshidrogenasa se reduce en la célula hospedadora por una o más modificaciones genéticas que reducen la expresión de o inactivan un gen que codifica una glicerolfosfato deshidrogenasa. Preferentemente, las modificaciones genéticas reducen o inactivan la expresión de cada copia endógena del gen que codifica una glicerolfosfato deshidrogenasa específica en el genoma de la célula. Una célula dada puede comprender múltiples copias del gen que codifica una glicerolfosfato deshidrogenasa específica con la misma secuencia de aminoácidos como resultado de di-, poli- o aneu-ploidía. En tales casos, preferentemente la expresión de cada copia del gen específico que codifica la glicerolfosfato deshidrogenasa se reduce o inactiva. Alternativamente, una célula puede contener varias (iso)enzimas diferentes con actividad de glicerolfosfato deshidrogenasa que se diferencian en secuencia de aminoácidos y que están cada una codificadas por un gen diferente. En tales casos, en algunas realizaciones de la invención, se prefiere que solo ciertos tipos de las isoenzimas se reduzcan o inactiven mientras que otros tipos siguen inafectados (véase más adelante). Preferentemente, el gen se inactiva por la delección de al menos parte del gen o por rotura 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, produciendo la delección o inactivación (parcial) una reducción de la expresión de la actividad de glicerolfosfato deshidrogenasa en la célula hospedadora.

Un gen preferido que codifica una glicerolfosfato deshidrogenasa cuya actividad va a reducirse o inactivarse en la célula de la invención es el gen *GPD2* de *S. cerevisiae* como se describe por Eriksson et al. (1995, Mol. Microbiol. 17: 95-107), que codifica la secuencia de aminoácidos de SEQ ID NO: 15 y ortólogos del mismo en otras especies. Por tanto, un gen que codifica una glicerolfosfato deshidrogenasa cuya actividad va a reducirse o inactivarse en la célula de la invención es preferentemente un gen que codifica una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos el 70, 75, 80, 85, 90, 95, 98 o 99 % de identidad de secuencia con SEQ ID NO: 15.

En una realización preferida de la invención, la célula hospedadora de la invención comprende una vía de respuesta de glicerol de alta osmolaridad funcional. Preferentemente, por tanto, solo la actividad del (de los) gen(es) que codifican una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos el 70 % de identidad de secuencia con SEQ ID NO: 15 se reducen o inactivan, mientras que al menos un gen endógeno que codifica una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos el 70, 75, 80, 85, 90, 95, 98 o 99 % de identidad de secuencia con SEQ ID NO: 16 es funcional. SEQ ID NO: 16 representa la secuencia de aminoácidos codificada por el gen *GPD1* de *S. cerevisiae* como se describe por Albertyn et al. (1994, Mol. Cell. Biol. 14: 4135-4144), que tiene el 69 % de identidad de aminoácidos con la glicerolfosfato deshidrogenasa *GPD2* de *S. cerevisiae*. El gen *GPD1* de *S. cerevisiae* es la glicerolfosfato deshidrogenasa inducida por el choque de *S. cerevisiae*, que es importante para el crecimiento bajo choque osmótico como puede producirse en condiciones de fermentaciones industriales. Su expresión está regulada, entre otros, por la vía de respuesta de glicerol de alta osmolaridad. Es, por tanto, ventajoso que una célula hospedadora de la invención tenga al menos una copia funcional de un gen endógeno que codifica una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos el 70, 75, 80, 85, 90, 95, 98 o 99 % de identidad de secuencia con SEQ ID NO: 16.

A pesar de lo anterior, los inventores han encontrado ahora sorprendentemente que la inactivación de la glicerolfosfato deshidrogenasa *GPD1* de *S. cerevisiae* tiene un efecto más ventajoso sobre la reducción de la producción de glicerol y el aumento del consumo de glicerol y acetato en comparación con la inactivación de la glicerolfosfato deshidrogenasa *GPD2* de *S. cerevisiae*. Por tanto, en una realización más preferida, la célula hospedadora de la invención comprende una modificación genética que reduce o inactiva la expresión de al menos el (los) gen(es) que codifican una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos el 70, 75, 80, 85, 90, 95, 98 o 99 % de identidad de secuencia con SEQ ID NO: 16 (*GPD1*).

En una realización adicional, la actividad de todos los genes en la célula hospedadora que codifican una glicerolfosfato deshidrogenasa se reducen o inactiva. En tales células, preferentemente todas las copias de genes endógenos que codifican una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos el 70, 75, 80, 85, 90, 95, 98 o 99 % de identidad de secuencia con SEQ ID NO: 15 o 16 se inactivan o al menos se reducen en expresión. Opcionalmente, la célula hospedadora no es una célula de levadura que comprenda un gen exógeno que codifica una enzima con la capacidad para convertir piruvato y coenzima-A en formiato y acetil-CoA. Preferentemente, la célula hospedadora no es una célula de levadura que comprenda una secuencia de nucleótidos que codifica una piruvato formiato liasa. Según las reivindicaciones, la célula hospedadora de levadura de la invención no es una célula de levadura que comprenda una modificación genética que reduce la actividad de formiato deshidrogenasa dependiente de NAD<sup>+</sup> específica en la célula.

En una realización preferida adicional, la célula hospedadora de la invención tiene al menos una de: a) la capacidad de isomerizar xilosa en xilulosa; y, b) la capacidad de convertir L-arabinosa en D-xilulosa 5-fosfato. Para a) la célula preferentemente tiene un gen de xilosa isomerasa exógeno funcional, gen que confiere a la célula la capacidad de isomerizar xilosa en xilulosa. Para b) la célula tiene preferentemente genes exógenos funcionales que codifican una L-arabinosa isomerasa, una L-ribulocinasa y una L-ribulosa-5-fosfato 4-epimerasa, genes que juntos confieren a la célula la capacidad de isomerizar convertir L-arabinosa en D-xilulosa 5-fosfato.

Células hospedadoras fúngicas que tienen la capacidad de isomerizar xilosa en xilulosa como se describe, por ejemplo, en el documento WO 03/0624430 y en el documento WO 06/009434. La capacidad de isomerizar xilosa en xilulosa se confiere preferentemente a la célula por transformación con una construcción de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica una xilosa isomerasa. Preferentemente, la célula adquiere así la capacidad de isomerizar directamente xilosa en xilulosa. Más preferentemente, la célula adquiere así la capacidad de crecer aeróbicamente y/o anaeróbicamente en xilosa como única fuente de energía y/o de carbono mediante la isomerización directa de xilosa en xilulosa (y metabolismo adicional de xilulosa). Se entiende en el presente documento que la isomerización directa de xilosa en xilulosa se produce en una única reacción catalizada por una xilosa isomerasa, a diferencia de la conversión de dos etapas de xilosa en xilulosa mediante un producto intermedio de xilitol como se cataliza por xilosa reductasa y xilitol deshidrogenasa, respectivamente.

Se han descrito en la materia varias xilosa isomerasas (y sus secuencias de aminoácidos y de nucleótidos codificantes) que pueden ser satisfactoriamente usadas para conferir a la célula de la invención la capacidad de isomerizar directamente la xilosa en xilulosa. Éstas incluyen las xilosa isomerasas de *Piromyces* sp. y de otros hongos anaerobios que pertenecen a las familias *Neocallimastix*, *Caecomyces*, *Piromyces* o *Ruminomyces* (documento WO 03/0624430), *Cyllamyces aberensis* (documento US 20060234364), *Orpinomyces* (Madhavan et al., 2008, DOI 10.1007/s00253-008-1794-6), la xilosa isomerasa del género bacteriano *Bacteroides*, que incluye, por ejemplo, *B. thetaiotaomicron* (documento WO 06/009434), *B. fragilis* y *B. uniformis* (documento WO 09/109633), la xilosa isomerasa de la bacteria anaerobia *Clostridium phytofermentans* (Brat et al., 2009, Appl. Environ. Microbiol. 75: 2304-2311), y las xilosas isomerasas de *Clostridium difficile*, *Ciona intestinalis* y *Fusobacterium mortiferum*.

Células hospedadoras fúngicas que tienen la capacidad de convertir L-arabinosa en D-xilulosa 5-fosfato como se describe, por ejemplo, en Wisselink et al. (2007, AEM Accepts, publicado en línea antes de impreso el 1 de junio de 2007; Appl. Environ. Microbiol. doi:10.1128/AEM,00177-07) y en el documento EP 1 499 708. La capacidad de convertir L-arabinosa en D-xilulosa 5-fosfato se confiere preferentemente a la célula por transformación con una construcción (construcciones) de ácidos nucleicos que comprenden secuencias de nucleótidos que codifican a) una

arabinosa isomerasa; b) una ribulocinasa, preferentemente una L-ribulocinasa a xilosa isomerasa; y c) una ribulosa-5-P-4-epimerasa, preferentemente una L-ribulosa-5-P-4-epimerasa. Preferentemente, en las células de la invención, la capacidad de convertir L-arabinosa en D-xilulosa 5-fosfato es la capacidad de convertir L-arabinosa en D-xilulosa 5-fosfato mediante las reacciones posteriores de 1) isomerización de arabinosa en ribulosa; 2) fosforilación de ribulosa en ribulosa 5-fosfato; y, 3) epimerización de ribulosa 5-fosfato en D-xilulosa 5-fosfato. Secuencias de nucleótidos adecuadas que codifican arabinosa isomerasas, ribulocinasas y ribulosa-5-P-4-epimerasas pueden obtenerse de *Bacillus subtilis*, *Escherichia coli* (véase, por ejemplo, el documento EP 1 499 708), lactobacilos, por ejemplo, *Lactobacillus plantarum* (véase, por ejemplo, Wisselink et al., arriba), o especies de *Clavibacter*, *Arthrobacter* y *Gramella*, de los que preferentemente *Clavibacter michiganensis*, *Arthrobacter aureescens* y *Gramella forsetii* (véase el documento WO2009/011591).

La célula hospedadora transformada de la invención comprende además preferentemente actividad de xilulosa cinasa de manera que la xilulosa isomerizada de xilosa puede ser metabolizada en piruvato. Preferentemente, la célula contiene actividad de xilulosa cinasa endógena. Más preferentemente, una célula de la invención comprende una modificación genética que aumenta la actividad de xilulosa cinasa específica. Preferentemente, la modificación genética produce la expresión en exceso de una xilulosa cinasa, por ejemplo, por 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 o puede ser una xilulosa cinasa que es heteróloga a la célula. Una secuencia de nucleótidos que puede usarse para la expresión en exceso de xilulosa cinasa en las células de la invención es, por ejemplo, el gen de xilulosa cinasa de *S. cerevisiae* (*XKS1*) como se describe por Deng y Ho (1990, Appl. Biochem. Biotechnol. 24-25: 193-199). Otra xilulosa cinasa preferida es una xilosa cinasa que está relacionada con la xilulosa cinasa de *Piromyces* (*xyiB*; véase el documento WO 03/0624430). Esta xilulosa cinasa de *Piromyces* está en realidad más relacionada con cinasa procariota que con todas las cinasas eucariotas conocidas, tales como la levadura cinasa. Las xilulosa cinasas eucariotas han sido indicadas como cinasas de azúcar no específicas, que tienen una amplia variedad de sustratos que incluyen xilulosa. A diferencia, se ha indicado que las xilulosa cinasas procariotas, con las que la cinasa de *Piromyces* está más estrechamente relacionada, son cinasas más específicas para xilulosa, es decir, que tienen una variedad de sustratos más estrecha. En las células de la invención, una xilulosa cinasa que va a expresarse en exceso se expresa en exceso al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 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 en estado estacionario de la actividad de la enzima, el nivel en estado estacionario de la proteína de la enzima, además de al nivel en estado estacionario del transcrito que codifica la enzima.

Una célula de la invención comprende además preferentemente una modificación genética que aumenta el flujo de la vía de pentosa fosfato como se describe en el documento WO 06/009434. En particular, la modificación genética produce un aumento de flujo de la vía de pentosa fosfato de parte no oxidativa. Una modificación genética que produce un aumento de flujo de la parte no oxidativa de la vía de pentosa fosfato se entiende en el presente documento que significa una modificación que aumenta el flujo al menos un factor 1,1, 1,2, 1,5, 2, 5, 10 o 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 pentosa fosfato puede medirse como se describe en el documento WO 06/009434.

Pueden introducirse modificaciones genéticas que aumentan el flujo de la vía de pentosa fosfato en las células de la invención de diversas formas. Éstas incluyen, por ejemplo, lograr niveles de actividad en estado estacionario más altos de xilulosa cinasa y/o una o más de las enzimas de la vía de pentosa fosfato de parte no oxidativa y/o un nivel en estado estacionario reducido de actividad de aldosa reductasa no específica. Estos cambios en los niveles de actividad en estado estacionario pueden efectuarse por selección de mutantes (espontánea o inducida 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 las enzimas o factores que regulan estos genes.

En una célula preferida de la invención, la modificación genética comprende la expresión en exceso de al menos una enzima de la vía de pentosa fosfato (de parte no oxidativa). Preferentemente, la enzima está seleccionada del grupo que consiste en las enzimas que codifican ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato 3-epimerasa, transcetolasa y transaldolasa. Pueden expresarse en exceso diversas combinaciones de enzimas de la vía de pentosa fosfato (de parte no oxidativa). Por ejemplo, las enzimas que se expresan en exceso pueden ser al menos las enzimas ribulosa-5-fosfato isomerasa y ribulosa-5-fosfato 3-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 3-epimerasa y transcetolasa; o al menos las enzimas ribulosa-5-fosfato 3-epimerasa y transaldolasa; o al menos las enzimas transcetolasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato 3-epimerasa, transcetolasa y transaldolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato 3-epimerasa, y transaldolasa; o al menos las enzimas ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato 3-epimerasa, y transcetolasa. En una realización de la invención, cada una de las enzimas ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato 3-epimerasa, transcetolasa y transaldolasa se expresan en exceso en la célula de la invención. Se prefiere una célula en la que la modificación genética comprenda al menos la expresión en exceso de la enzima transaldolasa. Es más preferida una célula en la que la modificación genética comprenda al menos la expresión en exceso de ambas enzimas transcetolasa y transaldolasa ya que una célula hospedadora tal ya es capaz de crecimiento anaerobio sobre xilosa.

En realidad, en algunas condiciones, los presentes inventores han encontrado que las células que expresan en exceso solo la transcetolasa y la transaldolasa ya tienen la misma velocidad de crecimiento anaerobio sobre xilosa que las células que expresan en exceso las cuatro enzimas, es decir, la ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato 3-epimerasa, transcetolasa y transaldolasa. Además, se prefieren las células de la invención que expresan en exceso ambas de las enzimas ribulosa-5-fosfato isomerasa y ribulosa-5-fosfato 3-epimerasa con respecto a las células que expresan en exceso solo la isomerasa o solo la 3-epimerasa ya que la expresión en exceso de solo una de estas enzimas puede producir desequilibrios metabólicos.

Existen diversos medios disponibles en la materia para la expresión en exceso de enzimas en las células de la invención. En particular, una enzima puede expresarse en exceso aumentando el número de copias del gen que codifica la enzima en la célula, por ejemplo, integrando copias adicionales del gen en el genoma de la célula, expresando el gen a partir de un vector de expresión multicopia episómico o introduciendo un vector de expresión episómico que comprende múltiples copias del gen. La secuencia codificante usada para la expresión en exceso de las enzimas es preferentemente homóloga a la célula hospedadora de la invención. Sin embargo, asimismo pueden aplicarse secuencias codificantes que son heterólogas a la célula hospedadora de la invención.

Alternativamente, la expresión en exceso de enzimas en las células de la invención puede lograrse usando un promotor que no es nativo para la secuencia que codifica la enzima que va a expresarse en exceso, es decir, un promotor que es heterólogo para la secuencia codificante con la que está operativamente unido. Aunque el promotor es preferentemente heterólogo para la secuencia codificante con la que está operativamente unido, también se prefiere que el promotor sea homólogo, es decir, endógeno para la célula de la invención. Preferentemente, el promotor heterólogo es capaz de producir un nivel en estado estacionario más alto 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 para la secuencia codificante, preferentemente en condiciones donde la xilosa o xilosa y glucosa están disponibles como fuentes de carbono, más preferentemente como fuentes de carbono principales (es decir, más del 50 % de la fuente de carbono disponible consiste en xilosa o xilosa y glucosa), lo más preferentemente como únicas fuentes de carbono.

Una célula preferida adicional de la invención comprende una modificación genética que reduce la actividad de aldosa reductasa no específica en la célula. Preferentemente, la actividad de aldosa reductasa no específica se reduce en la célula hospedadora por una o más modificaciones genéticas que reducen la expresión de o inactivan un gen que codifica una aldosa reductasa no específica. Preferentemente, las modificaciones genéticas reducen o inactivan la expresión de cada copia endógena de un gen que codifica una aldosa reductasa no específica que es capaz de reducir una aldopentosa, que incluye, xilosa, xilulosa y arabinosa, en el genoma de la célula. Una célula dada puede comprender múltiples copias de genes que codifican aldosa reductasas no específicas como resultado de di-, poli- o aneu-ploidía, y/o una célula puede contener varias (iso)enzimas diferentes con actividad de aldosa reductasa que se diferencian en la secuencia de aminoácidos y que están cada una codificada por un gen diferente. También en tales casos, preferentemente la expresión de cada gen que codifica una aldosa reductasa no específica se reduce o inactiva. Preferentemente, el gen se inactiva por delección de al menos parte del gen o por rotura 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 o inactivación (parcial) produce una reducción de la expresión de la actividad de aldosa reductasa no específica en la célula hospedadora. Una secuencia de nucleótidos que codifica una aldosa reductasa cuya actividad va a reducirse en la célula de la invención y secuencias de aminoácidos de tales aldosa reductasas se describen en el documento WO 06/009434 e incluyen, por ejemplo, los genes de aldosa reductasa (no específica) del gen *GRE3* de *S. cerevisiae* (Träff et al., 2001, Appl. Environm. Microbiol. 67: 5668-5674) y ortólogos del mismo en otras especies.

Una célula hospedadora transformada preferida adicional según la invención puede comprender además modificaciones genéticas que producen el elevado transporte de xilosa y/o arabinosa en la célula. La célula hospedadora transformada desvelada en el presente documento puede comprender además modificaciones genéticas que producen una o más de las características seleccionadas del grupo que consiste en (b) reducida sensibilidad a la represión de catabolitos; (c) elevada tolerancia a etanol, osmolaridad o ácidos orgánicos; y, (d) reducida producción de subproductos. Se entiende que los subproductos significan moléculas que contienen carbono distintas del producto de fermentación deseado e incluyen, por ejemplo, xilitol, arabinitol, glicerol y/o ácido acético. Cualquier modificación genética descrita en el presente documento puede introducirse por mutagénesis clásica y cribado y/o selección del mutante deseado, o simplemente por cribado y/o selección de los mutantes espontáneos con las características deseadas. Alternativamente, las modificaciones genéticas pueden consistir en la expresión en exceso de genes endógenos y/o la inactivación de genes endógenos. Genes cuya expresión en exceso se desea para el elevado transporte de arabinosa y/o xilosa en la célula se eligen preferentemente de genes que codifican un transportador de hexosa o pentosa. En *S. cerevisiae* y otras levaduras, estos genes incluyen *HXT1*, *HXT2*, *HXT4*, *HXT5*, *HXT7* y *GAL2*, de los que *HXT7*, *HXT5* y *GAL2* son los más preferidos (véase Sedlack y Ho, Yeast 2004; 21: 671-684). Otro transportador preferido para la expresión en levadura es el transportador de glucosa codificado por el gen *SUT1* de *P. stipitis* (Katahira et al., 2008, Enzyme Microb. Technol. 43: 115-119). Similarmente, pueden expresarse en exceso ortólogos de estos genes transportadores en otras especies. Otros genes que pueden expresarse en exceso en las células incluyen genes que codifican enzimas glucolíticas y/o enzimas etanológicas tales como alcohol deshidrogenasas. Genes endógenos preferidos para la inactivación incluyen genes de hexosa cinasa, por ejemplo, el gen *HXK2* de *S. cerevisiae* (véase Diderich et al., 2001, Appl. Environ. Microbiol. 67: 1587-

1593); los genes *MIG1* o *MIG2* de *S. cerevisiae*; genes que codifican enzimas implicadas en el metabolismo de glicerol tales como los genes glicerol-fosfato deshidrogenasa 1 y/o 2 de *S. cerevisiae*; u ortólogos (de hibridación) de estos genes en otras especies. Otras modificaciones adicionalmente preferidas de células hospedadoras para la fermentación de xilosa se describen en van Maris et al. (2006, Antonie van Leeuwenhoek 90:391-418), documentos  
 5 WO2006/009434, WO2005/023998, WO2005/111214 y WO2005/091733. Cualquiera de las modificaciones genéticas de las células como se describen en el presente documento se introducen o modifican preferentemente, en la medida de lo posible, por modificación genética de auto-clonación.

Una célula hospedadora preferida según la invención tiene la capacidad de crecer sobre al menos una de xilosa y arabinosa como fuente de carbono/energía, preferentemente como única fuente de carbono/energía, y  
 10 preferentemente en condiciones anaerobias, es decir, condiciones como se definen en el presente documento más adelante para el proceso de fermentación anaerobia. Preferentemente, cuando crece sobre xilosa como fuente de carbono/energía, la célula hospedadora esencialmente no produce xilitol, por ejemplo, el xilitol producido está por debajo del límite de detección o, por ejemplo, es inferior al 5, 2, 1, 0,5, o 0,3 % del carbono consumido en una base molar. Preferentemente, cuando crece sobre arabinosa como fuente de carbono/energía, la célula esencialmente no  
 15 produce arabinitol, por ejemplo, el arabinitol producido está por debajo del límite de detección o, por ejemplo, es inferior al 5, 2, 1, 0,5 o 0,3 % del carbono consumido en una base molar.

Una célula hospedadora preferida de la invención tiene la capacidad de crecer sobre una combinación de: a) al menos uno de una hexosa y una pentosa; b) ácido acético; y c) glicerol a una velocidad de al menos 0,01, 0,02, 0,05, 0,1, 0,2, 0,25 o 0,3 h<sup>-1</sup> en condiciones aerobias, o, más preferentemente, a una velocidad de al menos 0,005,  
 20 0,01, 0,02, 0,05, 0,08, 0,1, 0,12, 0,15 o 0,2 h<sup>-1</sup> en condiciones anaerobias. Por tanto, preferentemente, la célula hospedadora tiene la capacidad de crecer sobre al menos una de xilosa y arabinosa como única fuente de carbono/energía a una velocidad de al menos 0,01, 0,02, 0,05, 0,1, 0,2, 0,25 o 0,3 h<sup>-1</sup> en condiciones aerobias, o, más preferentemente, a una velocidad de al menos 0,005, 0,01, 0,02, 0,05, 0,08, 0,1, 0,12, 0,15 o 0,2 h<sup>-1</sup> en  
 25 condiciones anaerobias. Más preferentemente, la célula hospedadora tiene la capacidad de crecer sobre una mezcla de una hexosa (por ejemplo glucosa) y al menos una de xilosa y arabinosa (en una relación de peso 1:1) como única fuente de carbono/energía a una velocidad de al menos 0,01, 0,02, 0,05, 0,1, 0,2, 0,25 o 0,3 h<sup>-1</sup> en condiciones aerobias, o, más preferentemente, a una velocidad de al menos 0,005, 0,01, 0,02, 0,05, 0,08, 0,1, 0,12, 0,15 o 0,2 h<sup>-1</sup> en condiciones anaerobias.

En otro aspecto, la invención se refiere a un proceso de producción de etanol. El proceso comprende preferentemente la etapa de: a) fermentar un medio con una célula de levadura, por lo que el medio contiene o se  
 30 alimenta con: a) una fuente de al menos una de una hexosa y una pentosa; b) una fuente de ácido acético; y, c) una fuente de glicerol y por lo que la célula de levadura fermenta ácido acético, glicerol y al menos una de la hexosa y pentosa en etanol. La célula de levadura es una célula (hospedadora) como se ha definido anteriormente en el presente documento. El proceso comprende preferentemente una etapa adicional en la que se recupera el producto de fermentación, es decir, el etanol. El proceso pueden ser un proceso discontinuo, un proceso de lotes alimentados o un proceso continuo como son muy conocidos en la técnica. En el proceso de la invención, la fuente de glicerol puede ser cualquier fuente de carbono que tenga un estado más reducido que la glucosa. Una fuente de carbono que tiene un estado más reducido que la glucosa se entiende como una fuente de carbono de la que el estado de reducción promedio por moles de C (de los compuestos en su interior) es superior al estado de reducción por moles de C de glucosa. Ejemplos de fuentes de carbono que tienen un estado más reducido que la glucosa incluyen, por  
 35 ejemplo, alcoholes tales como propanol y butanol; polioles tales como 1,3-propanodiol, butanodiol, glicerol, manitol y xilitol.

En un proceso preferido, la fuente de hexosa comprende o consiste en glucosa. Preferentemente, la fuente de pentosa comprende o consiste en al menos una de xilosa y arabinosa. Preferentemente, el medio fermentado por las  
 45 células de la invención comprende o se alimenta con (fracciones de) biomasa hidrolizada que comprende al menos una al menos una de una hexosa y una pentosa tal como glucosa, xilosa y/o arabinosa. (Fracciones de) Biomasa hidrolizada que comprende las hexosas y pentosa normalmente también comprenderán ácido acético (o una sal del mismo). Un ejemplo de biomasa hidrolizada para ser fermentada en los procesos de la invención es, por ejemplo, biomasa lignocelulósica hidrolizada. Biomasa lignocelulósica se entiende en el presente documento como biomasa  
 50 de plantas que está compuesta por celulosa, hemicelulosa y lignina. Los polímeros de hidrato de carbono (celulosa y hemicelulosas) están fuertemente unidos a la lignina. Ejemplos de biomasa lignocelulósica que va a hidrolizarse para su uso en la presente invención incluyen residuos agrícolas (incluyendo hojas y tallos de maíz y bagazo de caña de azúcar), residuos de madera (incluyendo desechos de aserradero y de fábricas de papel) y residuos (municipales) de papel. Se conocen en sí en la técnica métodos de hidrólisis de biomasa tales como lignocelulosas e incluyen, por  
 55 ejemplo, ácidos, tales como ácido sulfúrico y enzimas tales como celulasas y hemicelulasas.

En el proceso de la invención, las fuentes de xilosa, glucosa y arabinosa pueden ser xilosa, glucosa y arabinosa como tales (es decir, como azúcares monoméricos) o pueden estar en forma de cualquier oligo- o polímero de hidrato de carbono que comprende unidades de xilosa, glucosa y/o arabinosa, tales como, por ejemplo, lignocelulosa, arabinanos, xilanos, celulosa, almidón y similares. Para la liberación de las unidades de xilosa,  
 60 glucosa y/o arabinosa de tales hidratos de carbono, pueden añadirse carbohidrasas apropiadas (tales como arabinasas, xilanasas, glucanasas, amilasas, celulasas, glucanasas y similares) al medio de fermentación o pueden ser producidas por la célula hospedadora modificada. En el último caso, la célula hospedadora modificada puede ser

genéticamente manipulada para producir y excretar tales carbohidrasas. Una ventaja adicional de uso de fuentes oligo- o poliméricas de glucosa es que permite mantener una concentración (más) baja de glucosa libre durante la fermentación, por ejemplo, usando cantidades limitantes de la velocidad de las carbohidrasas, preferentemente durante la fermentación. Esto, a su vez, prevendrá la represión de sistemas requeridos para el metabolismo y transporte de azúcares no de glucosa tales como xilosa y arabinosa. En un proceso preferido, la célula hospedadora modificada fermenta tanto la glucosa como al menos una de xilosa y arabinosa, preferentemente simultáneamente, en cuyo caso preferentemente se usa una célula hospedadora modificada que es insensible a la represión de glucosa para prevenir el crecimiento diaúxico. Además de una fuente de al menos una de xilosa y arabinosa (y glucosa) como fuente de carbono, el medio de fermentación comprenderá además el componente apropiado requerido para el crecimiento de la célula hospedadora modificada. Composiciones de medios de fermentación para el crecimiento de levaduras son muy conocidas en la técnica.

En el proceso de la invención, el medio comprende y/o se alimenta de una fuente de glicerol. El glicerol para uso en el proceso de la presente invención puede ser ventajosamente glicerol que se genera como un subproducto en la producción de biodiésel a partir de reacciones de transesterificación usando aceites vegetales o grasas animales y un alcohol.

El proceso de fermentación puede ser un proceso de fermentación aerobio o uno 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 menos de 5, 2,5 o 1 mmol/l/h, más preferentemente se consumen 0 mmol/l/h (es decir, el consumo de oxígeno no es detectable), y en el que las moléculas orgánicas sirven de tanto donante de electrones como aceptores de electrones. En ausencia de oxígeno, la NADH producida en la glicólisis y formación de biomasa no puede ser oxidada por la fosforilación oxidativa. Para resolver este problema, muchos microorganismos usan piruvato o uno de sus derivados como electrón y aceptor de hidrógeno regenerando así  $\text{NAD}^+$ . Así, en un proceso de fermentación anaerobio preferido se usa piruvato como electrón (y aceptor de hidrógeno) y se reduce a productos de fermentación tales como etanol, además de productos de fermentación no de etanol tales como ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, 1,3-propanodiol, butanoles (1-butanol, 2-butanol, isobutanol), productos derivados de isoprenoide. Se prefieren procesos anaerobios de la invención con respecto a los procesos aerobios debido a que los procesos anaerobios no requieren inversiones y energía para la aireación y además los procesos anaerobios producen rendimientos de producto más altos que los procesos aerobios.

Alternativamente, el proceso de fermentación de la invención puede realizarse en condiciones aerobias de oxígeno limitado. Preferentemente, en un proceso aerobio en condiciones de oxígeno limitado, la velocidad de consumo de oxígeno es al menos 5,5, más preferentemente al menos 6 e incluso más preferentemente al menos 7 mmol/l/h. En un proceso de fermentación aerobio de oxígeno limitado preferido de la invención, la célula de levadura de la invención consume menos del 30, 20, 18, 15, 12, 10, 8 o 5 % de la cantidad de oxígeno en una base molar de C relacionada con la fuente de carbono consumida durante la conversión de la fuente de carbono en el producto de fermentación. El coeficiente de conversión de oxígeno consumido con respecto al sustrato utilizado en una base molar de C ( $C_{OS}$ ) se entiende en el presente documento que significa moles de  $\text{O}_2$  usados por moles de C de la fuente de carbono consumida. Así, un proceso de la invención puede llevarse a cabo en condiciones anaerobias estrictas (es decir,  $C_{OS} = 0,0$ ), o el proceso de la invención puede llevarse a cabo en condiciones aerobias, preferentemente de oxígeno limitado, en las que  $C_{OS}$  es preferentemente inferior a 0,3, 0,2, 0,18, 0,15, 0,12, 0,1, 0,08 o 0,05.

El proceso de fermentación se realiza preferentemente a una temperatura que es óptima para las células modificadas de la invención. Así, para la mayoría de las células de levadura, el proceso de fermentación se realiza a una temperatura que es inferior a 42 °C, preferentemente inferior a 38 °C. Para células de levadura, el proceso de fermentación se realiza preferentemente a una temperatura que es inferior a 35, 33, 30 o 28 °C y a una temperatura que es superior a 20, 22 o 25 °C. El proceso de fermentación según la invención es un proceso para la producción de etanol, por lo que el proceso comprende la etapa de fermentar un medio con una célula de levadura según la invención como se ha descrito anteriormente, por lo que el medio contiene o se alimenta con: a) una fuente de al menos una de una hexosa y una pentosa; b) una fuente de ácido acético; y, c) una fuente de glicerol, por lo que la célula de levadura fermenta ácido acético, glicerol y al menos una de la hexosa y pentosa en etanol, y opcionalmente, b) recuperación del etanol. El medio de fermentación puede realizarse además como se ha descrito anteriormente. En el proceso, la productividad volumétrica de etanol es preferentemente al menos 0,5, 1,0, 1,5, 2,0, 2,5, 3,0, 5,0 o 10,0 g de etanol por litro por hora. El rendimiento de etanol sobre xilosa y/o glucosa y/o arabinosa y/o acetato y/o glicerol en el proceso es preferentemente al menos del 50, 60, 70, 80, 90, 95 o 98 %. El rendimiento de etanol se define en el presente documento como un porcentaje del máximo rendimiento teórico que, para xilosa, glucosa y arabinosa es 0,51 g de etanol por g de xilosa, glucosa o arabinosa. Para glicerol, el máximo rendimiento teórico es 0,50 g de etanol por g de glicerol y para ácido acético el máximo rendimiento teórico es 0,77 g de etanol por g de ácido acético.

En este documento y en sus reivindicaciones, el vector "comprender" y sus conjugaciones se usan en su sentido no limitante para significar que artículos que siguen a la palabra están incluidos, pero artículos no específicamente mencionados no están excluidos. Además, referencia a un elemento por el artículo indefinido "un" o "una" no excluye

la posibilidad de que más de uno del elemento esté presente, a menos que el contexto requiera claramente que haya uno y solo uno de los elementos. Así, el artículo indefinido "un" o "una" normalmente significa "al menos uno".

Los siguientes ejemplos se ofrecen para fines ilustrativos solo, y no pretenden limitar el alcance de la presente invención de ningún modo.

## 5 Descripción de las figuras

Figura 1. Se muestra la evolución de los niveles netos de glicerol (g/l) (es decir, producción menos consumo) con el tiempo (horas) para las cepas de *S. cerevisiae* RN1041, RN1041+pRN595, RN1186, RN1187, RN1188 y RN1189.

10 Figura 2. Se muestra la evolución de los niveles netos de ácido acético (g/l) (es decir, producción menos consumo) con el tiempo (horas) para las cepas de *S. cerevisiae* RN1041, RN1041+pRN595, RN1186, RN1187, RN1188 y RN1189.

## Ejemplos

### 1. Ensayos de actividad enzimática

15 Se prepararon extractos de células para ensayos de actividad a partir de cultivos discontinuos aerobios o anaerobios de crecimiento exponencial y se analizaron para el contenido de proteína como se describe por Abbot et al., (2009, Appl. Environ. Microbiol. 75: 2320-2325).

20 Se midió la actividad de acetaldehído deshidrogenasa dependiente de NAD<sup>+</sup> (EC 1.2.1.10) a 30 °C monitorizando la oxidación de NADH a 340 nm. La mezcla de reacción (volumen total 1 ml) contuvo tampón fosfato de potasio 50 mM (pH 7,5), NADH 0,15 mM y extracto de células. La reacción empezó mediante la adición de acetil-Coenzima A 0,5 mM.

25 Para la determinación de la actividad de glicerol 3-fosfato deshidrogenasa (EC 1.1.1.8), se prepararon extractos de células como se ha descrito anteriormente, excepto que el tampón fosfato se sustituyó por tampón trietanolamina (10 mM, pH 5). Se ensayaron las actividades de glicerol-3-fosfato deshidrogenasa en extractos de células a 30 °C como se ha descrito previamente (Blomberg y Adler, 1989, J. Bacteriol. 171: 1087-1092,9). Las velocidades de reacción fueron proporcionales a las cantidades de extracto de células añadidas.

30 Se midió la actividad de acetil-CoA sintasa (EC 6.2.1.1) como se describe por Frenkel y Kitchens (1977, J. Biol. Chem. 252: 504-507) que es una modificación del método de Webster (Webster, 1969, Methods Enzymol. 13: 375-381). La formación de NADH se mide espectrofotométricamente cuando la acetil-CoA producida se acopla con reacciones de citrato sintasa y malato deshidrogenasa. El sistema de ensayo contuvo Tris-Cl 100 mM (pH 7,6), MgCl<sub>2</sub> 10 mM, ATP 6 mM, malato 5 mM, NAD<sup>+</sup> 1 mM, NADH 0,1 mM, ditiotreitól o 2-mercaptoetanol 2,5 mM, coenzima A 0,2 mM, 25 µg de citrato sintasa (80 unidades/mg), 20 µg de malato deshidrogenasa (1000 unidades/mg) y acetato 10 mM y se midió la velocidad de reacción a 340 nm y se calculó a partir del coeficiente de extinción de NADH (6,22 x 10<sup>6</sup> cm<sup>2</sup>/mol).

35 Se mide la actividad de glicerol deshidrogenasa y dihidroxiacetona cinasa a 30 °C en extractos de células, esencialmente como se describe previamente (Gonzalez et al., 2008, Metab. Eng. 10, 234-245). Las actividades enzimáticas de glicerol deshidrogenasa y dihidroxiacetona cinasa se informan como µmoles de sustrato/min/mg de proteína de célula.

### 2. Construcción de cepas

40 Todas las modificaciones empiezan con la cepa fermentadora de xilosa y arabinosa RN1008 his<sup>-</sup>. RN1008 his<sup>-</sup>, también denominada en el presente documento RN1041, es una cepa fermentadora de arabinosa y xilosa basada en CEN.PK) con el genotipo:

Mat a, ura3-52, leu2-112, his3::loxP, gre3::loxP, loxP-Ptpi::TAL1, loxP-Ptpi::RKI1, loxP-Ptpi-TKL1, loxP-Ptpi-RPE1, delta::-LEU2, delta:: Padh1XKS1Tcyc1-URA3-Ptpi-xylA-Tcyc1, delta:: LEU2-AAAaraABD.

Mat a = tipo de apareamiento a

45 Mutaciones ura3-52, leu2-112, his3::loxP en los genes ura3, leu2 y his3, ura3 está complementado por la construcción de expresión en exceso de xylA-XKS, leu2 está complementado por la construcción de expresión en exceso de AraABD. his3 podrían usarse para la selección de plásmidos adicionales, RN1041 necesita histidina en el medio para el crecimiento.

50 gre3::loxP = delección del gen gre3 que codifica xilosa reductasa, queda el sitio loxP después de la eliminación del marcador.

loxP-Ptpi .....= expresión en exceso de la vía de pentosa fosfato het, queda el sitio loxP en la dirección 5' del promotor constitutivo después de la eliminación del marcador

delta:: = integración de la construcción después de la recombinación en las repeticiones terminales largas del retrotransposón Ty1.

- 5 AAAaraABD = genes *araA*, *araB* y *araD* de *Arthrobacter aureescens* optimizados en codones (véase el documento WO2009/011591)

#### Construcciones de delección para GPD1 y GPD2

- 10 La delección de GPD1 en RN1041 produce la cepa RN1197. La delección de GPD2 en RN1041 produce la cepa RN1198. En esta cepa, posteriormente se deleciona *gpd1* para producir la cepa RN1199. En estas cepas se introdujeron plásmidos para la expresión en exceso de los genes ACS (RN1200 a RN1207, Tabla 4) y genes adicionales como se indica en la Tabla 4.

#### *gpd1::hphMX*

- 15 Se usan cebadores *gpd1uf*, *gpd1ur*, *gpd1df* y *gpd1dr* para la amplificación de fragmentos de secuencias genómicas en la dirección 5' y en la dirección 3' del gen GPD1 para su inactivación. Tanto los fragmentos GPD1 en la dirección 5' como en 3' se clonan en un vector topo blunt (Invitrogen) dando *pGPD1up* y *pGPD1down*, respectivamente.

*gpd1uf*: AAGCTTGGTACCCGCTTGTCTCTCTCCCC (SEQ ID NO: 41)

*gpd1ur*: TCTAGACCAGCATTCAAGTGGCCGGA (SEQ ID NO: 42)

*gpd1df*: CGTACGAGTTGTTGAATGGCCAATCCGCT (SEQ ID NO: 43)

*gpd1dr*: CCATGGTACCGAGTGGTGTGTAACCACCCT (SEQ ID NO: 44)

- 20 *gpd1cf*: ACCAATACGTAAACGGGGCG (SEQ ID NO: 45)

*gpd1cr*: AATACACCCATACATACGGACGC (SEQ ID NO: 46)

- 25 Se construye el plásmido *pRN593* (SEQ ID NO: 40) por ligación del fragmento cortado con HindIII y XbaI de *pGPD1up* hasta el fragmento *hphMX* cortado con SpeI y BsrGI (colección de plásmidos C5YeastCompany) y el fragmento cortado con BsiWI y NcoI de *pGPD1down* en el vector topo T/A cortado con HindIII y NcoI (Invitrogen). Se corta el plásmido *pRN593* con KpnI para obtener el fragmento de delección para romper la copia genómica (SEQ ID NO: 17). La mezcla de fragmentos lineales se usa para la transformación de levadura. Los transformantes están seleccionados para resistencia a higromicina. La correcta integración produce la delección del marco de lectura abierto de GPD1. La integración se verifica por PCR con los cebadores *gpd1cf* y *gpd1cr*.

#### *gpd2::natMX*

- 30 Se usan cebadores *GPD2uf*, *GPD2ur*, *GPD2df* y *GPD2dr* para la amplificación de fragmentos de secuencias genómicas en la dirección 5' y en la dirección 3' del gen GPD2 para su inactivación. Se amplifica un fragmento de PCR de 407 pb en la dirección 5' con un sitio *AflI* en el extremo 3' (derivado de la secuencia de GPD2) y un sitio *BglII* en el extremo 5' (para el aislamiento de la construcción de delección) usando *GPD2uf*, *GPD2ur* y se clona en *pCR2.1* (topo T/A, Invitrogen).

- 35 *GPD2uf*: GGTACCAGATCTTTTGC GCGGAGGTGCCG (SEQ ID NO: 32)

*GPD2ur*: TCTAGACTTAAGGAATGTGTATCTTGTTAATCTTCTGACAGC (SEQ ID NO: 33)

Se amplifica un fragmento de PCR de 417 pb en la dirección 3' con un sitio *XhoI* en el extremo 5' y un sitio *BglII* en el extremo 3' usando *GPD2df* y *GPD2dr*.

*GPD2df*: CTCGAGATAGTCTACAACAACGTCCGCA (SEQ ID NO: 34)

- 40 *GPD2dr*: CCATGGAGATCTGCAGTGAAAAAGCTCGAAGAAACAGCT (SEQ ID NO: 35)

- 45 Para la construcción final, el plásmido que contiene el fragmento en la dirección 5' se corta con *AflI* y *KpnI*, el fragmento en la dirección 3' se corta con *XhoI* en *NcoI* y el marcador *natMX* (colección de plásmidos Royal Nedalco) se corta con *AflI* en *XhoI* y los fragmentos se ligan para producir el plásmido *pRN594* (SEQ ID NO: 36). *pRN594* se corta con *BglII* antes de la transformación en levadura. Los transformantes están seleccionados para resistencia a nourseotricina. La correcta integración se verifica por PCR.

#### Método de clonación para la expresión en exceso de los genes ACS1 y ACS2 de *Saccharomyces cerevisiae*

Se amplifica por PCR el marco de lectura abierto de ACS1 con los cebadores *acs1f* y *acs1r*.

acs1f: TTAAGCTTAAATGTCGCCCTCTGCCGT (SEQ ID NO: 47)

acs1r: AAGCGCGCTACAACCTTGACCGAATCAATTAGATGTCTACAATGCCAGGG (SEQ ID NO: 48)

5 Este fragmento de PCR se corta con las enzimas de restricción HindIII y BssHII y se liga al fragmento del promotor TEF1 cortado con Sall y HindIII (colección C5YeastCompany) y el fragmento del terminador ADH1 cortado con BssHII y BsiWI (colección C5YeastCompany). Este fragmento combinado se somete a PCR con cebadores específicos de promotor y terminador y se clona en el vector topo Blunt (InVitrogen) dando pACS1.

Se amplifica por PCR el marco de lectura abierto de ACS2 con los cebadores acs2f y acs2r.

acs2f: AACTGCAGAAAATGACAATCAAGGAACATAAAGTAGTTTATGAAGCTCA (SEQ ID NO: 49)

acs2r: ACGTCGACTATTTCTTTTTTTGAGAGAAAAATTGGTTCTCTACAGCAGA (SEQ ID NO: 50)

10 Este fragmento de PCR se corta con las enzimas de restricción PstI y Sall y se liga al fragmento del promotor PGK1 cortado con SpeI y PstI (colección C5YeastCompany) y el fragmento del terminador PGI1 cortado con XhoI y BsiWI (colección C5YeastCompany). Este fragmento combinado se somete a PCR con cebadores específicos de promotor y terminador y se clona en el vector topo Blunt (InVitrogen) dando el plásmido pACS2.

15 La construcción de expresión en exceso de ACS1 se corta de pACS1 con las enzimas de restricción Sall y BsiWI, la construcción de expresión en exceso de ACS2 se corta de pACS2 con las enzimas de restricción SpeI y BsiWI, el marcador KanMX se corta con BspEI y XbaI (colección de plásmidos C5YeastCompany). Estos fragmentos se ligan al plásmido pRS306 + 2 $\mu$  ORI (colección de plásmidos C5Yeast company) cortado con BspEI y XhoI dando el plásmido final pRN753 (SEQ ID NO: 51). Este plásmido se usa para transformar cepas de levadura como se indica en la Tabla 4 y los transformantes se seleccionan con resistencia a G418. La expresión en exceso se verifica por qPCR. Un plásmido alternativo que puede usarse para la expresión en exceso de ACS1 y ACS 2 es pRN500 (SEQ ID NO: 20).

#### Expresión de adhE de *E. coli*, ADH2 de *E. histolytica* o mphF de *E. coli*

25 Se añaden el promotor PGK1 (SpeI-PstI) y la secuencia terminadora ADH1 (AflII-NotI) a los fragmentos sintéticos optimizados en codones y se clonan en pRS303 con 2 $\mu$  ori cortado con SpeI y NotI y la construcción de expresión se clona en este vector. La expresión se verifica por qPCR. Secuencias optimizadas en codones para mphF de *E. coli* (SEQ ID NO: 2), adhE de *E. coli* (SEQ ID NO: 4) y ADH2 de *E. histolytica* (SEQ ID NO: 6) son como se indican en el listado de secuencias.

30 Para la expresión del gen mphF de *E. coli*, se ligaron un fragmento de promotor PGK1 de levadura (SpeI-PstI) y un fragmento de terminador ADH1 (AflII-NotI) (ambos de la colección de plásmidos de Nedalco) sobre el fragmento sintético optimizado en codones que codifica mphF de *E. coli* (SEQ ID NO: 2). Se cortó pRS 303 con 2 $\mu$  ori (= pRN347, colección de plásmidos de Royal Nedalco) con SpeI y NotI y la construcción de expresión de mphF se clonó en este vector para producir pRN558 (SEQ ID NO: 29).

35 Para la expresión del gen adhE de *E. coli*, un fragmento sintético optimizado en codones que codifica adhE de *E. coli* (SEQ ID NO: 4) se corta con XbaI y AflII y se liga en pRN558 cortado con XbaI y AflII (que sustituye al gen mphF de *E. coli* en pRN558) para producir pRN595 (SEQ ID NO: 30).

Para la expresión de adh2 de *Entamoeba histolytica*, un fragmento sintético optimizado en codones que codifica adh2 de *E. histolytica* (SEQ ID NO: 6) se corta con XbaI y AflII y se liga en pRN558 cortado con XbaI y AflII (que sustituye el gen mphF de *E. coli* en pRN558) para producir pRN596 (SEQ ID NO: 31).

40 pRN595 se usa para la construcción adicional de pRN957 y pRN977 (véase más adelante). Es evidente que pRN558 y pRN596 pueden usarse de la misma forma, sustituyendo así la expresión de adhE de *E. coli* con mphF de *E. coli* o adh2 de *E. histolytica*, respectivamente.

#### Expresión de gldA de *E. coli*

45 La construcción para la expresión en levadura de gldA de *E. coli* se hizo ligando un fragmento de promotor ACT1 de levadura (cortado con las enzimas de restricción SpeI y PstI), un ORF sintético (SEQ ID NO: 21), que codifica gldA de *E. coli* (cortado con PstI en BssHII) y un fragmento de terminador CYC1 de levadura (cortado con BssHII y BsiWI) juntos en pCRII blunt (Invitrogen) para dar pRNgldA (SEQ ID NO: 28).

#### Expresión en exceso de DAK1

50 Se realiza PCR en ADN genómico de *S. cerevisiae* con cebadores que introducen un sitio XbaI 5' de ATG y un sitio Sall 3' de TAA para producir el fragmento de SEQ ID NO: 22. Un fragmento de ADN que comprende el promotor TPI1 de *S. cerevisiae* se liga en la dirección 5' de ORF de DAK1 y el fragmento de ADN que comprende el fragmento de terminador PGI1 de *S. cerevisiae* se liga en la dirección 3' de ORF de DAK1 para producir pRNDAK (SEQ ID NO: 38).

Expresión de dhaK de *C. freundii*

5 La construcción para la expresión en levadura de *dhaK* de *Citrobacter freundii* se hizo ligando el fragmento de promotor TPI1 de levadura (cortado con las enzimas de restricción *XhoI* y *XbaI*), un ORF sintético (SEQ ID NO: 26), que codifica *dhaK* de *C. freundii* (cortado con *XbaI* y *SalI*) y un fragmento de terminador PG11 de levadura (cortado con *XhoI* y *BsiWI*) juntos en pCRII blunt (Invitrogen) para dar pRNdhaK (SEQ ID NO: 27).

Expresión en exceso de GUP1

10 Se realiza PCR en ADN genómico de *S. cerevisiae* con cebadores que introducen un sitio HindIII 5' de ATG y un sitio BamHI 3' de TAA para producir el fragmento de SEQ ID NO: 23. Un fragmento de ADN que comprende el promotor TDH3 de *S. cerevisiae* se liga en la dirección 5' de ORF de GUP1 y el fragmento de ADN que comprende el fragmento de terminador CYC1 de *S. cerevisiae* se liga en la dirección 3' de ORF de GUP1.

Expresión en exceso de FPS1

15 Se realiza PCR en ADN genómico de *S. cerevisiae* con cebadores que introducen un sitio NsiI 5' de ATG y un sitio BamHI 3' de TAA para producir el fragmento de SEQ ID NO: 24. Un fragmento de ADN que comprende el promotor de ADH1 (medio) de *S. cerevisiae* se liga en la dirección 5' de ORF de FSP1 y el fragmento de ADN que comprende el fragmento de terminador CYC1 de *S. cerevisiae* se liga en la dirección 3' de ORF de FSP1.

Construcción de pRN347 y la cepa de levadura RN1151

Se construye pRN347 clonando el origen de replicación 2 $\mu$  (que se amplificó por PCR a partir de pYES2) en pRS303 (con el gen HIS3 para complementación). RN1041 se transforma con el plásmido pRN347 para producir la cepa RN1151.

20 Cepas que expresan *gldA* de *E. coli* y *dhaK* de *C. freundii* o que expresan en exceso DAK1

Para la construcción de pRN957, la construcción de expresión *gldA* de *E. coli* se corta del plásmido pRN*gldA* con las enzimas de restricción *SpeI* y *BsiWI*. La construcción de expresión de *dhaK* de *C. freundii* se corta del plásmido pRN*dhaK* con las enzimas de restricción *BsiWI* y *XhoI*. Estos fragmentos se ligan en el plásmido pRN595 cortado con las enzimas de restricción *SpeI* y *SalI* dando pRN957 (SEQ ID NO: 37).

25 Para la construcción de pRN977, la construcción de expresión *gldA* de *E. coli* se corta del plásmido pRN*gldA* con las enzimas de restricción *SpeI* y *BsiWI*. La construcción de expresión de DAK1 se corta del plásmido pRN*DAK* con las enzimas de restricción *BsiWI* y *XhoI*. Estos fragmentos se ligan al plásmido pRN595 cortado con las enzimas de restricción *SpeI* y *SalI* dando pRN977 (SEQ ID NO: 39).

30 Los plásmidos pRN957 y pRN977 se usan para transformar RN1041, RN1197, RN1198 y RN1199 para dar cepas de levadura como se indica en la Tabla 4.

Tablas 4A y B: Visión general de cepas construidas

Tabla 4A	ACS1	ACS2	GPD1	GPD2	adhE	DAK1	Cf dhaK	Ec gidA	HIS3
Marcador de cepa	kanMX	kanMX	hphMX	natMX	HIS3	HIS3	HIS3	HIS3	
RN1041	wt	wt	wt	wt	ausente	wt	ausente	ausente	del
RN1151	wt	wt	wt	wt	ausente	wt	ausente	ausente	wt
RN1197	wt	wt	del	wt	ausente	wt	ausente	ausente	wt
RN1198	wt	wt	wt	del	ausente	wt	ausente	ausente	wt
RN1199	wt	wt	del	del	ausente	wt	ausente	ausente	wt
RN1200	up	up	wt	wt	expresión	up	ausente	expresión	
RN1201	up	up	wt	wt	expresión	wt	expresión	expresión	
RN1202	up	up	del	wt	expresión	up	ausente	expresión	
RN1203	up	up	del	wt	expresión	wt	expresión	expresión	
RN1204	up	up	wt	del	expresión	up	ausente	expresión	
RN1205	up	up	wt	del	expresión	wt	expresión	expresión	
RN1206	up	up	del	del	expresión	up	ausente	expresión	
RN1207	up	up	del	del	expresión	wt	expresión	expresión	
Tabla 4B	ACS1/2	gpd1Δ	gpd2Δ	adhE, gidA, DAK1	adhE, gidA, dhaK			HIS3	
RN1041									
RN1151								pRN347	
RN1197		pRN593						pRN347	
RN1198								pRN347	
RN1199		pRN593	pRN594					pRN347	
RN1200	pRN753			pRN957					

RN1201	pRN753				pRN977		
RN1202	pRN753	pRN593			pRN957		
RN1203	pRN753	pRN593			pRN977		
RN1204	pRN753		pRN594		pRN957		
RN1205	pRN753		pRN594		pRN977		
RN1206	pRN753	pRN593	pRN594		pRN957		
RN1207	pRN753	pRN593	pRN594		pRN977		

### 3. Fermentaciones anóxicas en medio de extracto de levadura-peptona estéril con cepas construidas en presencia y ausencia de glicerol y o ácido acético

5 Se obtuvo la prueba del principio de reducción simultánea de ácido acético y oxidación de glicerol usando un medio que contenía 1 % de extracto de levadura y 1 % de peptona. Los experimentos se realizaron en cultivo de quimiostato (1 litro de volumen de trabajo) a  $D = 0,05 \text{ h}^{-1}$  y el pH se mantuvo a 5,5 por adición automática de o bien KOH o bien  $\text{H}_2\text{SO}_4$ . Se añadieron glucosa (50 g/l) y xilosa (50 g/l) como fuente de carbono y de energía al medio de extracto de levadura-peptona. Para estos experimentos que demuestran la prueba del principio, no se incluyó arabinosa. Donde fuera relevante, se añadió ácido acético al medio de extracto de levadura-peptona a 4 g/l y glicerol a 10 g/l. La temperatura se mantuvo a 32 °C.

10 Se preparan precultivos de cepas inoculando un cultivo de reserva de glicerol congelado de la levadura en un medio YP (extracto de levadura al 1 % en p/v y peptona al 1 % en p/v) con adición de cada uno de los azúcares glucosa y xilosa (cada uno al 1 % p/v) a 32 °C y pH 5,5. Después de 24 h de incubación en condiciones óxicas en matraces oscilantes, se usan 50 ml de este cultivo para inocular los cultivos de quimiostato.

15 En el estado estacionario de las fermentaciones (cambios de 5 volúmenes), se tomó una muestra para el análisis del consumo de azúcar (glucosa y xilosa), consumo de ácido acético y metabolito (etanol y glicerol). Se monitorizan las concentraciones de etanol, glicerol y ácido acético por análisis de HPLC. Para determinar el consumo de azúcar, se determinan glucosa y xilosa por análisis de HPAEC (Dionex).

20 La cepa RN1151 no es capaz de alcanzar una situación de estado estacionario en el medio que contiene 4 g/l de ácido acético ni en presencia ni en ausencia de glicerol. Si no se añade ácido acético al medio, el organismo en el estado estacionario consumió toda la glucosa y xilosa (quedando menos de 1 g/l). No se consumió glicerol, sino que en su lugar se produjo.

25 Las cepas RN1200 y RN1201 se prueban similarmente en medios con ácido acético y ya sea con o sin glicerol añadido. Estas cepas rinden de forma claramente diferente de la cepa RN1151. En el medio que contiene glicerol, los azúcares glucosa y xilosa se consumen casi hasta la finalización (quedando menos de 1 g/l). Los niveles de ácido acético disminuyen a 0,5 g/l y las concentraciones de glicerol al final de la fermentación es 3 g/l en los tres casos. Las cantidades de etanol producidas por las cepas RN1200 y RN1201 oscilaron entre 43 y 47 g/l en diversos experimentos. En el medio que no contiene glicerol, pero que contiene 4 g/l de ácido acético, no se obtuvo estado estacionario estable. Las cepas no pueden crecer bajo estas condiciones. A partir de estos resultados, los presentes inventores llegan a la conclusión de que la expresión de los genes *gldA* y *adhE* de *E. coli* en combinación con la regulación por incremento de *DAK1* o expresión de *dhaK* de *C. freundii* tiene un profundo efecto sobre el rendimiento de las cepas. En presencia de glicerol, son capaces de consumir glicerol y ácido acético,

35 Las cepas RN1202 a RN1207 son similares a las cepas RN1200 y RN1201, excepto por el hecho de que los genes *GPD1* y/o *GPD2* han sido deletados. En el medio que contiene 4 g/l de ácido acético, los azúcares glucosa y xilosa son consumidos casi hasta la finalización (quedando menos de 1 g/l) si se añade glicerol al medio como es el caso para las cepas RN1200 y RN1201. Si no se añade glicerol, no se obtiene el estado estacionario.

### 4. Fermentaciones anóxicas con las cepas construidas en ácido acético que comprenden hidrolizados lignocelulósicos en presencia o ausencia de glicerol

40 El hidrolizado de fibra de maíz contiene: glucosa (38 g/l), xilosa (28 g/l), arabinosa (12 g/l) y ácido acético (4 g/l). Ha sido preparado tratando fibras de maíz a 160 °C y a pH 3,0 durante 20 minutos, seguido de hidrólisis enzimática por celulasas y hemicelulasas. Se añadió ácido acético a este hidrolizado produciendo una concentración total de ácido acético en el hidrolizado de 10 g/l. El pH de este hidrolizado enriquecido en ácido acético se restauró a pH = 4,5 por la adición de KOH. Se añadió extracto de levadura a este hidrolizado para alcanzar una concentración final de 5 g/l. En todos los experimentos posteriores, se empleó este hidrolizado enriquecido. El pH durante las fermentaciones se mantuvo a 6,5 por adición automática de cualquiera de KOH o  $\text{H}_2\text{SO}_4$ .

45 Se preparan precultivos de cepas inoculando un cultivo de reserva de glicerol congelado de la levadura en un medio YP (extracto de levadura al 1 % en p/v y peptona al 1 % en p/v) con adición de cada uno de los azúcares glucosa, xilosa y arabinosa (cada uno al 1 % p/v) a 32 °C y pH 5,5. Después de 24 h de incubación en condiciones óxicas en matraces oscilantes, se usan 50 ml de este cultivo para inocular los cultivos de fermentador. Las fermentaciones se realizan en una configuración de fermentación de lotes alimentados. El hidrolizado (ya sea con o sin glicerol añadido a 50 g/l) se bombea en el fermentador. Si no se añadió glicerol, entonces se añadieron 40 ml de agua. Durante las primeras 6 horas, el caudal para el hidrolizado se establece a una velocidad de 5 ml por hora. Durante las siguientes 6 horas, el caudal se establece a 10 ml por hora. Posteriormente, durante otras 43 horas, el caudal se establece a 20 ml por hora. El volumen total al final de la fermentación alcanza 1000 ml. Estas fermentaciones anóxicas de lotes alimentados se realizan a aproximadamente pH = 4,5 con agitación suave a 100 rpm. La temperatura durante las fermentaciones se establece a 32 °C. Para minimizar la infección, los hidrolizados se calientan durante 10 min a 105 °C antes de que se añadan las fermentaciones y el antibiótico kanamicina con una concentración final de 50 µg/ml.

Al final de las fermentaciones después de 55 h, se tomó una muestra para el análisis del consumo de azúcar (glucosa, xilosa y arabinosa), consumo de ácido acético y metabolito (etanol y glicerol). Se monitorizan las concentraciones de etanol, glicerol y ácido acético con el tiempo por análisis de HPLC. Para determinar el consumo de azúcar, se determinan glucosa, xilosa, y arabinosa por análisis de HPAEC (Dionex).

5 La cepa RN1151 (= RN1041 complementada con HIS3) se prueba en el hidrolizado ya sea con o sin glicerol añadido. En ambos casos, la concentración de glucosa al final de la serie de fermentación (55 h) es 35 g/l, mientras que la xilosa y arabinosa siguen a sus concentraciones iniciales de 28 y 12 g/l, respectivamente. Las cantidades de etanol producidas son 2 g/l y el ácido acético está presente a 9,5 g/l. No se detecta consumo de glicerol en el hidrolizado que contiene glicerol. La fermentación de los azúcares se detiene durante el transcurso de la operación de lotes alimentados debido a los niveles crecientes de ácido acético. Inicialmente, no está presente ácido acético en el fermentado, pero mientras que se bombea el hidrolizado que contiene niveles tóxicos de ácido acético, la concentración alcanza rápidamente niveles tóxicos.

15 Las cepas RN1200 y RN1201 se prueban similarmente en hidrolizado ya sea con o sin glicerol añadido. Estas cepas rinden de forma claramente diferente a la cepa RN1151. En el hidrolizado que contiene glicerol, los azúcares glucosa, xilosa y arabinosa se consumen hasta la finalización. Los niveles de ácido acético disminuyen a 2 g/l y las concentraciones de glicerol al final de la fermentación son 29,5 g/l en los tres casos. Las cantidades de etanol producidas por las cepas RN1200 y RN1201 son 51,7 y 52,2 g/l, respectivamente. En el hidrolizado que no contuvo glicerol, se consume considerablemente menos azúcar. Los niveles de xilosa y arabinosa son invariables a 28 y 12 g/l, respectivamente. La glucosa se consume, pero a un grado limitado solo. Al final de la fermentación, la concentración restante es 32 g/l en los tres casos, alcanzando el etanol una concentración de 3 g/l. La concentración de ácido acético disminuye a 9,1 g/l al final de la fermentación, mientras que se produce algo de glicerol (menos de 0,5 g/l). A partir de estos resultados, los presentes inventores llegan a la conclusión de que la expresión de los genes *gldA* y *adhE* de *E. coli* en combinación con la regulación por incremento de DAK1 o la expresión de *dhaK* de *C. freundii* tiene un profundo efecto sobre el rendimiento de las cepas. En presencia de glicerol, son capaces de consumir glicerol y ácido acético, y producir etanol adicional (en comparación con la cepa RN1151). En ausencia de glicerol, las cepas consumen algo de ácido acético. Pero durante la fermentación, el nivel de ácido acético aumenta a niveles tóxicos.

30 Las cepas RN1202 a RN1207 son similares a las cepas RN1200 y RN1201, excepto por el hecho de que los genes *GPD1* y/o *GPD2* han sido deletados. En el hidrolizado que contiene glicerol, los azúcares glucosa, xilosa y arabinosa se consumen hasta la finalización como era el caso para la cepa RN1200. Los niveles de ácido acético disminuyen similarmente a aproximadamente 2 g/l y las concentraciones de glicerol al final de la fermentación es 28 g/l en estos tres casos. Las cantidades de etanol producidas por las cepas RN1202, RN1203, RN1204, RN1205, RN1206 y RN1207 son 51,6, 52,9, 52,1, 52,5, 53,1 y 52,3 g/l, respectivamente. En el hidrolizado que no contiene glicerol, se consume considerablemente menos azúcar. Los niveles de xilosa y arabinosa son invariables a 28 y 12 g/l, respectivamente. La glucosa se consume pero a un grado limitado solo. Al final de la fermentación, la concentración restante en el hidrolizado no de glicerol para glucosa es 31 g/l en los tres casos, alcanzando el etanol una concentración de 3 g/l. La concentración de ácido acético disminuye a 9,1 g/l al final de la fermentación, mientras que se produce algo de glicerol (inferior a 0,5 g/l). A partir de estos resultados los presentes inventores llegan a la conclusión de que la delección de los genes *GPD1* y/o *GPD2* junto con las otras modificaciones en RN1202, RN1203, RN1204, RN1205, RN1206 y RN1207 produce cepas que pueden realizar las reacciones deseadas.

#### Materiales y métodos para los Ejemplos 5-8

##### Técnicas generales de biología molecular

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

##### Medios

50 Los medios usados en los experimentos fueron o bien medio YEP (10 g/l de extracto de levadura, 20 g/l de peptona) o bien medio YNB sólido (6,7 g/l de base de nitrógeno de levadura, 15 g/l de agar), complementados con azúcares como se indica en los ejemplos. Para el medio YEP sólido, se añadió 15 g/l de agar al medio líquido antes de la esterilización.

En los experimentos de AFM, se usó medio mineral. La composición del medio mineral se ha descrito por Verduyn et al. (*Yeast* (1992), Volumen 8, 501-517) y se complementó con 2,325 g/l de urea y azúcares como se indica en los ejemplos.

##### Transformación de células de levadura

La transformación de levadura se hizo según el método descrito por Schiestl y Gietz (*Current Genetics* (1989), Volumen 16, 339-346).

PCR de colonias

Se extrajo ADN genómico de colonias de levadura individuales para PCR según el método descrito por Looke et al. (BioTechniques (2011), Volumen 50, 325-328).

Procedimiento de AFM

- 5 El monitor de fermentación de alcohol (AFM; Halotec, Veenendaal, Los Países Bajos) es un birreactor en paralelo de laboratorio robusto y fácil de usar que permite comparaciones precisas de velocidades de conversión de carbono y rendimientos para seis fermentaciones anaerobias simultáneas.

10 El cultivo de partida del experimento de AFM contuvo 50 mg de levadura (peso seco). Para determinar esto, se hizo una curva de calibración de la cepa RN1041 de biomasa frente a DO700. Esta curva de calibración se usó en el experimento para determinar el volumen de cultivo celular necesario para 50 mg de levadura (peso seco).

Antes del inicio del experimento de AFM, se cultivaron precultivos como se indica en los ejemplos. Para cada cepa se midió la DO<sub>700</sub> y se inocularon 50 mg de levadura (peso seco) en 400 ml de medio mineral (Verduyn et al. (Yeast (1992), Volumen 8, 501-517), complementado con 2,325 g/l de urea y azúcares como se indica en los ejemplos.

Ensayo de actividad de glicerol deshidrogenasa

- 15 El método para la determinación del ensayo de actividad de glicerol deshidrogenasa fue adoptado de Lin y Magasanik (1960) J Biol Chem, 235:1820-1823.

Tabla 5: Condiciones de ensayo

Tampón carbonato/bicarbonato 1,0 M a pH 10	800 µl
Sulfato de amonio 1,0 M	33 µl
NAD+ 0,1 M	33 µl
Extracto libre de células	5 µl

- 20 Se preparó extracto libre de células recogiendo células por centrifugación. Las células se recogieron en la fase exponencial. El sedimento de células se lavó una vez con tampón carbonato/bicarbonato 1 M (pH 10) y se preparó un extracto libre de células en el mismo mediante la adición de perlas de vidrio y agitando con vórtex a velocidad máxima durante intervalos de 1 minuto hasta que las células se rompieron. Lo último se comprobó microscópicamente.

Experimentos en matraz oscilante

- 25 Se realizaron experimentos anaerobios en matraz oscilante como se indica en los ejemplos. Los experimentos típicos usan matraces Erlenmeyer de 100 ml con 25 ml de medio. Con el fin de garantizar las condiciones anaerobias, el matraz se cerró con un cierre hidráulico. Para cada momento de tiempo, se inoculó un matraz oscilante separado, omitiendo así la aireación durante el muestreo.

Cepas

- 30 La cepa parental usada en los experimentos descritos en los Ejemplos 5 a 8 es RN1041.

RN1041 se ha descrito en el documento WO 2012067510. Esta cepa tiene el siguiente genotipo:

*MAT a, ura3-52, leu2-112, his3::loxP, gre3::loxP, loxP-pTPI1::TAL1, loxP-pTPI1::RKI1, loxP-pTPI1-TKL1, loxP-pTPI1-RPE1, delta::pADH1-XKS1-tCYC1-LEU2, delta::URA3-pTPI1-xylA-tCYC1*

*MATa* = tipo de apareamiento a

- 35 Mutaciones *ura3-52*, *leu2-112*, *HIS3::loxP* en los genes *URA3*, *LEU2* y *HIS3*, respectivamente. La mutación *ura3-52* está complementada por el gen *URA3* en la construcción de expresión en exceso de *xylA*; la mutación *leu2-112* está complementada por el gen *LEU2* en la construcción de expresión en exceso de *XKS1*. La delección del gen *HIS3* produce una auxotrofia a la histidina. Por este motivo, RN1041 necesita histidina en el medio para el crecimiento.

- 40 *gre3::loxP* es una delección del gen *GRE3*, que codifica aldosa reductasa. El sitio *loxP* se deja atrás en el genoma *loxP* después de la eliminación del marcador.

loxP-pTPI1 designa la expresión en exceso de genes de, descritos en los experimentos en el presente documento, la vía de pentosa fosfato no oxidativa por sustitución del promotor nativo por el promotor del gen *TPI1*. El sitio loxP en la dirección 5' del promotor *TPI1* constitutivo fuerte sigue en el genoma después de la eliminación del marcador (Kuyper et al, FEMS Yeast Research 5 (2005) 925-934).

- 5 delta:: significa integración cromosómica de la construcción después de la recombinación en las repeticiones terminales largas del retrotransposón Ty1.

Ejemplo 5

Construcción de cepas

Se construyeron las siguientes cepas:

10 Tabla 6: Cepas construidas

RN1041	Cepa parental (véase anteriormente)
RN1067	RN1041 <i>gpd1::hphMX</i>
RN1068	RN1041 <i>gpd2::natMX</i>
RN1069	RN1041 <i>gpd1::hphMX gpd2::natMX</i>
RN1186	RN1041 + pRN977
RN1187	RN1067 + pRN977
RN1188	RN1068 + pRN977
RN1189	RN1069 + pRN977

La deleción del gen *GPD1* (*gpd1*) y/o el gen *GPD2* (*gpd2*) se produjo como se describe en el Ejemplo 2.

15 Las cepas RN1041, RN1067, RN1068 y RN1069 se transformaron con el plásmido pRN977. Este plásmido contiene las siguientes características: el gen *HIS3* para la selección de transformantes, el origen de replicación 2μ, el marcador de resistencia a ampicilina para la selección en *E. coli*, el gen *adhE* de *E. coli* bajo el control del promotor *PGK1* y el terminador *ADH1*, el gen *DAK1* de *S. cerevisiae* bajo el control del promotor *TPI1* y el terminador *PGI1* y el gen *gldA* de *E. coli*, bajo el control del promotor *ACT1* y el terminador *CYC1*. Todos los promotores y terminadores son de *S. cerevisiae*. La secuencia de plásmido pRN977 se explica en SEQ ID NO: 39.

20 Después de la transformación de cepas RN1041, RN1067, RN1068 y RN1069, cepas clínicas de colonias individuales se sometieron a análisis por PCR de colonias, con el fin de comprobar la presencia de pRN977 de plásmido. Se seleccionó una colonia representativa de cada transformación para experimentación adicional. Estas cepas seleccionadas se designan RN1186, RN1187, RN1188 y RN1189.

Similarmente, se generaron transformantes con las siguientes especificaciones:

Tabla 7: transformantes

RN1190	RN1041 + pRN957
RN1191	RN1067 + pRN957
RN1192	RN1068 + pRN957
RN1193	RN1069 + pRN957

25 El plásmido pRN957 es similar a pRN977; sin embargo, el gen *DAK1* de *S. cerevisiae* se ha sustituido por el gen *dhaK* de *Citrobacter freundii*. La secuencia de este plásmido, pRN957, se explica en SEQ ID NO: 37.

30 Como cepa de control, se transformó la cepa RN1041 con el plásmido pRN595 (RN1041+pRN595). Este plásmido, pRN595, es similar a pRN977; sin embargo, carece de los genes *gldA* y *DAK1*. La secuencia de plásmido pRN595 se explica en SEQ ID NO: 30.

Ejemplo 6Experimentos en matraz oscilante

5 Se probó el rendimiento de las cepas construidas en un experimento anaerobio en matraz oscilante. Para este fin, las células se cultivaron previamente en medio mineral (Verduyn) complementado con glucosa como fuente de carbono. Los cultivos se incubaron durante la noche en un agitador rotatorio a 280 rpm y 30 °C.

Se tomó una alícuota de las células de los cultivos durante la noche para inoculación de los cultivos anaerobios. La cantidad de células fue tal que el cultivo anaerobio tuvo una densidad óptica inicial a 600 nm de aproximadamente 0,1.

10 Composición de carbono del medio mineral: 2,5 % de glucosa, 2,5 % de xilosa, 1 % de glicerol y 2 g/l de HAc. El pH se ajustó a pH 4,5. Los matraces oscilantes se cerraron con un cierre hidráulico con el fin de garantizar condiciones anaerobias. Para cada momento de tiempo, se inoculó un matraz separado.

Los resultados del aumento neto o disminución de glicerol, después de 94 horas de fermentación, y el consumo de HAc, se indican en la tabla a continuación.

Tabla 8: Consumo de glicerol y HAc, y valores de producción de etanol por cepa

Cepa	Aumento neto (+) o disminución (-) de glicerol (en gramos/litro)	Consumo de HAc (en gramos/litro)	Título de etanol (en gramos por litro)
RN1041	+ 1,47	0,24	23,28
RN1186	- 1,20	0,99	25,32
RN1187	- 1,52	0,99	23,76
RN1188	- 0,80	0,97	25,01
RN1189	- 0,86	0,89	24,85
RN1190	- 0,47	0,71	24,38
RN1191	- 0,80	0,93	24,77
RN1192	+ 0,93	0,29	23,60
RN1193	- 0,84	0,92	24,93

15 La cepa indicada en la Tabla 8 como RN1041 se transformó con el plásmido pRS323, un vector de clonación estándar que contiene el gen *HIS3* y un origen de replicación 2 $\mu$ , complementando así la autotofía a la histidina.

Los resultados muestran:

- 20
- RN1041 produce glicerol, que tiene sentido ya que tanto los genes *GPD1* como *GPD2* son activos y *gldA* y *DAK1* no se expresan en exceso. Como *adhE* no se expresa en esta cepa, el consumo de HAc es bajo.
  - Las cepas RN1186 a RN1189 muestran consumo de glicerol y HAc, produciendo un aumento del título de etanol en comparación con RN1041.
  - Los experimentos con los transformantes RN1190, RN1191 y RN1193 muestran los mismos resultados, es decir, consumo de glicerol y acetato, sin embargo a un grado ligeramente menor. También aquí, el título de etanol es más alto en comparación con RN1041. El resultado de la cepa RN1192 es un artefacto, ya que la
- 25

La expresión en exceso de tanto una dihidroxiacetona cinasa homóloga como heteróloga, en combinación con la expresión en exceso de *gldA* y *adhE*, produce un consumo simultáneo de acetato y glicerol en condiciones anaerobias.

30

Ejemplo 7Experimentos de AFM

5 Se repitió el experimento descrito en el Ejemplo 6 en una configuración ligeramente diferente, es decir, el AFM (monitor de fermentación alcohólica), que permite la determinación de dióxido de carbono en línea, durante el experimento.

Las cepas probadas fueron RN1041, RN1041+pRN595, RN1186, RN1187, RN1188 y RN1189. La cepa RN1041 se transformó con el plásmido pRS323, un vector de clonación estándar que contiene el gen *HIS3* y un origen de replicación  $2\mu$ , complementando así la autotrofia a la histidina.

10 Las cepas se precultivaron durante la noche en medio mineral con 2 % de glucosa como fuente de carbono, en un agitador rotatorio a 280 rpm y 30 °C.

Las células se recogieron y se empezó un experimento de AFM como se ha descrito anteriormente.

Se tomaron muestras a intervalos regulares y se determinaron azúcares, etanol, glicerol y HAC por HPLC.

Los resultados se muestran en la tabla a continuación.

Tabla 9: Consumo de glicerol y HAC, y valores de producción de etanol por cepa en el tiempo = 112 horas.

Cepa	Aumento neto (-) o disminución (+) de glicerol (gramos por litro)	Consumo de HAC (gramos por litro)	Producción de etanol (gramos por litro)
RN1041	+ 1,4	0,1	24,0
RN1041+pRN595	+ 1,1	0,4	24,1
RN1186	- 0,3	0,7	25,5
RN1187	- 1,2	1,0	25,5
RN1188	- 0,3	0,7	25,2
RN1189	- 1,1	1,0	25,6

15 La evolución de los niveles de glicerol y HAC en el tiempo se muestran en las Figuras 1 y 2.

Las cepas RN1041 y RN1041+pRN595 están mostrando una producción neta de glicerol. Las cepas RN1186 y RN1188 están mostrando inicialmente producción de glicerol; sin embargo, después de aproximadamente 24 a 32 horas, comenzó el consumo de glicerol y continuó hasta que al final se observó consumo neto de glicerol.

20 Las cepas RN1187 y RN1189 no presentan la producción de glicerol inicial, como se observa con RN1186 y RN1188. Después de 24 horas, comienza el consumo de glicerol. El consumo de glicerol es significativamente más alto en estas cepas en comparación con RN1186 y RN1188. Estos resultados indican que la delección del gen *GPD1* produce mayor consumo de glicerol que la delección del gen *GPD2*.

25 La cepa RN1041+pRN595 está mostrando un mayor consumo de HAC que la cepa de referencia RN1041. RN1186 y RN1188 están presentando un mayor consumo de HAC que RN1041+pRN595. Este resultado indicó que el consumo de glicerol potenció el consumo de HAC. Este efecto es incluso más fuerte en las cepas RN1187 y RN1189.

Ejemplo 8Ensayo de actividad de glicerol deshidrogenasa

30 Se prepararon extractos libres de células (CFE) de la cepa RN1041 y RN1190 como se ha descrito anteriormente. Se realizó el ensayo de actividad de glicerol deshidrogenasa, adoptado del protocolo de Lin y Magasanik (1960) J Biol Chem, 235:1820-1823. Los resultados se muestran en la tabla a continuación.

Tabla 10: Ensayo de actividad de glicerol deshidrogenasa

Muestra	Cofactor	Aumento en A340/min
RN1041 5 µl de CFE	NAD+	0,00
RN1190 5 µl de CFE	NAD+	0,02
RN1041 20 µl de CFE	NAD+	0,00
RN1190 20 µl de CFE	NAD+	0,09
RN1041 5 µl de CFE	NADP+	0,00
RN1190 5 µl de CFE	NADP+	0,00
RN1041 20 µl de CFE	NADP+	0,00
RN1190 20 µl de CFE	NADP+	0,00

La cepa indicada en la Tabla 10 como RN1041 se transformó con el plásmido pRS323, un vector de clonación estándar que contenía el gen *HIS3* y un origen de replicación 2µ, complementando así la auxotrofia a la histidina.

- 5 Estos resultados indican que: a) *gldA* de *E. coli*, expresado en RN1190, es dependiente de NAD<sup>+</sup>, y b) que el aumento en la cantidad de CFE produjo un aumento proporcional de la velocidad de conversión de NAD<sup>+</sup>, y de ahí de glicerol en dihidroxiacetona.

LISTADO DE SECUENCIAS

<110> DSM IP Assets B.V.

10

<120> Cepas microbianas manipuladas para producir etanol a partir de hidratos de carbono, ácido acético y glicerol

<130> P6038681PCT

15

<150> EP 11191333

<151> 30-11-2011

<150>US 61/564.932

20

<151> 30-11-2011

<160> 51

<170> PatentIn versión 3.3

25

<210> 1

<211> 316

<212> PRT

<213> Escherichia coli

ES 2 648 865 T3

<400> 1

Met Ser Lys Arg Lys Val Ala Ile Ile Gly Ser Gly Asn Ile Gly Thr  
 1 5 10 15

Asp Leu Met Ile Lys Ile Leu Arg His Gly Gln His Leu Glu Met Ala  
 20 25 30

Val Met Val Gly Ile Asp Pro Gln Ser Asp Gly Leu Ala Arg Ala Arg  
 35 40 45

Arg Met Gly Val Ala Thr Thr His Glu Gly Val Ile Gly Leu Met Asn  
 50 55 60

Met Pro Glu Phe Ala Asp Ile Asp Ile Val Phe Asp Ala Thr Ser Ala  
 65 70 75 80

Gly Ala His Val Lys Asn Asp Ala Ala Leu Arg Glu Ala Lys Pro Asp  
 85 90 95

Ile Arg Leu Ile Asp Leu Thr Pro Ala Ala Ile Gly Pro Tyr Cys Val  
 100 105 110

Pro Val Val Asn Leu Glu Ala Asn Val Asp Gln Leu Asn Val Asn Met  
 115 120 125

Val Thr Cys Gly Gly Gln Ala Thr Ile Pro Met Val Ala Ala Val Ser  
 130 135 140

Arg Val Ala Arg Val His Tyr Ala Glu Ile Ile Ala Ser Ile Ala Ser



ES 2 648 865 T3

```

ctgcagtcta gatgtctaag agaaaggttg ctatcatcgg ttctggtaac atcgg tactg      60
acttgatgat caagatccta agacacggtc aacacttggga aatggctggt atggttggta      120
tcgaccaca atctgacggt ttggctagag ctagaagaat ggggtgtgct accaccacg      180
aagggtttat cggtttgatg aacatgccag aattcgctga catcgacatc gttttcgacg      240
ctacctctgc tgggtctcac gttaagaacg acgctgcttt gagagaagct aagccagaca      300
tcagattgat cgacttgacc ccagctgcta tcggtcata ctgtgttcca gttgttaact      360
tggaagctaa cgttgacca ttaaacgta acatggttac ctgtggtggt caagctacca      420
tccaatggt tgctgctggt tcaagagttg ctagagttca ctacgctgaa atcatcgctt      480
ctatcgcttc taagtctgct ggtccaggta ccagagctaa catcgacgaa ttcaccgaaa      540
ccacctctag ggctatcgaa gttgttggtg gtgctgctaa gggtaaggct atcatcgttt      600
tgaaccocagc tgaaccacca ttgatgatga gagacaccgt ttacgttttg tctgacgaag      660
catctcaaga cgacatcgaa gcttcaatca acgaaatggc tgaagctggt caagcatacg      720
ttccaggtta cagattgaag caaagagttc aattcgaagt tatcccaca gacaagccag      780
ttaacttgcc aggtgttggt caattctctg gtttgaagac cgctgtttgg ttggaagttg      840
aagggtgctgc tcaactactg ccagcttacg ctggtaactt ggacattatg acctcttctg      900
ctttggctac cgctgaaaag atggctcaat ctttggctag aaaggctggt gaagctgctt      960
aagcgcgc                                          968

```

<210> 3

5 <211> 891

<212> PRT

<213> Escherichia coli

<400> 3

10

ES 2 648 865 T3

Met Ala Val Thr Asn Val Ala Glu Leu Asn Ala Leu Val Glu Arg Val  
 1 5 10 15

Lys Lys Ala Gln Arg Glu Tyr Ala Ser Phe Thr Gln Glu Gln Val Asp  
 20 25 30

Lys Ile Phe Arg Ala Ala Ala Leu Ala Ala Ala Asp Ala Arg Ile Pro  
 35 40 45

Leu Ala Lys Met Ala Val Ala Glu Ser Gly Met Gly Ile Val Glu Asp  
 50 55 60

Lys Val Ile Lys Asn His Phe Ala Ser Glu Tyr Ile Tyr Asn Ala Tyr  
 65 70 75 80

Lys Asp Glu Lys Thr Cys Gly Val Leu Ser Glu Asp Asp Thr Phe Gly  
 85 90 95

Thr Ile Thr Ile Ala Glu Pro Ile Gly Ile Ile Cys Gly Ile Val Pro  
 100 105 110

Thr Thr Asn Pro Thr Ser Thr Ala Ile Phe Lys Ser Leu Ile Ser Leu  
 115 120 125

ES 2 648 865 T3

Lys Thr Arg Asn Ala Ile Ile Phe Ser Pro His Pro Arg Ala Lys Asp  
130 135 140

Ala Thr Asn Lys Ala Ala Asp Ile Val Leu Gln Ala Ala Ile Ala Ala  
145 150 155 160

Gly Ala Pro Lys Asp Leu Ile Gly Trp Ile Asp Gln Pro Ser Val Glu  
165 170 175

Leu Ser Asn Ala Leu Met His His Pro Asp Ile Asn Leu Ile Leu Ala  
180 185 190

Thr Gly Gly Pro Gly Met Val Lys Ala Ala Tyr Ser Ser Gly Lys Pro  
195 200 205

Ala Ile Gly Val Gly Ala Gly Asn Thr Pro Val Val Ile Asp Glu Thr  
210 215 220

Ala Asp Ile Lys Arg Ala Val Ala Ser Val Leu Met Ser Lys Thr Phe  
225 230 235 240

Asp Asn Gly Val Ile Cys Ala Ser Glu Gln Ser Val Val Val Val Asp  
245 250 255

Ser Val Tyr Asp Ala Val Arg Glu Arg Phe Ala Thr His Gly Gly Tyr  
260 265 270

Leu Leu Gln Gly Lys Glu Leu Lys Ala Val Gln Asp Val Ile Leu Lys  
275 280 285

Asn Gly Ala Leu Asn Ala Ala Ile Val Gly Gln Pro Ala Tyr Lys Ile  
290 295 300

Ala Glu Leu Ala Gly Phe Ser Val Pro Glu Asn Thr Lys Ile Leu Ile  
305 310 315 320

Gly Glu Val Thr Val Val Asp Glu Ser Glu Pro Phe Ala His Glu Lys  
325 330 335

Leu Ser Pro Thr Leu Ala Met Tyr Arg Ala Lys Asp Phe Glu Asp Ala  
340 345 350

Val Glu Lys Ala Glu Lys Leu Val Ala Met Gly Gly Ile Gly His Thr  
355 360 365

Ser Cys Leu Tyr Thr Asp Gln Asp Asn Gln Pro Ala Arg Val Ser Tyr  
370 375 380

ES 2 648 865 T3

Phe Gly Gln Lys Met Lys Thr Ala Arg Ile Leu Ile Asn Thr Pro Ala  
 385 390 395 400

Ser Gln Gly Gly Ile Gly Asp Leu Tyr Asn Phe Lys Leu Ala Pro Ser  
 405 410 415

Leu Thr Leu Gly Cys Gly Ser Trp Gly Gly Asn Ser Ile Ser Glu Asn  
 420 425 430

Val Gly Pro Lys His Leu Ile Asn Lys Lys Thr Val Ala Lys Arg Ala  
 435 440 445

Glu Asn Met Leu Trp His Lys Leu Pro Lys Ser Ile Tyr Phe Arg Arg  
 450 455 460

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

Arg Ala Leu Ile Val Thr Asp Arg Phe Leu Phe Asn Asn Gly Tyr Ala  
 485 490 495

Asp Gln Ile Thr Ser Val Leu Lys Ala Ala Gly Val Glu Thr Glu Val  
 500 505 510

Phe Phe Glu Val Glu Ala Asp Pro Thr Leu Ser Ile Val Arg Lys Gly  
 515 520 525

Ala Glu Leu Ala Asn Ser Phe Lys Pro Asp Val Ile Ile Ala Leu Gly  
 530 535 540

Gly Gly Ser Pro Met Asp Ala Ala Lys Ile Met Trp Val Met Tyr Glu  
 545 550 555 560

His Pro Glu Thr His Phe Glu Glu Leu Ala Leu Arg Phe Met Asp Ile  
 565 570 575

Arg Lys Arg Ile Tyr Lys Phe Pro Lys Met Gly Val Lys Ala Lys Met  
 580 585 590

Ile Ala Val Thr Thr Thr Ser Gly Thr Gly Ser Glu Val Thr Pro Phe  
 595 600 605

Ala Val Val Thr Asp Asp Ala Thr Gly Gln Lys Tyr Pro Leu Ala Asp  
 610 615 620

Tyr Ala Leu Thr Pro Asp Met Ala Ile Val Asp Ala Asn Leu Val Met  
 625 630 635 640

ES 2 648 865 T3

Asp Met Pro Lys Ser Leu Cys Ala Phe Gly Gly Leu Asp Ala Val Thr  
645 650 655

His Ala Met Glu Ala Tyr Val Ser Val Leu Ala Ser Glu Phe Ser Asp  
660 665 670

Gly Gln Ala Leu Gln Ala Leu Lys Leu Leu Lys Glu Tyr Leu Pro Ala  
675 680 685

Ser Tyr His Glu Gly Ser Lys Asn Pro Val Ala Arg Glu Arg Val His  
690 695 700

Ser Ala Ala Thr Ile Ala Gly Ile Ala Phe Ala Asn Ala Phe Leu Gly  
705 710 715 720

Val Cys His Ser Met Ala His Lys Leu Gly Ser Gln Phe His Ile Pro  
725 730 735

His Gly Leu Ala Asn Ala Leu Leu Ile Cys Asn Val Ile Arg Tyr Asn  
740 745 750

Ala Asn Asp Asn Pro Thr Lys Gln Thr Ala Phe Ser Gln Tyr Asp Arg  
755 760 765

Pro Gln Ala Arg Arg Arg Tyr Ala Glu Ile Ala Asp His Leu Gly Leu  
770 775 780

Ser Ala Pro Gly Asp Arg Thr Ala Ala Lys Ile Glu Lys Leu Leu Ala  
785 790 795 800

Trp Leu Glu Thr Leu Lys Ala Glu Leu Gly Ile Pro Lys Ser Ile Arg  
805 810 815

Glu Ala Gly Val Gln Glu Ala Asp Phe Leu Ala Asn Val Asp Lys Leu  
820 825 830

Ser Glu Asp Ala Phe Asp Asp Gln Cys Thr Gly Ala Asn Pro Arg Tyr  
835 840 845

Pro Leu Ile Ser Glu Leu Lys Gln Ile Leu Leu Asp Thr Tyr Tyr Gly  
850 855 860

Arg Asp Tyr Val Glu Gly Glu Thr Ala Ala Lys Lys Glu Ala Ala Pro  
865 870 875 880

Ala Lys Ala Glu Lys Lys Ala Lys Lys Ser Ala

885

890

<210> 4

5 <211> 2690

<212> ADN

<213> Artificial

ES 2 648 865 T3

<220>

<223> adhE sintético optimizado en codones

<400> 4

5

```

tctagaaaat ggctgttacc aacgttgctg aattgaacgc tttggttgaa agggttaaga      60
aggctcaaag agaatacgtt tctttcaccg aagaacaagt tgacaagatc ttcagagctg     120
ctgctttggc tgctgctgac gctagaatcc cattggctaa gatggctggt gctgaatctg     180
gtatgggtat cgttgaagac aaggttatca agaaccactt cgcttctgaa tacatctaca     240
acgcttacia ggacgaaaag acctgtgggtg ttttgtcaga agacgacacc ttcggtacca     300
tcaccatcgc tgaaccaatc ggtatcatct gtggatcgtg tccaaccacc aacccaacct     360
ctaccgctat cttcaagtct ttgatctctt tgaagaccag aaacgctatc atcttctctc     420
cacacccaag agctaaagac gctaccaaca aggctgctga catcgttttg caagctgcta     480
tcgctgctgg tgctccaaag gacttgatcg gttggatcga ccaaccatct gttgaattgt     540
ctaagcgttt gatgcaccac ccagacatca acttgatctt ggctaccggt ggtccaggta     600
tggttaaggc tgcttactct tctggtaagc cagctatcgg tgttggtgct ggtaacaccc     660
cagttggtat cgacgaaacc gctgacatca agagagctgt tgcttctggt ttgatgtcta     720
agaccttcca caacgggtgt atctgtgctt ctgaacaatc tgttgttgtt gttgactctg     780
tttacgacgc tgtagagaa agattcgcta cccacgggtg ttacttgttg caaggtaagg     840
aattgaaggc tgttcaagac gttatcttga agaacgggtgc tttgaacgct gctatcgttg     900
gtcaaccagc ttacaagatc gctgaattag ctggtttctc tgttccagaa aacaccaaga     960
tcttgatcgg tgaagttacc gttgttgacg aatctgaacc attcgtcac  gaaaagtgtg    1020
ctccaacctt ggctatgtac agagctaagg acttcgaaga cgctgttgaa aaagctgaaa    1080
agtgggttgc tatgggtggt attggtcaca cctcttgttt gtacaccgac caagacaacc    1140
aaccagctag agtttcttac ttcggcaca  agatgaagac cgctagaatc ttgatcaaca    1200
cccagcttc tcaaggtggt atcggtgact tgtacaactt caagttggct ccatctttga    1260
ccttgggttg tggttcttgg ggtggttaact ctatctctga aaacgttggg ccaaagcaact    1320
tgatcaacaa gaagaccgtt gctaagagag ctgaaaacat gttgtggcac aagttgccaa    1380
aatctatcta cttcagaaga gtttctttgc caatcgcttt ggacgaagtt atcaccgacg    1440
gtcacaagag agctttgatc gttaccgaca gattcttgtt caacaacggt tacgctgacc    1500
aaatcacctc tgttttgaag gctgctgggt ttgaaaccga agttttcttc gaagttgaag    1560

```

ES 2 648 865 T3

ctgacccaac cttgtctatc gttagaaagg gtgctgaatt ggctaactct ttcaagccag 1620  
 acgttatcat cgctttgggt ggtggttctc caatggacgc tgctaagatc atgtgggta 1680  
 tgtacgaaca cccagaaacc cacttogaag aattggcttt gagattcatg gacatcagaa 1740  
 agagaatcta caagttccca aagatgggtg ttaaggctaa gatgatcgct gttaccacca 1800  
 cctctggtag cggttctgaa gttaccccat tcgctgttgt taccgacgac gctaccggtc 1860  
 aaaagtaccc attggctgac tacgcttga cccagacat ggctatcggt gacgctaact 1920  
 tggttatgga catgccaaag tctttgtgtg ctttcgggtg tttggacgct gttaccacg 1980  
 ctatggaagc ttacgtttct gttttggctt ctgaattctc tgacggtaa gctttgcaag 2040  
 ctttgaagtt gttgaaggaa tacttgccag cttctacca cgaaggttct aagaaccag 2100  
 ttgctagaga aagagttcac tctgctgcta ccatcgctgg tatcgcttct gtaacgctt 2160  
 tcttgggtgt ttgtcactct atggctcaca agttgggttc tcaattccac atcccacag 2220  
 gtttgctaa cgctttgttg atctgtaacg ttatcagata caacgtaac gacaacccaa 2280  
 ccaagcaaac cgctttctct caatacgaca gaccacaagc tagaagaaga tacgctgaaa 2340  
 tcgctgacca cttgggtttg tctgctccag gtgacagaac cgctgcaaag atcgaaaagt 2400  
 tgttggcttg gttgaaaacc ttgaaggctg aattgggtat cccaaagtct atcagagaag 2460  
 ctggtgttca agaagctgac ttcttgcta acgttgaca gttgtctgaa gacgctttcg 2520  
 acgaccaatg taccggtgct aaccaagat acccattgat ctctgaattg aagcaaatct 2580  
 tgttgacac ctactacggt agagactacg ttgaaggatg aaccgctgct aagaaggaag 2640  
 ctgctccagc taaggetgaa aagaaggcta agaagtctgc ttagcttaag 2690

<210> 5

<211> 870

5 <212> PRT

<213> Entamoeba histolytica

<400> 5

Met Ser Thr Gln Gln Thr Met Thr Val Asp Glu His Ile Asn Gln Leu  
 1 5 10 15  
 Val Arg Lys Ala Gln Val Ala Leu Lys Glu Tyr Leu Lys Pro Glu Tyr  
 20 25 30  
 Thr Gln Glu Lys Ile Asp Tyr Ile Val Lys Lys Ala Ser Val Ala Ala  
 35 40 45  
 Leu Asp Gln His Cys Ala Leu Ala Ala Ala Ala Val Glu Glu Thr Gly  
 50 55 60

10

ES 2 648 865 T3

Arg Gly Ile Phe Glu Asp Lys Ala Thr Lys Asn Ile Phe Ala Cys Glu  
 65 70 75 80  
 His Val Thr His Glu Met Arg His Ala Lys Thr Val Gly Ile Ile Asn  
 85 90 95  
 Val Asp Pro Leu Tyr Gly Ile Thr Glu Ile Ala Glu Pro Val Gly Val  
 100 105 110  
 Val Cys Gly Val Thr Pro Val Thr Asn Pro Thr Ser Thr Ala Ile Phe  
 115 120 125  
 Lys Ser Leu Ile Ser Ile Lys Thr Arg Asn Pro Ile Val Phe Ser Phe  
 130 135 140  
 His Pro Ser Ala Leu Lys Cys Ser Ile Met Ala Ala Lys Ile Val Arg  
 145 150 155 160  
 Asp Ala Ala Ile Ala Ala Gly Ala Pro Glu Asn Cys Ile Gln Trp Ile  
 165 170 175  
 Glu Phe Gly Gly Ile Glu Ala Ser Asn Lys Leu Met Asn His Pro Gly  
 180 185 190  
 Val Ala Thr Ile Leu Ala Thr Gly Gly Asn Ala Met Val Lys Ala Ala  
 195 200 205  
 Tyr Ser Ser Gly Lys Pro Ala Leu Gly Val Gly Ala Gly Asn Val Pro  
 210 215 220  
 Thr Tyr Ile Glu Lys Thr Cys Asn Ile Lys Gln Ala Ala Asn Asp Val  
 225 230 235 240  
 Val Met Ser Lys Ser Phe Asp Asn Gly Met Ile Cys Ala Ser Glu Gln  
 245 250 255  
 Ala Ala Ile Ile Asp Lys Glu Ile Tyr Asp Gln Val Val Glu Glu Met  
 260 265 270  
 Lys Thr Leu Gly Ala Tyr Phe Ile Asn Glu Glu Glu Lys Ala Lys Leu  
 275 280 285  
 Glu Lys Phe Met Phe Gly Val Asn Ala Tyr Ser Ala Asp Val Asn Asn  
 290 295 300  
 Ala Arg Leu Asn Pro Lys Cys Pro Gly Met Ser Pro Gln Trp Phe Ala  
 305 310 315 320

ES 2 648 865 T3

Glu Gln Val Gly Ile Lys Val Pro Glu Asp Cys Asn Ile Ile Cys Ala  
 325 330 335  
 Val Cys Lys Glu Val Gly Pro Asn Glu Pro Leu Thr Arg Glu Lys Leu  
 340 345 350  
 Ser Pro Val Leu Ala Ile Leu Lys Ala Glu Asn Thr Gln Asp Gly Ile  
 355 360 365  
 Asp Lys Ala Glu Ala Met Val Glu Phe Asn Gly Arg Gly His Ser Ala  
 370 375 380  
 Ala Ile His Ser Asn Asp Lys Ala Val Val Glu Lys Tyr Ala Leu Thr  
 385 390 395  
 Met Lys Ala Cys Arg Ile Leu His Asn Thr Pro Ser Ser Gln Gly Gly  
 405 410 415  
 Ile Gly Ser Ile Tyr Asn Tyr Ile Trp Pro Ser Phe Thr Leu Gly Cys  
 420 425 430  
 Gly Ser Tyr Gly Gly Asn Ser Val Ser Ala Asn Val Thr Tyr His Asn  
 435 440 445  
 Leu Leu Asn Ile Lys Arg Leu Ala Asp Arg Arg Asn Asn Leu Gln Trp  
 450 455 460  
 Phe Arg Val Pro Pro Lys Ile Phe Phe Glu Pro His Ser Ile Arg Tyr  
 465 470 475 480  
 Leu Ala Glu Leu Lys Glu Leu Ser Lys Ile Phe Ile Val Ser Asp Arg  
 485 490 495  
 Met Met Tyr Lys Leu Gly Tyr Val Asp Arg Val Met Asp Val Leu Lys  
 500 505 510  
 Arg Arg Ser Asn Glu Val Glu Ile Glu Ile Phe Ile Asp Val Glu Pro  
 515 520 525  
 Asp Pro Ser Ile Gln Thr Val Gln Lys Gly Leu Ala Val Met Asn Thr  
 530 535 540  
 Phe Gly Pro Asp Asn Ile Ile Ala Ile Gly Gly Gly Ser Ala Met Asp  
 545 550 555 560  
 Ala Ala Lys Ile Met Trp Leu Leu Tyr Glu His Pro Glu Ala Asp Phe  
 565 570 575

ES 2 648 865 T3

Phe Ala Met Lys Gln Lys Phe Ile Asp Leu Arg Lys Arg Ala Phe Lys  
 580 585 590

Phe Pro Thr Met Gly Lys Lys Ala Arg Leu Ile Cys Ile Pro Thr Thr  
 595 600 605

Ser Gly Thr Gly Ser Glu Val Thr Pro Phe Ala Val Ile Ser Asp His  
 610 615 620

Glu Thr Gly Lys Lys Tyr Pro Leu Ala Asp Tyr Ser Leu Thr Pro Ser  
 625 630 635 640

Val Ala Ile Val Asp Pro Met Phe Thr Met Ser Leu Pro Lys Arg Ala  
 645 650 655

Ile Ala Asp Thr Gly Leu Asp Val Leu Val His Ala Thr Glu Ala Tyr  
 660 665 670

Val Ser Val Met Ala Asn Glu Tyr Thr Asp Gly Leu Ala Arg Glu Ala  
 675 680 685

Val Lys Leu Val Phe Glu Asn Leu Leu Lys Ser Tyr Asn Gly Asp Leu  
 690 695 700

Glu Ala Arg Glu Lys Met His Asn Ala Ala Thr Ile Ala Gly Met Ala  
 705 710 715 720

Phe Ala Ser Ala Phe Leu Gly Met Asp His Ser Met Ala His Lys Val  
 725 730 735

Gly Ala Ala Phe His Leu Pro His Gly Arg Cys Val Ala Val Leu Leu  
 740 745 750

Pro His Val Ile Arg Tyr Asn Gly Gln Lys Pro Arg Lys Leu Ala Met  
 755 760 765

Trp Pro Lys Tyr Asn Phe Tyr Lys Ala Asp Gln Arg Tyr Met Glu Leu  
 770 775 780

Ala Gln Met Val Gly Leu Lys Cys Asn Thr Pro Ala Glu Gly Val Glu  
 785 790 795 800

Ala Phe Ala Lys Ala Cys Glu Glu Leu Met Lys Ala Thr Glu Thr Ile  
 805 810 815

Thr Gly Phe Lys Lys Ala Asn Ile Asp Glu Ala Ala Trp Met Ser Lys

# ES 2 648 865 T3

			820					825					830				
Val	Pro	Glu	Met	Ala	Leu	Leu	Ala	Phe	Glu	Asp	Gln	Cys	Ser	Pro	Ala		
		835					840					845					
Asn	Pro	Arg	Val	Pro	Met	Val	Lys	Asp	Met	Glu	Lys	Ile	Leu	Lys	Ala		
	850					855					860						
Ala	Tyr	Tyr	Pro	Ile	Ala												
865					870												

<210> 6

<211> 2627

5 <212> ADN

<213> Artificial

<220>

<223> ADH2 sintético optimizado en codones

10

<400> 6

ES 2 648 865 T3

tctagaaaat	gtctacccaa	caaaccatga	ccgttgacga	acacatcaac	caattagtta	60
gaaaggctca	agttgctttg	aaggaatact	tgaagccaga	atacacccaa	gaaaagatcg	120
actacatcgt	taagaaggct	tctgttgctg	ctttggacca	acactgtgct	ttggctgctg	180
ctgctggtga	agaaaccggt	agaggatatct	togaagacaa	ggctaccaag	aacatcttcg	240
cttgtgaaca	cgttaccac	gaaatgagac	acgctaagac	cgttggtatc	atcaacgttg	300
acccattgta	cggtatcacc	gaaatcgctg	aaccagttgg	tgttgtttgt	ggtgttacc	360
cagttaccaa	cccaacctct	accgctatct	tcaagtcttt	gatctctatc	aagaccagaa	420
acccaatcgt	tttctcttct	caccatctcg	ctttgaagtg	ttctattatg	gctgctaaaa	480
tcgtagaga	cgctgctatc	gctgctggtg	ctccagaaaa	ctgtatccaa	tggatcgaat	540
tcggtggtat	cgaagcttct	aacaagttga	tgaaccacc	aggtgttgct	actatcttgg	600
ctaccggtgg	taacgctatg	gttaaggctg	catactcttc	tggttaagcca	gctttgggtg	660
ttggtgctgg	taacgttcca	acctacatcg	aaaagacctg	taacatcaag	caagctgcta	720
acgacgttgt	tatgtetaag	tctttcgaca	acggtatgat	ctgtgcttct	gaacaagctg	780
ctatcatcga	caaggaaatc	tacgaccaag	ttgttgaaga	aatgaagacc	ttgggtgctt	840
acttcatcaa	cgaagaagaa	aaggctaagt	tggaaaagtt	catgttcggt	gttaacgctt	900
actctgctga	cgttaacaac	gctagattga	acccaaagtg	tccaggtatg	tctccacaat	960
ggttcgctga	acaagttggt	atcaaggtag	cagaagactg	taacatcatc	tgtgctgttt	1020
gtaaggaagt	tggtccaaac	gaaccattga	ccagagaaaa	gttgtctcca	gttttggcta	1080
tcttgaaagc	tgaaaacacc	caagacggta	tcgacaaggc	tgaagctatg	gttgaattta	1140
acggtagagg	tcactctgct	gctatccact	ctaacgacaa	ggctgttggt	gaaaagtacg	1200

ES 2 648 865 T3

ctttgacat gaaggcttgt agaattctgc acaacacccc atcttctcaa ggtggtatcg 1260  
 gttctatcta caactacatc tggccatcctt tcaccttggg ttgtggttct tacgggtgta 1320  
 actctgtttc tgctaacggt acctaccaca acttggtgaa catcaagaga ttggctgaca 1380  
 gaagaaacaa cttgcaatgg ttcagagttc caccaaagat cttcttcgaa ccacactcta 1440  
 tcagatactt ggctgaattg aaggaattgt ctaagatctt catcgtttct gacagaatga 1500  
 tgtacaagtt gggttacggt gacagagtta tggacgtttt gaagagaaga tctaacgaag 1560  
 ttgaaatcga aatcttcacg gacgttgaac cagaccatc tatccaaacc gttcaaaagg 1620  
 gtttggtctg tatgaacacc ttcggtccag acaacatcat cgctatcggt ggtggttctg 1680  
 ctatggacgc tgctaagatc atgtggttgt tgtacgaaca cccagaagct gacttcttcg 1740  
 ctatgaagca aaagttcatc gacttgagaa agagagcttt caagttccca accatgggta 1800  
 agaaggctag attgatctgt atcccaacca cctctggtag cggttctgaa gttacccat 1860  
 tcgctgttat ctctgaccac gaaaccggta agaagtaccg attggtgac tactctttga 1920  
 ccccatctgt tgctatcgtt gacccaatgt tcaccatgtc tttgccaaag agagctatcg 1980  
 ctgacaccgg tttggacggt ttggttcacg ctaccgaagc ttacgttctt gttatggcta 2040  
 acgaatacac cgacggtttg gctagagaag ctgttaagtt ggtttttgaa aacttggtga 2100  
 agtcttacia cggtgacttg gaagctagag aaaagatgca caacgctgct accatcgctg 2160  
 gtatggcttt cgcttctgct ttcttgggta tggaccactc tatggctcac aaggttggg 2220  
 ctgctttcca cttgccacac ggtagatgtg ttgctgtttt gttgccacac gttatcagat 2280  
 acaacgggtca aaagccaaga aagttggcta tgtggccaaa gtacaacttc tacaaggctg 2340  
 accaaagata catggaattg gctcaaattg ttggtttgaa gtgtaacacc ccagctgaag 2400  
 gtgttgaagc tttcgctaag gcttgtgaag aattgatgaa ggctaccgaa accatcaccg 2460  
 gtttcaagaa ggctaacatc gacgaagctg cttggatgtc taaggttoca gaaatggctt 2520  
 tgttggcttt cgaagaccaa tgttctccag ctaaccacaa agttccaatg gtttaaggaca 2580  
 tggaaaagat cttgaaggct gcttactacc caatcgctta gcttaag 2627

<210> 7

<211> 367

5 <212> PRT

<213> Escherichia coli

<400> 7

Met Asp Arg Ile Ile Gln Ser Pro Gly Lys Tyr Ile Gln Gly Ala Asp  
 1 5 10 15

Val Ile Asn Arg Leu Gly Glu Tyr Leu Lys Pro Leu Ala Glu Arg Trp  
 20 25 30

10

ES 2 648 865 T3

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

ES 2 648 865 T3

275 280 285

Pro Val Glu Glu Ile Glu Thr Val Ala Ala Leu Ser His Ala Val Gly  
 290 295 300

Leu Pro Ile Thr Leu Ala Gln Leu Asp Ile Lys Glu Asp Val Pro Ala  
 305 310 315 320

Lys Met Arg Ile Val Ala Glu Ala Ala Cys Ala Glu Gly Glu Thr Ile  
 325 330 335

His Asn Met Pro Gly Gly Ala Thr Pro Asp Gln Val Tyr Ala Ala Leu  
 340 345 350

Leu Val Ala Asp Gln Tyr Gly Gln Arg Phe Leu Gln Glu Trp Glu  
 355 360 365

<210> 8

<211> 584

5

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 8

Met Ser Ala Lys Ser Phe Glu Val Thr Asp Pro Val Asn Ser Ser Leu  
 1 5 10 15

Lys Gly Phe Ala Leu Ala Asn Pro Ser Ile Thr Leu Val Pro Glu Glu  
 20 25 30

Lys Ile Leu Phe Arg Lys Thr Asp Ser Asp Lys Ile Ala Leu Ile Ser  
 35 40 45

Gly Gly Gly Ser Gly His Glu Pro Thr His Ala Gly Phe Ile Gly Lys  
 50 55 60

Gly Met Leu Ser Gly Ala Val Val Gly Glu Ile Phe Ala Ser Pro Ser  
 65 70 75 80

Thr Lys Gln Ile Leu Asn Ala Ile Arg Leu Val Asn Glu Asn Ala Ser  
 85 90 95

Gly Val Leu Leu Ile Val Lys Asn Tyr Thr Gly Asp Val Leu His Phe  
 100 105 110

Gly Leu Ser Ala Glu Arg Ala Arg Ala Leu Gly Ile Asn Cys Arg Val  
 115 120 125

Ala Val Ile Gly Asp Asp Val Ala Val Gly Arg Glu Lys Gly Gly Met

10



ES 2 648 865 T3

Phe Asp Lys Phe Ala Glu Trp Met Lys Ser Gly Ala Glu Gln Val Ile  
 385 390 395 400

Lys Ser Glu Pro His Ile Thr Glu Leu Asp Asn Gln Val Gly Asp Gly  
 405 410 415

Asp Cys Gly Tyr Thr Leu Val Ala Gly Val Lys Gly Ile Thr Glu Asn  
 420 425 430

Leu Asp Lys Leu Ser Lys Asp Ser Leu Ser Gln Ala Val Ala Gln Ile  
 435 440 445

Ser Asp Phe Ile Glu Gly Ser Met Gly Gly Thr Ser Gly Gly Leu Tyr  
 450 455 460

Ser Ile Leu Leu Ser Gly Phe Ser His Gly Leu Ile Gln Val Cys Lys  
 465 470 475 480

Ser Lys Asp Glu Pro Val Thr Lys Glu Ile Val Ala Lys Ser Leu Gly  
 485 490 495

Ile Ala Leu Asp Thr Leu Tyr Lys Tyr Thr Lys Ala Arg Lys Gly Ser  
 500 505 510

Ser Thr Met Ile Asp Ala Leu Glu Pro Phe Val Lys Glu Phe Thr Ala  
 515 520 525

Ser Lys Asp Phe Asn Lys Ala Val Lys Ala Ala Glu Glu Gly Ala Lys  
 530 535 540

Ser Thr Ala Thr Phe Glu Ala Lys Phe Gly Arg Ala Ser Tyr Val Gly  
 545 550 555 560

Asp Ser Ser Gln Val Glu Asp Pro Gly Ala Val Gly Leu Cys Glu Phe  
 565 570 575

Leu Lys Gly Val Gln Ser Ala Leu  
 580

<210> 9

<211> 591

5 <212> PRT

<213> *Saccharomyces cerevisiae*

<400> 9

10 Met Ser His Lys Gln Phe Lys Ser Asp Gly Asn Ile Val Thr Pro Tyr  
 1 5 10 15

ES 2 648 865 T3

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

ES 2 648 865 T3

Met Gly Gly Met Ser Asn Leu Glu Leu Gly Tyr Ala Ala Glu Val Ile  
 275 280 285

Ser Glu Gln Leu Ile Asp Lys Tyr Gln Ile Val Pro Lys Arg Thr Ile  
 290 295 300

Thr Gly Ala Phe Ile Thr Ala Leu Asn Gly Pro Gly Phe Gly Ile Thr  
 305 310 315 320

Leu Met Asn Ala Ser Lys Ala Gly Gly Asp Ile Leu Lys Tyr Phe Asp  
 325 330 335

Tyr Pro Thr Thr Ala Ser Gly Trp Asn Gln Met Tyr His Ser Ala Lys  
 340 345 350

Asp Trp Glu Val Leu Ala Lys Gly Gln Val Pro Thr Ala Pro Ser Leu  
 355 360 365

Lys Thr Leu Arg Asn Glu Lys Gly Ser Gly Val Lys Ala Asp Tyr Asp  
 370 375 380

Thr Phe Ala Lys Ile Leu Leu Ala Gly Ile Ala Lys Ile Asn Glu Val  
 385 390 395 400

Glu Pro Lys Val Thr Trp Tyr Asp Thr Ile Ala Gly Asp Gly Asp Cys  
 405 410 415

Gly Thr Thr Leu Val Ser Gly Gly Glu Ala Leu Glu Glu Ala Ile Lys  
 420 425 430

Asn His Thr Leu Arg Leu Glu Asp Ala Ala Leu Gly Ile Glu Asp Ile  
 435 440 445

Ala Tyr Met Val Glu Asp Ser Met Gly Gly Thr Ser Gly Gly Leu Tyr  
 450 455 460

Ser Ile Tyr Leu Ser Ala Leu Ala Gln Gly Val Arg Asp Ser Gly Asp  
 465 470 475 480

Lys Glu Leu Thr Ala Glu Thr Phe Lys Lys Ala Ser Asn Val Ala Leu  
 485 490 495

Asp Ala Leu Tyr Lys Tyr Thr Arg Ala Arg Pro Gly Tyr Arg Thr Leu  
 500 505 510

Ile Asp Ala Leu Gln Pro Phe Val Glu Ala Leu Lys Ala Gly Lys Gly  
 515 520 525

ES 2 648 865 T3

Pro Arg Ala Ala Ala Gln Ala Ala Tyr Asp Gly Ala Glu Lys Thr Arg  
 530 535 540

Lys Met Asp Ala Leu Val Gly Arg Ala Ser Tyr Val Ala Lys Glu Glu  
 545 550 555 560

Leu Arg Lys Leu Asp Ser Glu Gly Gly Leu Pro Asp Pro Gly Ala Val  
 565 570 575

Gly Leu Ala Ala Leu Leu Asp Gly Phe Val Thr Ala Ala Gly Tyr  
 580 585 590

<210> 10

<211> 560

5

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 10

Met Ser Leu Ile Ser Ile Leu Ser Pro Leu Ile Thr Ser Glu Gly Leu  
 1 5 10 15

Asp Ser Arg Ile Lys Pro Ser Pro Lys Lys Asp Ala Ser Thr Thr Thr  
 20 25 30

Lys Pro Ser Leu Trp Lys Thr Thr Glu Phe Lys Phe Tyr Tyr Ile Ala  
 35 40 45

Phe Leu Val Val Val Pro Leu Met Phe Tyr Ala Gly Leu Gln Ala Ser  
 50 55 60

Ser Pro Glu Asn Pro Asn Tyr Ala Arg Tyr Glu Arg Leu Leu Ser Gln  
 65 70 75 80

Gly Trp Leu Phe Gly Arg Lys Val Asp Asn Ser Asp Ser Gln Tyr Arg  
 85 90 95

Phe Phe Arg Asp Asn Phe Ala Leu Leu Ser Val Leu Met Leu Val His  
 100 105 110

Thr Ser Ile Lys Arg Ile Val Leu Tyr Ser Thr Asn Ile Thr Lys Leu  
 115 120 125

Arg Phe Asp Leu Ile Phe Gly Leu Ile Phe Leu Val Ala Ala His Gly  
 130 135 140

Val Asn Ser Ile Arg Ile Leu Ala His Met Leu Ile Leu Tyr Ala Ile  
 145 150 155 160

10

ES 2 648 865 T3

Ala His Val Leu Lys Asn Phe Arg Arg Ile Ala Thr Ile Ser Ile Trp  
165 170 175

Ile Tyr Gly Ile Ser Thr Leu Phe Ile Asn Asp Asn Phe Arg Ala Tyr  
180 185 190

Pro Phe Gly Asn Ile Cys Ser Phe Leu Ser Pro Leu Asp His Trp Tyr  
195 200 205

Arg Gly Ile Ile Pro Arg Trp Asp Val Phe Phe Asn Phe Thr Leu Leu  
210 215 220

Arg Val Leu Ser Tyr Asn Leu Asp Phe Leu Glu Arg Trp Glu Asn Leu  
225 230 235 240

Gln Lys Lys Lys Ser Pro Ser Tyr Glu Ser Lys Glu Ala Lys Ser Ala  
245 250 255

Ile Leu Leu Asn Glu Arg Ala Arg Leu Thr Ala Ala His Pro Ile Gln  
260 265 270

Asp Tyr Ser Leu Met Asn Tyr Ile Ala Tyr Val Thr Tyr Thr Pro Leu  
275 280 285

Phe Ile Ala Gly Pro Ile Ile Thr Phe Asn Asp Tyr Val Tyr Gln Ser  
290 295 300

Lys His Thr Leu Pro Ser Ile Asn Phe Lys Phe Ile Phe Tyr Tyr Ala  
305 310 315 320

Val Arg Phe Val Ile Ala Leu Leu Ser Met Glu Phe Ile Leu His Phe  
325 330 335

Leu His Val Val Ala Ile Ser Lys Thr Lys Ala Trp Glu Asn Asp Thr  
340 345 350

Pro Phe Gln Ile Ser Met Ile Gly Leu Phe Asn Leu Asn Ile Ile Trp  
355 360 365

Leu Lys Leu Leu Ile Pro Trp Arg Leu Phe Arg Leu Trp Ala Leu Leu  
370 375 380

Asp Gly Ile Asp Thr Pro Glu Asn Met Ile Arg Cys Val Asp Asn Asn  
385 390 395 400

Tyr Ser Ser Leu Ala Phe Trp Arg Ala Trp His Arg Ser Tyr Asn Lys

ES 2 648 865 T3

405 410 415

Trp Val Val Arg Tyr Ile Tyr Ile Pro Leu Gly Gly Ser Lys Asn Arg  
420 425 430

Val Leu Thr Ser Leu Ala Val Phe Ser Phe Val Ala Ile Trp His Asp  
435 440 445

Ile Glu Leu Lys Leu Leu Leu Trp Gly Trp Leu Ile Val Leu Phe Leu  
450 455 460

Leu Pro Glu Ile Phe Ala Thr Gln Ile Phe Ser His Tyr Thr Asp Ala  
465 470 475 480

Val Trp Tyr Arg His Val Cys Ala Val Gly Ala Val Phe Asn Ile Trp  
485 490 495

Val Met Met Ile Ala Asn Leu Phe Gly Phe Cys Leu Gly Ser Asp Gly  
500 505 510

Thr Lys Lys Leu Leu Ser Asp Met Phe Cys Thr Val Ser Gly Phe Lys  
515 520 525

Phe Val Ile Leu Ala Ser Val Ser Leu Phe Ile Ala Val Gln Ile Met  
530 535 540

Phe Glu Ile Arg Glu Glu Glu Lys Arg His Gly Ile Tyr Leu Lys Cys  
545 550 555 560

<210> 11

<211> 609

5

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 11

Met Ser Met Leu Arg Ile Trp Ser Cys Ile Val His Phe Phe Ser Val  
1 5 10 15

Gln Ala Leu Asp Ser Arg Ile Lys Pro Asp Ile Glu Phe Lys Arg Arg  
20 25 30

Gln Arg Ile Phe Ile Asn Ser Ser Lys Glu Glu Asn Gly Ser Ser Ser  
35 40 45

Ser Ala Val Thr Val Thr Arg Asn Pro Val Leu Ser Ser Asn Ser Pro  
50 55 60

Ser Pro Pro Leu Trp Asn Thr Trp Glu Phe Arg Leu Tyr Tyr Leu Ala

10



ES 2 648 865 T3

Gln Asp Tyr Asn Phe Ile Asn Phe Ile Ala Tyr Ile Thr Tyr Ala Pro  
 325 330 335

Leu Phe Leu Val Gly Pro Ile Ile Thr Phe Asn Asp Tyr Leu Tyr Gln  
 340 345 350

Ser Glu Asn Lys Leu Pro Ser Leu Thr Lys Lys Asn Ile Gly Phe Tyr  
 355 360 365

Ala Leu Lys Val Phe Ser Ser Leu Leu Leu Met Glu Ile Ile Leu His  
 370 375 380

Tyr Ile Tyr Val Gly Ala Ile Ala Arg Thr Lys Ala Trp Asn Asn Asp  
 385 390 395 400

Thr Pro Leu Gln Gln Ala Met Ile Ala Leu Phe Asn Leu Asn Ile Met  
 405 410 415

Tyr Leu Lys Leu Leu Ile Pro Trp Arg Leu Phe Arg Leu Trp Ala Met  
 420 425 430

Val Asp Gly Ile Asp Ala Pro Glu Asn Met Leu Arg Cys Val Asp Asn  
 435 440 445

Asn Tyr Ser Thr Val Gly Phe Trp Arg Ala Trp His Thr Ser Phe Asn  
 450 455 460

Lys Trp Val Ile Arg Tyr Ile Tyr Val Pro Phe Gly Gly Ser Asn Asn  
 465 470 475 480

Lys Ile Leu Thr Ser Phe Ala Val Phe Ser Phe Val Ala Ile Trp His  
 485 490 495

Asp Ile Gln Leu Arg Val Leu Phe Trp Gly Trp Leu Thr Val Leu Leu  
 500 505 510

Leu Leu Gly Glu Thr Tyr Ile Thr Asn Cys Phe Ser Arg Tyr Arg Phe  
 515 520 525

Arg Ser Trp Tyr Arg Phe Val Cys Gly Ile Gly Ala Ala Ile Asn Ile  
 530 535 540

Cys Met Met Met Ile Ile Asn Val Tyr Gly Phe Cys Leu Gly Ala Glu  
 545 550 555 560

Gly Thr Lys Leu Leu Leu Lys Gly Ile Phe Asn Asn Ser His Ser Pro  
 565 570 575

Glu Phe Leu Thr Ala Val Met Val Ser Leu Phe Ile Ala Val Gln Val  
 580 585 590

Met Phe Glu Ile Arg Glu Glu Glu Lys Arg His Gly Ile Asn Leu Lys  
 595 600 605

Cys

ES 2 648 865 T3

<210> 12

<211> 669

<212> PRT

<213> Saccharomyces cerevisiae

5

<400> 12

```

Met Ser Asn Pro Gln Lys Ala Leu Asn Asp Phe Leu Ser Ser Glu Ser
 1                               5 10 15

Val His Thr His Asp Ser Ser Arg Lys Gln Ser Asn Lys Gln Ser Ser
 20 25 30

Asp Glu Gly Arg Ser Ser Ser Gln Pro Ser His His His Ser Gly Gly
 35 40 45

Thr Asn Ser Asn Asn
 50 55 60

Asn Asn Asn Gly Asn Asp Gly Gly Asn Asp Asp Asp Tyr Asp Tyr Glu
 65 70 75 80

Met Gln Asp Tyr Arg Pro Ser Pro Gln Ser Ala Arg Pro Thr Pro Thr
 85 90 95

Tyr Val Pro Gln Tyr Ser Val Glu Ser Gly Thr Ala Phe Pro Ile Gln
 100 105 110

Glu Val Ile Pro Ser Ala Tyr Ile Asn Thr Gln Asp Ile Asn His Lys
 115 120 125

Asp Asn Gly Pro Pro Ser Ala Ser Ser Asn Arg Ala Phe Arg Pro Arg
 130 135 140

Gly Gln Thr Thr Val Ser Ala Asn Val Leu Asn Ile Glu Asp Phe Tyr
 145 150 155 160

Lys Asn Ala Asp Asp Ala His Thr Ile Pro Glu Ser His Leu Ser Arg
 165 170 175
    
```

ES 2 648 865 T3

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

ES 2 648 865 T3

Cys Gly Ala Met Leu Gln Ala Gly Thr Phe Ala Leu Thr Asp Pro Tyr  
435 440 445

Thr Cys Leu Ser Ser Asp Val Phe Pro Leu Met Met Phe Ile Leu Ile  
450 455 460

Phe Ile Ile Asn Ala Ser Met Ala Tyr Gln Thr Gly Thr Ala Met Asn  
465 470 475 480

Leu Ala Arg Asp Leu Gly Pro Arg Leu Ala Leu Tyr Ala Val Gly Phe  
485 490 495

Asp His Lys Met Leu Trp Val His His His His Phe Phe Trp Val Pro  
500 505 510

Met Val Gly Pro Phe Ile Gly Ala Leu Met Gly Gly Leu Val Tyr Asp  
515 520 525

Val Cys Ile Tyr Gln Gly His Glu Ser Pro Val Asn Trp Ser Leu Pro  
530 535 540

Val Tyr Lys Glu Met Ile Met Arg Ala Trp Phe Arg Arg Pro Gly Trp  
545 550 555 560

Lys Lys Arg Asn Arg Ala Arg Arg Thr Ser Asp Leu Ser Asp Phe Ser  
565 570 575

Tyr Asn Asn Asp Asp Asp Glu Glu Phe Gly Glu Arg Met Ala Leu Gln  
580 585 590

Lys Thr Lys Thr Lys Ser Ser Ile Ser Asp Asn Glu Asn Glu Ala Gly  
595 600 605

Glu Lys Lys Val Gln Phe Lys Ser Val Gln Arg Gly Lys Arg Thr Phe  
610 615 620

Gly Gly Ile Pro Thr Ile Leu Glu Glu Glu Asp Ser Ile Glu Thr Ala  
625 630 635 640

Ser Leu Gly Ala Thr Thr Thr Asp Ser Ile Gly Leu Ser Asp Thr Ser  
645 650 655

Ser Glu Asp Ser His Tyr Gly Asn Ala Lys Lys Val Thr  
660 665

<210> 13

<211> 713

5 <212> PRT

<213> Saccharomyces cerevisiae

<400> 13

ES 2 648 865 T3

Met Ser Pro Ser Ala Val Gln Ser Ser Lys Leu Glu Glu Gln Ser Ser  
 1 5 10 15

Glu Ile Asp Lys Leu Lys Ala Lys Met Ser Gln Ser Ala Ala Thr Ala  
 20 25 30

Gln Gln Lys Lys Glu His Glu Tyr Glu His Leu Thr Ser Val Lys Ile  
 35 40 45

Val Pro Gln Arg Pro Ile Ser Asp Arg Leu Gln Pro Ala Ile Ala Thr  
 50 55 60

His Tyr Ser Pro His Leu Asp Gly Leu Gln Asp Tyr Gln Arg Leu His  
 65 70 75 80

Lys Glu Ser Ile Glu Asp Pro Ala Lys Phe Phe Gly Ser Lys Ala Thr  
 85 90 95

Gln Phe Leu Asn Trp Ser Lys Pro Phe Asp Lys Val Phe Ile Pro Asp  
 100 105 110

Pro Lys Thr Gly Arg Pro Ser Phe Gln Asn Asn Ala Trp Phe Leu Asn  
 115 120 125

Gly Gln Leu Asn Ala Cys Tyr Asn Cys Val Asp Arg His Ala Leu Lys  
 130 135 140

Thr Pro Asn Lys Lys Ala Ile Ile Phe Glu Gly Asp Glu Pro Gly Gln  
 145 150 155 160

Gly Tyr Ser Ile Thr Tyr Lys Glu Leu Leu Glu Glu Val Cys Gln Val  
 165 170 175

Ala Gln Val Leu Thr Tyr Ser Met Gly Val Arg Lys Gly Asp Thr Val  
 180 185 190

Ala Val Tyr Met Pro Met Val Pro Glu Ala Ile Ile Thr Leu Leu Ala  
 195 200 205

Ile Ser Arg Ile Gly Ala Ile His Ser Val Val Phe Ala Gly Phe Ser  
 210 215 220

Ser Asn Ser Leu Arg Asp Arg Ile Asn Asp Gly Asp Ser Lys Val Val



ES 2 648 865 T3

Pro Leu Ala Gly Gly Val Thr Pro Met Lys Pro Gly Ser Ala Ser Phe  
 485 490 495

Pro Phe Phe Gly Ile Asp Ala Val Val Leu Asp Pro Asn Thr Gly Glu  
 500 505 510

Glu Leu Asn Thr Ser His Ala Glu Gly Val Leu Ala Val Lys Ala Ala  
 515 520 525

Trp Pro Ser Phe Ala Arg Thr Ile Trp Lys Asn His Asp Arg Tyr Leu  
 530 535 540

Asp Thr Tyr Leu Asn Pro Tyr Pro Gly Tyr Tyr Phe Thr Gly Asp Gly  
 545 550 555 560

Ala Ala Lys Asp Lys Asp Gly Tyr Ile Trp Ile Leu Gly Arg Val Asp  
 565 570 575

Asp Val Val Asn Val Ser Gly His Arg Leu Ser Thr Ala Glu Ile Glu  
 580 585 590

Ala Ala Ile Ile Glu Asp Pro Ile Val Ala Glu Cys Ala Val Val Gly  
 595 600 605

Phe Asn Asp Asp Leu Thr Gly Gln Ala Val Ala Ala Phe Val Val Leu  
 610 615 620

Lys Asn Lys Ser Ser Trp Ser Thr Ala Thr Asp Asp Glu Leu Gln Asp  
 625 630 635 640

Ile Lys Lys His Leu Val Phe Thr Val Arg Lys Asp Ile Gly Pro Phe  
 645 650 655

Ala Ala Pro Lys Leu Ile Ile Leu Val Asp Asp Leu Pro Lys Thr Arg  
 660 665 670

Ser Gly Lys Ile Met Arg Arg Ile Leu Arg Lys Ile Leu Ala Gly Glu  
 675 680 685

Ser Asp Gln Leu Gly Asp Val Ser Thr Leu Ser Asn Pro Gly Ile Val  
 690 695 700

Arg His Leu Ile Asp Ser Val Lys Leu  
 705 710

<210> 14

<211> 683

5 <212> PRT

<213> Saccharomyces cerevisiae

<400> 14

ES 2 648 865 T3

Met Thr Ile Lys Glu His Lys Val Val Tyr Glu Ala His Asn Val Lys  
 1 5 10 15

Ala Leu Lys Ala Pro Gln His Phe Tyr Asn Ser Gln Pro Gly Lys Gly  
 20 25 30

Tyr Val Thr Asp Met Gln His Tyr Gln Glu Met Tyr Gln Gln Ser Ile  
 35 40 45

Asn Glu Pro Glu Lys Phe Phe Asp Lys Met Ala Lys Glu Tyr Leu His  
 50 55 60

Trp Asp Ala Pro Tyr Thr Lys Val Gln Ser Gly Ser Leu Asn Asn Gly  
 65 70 75 80

Asp Val Ala Trp Phe Leu Asn Gly Lys Leu Asn Ala Ser Tyr Asn Cys  
 85 90 95

Val Asp Arg His Ala Phe Ala Asn Pro Asp Lys Pro Ala Leu Ile Tyr  
 100 105 110

Glu Ala Asp Asp Glu Ser Asp Asn Lys Ile Ile Thr Phe Gly Glu Leu  
 115 120 125

Leu Arg Lys Val Ser Gln Ile Ala Gly Val Leu Lys Ser Trp Gly Val  
 130 135 140

Lys Lys Gly Asp Thr Val Ala Ile Tyr Leu Pro Met Ile Pro Glu Ala  
 145 150 155 160

Val Ile Ala Met Leu Ala Val Ala Arg Ile Gly Ala Ile His Ser Val  
 165 170 175

Val Phe Ala Gly Phe Ser Ala Gly Ser Leu Lys Asp Arg Val Val Asp  
 180 185 190

Ala Asn Ser Lys Val Val Ile Thr Cys Asp Glu Gly Lys Arg Gly Gly  
 195 200 205

Lys Thr Ile Asn Thr Lys Lys Ile Val Asp Glu Gly Leu Asn Gly Val  
 210 215 220

Asp Leu Val Ser Arg Ile Leu Val Phe Gln Arg Thr Gly Thr Glu Gly  
 225 230 235 240

ES 2 648 865 T3

Ile Pro Met Lys Ala Gly Arg Asp Tyr Trp Trp His Glu Glu Ala Ala  
 245 250 255

Lys Gln Arg Thr Tyr Leu Pro Pro Val Ser Cys Asp Ala Glu Asp Pro  
 260 265 270

Leu Phe Leu Leu Tyr Thr Ser Gly Ser Thr Gly Ser Pro Lys Gly Val  
 275 280 285

Val His Thr Thr Gly Gly Tyr Leu Leu Gly Ala Ala Leu Thr Thr Arg  
 290 295 300

Tyr Val Phe Asp Ile His Pro Glu Asp Val Leu Phe Thr Ala Gly Asp  
 305 310 315 320

Val Gly Trp Ile Thr Gly His Thr Tyr Ala Leu Tyr Gly Pro Leu Thr  
 325 330 335

Leu Gly Thr Ala Ser Ile Ile Phe Glu Ser Thr Pro Ala Tyr Pro Asp  
 340 345 350

Tyr Gly Arg Tyr Trp Arg Ile Ile Gln Arg His Lys Ala Thr His Phe  
 355 360 365

Tyr Val Ala Pro Thr Ala Leu Arg Leu Ile Lys Arg Val Gly Glu Ala  
 370 375 380

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

Gly Glu Pro Ile Ser Pro Asp Leu Trp Glu Trp Tyr His Glu Lys Val  
 405 410 415

Gly Asn Lys Asn Cys Val Ile Cys Asp Thr Met Trp Gln Thr Glu Ser  
 420 425 430

Gly Ser His Leu Ile Ala Pro Leu Ala Gly Ala Val Pro Thr Lys Pro  
 435 440 445

Gly Ser Ala Thr Val Pro Phe Phe Gly Ile Asn Ala Cys Ile Ile Asp  
 450 455 460

Pro Val Thr Gly Val Glu Leu Glu Gly Asn Asp Val Glu Gly Val Leu  
 465 470 475 480

Ala Val Lys Ser Pro Trp Pro Ser Met Ala Arg Ser Val Trp Asn His  
 485 490 495

ES 2 648 865 T3

His Asp Arg Tyr Met Asp Thr Tyr Leu Lys Pro Tyr Pro Gly His Tyr  
 500 505 510

Phe Thr Gly Asp Gly Ala Gly Arg Asp His Asp Gly Tyr Tyr Trp Ile  
 515 520 525

Arg Gly Arg Val Asp Asp Val Val Asn Val Ser Gly His Arg Leu Ser  
 530 535 540

Thr Ser Glu Ile Glu Ala Ser Ile Ser Asn His Glu Asn Val Ser Glu  
 545 550 555 560

Ala Ala Val Val Gly Ile Pro Asp Glu Leu Thr Gly Gln Thr Val Val  
 565 570 575

Ala Tyr Val Ser Leu Lys Asp Gly Tyr Leu Gln Asn Asn Ala Thr Glu  
 580 585 590

Gly Asp Ala Glu His Ile Thr Pro Asp Asn Leu Arg Arg Glu Leu Ile  
 595 600 605

Leu Gln Val Arg Gly Glu Ile Gly Pro Phe Ala Ser Pro Lys Thr Ile  
 610 615 620

Ile Leu Val Arg Asp Leu Pro Arg Thr Arg Ser Gly Lys Ile Met Arg  
 625 630 635 640

Arg Val Leu Arg Lys Val Ala Ser Asn Glu Ala Glu Gln Leu Gly Asp  
 645 650 655

Leu Thr Thr Leu Ala Asn Pro Glu Val Val Pro Ala Ile Ile Ser Ala  
 660 665 670

Val Glu Asn Gln Phe Phe Ser Gln Lys Lys Lys  
 675 680

<210> 15

<211> 384

5 <212> PRT

<213> Saccharomyces cerevisiae

<400> 15

Met Thr Ala His Thr Asn Ile Lys Gln His Lys His Cys His Glu Asp  
 1 5 10 15

His Pro Ile Arg Arg Ser Asp Ser Ala Val Ser Ile Val His Leu Lys  
 20 25 30

10

ES 2 648 865 T3

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

ES 2 648 865 T3

275 280 285

Lys Val Glu Thr Tyr Tyr Gln Glu Ser Ala Gly Val Ala Asp Leu Ile  
 290 295 300

Thr Thr Cys Ser Gly Gly Arg Asn Val Lys Val Ala Thr Tyr Met Ala  
 305 310 315 320

Lys Thr Gly Lys Ser Ala Leu Glu Ala Glu Lys Glu Leu Leu Asn Gly  
 325 330 335

Gln Ser Ala Gln Gly Ile Ile Thr Cys Arg Glu Val His Glu Trp Leu  
 340 345 350

Gln Thr Cys Glu Leu Thr Gln Glu Phe Pro Ile Ile Arg Gly Ser Leu  
 355 360 365

Pro Asp Ser Leu Gln Gln Arg Pro His Gly Arg Pro Thr Gly Asp Asp  
 370 375 380

<210> 16

<211> 391

5

<212> PRT

<213> *Saccharomyces cerevisiae*

<400> 16

Met Ser Ala Ala Ala Asp Arg Leu Asn Leu Thr Ser Gly His Leu Asn  
 1 5 10 15

Ala Gly Arg Lys Arg Ser Ser Ser Ser Val Ser Leu Lys Ala Ala Glu  
 20 25 30

Lys Pro Phe Lys Val Thr Val Ile Gly Ser Gly Asn Trp Gly Thr Thr  
 35 40 45

Ile Ala Lys Val Val Ala Glu Asn Cys Lys Gly Tyr Pro Glu Val Phe  
 50 55 60

Ala Pro Ile Val Gln Met Trp Val Phe Glu Glu Glu Ile Asn Gly Glu  
 65 70 75 80

Lys Leu Thr Glu Ile Ile Asn Thr Arg His Gln Asn Val Lys Tyr Leu  
 85 90 95

Pro Gly Ile Thr Leu Pro Asp Asn Leu Val Ala Asn Pro Asp Leu Ile  
 100 105 110

Asp Ser Val Lys Asp Val Asp Ile Ile Val Phe Asn Ile Pro His Gln

10

ES 2 648 865 T3

		115						120						125			
Phe	Leu	Pro	Arg	Ile	Cys	Ser	Gln	Leu	Lys	Gly	His	Val	Asp	Ser	His		
	130					135					140						
Val	Arg	Ala	Ile	Ser	Cys	Leu	Lys	Gly	Phe	Glu	Val	Gly	Ala	Lys	Gly		
145					150					155					160		
Val	Gln	Leu	Leu	Ser	Ser	Tyr	Ile	Thr	Glu	Glu	Leu	Gly	Ile	Gln	Cys		
				165					170					175			
Gly	Ala	Leu	Ser	Gly	Ala	Asn	Ile	Ala	Thr	Glu	Val	Ala	Gln	Glu	His		
			180					185					190				
Trp	Ser	Glu	Thr	Thr	Val	Ala	Tyr	His	Ile	Pro	Lys	Asp	Phe	Arg	Gly		
		195					200					205					
Glu	Gly	Lys	Asp	Val	Asp	His	Lys	Val	Leu	Lys	Ala	Leu	Phe	His	Arg		
	210					215					220						
Pro	Tyr	Phe	His	Val	Ser	Val	Ile	Glu	Asp	Val	Ala	Gly	Ile	Ser	Ile		
225					230					235					240		
Cys	Gly	Ala	Leu	Lys	Asn	Val	Val	Ala	Leu	Gly	Cys	Gly	Phe	Val	Glu		
				245					250					255			
Gly	Leu	Gly	Trp	Gly	Asn	Asn	Ala	Ser	Ala	Ala	Ile	Gln	Arg	Val	Gly		
			260					265					270				
Leu	Gly	Glu	Ile	Ile	Arg	Phe	Gly	Gln	Met	Phe	Phe	Pro	Glu	Ser	Arg		
		275					280					285					
Glu	Glu	Thr	Tyr	Tyr	Gln	Glu	Ser	Ala	Gly	Val	Ala	Asp	Leu	Ile	Thr		
	290					295						300					
Thr	Cys	Ala	Gly	Gly	Arg	Asn	Val	Lys	Val	Ala	Arg	Leu	Met	Ala	Thr		
305					310					315					320		
Ser	Gly	Lys	Asp	Ala	Trp	Glu	Cys	Glu	Lys	Glu	Leu	Leu	Asn	Gly	Gln		
				325					330					335			
Ser	Ala	Gln	Gly	Leu	Ile	Thr	Cys	Lys	Glu	Val	His	Glu	Trp	Leu	Glu		
			340					345					350				
Thr	Cys	Gly	Ser	Val	Glu	Asp	Phe	Pro	Leu	Phe	Glu	Ala	Val	Tyr	Gln		
		355					360					365					
Ile	Val	Tyr	Asn	Asn	Tyr	Pro	Met	Lys	Asn	Leu	Pro	Asp	Met	Ile	Glu		
	370					375						380					
Glu	Leu	Asp	Leu	His	Glu	Asp											
385					390												

ES 2 648 865 T3

<211> 2573

<212> ADN

<213> Artificial

5

<220>

<223> Rotura de GDP1

<400> 17

```

ggtagccccg ccttgcttct ctccccttcc ttttcttttt ccagttttcc ctattttgtc      60
cctttttccg cacaacaagt atcagaatgg gttcatcaaa tctatccaac ctaattcgca      120
cgtagactgg cttggatttg gcagtttcgt agttatata atactaccat gagtgaaact      180
gttacgttac cttaaattct ttctcccttt aattttcttt tatcttactc tcctacataa      240
gacatcaaga aacaattgta tattgtacac cccccctc cacaacaca aatattgata      300
atataaagat gtctgctgct gctgatagat taaacttaac ttccggccac ttgaatctag      360
atgcatgctc gagcggccgc cagtgtgatg gatatctgca gaattcgccc ttttgggccc      420
tgtacaccta ggatccgtcg acactggatg gcggcgttag tatcgaatcg acagcagtat      480
agcgaccagc attcacatac gattgacgca tgatattact ttctgcgcac ttaacttcgc      540
atctgggcag atgatgtcga ggcgaaaaaa aatataaatc acgctaacat ttgattaaaa      600
tagaacaact acaatataaa aaaactatac aatgacaag ttcttgaaaa caagaatctt      660
tttattgtca gtactgatta ttcttttggc ctcgacgag tgctggggcg tcggtttcca      720
ctatcggcga gtactttctac acagccatcg gtccagacgg ccgcgcttct gcgggcgatt      780
tgtgtacgcc cgacagtccc ggctccggat cggacgattg cgtcgcatcg accctgcgcc      840
caagctgcat catcgaaatt gccgtcaacc aagctctgat agagttggtc aagaccaatg      900
cggagcatat acgcccggag ccgcggcgat cctgcaagct ccggatgcct ccgctcgaag      960
tagcgcgtct gctgctccat acaagccaac cacggcctcc agaagaagat gttggcgacc     1020
tcgtattggg aatccccgaa catcgctctg ctccagtcaa tgaccgctgt tatgoggcca     1080
ttgtccgtca ggacattggt ggagccgaaa tccgcgtgca cgaggtgccg gacttcgggg     1140
cagtcctcgg cccaaagcat cagctcatcg agagcctcgc cgacggacgc actgacggtg     1200
tcgtccatca cagtttgcca gtgatacaca tggggatcag caatcgcgca tatgaaatca     1260
cgccatgtag tgtattgacc gattccttgc ggtccgaatg ggccgaacct gctcgtctgg     1320
ctaagatcgg ccgcagcgat cgcattcatg gcctccgca ccggctgcag aacagcgggc     1380

```

10

ES 2 648 865 T3

agttcggttt caggcaggtc ttgcaacgtg acaccctgtg cacggcggga gatgcaatag 1440  
 gtcaggctct cgctgaattc cccaatgtca agcacttccg gaatcgggag cgcggccgat 1500  
 gcaaagtgcc gataaacata acgatctttg tagaaacat cggcgcagct atttaccgc 1560  
 aggacatatc cacgccctcc tacatogaag ctgaaagcac gagattcttc gcctccgag 1620  
 agctgcatca ggtcggagac gctgtcgaac ttttcgatca gaaacttctc gacagacgtc 1680  
 gcggtgagtt caggcttttt acccatgggt gtttatgttc ggatgtgatg tgagaactgt 1740  
 atcctagcaa gattttaaaa ggaagtatat gaaagaagaa cctcagtggc aatcctaac 1800  
 cttttatatt tctctacagg ggccggcgtt ggggacaatt cacgcgtctg tgaggggagc 1860  
 gtttccctgc tcgcaggtct gcagcgagga gccgtaattt ttgcttcgag cctgcccggc 1920  
 atcaaatgt atggatgcaa atgattatac atggggatgt atgggctaaa tgtacgggag 1980  
 acagtcacat catgccctg agctgcgcac gtcaagactg tcaaggaggg tattctgggc 2040  
 ctccatgtcg ctggccgggt gaccggcggg ggacgaggcc ttaagttcga acgtacggag 2100  
 ttgttgaatg gccaatccgc tcaaggttta attacctgca aagaagttca cgaatggtt 2160  
 gaaacatgtg gctctgtcga agacttccca ttatttgaag ccgtatacca aatcgtttac 2220  
 aacaactacc caatgaagaa cctgccggac atgattgaag aattagatct acatgaagat 2280  
 tagatttatt ggagaaagat aacatatcat actttcccc acttttttcg aggctcttct 2340  
 atatcatatt cataaattag cattatgtca tttctcataa ctactttatc acgttagaaa 2400  
 ttacttatta ttattaaatt aatacaaat ttagtaacca aataaatata aataaatatg 2460  
 tatatttaa ttttaaaaa aaaatcctat agagcaaaa gattttccat tataatatta 2520  
 gctgtacacc tcttccgcat ttttgaggg tggttacaac accactcggg acc 2573

<210> 18

<211> 2058

5

<212> ADN

<213> Artificial

<220>

<223> Rotura de GPD2

10

<400> 18

ES 2 648 865 T3

agatcttttg cggcgagggtg cccgatgggtt gctgagggga agagtgttta gcttacggac 60  
 ctattgccat tgttattccg attaactctat tgttcagcag ctcttctcta ccctgtcatt 120  
 ctagtatttt tttttttttt ttttggtttt actttttttt cttcttgect ttttttcttg 180  
 ttactttttt tctagttttt tttccttcca ctaagctttt tccttgattt atccttgggt 240  
 tcttctttct actccttttag atttttttt tataatataa tttttaagtt tatgtatttt 300  
 ggtagattca attctcttcc cctttccttt tccttcgctc cccttcctta tcaatgcttg 360  
 ctgtcagaag attaacaaga tacacattcc ttaaggcctc gtccccgccg ggtcaccccg 420  
 ccagcgacat ggaggcccag aataccctcc ttgacagtct tgacgtgccc agctcagggg 480  
 catgatgtga ctgtcgcccg tacatcttagc ccatacatcc ccattgtataa tcatttgcatt 540  
 ccatacattt tgatggccgc acggcgcgaa gcaaaaatta cggctcctcg ctgcagacct 600  
 gcgagcaggg aaacgctccc ctcacagacg cgttgaattg tccccacgcc gcgcccctgt 660  
 agagaaatat aaaagggttag gatttgccac tgaggttctt ctttcatata ctccctttta 720  
 aaatcttgcct aggatacagt tctcacatca catccgaaca taaacaacca tgtaaaatga 780  
 ccactcttga cgacacggct taccggtagc gcaccagtgt cccgggggac gccgagggca 840  
 tcgaggcact ggatgggtcc ttcaccaccg acaccgtctt ccgcgtcacc gccaccgggg 900  
 acggcttcac cctgcgggag gtgcccgttg acccgcccct gaccaaggtg ttccccgacg 960  
 acgaatcggg cgacgaatcg gacgccgggg aggacggcga cccggactcc cggacgttcc 1020  
 tcgcgtacgg ggacgacggc gacctggcgg gcttcgtggt cgtctcgtac tccggctgga 1080  
 accgcccgtt gaccgtcagc gacatcgagg tcgccccgga gcaccggggg cacggggctc 1140  
 ggcgcgcgct gatggggctc gcgacggagt tcgcccgcga gcggggcgcc gggcacctct 1200  
 ggctggagggt caccaacgtc aacgcaccgg cgatccacgc gtaccggcgg atggggttca 1260  
 ccctctgcgg cctggacacc gccctgtacg acggcaccgc ctcgacggc gagcagggc 1320  
 tctacatgag catgccctgc ccctagtact gacaataaaa agattcttgt tttcaagaac 1380  
 ttgtcatttg tatagttttt ttatattgta gttgttctat tttaatcaaa tgttagcgtg 1440  
 atttatattt tttttcgcct cgacatcctc tgcccagatg ogaagttaag tgcgcagaaa 1500  
 gtaatatcat gcgtcaatcg tatgtgaatg ctggtcgcta tactgctgct gattcgatac 1560  
 taacgcccgc atccagtgtc gacggatcct aggtgtacag ggccccaaaag ggcgaattct 1620  
 gcagatatcc atcacactgg cggccgctcg aggatagtct acaacaacgt ccgcatggaa 1680  
 gacctaccgg agatgattga agagctagac atcgatgacg aatagacact ctccccccc 1740  
 ctccccctct gatctttcct gttgcctctt tttccccca ccaatttatc attatacaca 1800  
 agttctacaa ctactactag taacattact acagttatta taattttcta ttctcttttt 1860  
 cttaagaat ctatcattaa cgtaatttc tataatatac taactacatc tatacacgct 1920  
 attatcgttt acatatcaca tcaccgttaa tgaaagatac gacaccctgt aactaacac 1980  
 aattaaataa togccataac cttttctggt atctatagcc cttaaagctg tttcttcgag 2040  
 ctttttcaact gcagatct 2058

ES 2 648 865 T3

<210> 19  
 <211> 816  
 <212> ADN  
 <213> Saccharomyces cerevisiae

5

<400> 19

```

ccatatgatac atgtgtcgtc gcacacatat atatatgcct gtatgtgtca gcactaaagt      60
tgcctggcca tccacgctat atatacacgc ctggcgggatac tgctcgagga ttgcctacgc      120
gtgggcttga tccaccaacc aacgctcgcc aatgaactg gcgctttggt cttctgccat      180
cgtccgtaaa ccccgcccaa agagaccgga aagatcggtg aaaacatctt gatcttgctc      240
ccgggaattt tagattcagg taggaaattg attacatcaa tactgttacc ctgaatcata      300
ttcgacgatg tcgtctcaca cggaatatata attcatttct tggttttcca aaaaaatttt      360
catttttttt cacttttttg tttcgtcttc cttttttttt tttttttttt attttttttc      420
ctgtgttcac cttttttttt ttcagttgac atctttctgc attcttttct gtgttttttt      480
tttttttttt cgtttttcca ttgttcgttc gttgcctggt ttttcgccct attgttctcg      540
agcctaaaaa ttttttcctt tcctgctttc ctttcttcgt tcaaagtttc ctattccatt      600
gttctctttg gtaaaactcat tgttgcgga actcagatat attcaggtca atttactgta      660
cttcaattga cttttttctt gaaatttcaa cttgcctttt caactgttc ttctttttta      720
atcttattct acactttagt tcccttacct tgttcctaata tattgtctag caaaaagaaa      780
acatacacct atttcattca cacactgcag aaaatg                                     816
    
```

10 <210> 20  
 <211> 12350  
 <212> ADN  
 <213> Artificial

15 <220>  
 <223> pRN500

<400> 20

# ES 2 648 865 T3

tcgatgataa gctgtcaaag atgagaatta attccacgga ctatagacta tactagatac	60
tccgtctact gtacgataca cttccgctca ggtccttgtc ctttaacgag gcottaccac	120
tcttttgta ctctattgat ccagctcagc aaaggcagtg tgatctaaga ttctatcttc	180
gcgatgtagt aaaactagct agaccgagaa agagactaga aatgcaaaag gcacttctac	240
aatggctgcc atcattatta tccgatgtga cgctgcagct tctcaatgat attcgaatac	300
gctttgagga gatacagcct aatatccgac aaactgtttt acagatttac gatcgtactt	360
gttaccocatc attgaatfff gaacatccga acctgggagt tttccctgaa acagatagta	420
tatttgaacc tgtataataa tatatagtct agcgctttac ggaagacaat gtatgtatff	480
cggttcctgg agaaactatt gcatctattg cataggtaat cttgcacgtc gcatccccgg	540
ttcattttct gcgtttccat cttgcacttc aatagcatat ctttgttaac gaagcatctg	600
tgcttcattt tgtagaacia aaatgcaacg cgagagcgct aatffffcaa acaagaatc	660
tgagctgcat ttttacagaa cagaaatgca acgcgaaagc gctatffffac caacgaagaa	720

ES 2 648 865 T3

tctgtgcttc atttttgtaa aacaaaaatg caacgcgacg agagcgctaa tttttcaaac 780  
aaagaatctg agctgcattt ttacagaaca gaaatgcaac gcgagagcgc tattttacca 840  
acaaagaatc tatacttctt ttttgttcta caaaaatgca tcccagagagc gctatttttc 900  
taacaaagca tottagatta ctttttttct cctttgtgog ctctataatg cagtctcttg 960  
ataacttttt gcactgtagg tccgttaagg ttagaagaag gctactttgg tgtctatttt 1020  
ctcttccata aaaaaagcct gactccactt cccgcgttta ctgattacta gcgaagctgc 1080  
gggtgcattt tttcaagata aaggcatccc cgattatatt ctataccgat gtggattgcg 1140  
catactttgt gaacagaaaag tgatagcgtt gatgattctt cattgggtcag aaaattatga 1200  
acggtttctt ctattttgc tctatatact acgtatagga aatgtttaca ttttcgtatt 1260  
gttttcgatt cactctatga atagttctta ctacaatfff tttgtctaaa gagtaatact 1320  
agagataaac ataaaaaatg tagaggtcga gtttagatgc aagttcaagg agcgaaggt 1380  
ggatgggtag gttatatagg gatatagcac agagatata agcaaagaga tacttttgag 1440  
caatgtttgt ggaagcggta ttgcgaatgc cggctttccc cgtcaagctc taaatcgggg 1500  
gctcccttta gggttccgat ttagtgcttt acggcacctc gacccccaaa aacttgatta 1560  
gggtgatggg tcacgtagtg ggccatcgcc ctgatagacg gtttttcgcc ctttgacgtt 1620  
ggagtccacg ttctttaata gtggactctt gttccaaact ggaacaacac tcaaccctat 1680  
ctcggctctat tcttttgatt tataagggat tttgccgatt tcggcctatt ggttaaaaaa 1740  
tgagctgatt taacaaaaat ttaacggaa ttttaacaaa atattaacgt ttacaatttc 1800  
ctgatcgggg attttctct tacgcatctg tgcggtatff cacaccgag ggtaataact 1860  
gatataatta aattgaagct ctaatttgag agtttagtat acatgcattt acttataata 1920  
cagtttttta gttttgctgg ccgcatcttc tcaaataatgc tcccagcct gcttttctgt 1980  
aacgttcacc ctctacctta gcaccccttc cctttgcaa tagtctctt ccaacaataa 2040  
taatgtcaga tcctgtagag accacatcat ccacggttct atactgttga cccaatgctt 2100  
ctccctgtc atctaaaacc acaccgggtg tcataatcaa ccaatcgtaa ccttcatctc 2160  
ttccacccat gtctctttga gcaataaagc cgataacaaa atctttgtcg ctcttcgcaa 2220  
tgtcaacagt acccttagta tattctccag tagatagga gcccttgcac gacaattctg 2280  
ctaacatcaa aaggcctcta ggttctttg ttacttcttc tgcgcctgc ttcaaacgc 2340  
taacaatacc tgggccacc acaccgtgtg cattcgtaat gtctgccat tctgtatctc 2400  
tgtatacacc cgcagagtac tgcaatttga ctgtattacc aatgtcagca aattttctgt 2460  
cttcgaagag taaaaaattg tacttgccgg ataatgcctt tagcggctta actgtgcct 2520  
ccatggaaaa atcagtaag atateccat gtgttttag taaacaaatt ttgggacct 2580  
atgcttcaac taactccagt aattccttgg tggtagaac atccaatgaa gcacacaagt 2640

ES 2 648 865 T3

ttgtttgctt ttcgtgcatg atattaata gcttggcagc aacaggacta ggatgagtag 2700  
 cagcacgttc cttatatgta gctttcgaca tgatttatct tcgtttcctg caggtttttg 2760  
 ttctgtgcag ttgggttaag aatactgggc aatttcatgt ttcttcaaca ctacatatgc 2820  
 gtatatatac caatctaagt ctgtgctcct tccttcgttc ttccttctgt tcggagatta 2880  
 ccgaatcaaa aaaatttcaa agaaaccgaa atcaaaaaaa agaataaaaa aaaatgatg 2940  
 aattgaattg aaaagcgtgg tgcactctca gtacaatctg ctctgatgcc gcatagtaa 3000  
 gccagccccg acaccgccca acaccgctg acgcgccctg acgggcttgt ctgctcccgg 3060  
 catccgctta cagacaagct gtgaccgtct ccgggagctg catgtgtcag aggttttcac 3120  
 cgtcatcacc gaaacgcgcg agacgaaagg gcctcgtgat acgcctatct ttataggtta 3180  
 atgtcatgat aataatggtt tcttagacgt cagggtggcac ttttcgggga aatgtgcgcg 3240  
 gaaccctat ttgtttatct ttctaaatac attcaaatac gtatccgctc atgagacaat 3300  
 aaccctgata aatgcttcaa taatattgaa aaaggaagag tatgagtatt caacatttcc 3360  
 gtgtcgcctt tattcccttt tttgcggcat tttgccttcc tgtttttgct caccagaaa 3420  
 cgctggtgaa agtaaaagat gctgaagatc agttgggtgc acgagtgggt tacatcgaac 3480  
 tggatctcaa cagcggtaag atccttgaga gttttcgccc cgaagaacgt tttccaatga 3540  
 tgagcacttt taaagttctg ctatgtggcg cggattatc ccgtattgac gccgggcaag 3600  
 agcaactcgg tcgccgata cactattctc agaatgactt ggttgagtac tcaccagtca 3660  
 cagaaaagca tcttacggat ggcatgacag taagagaatt atgcagtgct gccataacca 3720  
 tgagtgataa cactgcggcc aacttacttc tgacaacgat cggaggaccg aaggagctaa 3780  
 ccgctttttt tcacaacatg ggggatcatg taactcgctt tgatcgttgg gaaccggagc 3840  
 tgaatgaagc cataccaaac gacgagcgtg acaccacgat gcctgtagca atggcaacaa 3900  
 cgttgcgcaa actattaact ggcgaactac ttactctagc ttcccggcaa caattaatag 3960  
 actggatgga ggcggataaa gttgcaggac cacttctgcg ctcgccctt ccggtggct 4020  
 ggtttattgc tgataaatct ggagccgggt agcgtgggtc tcgcggtatc attgcagcac 4080  
 tggggccaga tggttaagccc tcccgtatcg tagttatcta cacgacgggc agtcaggcaa 4140  
 ctatggatga acgaaataga cagatcgctg agataggtgc ctactgatt aagcattggt 4200  
 aactgtcaga ccaagtttac tcatatatac tttagattga tttaaaactt catttttaat 4260  
 ttaaaaggat ctaggatgaag atcctttttg ataatctcat gaccaaaatc ccttaacgtg 4320  
 agttttcgtt cactgagcg tcagaccocg tagaaaagat caaaggatct tcttgagatc 4380  
 cttttttctt gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta ccagcgggtg 4440  
 tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc ttcagcagag 4500

ES 2 648 865 T3

cgcagatacc aaatactgtc cttctagtgt agccgtagtt aggccaccac ttcaagaact 4560  
 ctgtagcacc gcctacatac ctogctctgc taatcctgtt accagtggct gctgccagtg 4620  
 gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat aaggcgcagc 4680  
 ggtcgggctg aacgggggggt tcgtgcacac agcccagett ggagcgaacg acctacaccg 4740  
 aactgagata cctacagcgt gagcattgag aaagcggcac gcttcccga gggagaaagg 4800  
 cggacaggta tccgtaagc ggcagggtcg gaacaggaga gcgcacgagg gagcttccag 4860  
 gggggaacgc ctggtatcct tatagtcctg tcgggtttcg ccacctctga cttgagcgtc 4920  
 gatttttgtg atgctcgtca ggggggcca gcctatggaa aaacgccagc aacgcggcct 4980  
 ttttacgggt cctggccttt tgctggcctt ttgctcacat gttctttcct gcgttatccc 5040  
 ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct cgcgcagcc 5100  
 gaacgaccga gcgcagcagc tcagtgagcg aggaagcggg agagcgccca atacgaaaac 5160  
 cgcctctccc cgcgcgttgg ccgattcatt aatgcagctg gcacgacagc tttcccgact 5220  
 ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta cctcactcat taggcacccc 5280  
 aggctttaca ctttatgctt ccggctccta tgttgtgtgg aattgtgagc ggataacaat 5340  
 ttcacacagg aaacagctat gaccatgatt acgccaagct cggaaattaac cctcactaaa 5400  
 gggaaacaaa gctgggtacc gggccccccc tcgaccggtc ccacacacca tagcttcaaa 5460  
 atgtttctac tcctttttta ctcttcaga ttttctcggg ctccgcgcac cgcctacca 5520  
 cttcaaaaaca cccaagcaca gcatactaaa tttcccctct ttcttctct aggggtgctg 5580  
 taattaccgc tactaaaggt ttggaaaaga aaaaagagac cgcctcgttt ctttttcttc 5640  
 gtcgaaaaag gcaataaaaa tttttatcac gtttcttttt cttgaaaatt ttttttttg 5700  
 attttttct ctttcgatga cctcccattg atatttaagt taataaacgg tcttcaattt 5760  
 ctcaagtttc agtttcattt ttcttgttct attacaactt tttttacttc ttgctcatta 5820  
 gaaagaaagc atagcaatct aatctaagct gcagaagctt aaaatgtcgc cctctgccgt 5880  
 acaatcatca aaactagaag aacagtcaag tgaaattgac aagttgaaag caaaaatgtc 5940  
 ccagctctgcc gccactgcgc agcagaagaa ggaacatgag tatgaacatt tgacttcggg 6000  
 caagatcgtg ccacaacggc ccatctcaga tagactgcag cccgcaattg ctaccacta 6060  
 ttctccacac ttggacgggt tgcaggacta tcagcgttg cacaaggagt ctattgaaga 6120  
 ccctgctaag ttcttcgggt ctaaagetac ccaatttta aactggctca agccattcga 6180  
 taagggtgtc atcccagacc ctaaaaacggg caggccctcc ttccagaaca atgcatggtt 6240  
 cctcaacggc caattaaacg cctgttaca ctgtgttgac agacatgcct tgaagactcc 6300  
 taacaagaaa gccattattt tcgaaggtga cgagcctggc caaggetatt ccattaccta 6360  
 caaggaacta cttgaagaag tttgtcaagt ggcacaagtg ctgacttact ctatgggctg 6420

ES 2 648 865 T3

tcgcaagggc gatactgttg ccgtgtacat gcctatggtc ccagaagcaa tcataacctt 6480  
 gttggccatt tcccgtatcg gtgccattca ctccgtagtc tttgccgggt tttcttccaa 6540  
 ctccctgaga gatcgtatca acgatgggga ctctaaagt gtcacacta cagatgaatc 6600  
 caacagagggt ggtaaagtca ttgagactaa aagaattggt gatgacgcgc taagagagac 6660  
 cccagcgtg agacacgtct tggtttatag aaagaccaac aatccatctg ttgctttcca 6720  
 tgccccaga gatttgatt gggcaacaga aaagaagaaa tacaagacct actatccatg 6780  
 cacaccggt gattctgagg atccattatt ctgttgat acgtctggtt ctactggtgc 6840  
 ccccaagggt gttcaacatt ctaccgcagg ttaactgctg ggagcttgtg tgaccatgcg 6900  
 ctacactttt gacactcacc aagaagacgt tttcttcaca gctggagaca ttggctggat 6960  
 tacaggccac acttatgtgg tttatggtcc ctactatat ggttggtcca ctttggctctt 7020  
 tgaagggact cctgcgtacc caaattactc ccgttattgg gatattattg atgaacacaa 7080  
 agtcacccaa ttttatgttg cgccaactgc tttgcgtttg ttgaaaagag ctgggtgattc 7140  
 ctacatcgaa aatcattcct taaaatcttt gcgttgcttg ggttcggtcg gtgagccaat 7200  
 tgctgctgaa gtttgggagt ggtactctga aaaaataggt aaaaatgaaa tccccattgt 7260  
 agacacctac tggcaaacag aatctggttc gcatctggtc accccgctgg ctgggtggtg 7320  
 tacaccaatg aaaccgggtt ctgcctcatt ccccttcttc ggtattgatg cagttgttct 7380  
 tgaccctaac actggtgaag aacttaacac cagccacgca gaggggtgcc ttgccgcaa 7440  
 agctgcatgg ccacatttg caagaactat ttggaaaaat catgataggat atctagacac 7500  
 ttatttgaac ccttaccctg gctactattt cactggtgat ggtgctgcaa aggataagga 7560  
 tggttatatac tggatttttg gtcgtgtaga cgatgtggtg aacgtctctg gtcaccgtct 7620  
 gtctaccgct gaaattgagg ctgctattat cgaagatcca attgtggccg agtgtgctgt 7680  
 tgtcggattc aacgatgact tgactggtca agcagttgct gcatttggg tgttgaaaaa 7740  
 caaatctagt tggtcaccg caacagatga tgaattacaa gatatcaaga agcatttgggt 7800  
 ctttactggt agaaaagaca tcgggccatt tgccgcacca aaattgatca ttttagtgga 7860  
 tgacttgccc aagacaagat ccggcaaaat tatgagacgt attttaagaa aaatcctagc 7920  
 aggagaaagt gaccaactag gcgacgtttc tacattgtca aaccctggca ttgtagaca 7980  
 tctaattgat tcggtcaagt tgtaacttaa gcgcgcgaat ttcttatgat ttatgatatt 8040  
 tattattaaa taagttataa aaaaaataag tgtatacaaa ttttaaagt actcttaggt 8100  
 tttaaaacga aaattcttat tcttgagtaa ctcttctctg taggtcaggt tgctttctca 8160  
 ggtatagcat gaggtcgtc ttattgacca cacctctacc ggcacgcca gcaaatgcct 8220  
 gcaaatcgt ccccatctca cccaattgta gatatgctaa ctccagcaat gagttgatga 8280

ES 2 648 865 T3

atctcgggtgt gtatthttatg tcttcagagg acaacacctg ttgtaatcgt tcttccacac 8340  
 gtacgthttta aacagttgat gagaaccttt ttogcaagtt caaggtgctc taatthtttaa 8400  
 aatthtttact tttcgcgaca caataaagtc ttcacgacgc taaactatta gtgcacataa 8460  
 tgtagtact tggacgctgt tcaataatgt ataaaattta tttcctttgc attacgtaca 8520  
 ttatataacc aaatcttaaa aatatagaaa tatgatatgt gtataataat ataagcaaaa 8580  
 tttacgtatc tttgcttata atatagcttt aatgthcttt aggtatata ttaagagcga 8640  
 tttgtctcga cttatthctt tthttgagag aaaaattggt tctctacagc agaatgatg 8700  
 gcaggtacaa cttctgggtt ggccaaagta gttaggtcac ctagctgttc ggcttogtta 8760  
 gaagcaacct ttcttagaac tcttctcata atctthctg acctgttct tggtagatct 8820  
 ctaactagaa taatgthttt tggtgaggcg aaaggaccaa tctcaccct aacttgtaag 8880  
 atcaatctc tacgtaaatt atctggtgtg atgtgttctg catcacctc agtagcgtt 8940  
 tthtttagat aacctcttt tagggaaca tatgcaacga cgtthtgacc ggtcaattca 9000  
 tctggaatac cgacaacagc agcttccgag acgthttctg gatttgagat agatgcttca 9060  
 atthctgatg tggataatct atgaccgga acatthcaa cgtcgtcaac tctaccctg 9120  
 atccagtagt aacctcatg atctctacca gcacctcac ctgtgaaata gtgaccagga 9180  
 taagthttca agtaagtatc catgtaacgg tctggtggt tccaaacaga tctagccatt 9240  
 gatggccatg gtgatttaac ggcaaggaca ccttcgacat cattacctc taatccaca 9300  
 cctgtaacag ggtcaatgat acaagcgtta ataccaaaga atggcacggt agcagaacca 9360  
 ggtthttgtt ggacagcacc tgccaaagga gcaatthaaat gagaaccaga ctctgthttgc 9420  
 cacatagtgt cacaaatgac acagthtttg ttaccactt tthctatgata ccattcccat 9480  
 aagtctggag agattggttc accgacgga cccaagacac gtaatgagga agtgcataat 9540  
 ttggcaatth cggcttcacc tacacgtthg atthaatctta aagcagthtg agccacatag 9600  
 aatgggttag ccttgtagc tgggataatt ctccaatata taccataatc tgggtaggca 9660  
 ggagtggatt cgaaaattat tgaggcgtta cccaagthta atggaccata tagagcatag 9720  
 gtgtgaccog tgatccagcc gacgtcacog gcagtgaaga gaacatctc tgggtgaata 9780  
 tcaaaaacgt atctagthgt taaagcggca cctaataaat aaccacctgt agtgtgaacg 9840  
 acacctthg gagaaccagt ggaaccgga gtgtataata aaaatagagg atcttcagcg 9900  
 tcacatgaaa caggaggtag gtaagthctc tgcttagcgg cctcctcatg ccaccagtaa 9960  
 tctctaccog ccttcattgg aataccttca gtaccagthc tthggaaaac caagatacgg 10020  
 gaaaccaaath cgactccgth caaaccttgc tcaacaatth tthtagthgt gatgthctta 10080  
 ccacctctth taccttcatc acaagthgat accactthg aatthagcgc aacgacacga 10140  
 tctthcaacg aaccagcggga gaaccagca aagacaacag agtgaatagc accaatacga 10200

ES 2 648 865 T3

gccacagcca acatagcaat gaccgcttct ggaatcattg gcaaatagat agccactgtg 10260  
tcacctttct taacgcccc a gctttttaag acaccagcga tttgggaaac ttttctgagt 10320  
aattcaccaa atgtgatgat tttgttgcg gattcgtcat cagcttcata gatcaaagct 10380  
ggcttgcgg gattagcaaa ggcatgtctg tcaacacaat tgtatgatgc attcaattta 10440  
ccgttcaaaa accatgcaac atcaccattg ttcaatgaac cagattgaac tttgggtgat 10500  
ggagcatccc aatgcaagta ttccttagcc atcttatcaa agaatttttc tggctcattg 10560  
atagattggt gatacatttc ttgataatgt tgcatatcag taacgtaacc cttgccgggt 10620  
tggctgtgtg aaaaatgttg aggagcctta agagccttta cgttgcgagc ttcataaact 10680  
actttatggt ccttgattgt cattttctgc agtctagata tatttgttgt aaaaagtaga 10740  
taattacttc cttgatgatc tgtaaaaaag agaaaaagaa agcatctaag aacttgaaaa 10800  
actacgaatt agaaaagacc aaatatgtat ttcttgcatt gaccaattta tgcaagtta 10860  
tatatatgta aatgtaagtt tcacgaggtt ctactaaact aaaccacccc cttgggttaga 10920  
agaaaagagt gtgtgagaac aggctgttgt tgtcacacga ttcggacaat tctgtttgaa 10980  
agagagagag taacagtacg atcgaacgaa ctttgcctctg gagatcacag tgggcatcat 11040  
agcatgtggt actaaaccct ttcccgccat tccagaacct tgcattgctt gttacaaaac 11100  
ctgtgagccg tcgctaggac cttgttgtgt gacgaaattg gaagctgcaa tcaataggaa 11160  
gacaggaagt cgagcgtgtc tgggtttttt cagttttgtt ctttttgcaa acaaatcacg 11220  
agcgacggta atttctttct cgataagagg ccacgtgctt tatgagggtta acatcaattc 11280  
aagaaggagg gaaacacttc cttttctggt ccctgataat agtatgaggg tgaagccaaa 11340  
ataaaggatt cgcgccccaa tcggcatctt taaatgcagg tatgcgatag ttcctcactc 11400  
tttcttact cacgagtaat tcttgcaaat gcctattatg cagatgttat aatatctgtg 11460  
cgtcttgagt tgaagagctc gagactagat gcatgctcga gcgcccgcca gtgtgatgga 11520  
tatctgcaga attgccectt ttgggcccctg tacacctagg atccgtcgac actggatggc 11580  
ggcgtagta tcgaatcgac agcagtatag cgaccagcat tcacatacga ttgacgcagt 11640  
atattacttt ctgcgcactt aacttcgcat ctgggcagat gatgtcgagg cgaaaaaaaa 11700  
tataaatcac gctaacattt gattaaaata gaacaactac aatataaaaa aactatacaa 11760  
atgacaagtt cttgaaaaca agaattcttt tattgtcagt actgattaga aaaactcatc 11820  
gagcatcaaa tgaactgca atttattcat atcaggatta tcaataccat atttttgaaa 11880  
aagccgtttc tghtaatgaag gagaaaactc accgaggcag ttccatagga tggcaagatc 11940  
ctggtatcgg tctgcgattc cgactcgtcc aacatcaata caacctatta atttcccctc 12000  
gtcaaaaata aggttatcaa gtgagaaatc accatgagtg acgactgaat ccggtgagaa 12060  
tggcaaaagc ttatgcattt ctttccagac ttgttcaaca ggccagccat tacgctcgtc 12120  
atcaaaatca ctgcgatcaa ccaaaccgtt attcattcgt gattgcgcct gagcgagacg 12180  
aaatacgcga tcgctgttaa aaggacaatt acaaacagga atcgaatgca accggcgcag 12240  
gaacactgcc agcgcatcaa caatattttc acctgaatca ggatattctt ctaataacctg 12300  
gaatgctggt ttgccgggga tcgcagtggt gagtaacctat gcatcatcag 12350

ES 2 648 865 T3

<210> 21

<211> 1120

<212> ADN

<213> Artificial

5

<220>

<223> secuencia de nucleótidos que codifica gldA de E. coli optimizado en codones para levadura

<400> 21

10

```

ctgcagaaaa tggacagaat catccaatct ccaggtaagt acatccaagg tgctgacggt      60
atcaacagat tgggtgaata cttgaagcca ttggctgaaa gatggttggg tgttgggtgac     120
aagttcgttt tgggtttcgc tcaatctacc gttgaaaagt ctttcaagga cgctgggttg      180
gttggtgaaa tcgctccatt cgggtggtgaa tgttctcaaa acgaaatoga cagattgaga     240
ggtatcgctg aaaccgctca atgtggtgct atcttgggta tcggtggtgg taagaccttg      300
gacaccgcta aggctttggc tcacttcatg ggtgttccag ttgctatcgc tccaaccatc     360
gcttctaccg acgctccatg ttctgctttg tctgttatct acaccgacga aggtgaattc     420
gacagatact tgttggtgcc aaacaacca aacatgggta tcggtgacac caagatcggt      480
gctggtgctc cagctagatt gttggcagct ggtatcggtg acgctttggc tacctggttc     540
gaagctagag cttgttctag atctggtgct accaccatgg ctggtggtaa gtgtacccaa     600
gctgctttgg ctttggctga attgtgttac aacaccttgt tggagaagg tgaaaaggct      660
atgttggctg ctgaacaaca cgttgttacc ccagctttgg aaagagttat cgaagctaac     720
acctacttgt ctggtgttgg tttcgaatct ggtggtttgg ctgctgctca cgctgttcac     780
aacggtttga ccgctatccc agacgctcac cactactacc acggtgaaaa ggttgctttc     840
ggtaccttga cccaattggt tttggaaaac gctccagttg aagaaatcga aaccgttgct     900
gctttgtctc acgctgttgg tttgccaatc accttggctc aattggacat caaggaagac     960
gttccagcta agatgagaat cgttctgtaa gctgcttctg ctgaaggtga aaccatccac    1020
aacatgccag gtggtgctac cccagaccaa gtttacgctg ctttgttggg tgctgaccaa    1080
tacgggtcaaa gattcctaca agaatgggaa taaggcgcgc      1120

```

<210> 22

<211> 1450

<212> ADN

<213> Artificial

15

<220>

ES 2 648 865 T3

<223> Fragmento de PCR DAK1

<400> 22

tctagaaaat	gtccgctaaa	tcgtttgaag	tcacagatcc	agtcaattca	agtctcaaag	60
ggtttgccct	tgctaacccc	tccattacgc	tggtccctga	agaaaaaatt	ctcttcagaa	120
agaccgattc	cgacaagatc	gcattaatth	ctgggtgtgg	tagtggacat	gaacctacac	180
acgccggtht	cattggtaag	ggtatgttga	gtggcgccgt	ggttggcgaa	atthtttgcac	240
ccccttcaac	aaaacagatt	ttaaatgcaa	tccgthtagt	caatgaaaat	gcgtctggcg	300
thttattgat	tgtgaagaac	tacacaggtg	atgthttgca	thttgtctg	tccgctgaga	360
gagcaagagc	cttgggtatt	aactgcccgc	ttgctgtcat	aggtgatgat	gttgcagttg	420
gcagagaaaa	gggtggtatg	gttggtagaa	gagcattggc	aggtaccgth	ttggttcata	480
agattgttag	tgcttccgca	gaagaatatt	ctagtaagta	tggcttagac	ggtacagcta	540
aagtggctaa	aattatcaac	gacaatthtg	tgaccattgg	atcttctth	gaccattgta	600
aagthctcgg	caggaaaatc	gaaagtgaat	taaacgaaaa	acaaatggaa	ttgggtatgg	660
gtattcataa	cgaacctggt	gtgaaagtht	tagaccctat	tccttctacc	gaagacttga	720
tctccaagta	tatgctacca	aaactatthg	atccaaacga	taaggataga	gctthttgtaa	780
agthtgatga	agatgatgaa	gthgtcttgt	tagthtaacaa	tctcggcggth	gthttctaatt	840
thgttattag	thctatcact	tccaaaacta	cggatthctt	aaaggaaaat	tacaacataa	900
ccccgthtca	aacaatthgct	ggcacattga	tgacctctt	caatgthaat	gggttcagta	960
tcacattact	aaacgccact	aaggctacaa	aggctthgca	atctgathth	gaggagatca	1020
aatcagthact	agactthgtg	aacgcathth	cgaacgcacc	gggctggcca	attgcagatt	1080
thgaaaagac	thctgccccca	tctgttaacg	atgactthgt	acataatgaa	gtaacagcaa	1140
aggccgthcg	tacctatgac	thtgacaagt	thgctgagth	gatgaagagth	ggtgctgaac	1200
aagthtatcaa	gagcgaaccg	cacattacgg	aactagacaa	tcaagthggth	gatggthgatt	1260
gtgthtacac	thtagthggca	ggagthaaag	gcatcaccga	aaacctgac	aagctgthcga	1320
aggactcatt	atctcagggc	gthgccccaaa	thtcagathth	cathggaaggc	tcaatgggag	1380
gtactthctgg	tggthttatat	tctattctth	tgtcgggtht	thcacacggga	thaatthcagg	1440
5	thtgthaaatc					1450

<210> 23

<211> 1698

<212> ADN

10 <213> Artificial

<220>

ES 2 648 865 T3

<223> Fragmento de PCR GUP1

<400> 23

```

gaattcaaaa tgtcgctgat cagcatcctg tctcccctaa ttacttccga gggcttagat      60
tcaagaatca aaccttcacc aaaaaaggat gcctctacta ccactaagcc atcaactatgg      120
aaaactactg agttcaaatt ctactacatt gcatttctgg tcgtggttcc cttgatgttc      180
tatgctgggt tacaagctag ttcgcccgaa aatccaaact atgcaagata cgaacgtctc      240
ctatctcaag gttggttatt tggcagaaaa gtagacaata gtgattctca atataggttt      300
ttcagggaca attttgcgct attgtcagtt ttaatgctag tccacacttc tataaaacgc      360
attgtacttt attcaacaaa tatcactaaa ttgaggtttg atctgatatt tggtttgatc      420
tttttagtgg cgcctcatgg tgtcaattcg ataagaattt tagcccatat gctaatttta      480
tatgccatcg cccatgtact aaagaacttt agaagaatag ccaccatcag catttggatt      540
tatggtatth ctacgctttt tattaacgac aacttcagag catatccatt tggtaatatt      600
tgctcttttt taagcccatt ggaccattgg tatagaggta tcattccaag atgggatgtc      660
tttttcaatt ttactctttt gagagtctta agttacaact tggacttctt agagaggtgg      720
gagaatthac aaaagaagaa aagtccatcc tatgaatcaa aagaagctaa atcagccatt      780
ttgctcaatg aacgtgctag attaactgct gcacacccca tacaggacta cagcttaatg      840
aattatattg catatgttac ttacacgcca cttttcattg cgggccccat tataacattc      900
aatgattatg tttaccaatc gaaacatacc ttgccatcaa taaatttcaa attcattttt      960
tactatgcgg tgagattcgt tattgctctc ttatctatgg agttcatttt acactttctc     1020
cacgttgggt caatctcaaa aaccaaagcg tgggaaaatg acacaccttt ccagatttcc     1080
atgattggct tatttaattt gaatattatt tggctaaaac tactgattcc gtggaggctg     1140
tttaggctgt gggctttgct agacggaatc gatacacctg aaaatatgat cagggtgtgtt     1200
gataacaatt acagttcact agcattctgg agagcttggc atagaagcta caataagtgg     1260
gttgtccgth acatataat tctcttaggt ggttcaaaaa atagagtttt gacatcacta     1320
gcagtctttt ccttcgtagc tatatggcat gacatcgaac taaagttatt attatggggg     1380
tggctaatag ttttgttcct cttaccagaa atttttgcta cccaaatttt ctctcattat     1440
accgacgcag tctggtacag acacgthtgc gctgtcggtg ctgttttcaa catatgggth     1500
atgatgatcg ctaatctttt tggattctgc ttgggctctg acggtactaa aaaattacta     1560
agcgatatgt tctgtaccgt atctggthtc aaatttgtaa ttttggcaag cgttagthta     1620
ttcatcgcag tacaataaat gthtgaaatc agagaagaag aaaagaggca cggaaattac     1680
ctaaaatgct gaggatcc

```

5

<210> 24

<211> 3037

ES 2 648 865 T3

<212> ADN

<213> Artificial

<220>

5 <223> Fragmento de PCR FPS1

<400> 24

```

aagagctccg gactagtcgt acgaattcta tccttttggt gtttccgggt gtacaatatg      60
gacttcctct tttctggcaa ccaaaccocat acatcgggat tcctataata ccttoggttg      120
tctccctaac atgtaggtgg cggagggggag atatacaata gaacagatac cagacaagac      180
ataatgggct aaacaagact acaccaatta cactgcctca ttgatggtgg tacataacga      240
actaatactg tagccctaga cttgatagcc atcatcatat cgaagtttca ctaccctttt      300
tccatttgcc atctattgaa gtaataatag gcgcatgcaa cttcttttct ttttttttct      360
tttctctctc ccccgttggt gtctcaccat atccgcaatg acaaaaaaat gatggaagac      420
actaaaggaa aaaattaacg acaaagacag caccaacaga tgtcgttggt ccagagctga      480
tgaggggtat ctcgaagcac acgaaacttt ttccttcctt cattcacgca cactactctc      540
taatgagcaa cggatacggc ccttccttcc agttacttga atttgaata aaaaaagttt      600
gctgtcttgc tatcaagtat aaatagacct gcaattatta atcttttggt tcctogtcat      660
tgttctcgtt ccctttcttc cttgtttctt tttctgcaca atatttcaag ctataccaag      720
catacaatca actccagctg cattaaaatg agtaatcttc aaaaagctct aaacgacttt      780
ctgtccagtg aatctgttca tacacatgat agttctagga aacaatctaa taagcagtca      840
tccgacgaag gacgctcttc atcacaacct tcacatcatc actctggtgg tactaacaac      900
aataataaca ataataataa taataataac agtaacaaca acaacaacgg caacgatggg      960
ggaatgatg acgactatga ttatgaaatg caagattata gaccttctcc gcaaagtgcg     1020
cggcctactc ccacgtatgt tccacaatat tctgtagaaa gtgggactgc tttcccgatt     1080
caagaggtta ttocctagcgc atacattaac acacaagata taaaccataa agataacggt     1140
ccgccgagtg caagcagtaa tagagcattc aggcctagag ggcagaccac agtgtcggcc     1200
aacgtgctta acattgaaga tttttacaaa aatgcagacg atgcgcatac catcccggag     1260
tcacatttat cgagaaggag aagtaggtcg agggctacga gtaatgctgg gcacagtgcc     1320
aatacaggcg ccacgaatgg caggactact ggtgcccaa ctaatatgga aagcaatgaa     1380
tcaccacgta agtccccat tatggtgaag ccaaagacat tataccagaa ccctcaaaca     1440
cctacagtct tgccctccac ataccatcca attaataaat ggtcttccgt caaaaacact     1500
tatttgaagg aattttttagc cgagtttatg ggaacaatgg ttatgattat tttcggtagt     1560
gctgttgttt gtcaggtaa tggtgctggg aaaatacagc aggacaatth caacgtggct     1620
ttggataacc ttaacgttac cgggtcttct gcagaaacga tagacgctat gaagagttta     1680

```

ES 2 648 865 T3

```

acatccttgg tttcatccgt tgcgggoggt acctttgatg atgtggcatt gggctgggct 1740
gctgccgtgg tgatgggcta tttctgcgct ggtggtagtg ccatctcagg tgctcatttg 1800
aatccgtcta ttacattagc caatttggtg tatagagggt tccccctgaa gaaagttcct 1860
tattactttg ctggacaatt gatcggtgcc ttcacaggcg ctttgatcct gtttatttgg 1920
tacaaaaggg tgttacaaga ggcatatagc gattgggtga tgaatgaaag tgttgcgggga 1980
atgttttgcg tttttccaaa gccttatcta agttcaggac ggcaattttt ttccgaattt 2040
ttatgtggag ctatgttaca agcaggaaca tttgcgctga ccgatcctta tacgtgtttg 2100
tcctctgatg ttttcccatt gatgatgttt attttgattt tcattatcaa tgcttccatg 2160
gcttatcaga caggtacagc aatgaatttg gctcgtgatc tgggcccacg tcttgcaacta 2220
tatgcagttg gatttgatca taaaatgctt tgggtgcatc atcatcattt cttttggggt 2280
cccattgtag gccatttat tgggtgcgta atgggggggt tggtttacga tgtctgtatt 2340
tatcagggtc atgaatctcc agtcaactgg tctttaccag tttataagga aatgattatg 2400
agagcctggt ttagaaggcc tggttggaag aagagaaata gagcaagaag aacatcggac 2460
ctgagtgact tctcatacaa taacgatgat gatgaggaat ttggagaaag aatggctcct 2520
caaaaagcaa agaccaagtc atctatttca gacaacgaaa atgaagcagg agaaaagaaa 2580
gtgcaattta aatctgttca gcgcggcaaa agaacgtttg gtggtatacc aacaattcct 2640
gaagaagaag attccattga aactgcttcg ctaggtgcga cgacgactga ttctattggg 2700
ttatccgaca catcatcaga agattcgcac tatggtaatg ctaagaaggt aacatgagga 2760
tccccctttc ctttgtcgat atcatgtaat tagttatgtc acgcttacat tcacgccctc 2820
ctcccacatc cgctctaacc gaaaaggaag gagttagaca acctgaagtc taggtcccta 2880
tttatttttt ttaatagtta tgttagtatt aagaacgta tttatatttc aaatttttct 2940
ttttttctg tacaaacgcg tgtacgcatg taacattata ctgaaaacct tgcttgagaa 3000
ggttttggga cgctcgaagg cttcctaggc tcgagtt 3037

```

<210> 25

<211> 552

5 <212> PRT

<213> Citrobacter freundii

<400> 25

```

Met Ser Gln Phe Phe Phe Asn Gln Arg Thr His Leu Val Ser Asp Val
1           5           10           15

```

```

Ile Asp Gly Thr Ile Ile Ala Ser Pro Trp Asn Asn Leu Ala Arg Leu
20           25           30

```

10

ES 2 648 865 T3

Glu Ser Asp Pro Ala Ile Arg Ile Val Val Arg Arg Asp Leu Asn Lys  
 35 40 45  
 Asn Asn Val Ala Val Ile Ser Gly Gly Gly Ser Gly His Glu Pro Ala  
 50 55 60  
 His Val Gly Phe Ile Gly Lys Gly Met Leu Thr Ala Ala Val Cys Gly  
 65 70 75 80  
 Asp Val Phe Ala Ser Pro Ser Val Asp Ala Val Leu Thr Ala Ile Gln  
 85 90 95  
 Ala Val Thr Gly Glu Ala Gly Cys Leu Leu Ile Val Lys Asn Tyr Thr  
 100 105 110  
 Gly Asp Arg Leu Asn Phe Gly Leu Ala Ala Glu Lys Ala Arg Arg Leu  
 115 120 125  
 Gly Tyr Asn Val Glu Met Leu Ile Val Gly Asp Asp Ile Ser Leu Pro  
 130 135 140  
 Asp Asn Lys His Pro Arg Gly Ile Ala Gly Thr Ile Leu Val His Lys  
 145 150 155 160  
 Ile Ala Gly Tyr Phe Ala Glu Arg Gly Tyr Asn Leu Ala Thr Val Leu  
 165 170 175  
 Arg Glu Ala Gln Tyr Ala Ala Asn Asn Thr Phe Ser Leu Gly Val Ala  
 180 185 190  
 Leu Ser Ser Cys His Leu Pro Gln Glu Ala Asp Ala Ala Pro Arg His  
 195 200 205  
 His Pro Gly His Ala Glu Leu Gly Met Gly Ile His Gly Glu Pro Gly  
 210 215 220  
 Ala Ser Val Ile Asp Thr Gln Asn Ser Ala Gln Val Val Asn Leu Met  
 225 230 235 240  
 Val Asp Lys Leu Met Ala Ala Leu Pro Glu Thr Gly Arg Leu Ala Val  
 245 250 255  
 Met Ile Asn Asn Leu Gly Gly Val Ser Val Ala Glu Met Ala Ile Ile  
 260 265 270  
 Thr Arg Glu Leu Ala Ser Ser Pro Leu His Pro Arg Ile Asp Trp Leu  
 275 280 285

ES 2 648 865 T3

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

Leu Thr Ala Ile Val Leu Glu Glu Ser Ile Glu Lys Ala Leu Leu Thr  
 305 310 315 320

Glu Val Glu Thr Ser Asn Trp Pro Thr Pro Val Pro Pro Arg Glu Ile  
 325 330 335

Ser Cys Val Pro Ser Ser Gln Arg Ser Ala Arg Val Glu Phe Gln Pro  
 340 345 350

Ser Ala Asn Ala Met Val Ala Gly Ile Val Glu Leu Val Thr Thr Thr  
 355 360 365

Leu Ser Asp Leu Glu Thr His Leu Asn Ala Leu Asp Ala Lys Val Gly  
 370 375 380

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

Ser Leu Leu His Arg Gln Gln Leu Pro Leu Asp Asn Leu Ala Thr Leu  
 405 410 415

Phe Ala Leu Ile Gly Glu Arg Leu Thr Val Val Met Gly Gly Ser Ser  
 420 425 430

Gly Val Leu Met Ser Ile Phe Phe Thr Ala Ala Gly Gln Lys Leu Glu  
 435 440 445

Gln Gly Ala Ser Val Ala Glu Ser Leu Asn Thr Gly Leu Ala Gln Met  
 450 455 460

Lys Phe Tyr Gly Gly Ala Asp Glu Gly Asp Arg Thr Met Ile Asp Ala  
 465 470 475 480

Leu Gln Pro Ala Leu Thr Ser Leu Leu Thr Gln Pro Gln Asn Leu Gln  
 485 490 495

Ala Ala Phe Asp Ala Ala Gln Ala Gly Ala Glu Arg Thr Cys Leu Ser  
 500 505 510

Ser Lys Ala Asn Ala Gly Arg Ala Ser Tyr Leu Ser Ser Glu Ser Leu  
 515 520 525

Leu Gly Asn Met Asp Pro Gly Ala His Ala Val Ala Met Val Phe Lys  
 530 535 540

Ala Leu Ala Glu Ser Glu Leu Gly  
 545 550

<210> 26

5 <211> 1678

<212> ADN

ES 2 648 865 T3

<213> Artificial

<220>

<223> secuencia de nucleótidos que codifica dhaK de *C. freundii* optimizado en codones para levadura

5

<400> 26

```

tctagaaaaa tgtctcaatt cttcttcaac cagagaaccc acttggtttc tgacggtatc      60
gacggtgcta tcatcgcttc accatggaac aatttggcta gattggaatc tgaccagct      120
atcagaatcg ttgttagaag agacttgaac aagaacaacg ttgctgttat ctctgggtgt      180
ggttctggtc acgaaccagc tcacgttggc ttcacggta agggatggt gaccgctgct      240
gtttggtg acgttttcgc ttctccatct gttgacgctg ttttgactgc tatccaagct      300
gttaccgggtg aagctggttg tttgtgatc gttagaact acaccggtga cagattgaac      360
ttcggtttgg ctgctgaaaa ggctagaaga ttgggttaca acggtgaaat gttgatcgtt      420
ggtgacgaca tctctttgcc agacaacaag caccacaagag gtatcgctgg taccatcttg      480
gttcacaaga tcgctgggta cttcgctgaa agagggttaca acttagctac cgttttgaga      540
gaagctcaat acgctgcttc taacaccttc tctttgggtg ttgctttgct ttcttctcac      600
ttgccacaag aaaccgacgc tgctccaaga caccaccag gtcacgctga atgggtatg      660
ggtatccacg gtgaaccagg tgcttctggt atcgacacc aaaactctgc tcaagttggt      720
aacttgatgg ttgacaagtt gttggctgct ttgccagaaa cggtagatt ggctggtatg      780
atcaacaact tgggtggtgt ttctgttctg gaaatggcta tcatcaccag agaattggct      840
tcttctccat tgcaactcaag aatcgactgg ttgatcggtc cagcttcttt ggtaaccgct      900
ttggacatga agggtttctc tttgaccgct atcgttttgg aagaatctat cgaaaaggct      960
ttgttgaccg aagttgaaac ctctaactgg ccaaccccag ttccaccaag agaatacacc     1020
tgtgttgttt cttctcacgc ttctgctaga gttgaattcc aaccatctgc taacgctttg     1080
gttgctggta tcggtgaatt ggttaccgct accttgtctg acttgaaac cacttgaac     1140
gctttggacg ctaaggttgg tgacggtgac accggttcta ccttcgctgc tgctgctaga     1200
gaaatcgctt ctttgttgca cagacaacaa ttgccattga acaacttggc taccttgttc     1260
gctttgatcg gtgaaagatt gaccgttgtt atgggtggtt cttctggtgt tttgatgtct     1320
atcttcttca cagctgctgg tcaaaagttg gaacaagggtg ctaacgttgt tgaagctttg     1380
aacaccggtt tggctcaaat gaagttctac ggtggtgctg acgaagggtga cagaaccatg     1440
atcgacgctt tgcaaccagc tttgacctct ttggtggctc aaccaagaa cttgcaagct     1500
gctttogacg ctgctcaagc tgggtctgaa agaacctgtt tgtcttctaa ggctaacgct     1560
ggtagagctt cttacttgtc ttctgaatct ttggtgggta acatggacc aggtgctcaa     1620
agattggcta tggttttcaa ggctttggct gaatctgaat tgggttaata aggtcgac     1678

```

10

ES 2 648 865 T3

<210> 27

<211> 2951

<212> ADN

<213> Artificial

5

<220>

<223> pRNdhak

<400> 27

10

```

ggatccacta gtaacggcgc ccagtggtgct ggaattcgcc cttctcgagc ttaagacgcg      60
tttcttcttc agattccctc atggagaaag tgcggcagat gtatatgaca gactgcgcag      120
ttccaagag actttattca ggcacttcca tgataggcaa gagagaagac ccagagatgt      180
tgttgtccta gttacacatg gtatttattc cagagtattc ctgatgaaat ggttttagatg      240
gacatacgaa gagtttgaat cgtttaccaa tgttcctaac gggagcgtaa tggatgatgga      300
actggacgaa tccatcaata gatacgtcct gaggaccgtg ctacccaaat ggactgattg      360
tgaggggagac ctaactacat agtgtttaa gattacggat atttaactta cttagaataa      420
tgccatTTTT ttgagttata ataatcctac gttagtgtga gcgggattta aactgtgagg      480
acctaatac attcagacac ttctgcggta tcaccctact tattcccttc gagattatat      540
ctaggaaccc atcaggttgg tggaagatta cccgttctaa gacttttcag ctctctctat      600
tgatgttaca cctggacacc cctttctggt catccagttt ttaatcttca gtggcatgtg      660
agattctccg aaattaatta aagcaatcac acaattctct cggataaccac ctccggttgaa      720
actgacaggt ggtttgttac gcatgcta atgcaaaggagc ctatatacct ttggctcggc      780
tgctgtaaca gggaaataaa agggcagcat aatttaggag tttagtgaac ttgcaacatt      840
tactatTTTT ccttcttacg taaatatttt tctttttaat tctaaatcaa tctttttcaa      900
ttttttgttt gtattctttt cttgcttaaa tctataacta caaaaaacac atacataaat      960
ctagaaaaat gtctcaattc ttcttcaacc agagaacca cttggtttct gacgttatcg     1020
acgggtgctat catcgttca ccatggaaca atttggctag attggaatct gaccagccta     1080
tcagaatcgt tgttagaaga gacttgaaca agaacaacgt tgctgttatc tctggtggtg     1140
gttctgggtca cgaaccagct cacgttgggt tcatcggtaa gggatggtg accgctgctg     1200
tttgtggtga cgttttcgct tctccatctg ttgacgctgt tttgactgct atccaagctg     1260
ttaccgggtga agctggttgt ttgtgatcgc ttaagaacta caccggtgac agattgaact     1320

```

ES 2 648 865 T3

tcggtttggc tgctgaaaag gctagaagat tgggttaciaa cgttgaaatg ttgatcgttg 1380  
 gtgacgacat ctctttgcc aacaacaagc acccaagagg taccgctggg accatcttgg 1440  
 ttcaacaagat cgctgggttac ttcgctgaaa gaggttaciaa cttagctacc gttttgagag 1500  
 aagctcaata cgctgcttct aacaccttct ctttgggtgt tgctttgtct tcttgtcact 1560  
 tgccacaaga aaccgacgct gctccaagac accaccagcgc tcacgctgaa ttgggtatgg 1620  
 gtatccacgg tgaaccaggt gcttctgtta tcgacaccca aaactctgct caagttgtta 1680  
 acttgatggg tgacaagttg ttggctgctt tgccagaaac cggtagattg gctgttatga 1740  
 tcaacaactt ggggtgtgtt tctgttctgt aaatggctat catcaccaga gaattggctt 1800  
 cttctccatt gcactcaaga atcgactggg tgatcgggtc agcttctttg gtaaccgctt 1860  
 tggacatgaa gggtttctct ttgaccgcta tcgttttgga agaattctatc gaaaaggctt 1920  
 tgttgaccga agttgaaacc tctaactggc caacccagcgc tccaccaaga gaaatcacct 1980  
 gtgttgtttc ttctcacgct tctgctagag ttgaattcca accatctgct aacgctttgg 2040  
 ttgctgggat cgttgaattg gttaccgcta ccttgtctga cttggaaacc cacttgaacg 2100  
 ctttgacgc taaggttggg gacggtgaca ccggttctac cttcgctgct gctgctagag 2160  
 aaatcgcttc tttgttgac agacaacaat tgccattgaa caacttggct accttgttcg 2220  
 ctttgatcgg tgaagattg accgttgta tgggtgggtc ttctgggtgt ttgatgtcta 2280  
 tcttctcac cgctgctggg caaaagtgg aacaagggtc taacgttgtt gaagctttga 2340  
 acaccggtt gggtcaaatg aagttctacg gtgggtgctga ogaagggtgac agaaccatga 2400  
 tcgacgctt gcaaccagct ttgacctt ttgtggctca accaaagaac ttgcaagctg 2460  
 cttcgacgc tgctcaagct ggtgctgaaa gaacctgtt gtcttctaag gctaacgctg 2520  
 gttagacttc ttacttctct tctgaatctt tgttgggtaa catggacca ggtgctcaaa 2580  
 gatggctat ggttttcaag gctttggctg aatctgaatt gggtaataa ggtcgagaca 2640  
 aatcgctctt aaatatatac ctaaagaaca ttaaagctat attataagca aagatacgt 2700  
 aattttgctt atattattat acacatatca tatttctata tttttaagat ttggttatat 2760  
 aatgtacgta atgcaaagga aataaattt atacattatt gaacagcgtc caagtaacta 2820  
 cattatgtgc actaatagtt tagcgtcgtg aagactttat tgtgtcgcga aaagtaaaaa 2880  
 ttttaaaaat tagagcacct tgaacttgcg aaaaagggtc tcatcaactg tttaaaacgt 2940  
 acgaagctta a 2951

<210> 28

<211> 1949

5

<212> ADN

<213> Artificial

<220>

<223> pRNglDA

ES 2 648 865 T3

<400> 28

actagtaaat gtgtggggaa gcgggtaagc tgccacagca attaatgcac aacatttaac 60  
ctacattcct ccttatcgga tcctcaaaac ccttaaaaac atatgcctca ccctaacata 120  
ttttccaatt aaccctcaat atttctctgt caccocggcct ctattttcca ttttcttctt 180  
taccocgccac gogttttttt ctttcaaatt tttttcttcc ttcttctttt tcttocacgt 240  
cctcttgcat aaataaataa accgttttga aaccaaaactc gcctctctct ctcccttttg 300  
aaatattttt gggtttgttt gatcctttcc ttccaatct ctcttgttta atatatattc 360  
atztatatca cgctctcttt ttatcttcc ttttttctc tctcttgat tcttccctcc 420  
cctttctact caaaccaaga agaaaaagaa aagggtcaatc tttgttaaag aataggatct 480  
tctactacat cagcttttag atttttcacc cttactgctt ttttcttccc aagatcgaaa 540  
atttactgaa ttaactgcag aaaatggaca gaatcatcca atctccaggt aagtacatcc 600  
aagggtgctga cgttatcaac agattgggtg aatacttgaa gccattggct gaaagatggt 660  
tggttggttg tgacaagttc gttttgggtt tcgctcaatc taccgttgaa aagtctttca 720  
aggacgctgg tttgggttgt gaaatcgctc cattcggtgg tgaatgttct caaacgaaa 780  
tcgacagatt gagaggtatc gctgaaaccg ctcaatgtgg tgctatcttg ggtatcggtg 840  
gtggtaagac cttggacacc gctaaggctt tggtcactt catgggtggt ccagttgcta 900  
tcgctccaac catcgcttct accgacgctc catgttctgc tttgtctggt atctacaccg 960  
acgaaggtga attcgacaga tacttggtgt tgccaaacia cccaacatg gttatcggtg 1020  
acaccaagat cgttgctggt gctccagcta gattgttggc agctggtatc ggtgacgctt 1080  
tggtacctg gttcgaagct agagcttgtt ctagatctgg tgctaccacc atggctggtg 1140  
gtaagtgtac ccaagctgct ttggttttg ctgaattgtg ttacaacacc ttgttggaag 1200  
aagggtgaaa ggctatgttg gctgctgaac aacacgttgt taccocagct ttggaaagag 1260  
ttatogaagc taacacctac ttgtctggtg ttggtttcga atctggtggt ttggctgctg 1320  
ctcacgctgt tcacaacggt ttgaccgcta tcccagacgc tcaccactac taccacggtg 1380  
aaaaggttgc tttcggtagc ttgaccat tggttttga aaacgctcca gttgaagaaa 1440  
tcgaaaccgt tgctgctttg tctcacgctg ttggtttgcc aatcaccttg gctcaattgg 1500  
acatcaagga agacgttcca gctaagatga gaatcgttgc tgaagctgct tgtgctgaag 1560  
gtgaaacat ccacaacatg ccagggtggt ctaccccaga ccaagtttac gctgctttgt 1620  
tggttgctga ccaatacggc caaagattcc tacaagaatg ggaataaggc gcgccccttt 1680  
tcctttgtcg atatcatgta attagttatg tcacgcttac attcacgccc tcctcccaca 1740  
tccgctctaa ccgaaaagga aggagttaga caacctgaag tctaggtccc tattttatttt 1800  
ttttaatagt tatgtagta ttaagaacgt tatttatatt tcaaattttt cttttttttc 1860  
tgtacaaacg cgtgtacgca tgtaacatta tactgaaaac cttgcttgag aaggtttttg 1920  
gacgctcgaa ggcttcctag gcgtacggt 1949

ES 2 648 865 T3

<210> 29

<211> 7690

<212> ADN

<213> Artificial

5

<220>

<223> pRN558

<400> 29

10

```

tgcgcgcttt cggatgatgac ggtgaaaacc tctgacacat gcagctcccg gagacgggtca      60
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg      120
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc      180
accataattc cgttttaaga gcttgggtgag cgctaggagt cactgccagg tatcgtttga      240
acacggcatt agtcagggaa gtcataacac agtcctttcc cgcaattttc tttttctatt      300
actcttggcc tcctctagta cactctatat ttttttatgc ctcgtaatg attttcattt      360
ttttttttcc acctagcggg tgactctttt tttttcttag cgattggcat tatcacataa      420
tgaattatac attatataaa gtaatgtgat ttcttcgaag aatatactaa aaaatgagca      480
ggcaagataa acgaaggcaa agatgacaga gcagaaagcc ctagtaaagc gtattacaaa      540
tgaaccaaac attcagattg cgatctcttt aaaggggtgt cccctagcga tagagcactc      600
gatcttccca gaaaaagagg cagaagcagt agcagaacag gccacacaat cgcaagtgat      660
taacgtccac acaggtatag ggtttctgga ccatatgata catgctctgg ccaagcattc      720
cggctgggtc ctaatcgttg agtgcatggg tgacttacac atagacgacc atcacaccac      780
tgaagactgc gggattgctc tcggtcaagc ttttaaagag gccctactgg cgcgtggagt      840
aaaaagggtt ggatcaggat ttgccccttt ggatgaggca ctttccagag cggtggtaga      900
tctttcgaac aggccgtacg cagttgtcga acttggtttg caaagggaga aagtaggaga      960
tctctcttgc gagatgatcc cgcattttct tgaaagcttt gcagaggcta gcagaattac     1020
cctccacggt gattgtctgc gaggcaagaa tgatcatcac cgtagtgaga gtgcggtcaa     1080
ggctcttgcg gttgccataa gagaagccac ctgcccacat ggtaccaacg atgttccctc     1140
caccaaaggt gttcttatgt agtgacaccg attatttaaa gctgcagcat acgatataata     1200
tacatgtgta tatatgtata cctatgaatg tcagtaagta tgtatacga cagtatgata     1260
ctgaagatga caaggtaatg catcattcta tacgtgtcat tctgaacgag gcgcgctttc     1320
cttttttctt tttgcttttt cttttttttt ctcttgaact cgacggatca tatgcggtgt     1380
gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggaaattgta aacgtaata     1440

```

ES 2 648 865 T3

ttttgttaaa attcgcgtta aatTTTTgtt aaatcagctc atTTTTtaac caataggccg 1500  
 aaatcgcaaa aatcccttat aaatcaaaag aatagaccga gatagggttg agtggtgttc 1560  
 cagtttgcaa caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa 1620  
 ccgtctatca gggcgatggc ccactacgtg aaccatcacc ctaatcaagt tttttggggt 1680  
 cgaggtgccg taaagcacta aatcggaacc ctaaaggag cccccgattt agagcttgac 1740  
 ggggaaaacc ggcattgcga ataccgcttc cacaaacatt gctcaaaagt atctctttgc 1800  
 tataatatctc tgtgctatat ccctatataa cctaccatc cacctttcgc tccttgaact 1860  
 tgcactataa ctgcacctct acattTTTTa tgtttatctc tagtattact ctttagacaa 1920  
 aaaaattgta gtaagaacta ttcatagagt gaatcgaaaa caatacgaaa atgtaaacat 1980  
 ttctatacag tagtatatag agacaaaata gaagaaaccg ttcataattt tctgaccaat 2040  
 gaagaatcat caacgctatc actttctggt cacaaagtat gcgcaatcca catcggata 2100  
 gaatataatc ggggatgcct ttatcttgaa aaaatgcacc cgcagcttcg ctagtaatca 2160  
 gtaaaccggg gaagtggagt caggcttttt ttatggaaga gaaaatagac accaaagtag 2220  
 ccttcttcta acctaacgg acctacagtg caaaaagtta tcaagagact gcattataga 2280  
 gcgcacaaag gagaaaaaaa gtaatctaag atgctttgtt agaaaaatag cgctctcggg 2340  
 atgcattttt gtagaacaaa aaagaagtat agattctttg ttggtaaaaat agcgcctctc 2400  
 cgttgcatth ctgttctgta aaaatgcagc tcagattctt tgtttgaaaa attagcgtc 2460  
 tcgtcgcgtt gcatttttgt tttacaaaaa tgaagcacag attcttcggt ggtaaaatag 2520  
 cgctttcggg ttgcatttct gttctgtaaa aatgcagctc agattctttg tttgaaaaat 2580  
 tagcgtctc gcgttgcatth tttgttctac aaaatgaagc acagatgctt cgtaacaaa 2640  
 gatatgctat tgaagtgcaa gatggaacg cagaaaatga accggggatg cgacgtgcaa 2700  
 gattacctat gcaatagatg caatagtthc tccaggaacc gaaatacata cattgtcttc 2760  
 cgtaaagcgc tagactatat attattatac aggttcaaat atactatctg tttcagggaa 2820  
 aactcccagg ttcggtggtt caaaattcaa tgatgggtaa caagtacgat cgtaaatctg 2880  
 taaaacagtt tgtcggatat taggctgtat ctctcaaaag cgtattcgaa tatcattgag 2940  
 aagctgcagc gtcacatcgg ataataatga tggcagccat tgtagaagtg ccttttgcath 3000  
 ttctagtctc tttctcggc tagctagtht tactacatcg cgaagataga atcttagatc 3060  
 aactgcctt tgcagcgtg gatcaataga gtaacaaaag agtggttaagg cctcgttaaa 3120  
 ggacaaggac ctgagcggaa gtgtatcgta cagtagacgg agtatctagt atagtctata 3180  
 gtccgtggaa ttaattctca tctttgacag cttatcatcg ataacggga gctagcatgc 3240  
 ggccgccagt gtgatggata tctgcagaat tcgccctttt aagcttcgta cgtgtggaag 3300

ES 2 648 865 T3

aacgattaca acaggtgttg tcctctgagg acataaaata cacaccgaga ttcatacaact 3360  
cattgctgga gtttagcatat ctacaattgg gtgaaatggg gagcgtattg caggcatttg 3420  
ctcggcatgc cggtagaggt gtggtcaata agagcgacct catgctatac ctgagaaagc 3480  
aacctgacct acaggaaga gttactcaag aataagaatt ttcgttttaa aacctaaagag 3540  
tcactttaaa atttgtatac acttattttt ttataactt atttaataat aaaaatcata 3600  
aatcataaga aattcgcgcg ctttaagcagc ttcaccagcc tttctagcca aagattgagc 3660  
catcttttca gcggtagcca aagcagaaga ggtcataatg tccaagttae cagcgtaaagc 3720  
tggcaagtag tgagcagcac cttcaacttc caaccaaaca gcggtcttca aaccagagaa 3780  
ttgaccaaca cctggcaagt taactggctt gtcttggggg ataacttcga attgaactct 3840  
ttgcttcaat ctgtaacctg gaacgtatgc ttgaacagct tcagccattt cgttgattga 3900  
agcttcgatg tcgtcttgag atgcttcgtc agacaaaacg taaacgggtg ctctcatcat 3960  
caatggtggg tcagctgggt tcaaaacgat gatagcetta cccttagcag caccaccaac 4020  
aacttcgata gccctagagg tggtttcggt gaattcgtcg atgtagctc tggtagctgg 4080  
accagcagac ttagaagcga tagaagcgat gatttcagcg tagtgaactc tagcaactct 4140  
tgaaacagca gcaaccattg ggatggtagc ttgaccacca caggttaacca tgttaacggt 4200  
taattggtca acgtagctt ccaagttaac aactggaaca cagtatggac cgatagcagc 4260  
tggggtcaag tcgatcaatc tgatgtctgg cttagcttct ctcaaagcag cgtcgttctt 4320  
aacgtgagca ccagcagagg tagcgtcga aacgatgtcg atgtcagcga attctggcat 4380  
gttcatcaaa ccgataaac cttcgtgggt ggtagcaaca cccattcttc tagctctagc 4440  
caaaccgtca gattgtgggt cgataccaac cataacagcc atttccaagt gttgaccgtg 4500  
tcttaggatc ttgatcatca agtcagtacc gatgttacca gaaccgatga tagcaacctt 4560  
tctcttagac atctgcagtc tagatatatt tgttgtaaaa agtagataat tacttccttg 4620  
atgatctgta aaaaagagaa aaagaaagca tctaagaact tgaaaaacta cgaattagaa 4680  
aagaccaa atgtatttct tgcattgacc aatttatgca agtttatata tatgtaaatg 4740  
taagttcac gaggttctac taaactaac cacccttg gttagaagaa aagagtgtgt 4800  
gagaacaggc tgttgtgtc acacgattcg gacaattctg tttgaaagag agagagtaac 4860  
agtacgatcg aacgaacttt gctctggaga tcacagtggg catcatagca tgtggtacta 4920  
aacccttcc cgcattcca gaacctcga ttgcttghta caaacctgt gagccgtcgc 4980  
taggaccttg ttgtgtgacg aaattggaag ctgcaatcaa taggaagaca ggaagtcgag 5040  
cgtgtctggg tttttcagc tttgttctt ttgcaacaa atcacgagcg acggtaat 5100  
ctttctgat aagaggccac gtgcttatg agggtaacat caattcaaga aggagggaaa 5160  
cacttcctt ttctggccct gataatagta tgagggtgaa gccaaaataa aggattcgcg 5220

ES 2 648 865 T3

cccaaatcgg catctttaa tgcaggtatg cgatagttcc tcaactcttc cttactcagc 5280  
 agtaattctt gcaaatgcct attatgcaga tgttataata tctgtgcgtc ttgagttgaa 5340  
 gagctcgaga ctagtggatc ccccgggctg caggaattcg atatcaagct tatcgatacc 5400  
 gtcgacctcg agggggggcc cggtaaccag cttttgttcc ctttagtgag ggtaattcc 5460  
 gagcttggcg taatcatggt catagctggt tcctgtgtga aattgttacc cgctcacaat 5520  
 tccacacaac ataggagccg gaagcataaa gtgtaaagcc tgggggcctc aatgagtgag 5580  
 gtaactcaca ttaattgcgt tgcgctcact gcccgcttc cagtcgggaa acctgtcgtg 5640  
 ccagctgcat taatgaatcg gccaacgccc ggggagagcc ggtttgcgta ttgggcgctc 5700  
 ttccgcttcc tcgctcactg actcgtcgtg ctccggtcgtt cggctgcggc gagcggatc 5760  
 agctcactca aaggcggtaa tacggttacc cacagaatca ggggataacg caggaaagaa 5820  
 catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt 5880  
 tttccatagg ctccggcccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg 5940  
 gcgaaacccg acaggactat aaagatacca ggcgttcccc cctggaagct ccctcgtcgc 6000  
 ctctcctggt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag 6060  
 cgtggcgtt tctcaatgct caccgtgtag gtatctcagt tcggtgtagg tcgttcgctc 6120  
 caagctgggc tgtgtgcacg aacccccctg tcagccccgac cgctgcgcct tatccggtaa 6180  
 ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccaactggcag cagccactgg 6240  
 taacaggatt agcagagcga ggtatgtagg cgggtctaca gagttcttga agtggtggcc 6300  
 taactacggc tacaactagaa ggacagtatt tggatctctg gctctgctga agccagttac 6360  
 cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcgggtg 6420  
 ttttttgtt tgcaagcagc agattacgcg cagaaaaaaaa ggatctcaag aagatccttt 6480  
 gatcttttct acggggtctg acgctcagtg gaacgaaaaac tcacgttaag ggattttggt 6540  
 catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa 6600  
 atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga 6660  
 ggcacctatc tcagcgatct gtctatctcg ttcattccata gttgcctgac tgcccgtcgt 6720  
 gtagataact acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg 6780  
 agaccacgc tcaccggctc cagatttacc agcaataaac cagccagccg gaagggccga 6840  
 gcgcagaagt ggtcctgcaa ctttatccgc ctccatccag tctattaatt gttgccggga 6900  
 agctagagta agtagttcgc cagttaatag tttgcgcaac gttgttgcca ttgctacagg 6960  
 catcgtggtg tcacgctcgt cgtttggtat ggcttcattc agctccggtt cccaacgatc 7020  
 aaggcgagtt acatgatccc ccatgttggt aaaaaagcg gttagctcct tcggctcctc 7080

ES 2 648 865 T3

gatcgttgtc agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca 7140  
taattctctt actgtcatgc catccgtaag atgcttttct gtgactgggtg agtactcaac 7200  
caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccg cgtcaatacg 7260  
ggataatacc gcgccacata gcagaacttt aaaagtgtc atcattggaa aacgttcttc 7320  
ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc agttcgatgt aacccactcg 7380  
tgcacccaac tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac 7440  
aggaaggcaa aatgccgcaa aaaagggaaat aagggcgaca cggaaatgtt gaatactcat 7500  
actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata 7560  
catatgtgaa tgtattaga aaaataaaca aataggggtt ccgcgcacat ttccccgaaa 7620  
agtgccacct gacgtcttat tatcatgaca ttaacctata aaaataggcg tatcacgagg 7680  
ccctttcgtc 7690

<210> 30

<211> 9416

5

<212> ADN

<213> Artificial

<220>

<223> pRN595

10

<400> 30

# ES 2 648 865 T3

gacgaaaggg cctcgtgata cgcctatfff tataggftra tgfcafgata ataagacgfc	60
aggtggcact fffcggggaa atgtgcgcgg aaccocfaff fgtffaffff fcfaaafaca	120
ffcaaatatg fatccgcfca fgagacaata acccfgataa atgcfcfcaaf aafaffgaaa	180
aaggaagagt atgagfaffc aacaffffccg fgfcgcfccff affcccffff ffggcgcaff	240
ffgcfccffcf gfffffgcfc acccagaaaac gcfggfgaaa gfaaaagafg cfgaagafca	300
gffgggfgca cagagfggff acafcfgaacf ggaacfcaac agcggfaaga fcfcfgaag	360
ffffcgcccc gaagaacfff ffcaafgaf gagcacffff aaagffcfgc fatgfgggcg	420
ggaaffafcc cgaaffgacg ccgggcaaga gcaacfcfgg cggcfcaafc affaffcfca	480
gaafgacfff gffgagfaff caccagfcac agaaaagcaaf cffacfgaaf gcafgacagf	540
aagagaafaa fgcafgcfcg ccafaaccaf gafgfaaac affcgggcca afffaffffcf	600
gacaacfgafc gggaggaccga aggagcfaac cgcfffffff cacaacfaff gggafcafgf	660
aacfcfccff gacfgtfggg aaccggagcf gaafgaagcc afaccaaaaf acgagcfgfa	720
caccacfgaaf ccfgtagcaa fggaacaac gffgcfcaaa cfaffaacf gcfgaacfaff	780
faffcfafgcf fcccggaac aaffaafaga cfggafggag gcfgaafaaag ffgcaggacc	840
affcfcfgcfc fcggcccfcc cggcfggcfg gffffaffcf gafaaacf gacccggfga	900
gcfgfggffcf cggggaafca ffgcagcaff ggggcfcaaf ggaagcccf cccgaaffcf	960

ES 2 648 865 T3

agttatctac acgacgggca gtcaggcaac tatggatgaa cgaaatagac agatcgtga 1020  
 gataggtgcc tcaactgatta agcattggta actgtcagac caagtttact catatatact 1080  
 ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tcctttttga 1140  
 taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccg 1200  
 agaaaagatc aaagatcctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca 1260  
 aacaaaaaaaa ccaccgctac cagcggtggt ttgtttgccc gatcaagagc taccaactct 1320  
 ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgt 1380  
 gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 1440  
 aatcctgtta ccagtggctg ctgccagtgg cgataagtcc tgtcttaccg ggttggaactc 1500  
 aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggt cgtgcacaca 1560  
 gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agcattgaga 1620  
 aagcgcacg ctccccaag ggagaaaggc ggacaggtat ccgtaagcg gcagggctcg 1680  
 aacaggagag cgcacgagg agcttccagg ggggaacgcc tggatcttt atagtctgt 1740  
 cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggccgag 1800  
 cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 1860  
 tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 1920  
 tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcagag cagtgagcga 1980  
 ggaagcggaa gagcgcctaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 2040  
 atgcagctgg cacgacaggt ttcccactg gaaagcgggc agtgagcgca acgcaattaa 2100  
 tgtgagttac ctcaactcatt aggcacccca ggctttacac tttatgcttc cggctcctat 2160  
 gttgtgtgga atgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 2220  
 cgccaagctc ggaattaacc ctcaactaaag ggaacaaaag ctgggtaccg ggccccccct 2280  
 cgaggtcgac ggtatcgata agcttgatat cgaattcctg cagcccgggg gatccactag 2340  
 tctcgagctc ttcaactcaa gacgcacaga tattataaca tctgcataat aggcatttgc 2400  
 aagaattact cgtgagtaag gaaagagtga ggaactatcg catacctgca tttaaagatg 2460  
 ccgatttggg cgcgaatcct ttattttggc ttcaccctca tactattatc agggccagaa 2520  
 aaaggaagtg tttccctcct tcttgaattg atgttaccct cataaagcac gtggcctctt 2580  
 atcgagaag aaattaccgt cgctcgtgat ttgtttgcaa aaagaacaaa actgaaaaaa 2640  
 ccagacacg ctcgacttcc tgtcttcta ttgattgag cttccaattt cgtcacacaa 2700  
 caagtccta cgcacggctc acaggttttg taacaagcaa tcgaaggctc tggaatggcg 2760  
 ggaaaggggt tagtaccaca tgctatgat cccactgtga tctccagagc aaagttcgtt 2820

ES 2 648 865 T3

cgatcgtact gttactctct ctctttcaaa cagaattgtc cgaatcgtgt gacaacaaca 2880  
 gcctgttctc acacactctt ttcttttaac caagggggtg gtttagttta gtagaacctc 2940  
 gtgaaactta cattedacata tatataaact tgcataaatt ggtcaatgca agaaatacat 3000  
 atttggtctt ttctaattcg tagtttttca agttcttaga tgctttcttt ttctcttttt 3060  
 tacagatcat caaggaagta attatctact ttttacaaca aatatactta gaaaatggct 3120  
 gttaccaacg ttgctgaatt gaacgctttg gttgaaaggg ttaagaaggc tcaagagaaa 3180  
 tacgcttctt tcaccecaaga acaagttgac aagatcttca gagctgctgc tttggctgct 3240  
 gctgacgcta gaatcccatt ggctaagatg gctggtgctg aatctggtat gggatctgct 3300  
 gaagacaagg ttatcaagaa ccacttcgct tctgaatata tctacaacgc ttacaaggac 3360  
 gaaaagacct gtggtgtttt gtcagaagac gacaccttgc gtaccatcac catcgctgaa 3420  
 ccaatcggta tcatctgtgg tatcgttcca accaccaacc caacctctac cgctatcttc 3480  
 aagtctttga tctctttgaa gaccagaaac gctatcatct tctctccaca cccaagagct 3540  
 aaagacgcta ccaacaaggc tgctgacatc gttttgcaag ctgctatcgc tgctgggtgct 3600  
 ccaaaggact tgatcggttg gatcgaccaa ccatctgttg aattgtctaa cgctttgatg 3660  
 caccacccag acatcaactt gatcttggct accgggtgct caggtatggt taaggctgct 3720  
 tactcttctg gtaagccagc tatcgggtgt ggtgctggtg acaccccagt tgttatcgac 3780  
 gaaaccgctg acatcaagag agctgttgct tctgttttga tgtctaagac ctctgacaa 3840  
 ggtgttatct gtgcttctga acaatctgtt gttggtgttg actctgttta cgacgctgct 3900  
 agagaaagat tcgctaccca cgggtggttac ttggtgcaag gtaaggaatt gaaggctgct 3960  
 caagacgta tcttgaagaa cgggtgctttg aacgctgcta tcggttggtca accagcttac 4020  
 aagatcgtg aattagctgg tttctctggt ccagaaaaa ccaagatctt gatcggtgaa 4080  
 gttaccgctg ttgacgaatc tgaaccattc gctcacgaaa agttgtctcc aaccttggct 4140  
 atgtacagag ctaaggactt cgaagacgct gttgaaaaag ctgaaaagtt ggttgcctatg 4200  
 ggtggtattg gtcacacctc ttgtttgatc accgaccaag acaaccaacc agctagagtt 4260  
 tcttacttctg gtcaaaaagat gaagaccgct agaactctga tcaacacccc agcttctcaa 4320  
 ggtggtatcg gtgacttgta caacttcaag ttggctccat ctttgacctt gggttgtggt 4380  
 tcttgggggtg gtaactctat ctctgaaaac gttggtccaa agcacttgat caacaagaag 4440  
 accgttgcta agagagctga aaacatgttg tggcacaagt tgccaaaatc tatctacttc 4500  
 agaagagggt ctttgccaat cgctttggac gaagttatca ccgacggtca caagagagct 4560  
 ttgatcgtta ccgacagatt cttgttcaac aacggttacg ctgaccaaata cacctctgct 4620  
 ttgaaggctg ctgggtgtga aaccgaagtt ttcttcgaag ttgaagctga cccaaccttg 4680  
 tctatcgtta gaaagggtgc tgaattggct aactctttca agccagacgt tatcatcgtc 4740

ES 2 648 865 T3

ttgggtggtg gttctccaat ggacgctgct aagatcatgt gggttatgta cgaacaccca 4800  
 gaaaccact tcgaagaatt ggctttgaga ttcattgaca tcagaaagag aatctacaag 4860  
 ttcccaaaga tgggtgtaa ggctaagatg atcgctgta ccaccacctc tggtagcgg 4920  
 tctgaagtta cccattcgc tgttgttacc gacgacgcta cccgtcaaaa gtaccattg 4980  
 gctgactacg ctttgacccc agacatggct atcgttgacg ctaacttgg 5040  
 ccaaagtctt tgtgtgcttt cgggtggttt gacgctgta cccacgctat ggaagcttac 5100  
 gtttctgttt tggcttctga attctctgac ggtcaagctt tgcaagcttt gaagttgtt 5160  
 aaggaatact tgcagcttc ttaccacgaa ggttctaaga acccagttgc tagagaaaga 5220  
 gttcactctg ctgctacat cgctggatc gctttcgcta acgctttctt ggggtgtttgt 5280  
 cactctatgg ctcaaacgtt gggttctcaa ttccacatcc cacacggttt ggctaacgct 5340  
 ttgttgatct gtaacgttat cagatacaac gctaacgaca acccaaccaa gcaaaccgct 5400  
 ttctctcaat acgacagacc acaagctaga agaagatacg ctgaaatcgc tgaccattg 5460  
 ggtttgtctg ctccaggtga cagaaccgct gcaaagatcg aaaagtgtt ggcttggtt 5520  
 gaaaccttga aggctgaatt gggatccca aagtctatca gagaagctgg tgttcaagaa 5580  
 gctgacttct tggctaactg tgacaagttg tctgaagacg ctttcgacga ccaatgtacc 5640  
 ggtgctaacc caagataccc attgatctct gaattgaagc aatcttgtt ggacacctac 5700  
 tacggtagag actacgttga aggtgaaacc gctgctaaga aggaagctgc tccagctaag 5760  
 gctgaaaaga aggctaagaa gtctgcttag ctttaagcgc cgaatttctt atgatttatg 5820  
 atttttatta ttaataagt tataaaaaa ataagtgtat acaaatttta aagtgactct 5880  
 taggttttaa aacgaaaatt cttattcttg agtaactctt tcctgtaggt caggttgctt 5940  
 tctcaggtat agcatgaggt cgctcttatt gaccacacct ctaccggcat gccgagcaa 6000  
 tgctgcaaa tcgctccca ttccaccaa ttgtagatat gctaactcca gcaatgagtt 6060  
 gatgaatctc ggtgtgtatt ttatgtctc agaggacaac acctgttgta atcgttcttc 6120  
 cacacgtacg aagcttaaaa gggcgaattc tgcagatata catcacactg gcggccgcat 6180  
 gctagctccg gattatcgat gataagctgt caaagatgag aattaattcc acggactata 6240  
 gactatacta gatactccgt ctactgtacg atacacttcc gctcaggtcc ttgtcctta 6300  
 acgaggcctt accactctt tgttactcta ttgatccagc tcagcaaagg cagtgtgatc 6360  
 taagattcta tcttcgcat gtagtaaac tagctagacc gagaagaga ctagaaatgc 6420  
 aaaaggcact tctacaatgg ctgccatcat tattatccga tgtgacgctg cagcttctca 6480  
 atgatattcg aatacgttt gaggagatac agcctaatat ccgacaaact gttttacaga 6540  
 ttacgatcg tacttgttac ccatcattga atttgaaca tccgaacctg ggagttttcc 6600

ES 2 648 865 T3

ctgaaacaga tagtatattt gaacctgtat aataatatat agtctagcgc tttacggaag 6660  
 acaatgtatg tatttcggtt cctggagaaa ctattgcatc tattgcatag gtaatcttgc 6720  
 acgtcgcate cccggttcat tttctgcgtt tccatcttgc acttcaatag catatctttg 6780  
 ttaacgaagc atctgtgctt cattttgtag aacaaaaatg caacgcgaga gcgctaattt 6840  
 ttcaaaacaaa gaatctgagc tgcattttta cagaacagaa atgcaacgcg aaagcgctat 6900  
 tttaccaacg aagaatctgt gcttcatttt tgtaaaacaa aaatgcaacg cgacgagagc 6960  
 gctaattttt caaacaaga atctgagctg catttttaca gaacagaaat gcaacgcgag 7020  
 agcgctattt taccaacaaa gaatctatac ttcttttttg ttctacaaa atgcatcccg 7080  
 agagcgctat ttttctaaca aagcatctta gattactttt tttctccttt gtgcgctcta 7140  
 taatgcagtc tcttgataac tttttgcact gtaggtccgt taagggttaga agaaggctac 7200  
 tttggtgtct attttctctt ccataaaaa agcctgactc cacttcccgc gtttactgat 7260  
 tactagcga gctgcgggtg cattttttca agataaaggc atccccgatt atattctata 7320  
 ccgatgtgga ttgcgcatac tttgtgaaca gaaagtata gcgttgatga ttcttcattg 7380  
 gtcagaaaat tatgaacggt ttcttctatt ttgtctctat atactacgta taggaaatgt 7440  
 ttacattttc gtattgtttt cgattcactc tatgaatagt tcttactaca atttttttgt 7500  
 ctaaagagta atactagaga taaacataaa aaatgtagag gtcgagttta gatgcaagtt 7560  
 caaggagcga aagggtgatg ggtaggttat atagggatat agcacagaga tatatagcaa 7620  
 agagatactt ttgagcaatg tttgtggaag cggatttcgc aatgccggct ttccccgtca 7680  
 agctctaaat cgggggctcc ctttaggggt ccgatttagt gctttacggc acctcgacc 7740  
 caaaaaactt gattaggggt atggttcaag tagtgggcca tcgccctgat agacggtttt 7800  
 tcgccctttg acggttgagt ccacgttctt taatagtgga ctcttgttcc aaactggaac 7860  
 aacactcaac cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc 7920  
 ctattggtta aaaaatgagc tgatttaaca aaaatttaac gcgaattttt acaaaatatt 7980  
 aacgtttaca atttctgat gcggtatfff ctcttaccgc atctgtgcgg tatttcacac 8040  
 cgcataatgat ccgtcgagtt caagagaaaa aaaaagaaaa agcaaaaaaga aaaaaggaaa 8100  
 gcgcgcctcg ttcagaatga cacgtataga atgatgcatt accttgtcat cttcagtatc 8160  
 atactgttcg tatacatact tactgacatt cataggtata catatataca catgtatata 8220  
 tatcgtatgc tgcagcttta aataatcggg gtcactacat aagaacacct ttggtggagg 8280  
 gaacatogtt ggtaccattg ggcgaggtgg cttctcttat ggcaaccgca agagccttga 8340  
 acgcaactctc actacggtga tgatcattct tgccctgcag acaatcaacg tggagggtaa 8400  
 ttctgctagc ctctgcaaa ctttcaagaa aatgcgggat catctcgaa gagagatctc 8460  
 ctactttctc cctttgcaaa ccaagttcga caactgcgta cggcctgttc gaaagatcta 8520

ES 2 648 865 T3

ccaccgctct ggaaagtgcc tcatccaaag gcgcaaatcc tgatccaaac ctttttactc 8580  
cacgcgccag tagggcctct ttaaaagctt gaccgagagc aatcccgcag tcttcagtgg 8640  
tgtgatggtc gtctatgtgt aagtcaccaa tgcactcaac gattagcgac cagccggaat 8700  
gcttggccag agcatgtatc atatggtcca gaaaccctat acctgtgtgg acgttaatca 8760  
cttgcgattg tgtggcctgt tctgctactg cttctgcctc tttttctggg aagatcgagt 8820  
gctctatcgc taggggacca ccctttaaag agatcgcaat ctgaatcttg gtttcatttg 8880  
taatacgcctt tactagggct ttctgctctg tcatctttgc cttegtttat cttgcctgct 8940  
catttttttag tatattcttc gaagaaatca cttacttta tataatgtat aattcattat 9000  
gtgataatgc caatcgctaa gaaaaaaaa gagtcatccg ctaggaggaa aaaaaaaaaat 9060  
gaaaatcatt accgaggcat aaaaaaatat agagtgtact agaggaggcc aagagtaata 9120  
gaaaagaaa attgcgggaa aggactgtgt tatgacttcc ctgactaatg ccgtgttcaa 9180  
acgatacctg gcagtgactc ctagcgctca ccaagctctt aaaacggaat tatgggtgcac 9240  
tctcagtaca atctgctctg atgccgcata gtttaagccag ccccgacacc cgccaacacc 9300  
cgctgacgcg ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac 9360  
cgtctccggg agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcga 9416

<210> 31

<211> 9352

5

<212> ADN

<213> Artificial

<220>

<223> pRN596

10

<400> 31

ES 2 648 865 T3

gacgaaaggg cctcgtgata cgcctatfff tataggfmaa tgfcatgata ataagacgfc	60
aggtggcact fffcggggaa atgtgcgcgg aaccocfaff fgtffaffff fcfaaafaca	120
ffcaaatatg fatccgcfca fgagacaata acccfataa atgcfafcaaf aafaffgaaa	180
aaggaagagt atgagfaffc aacaffffccg fgtcgcfcff affccfffff ffgggcatt	240
ffcfccffcf gfffffgcf acccagaaa acffgfgaaa gfaaaagafg cfgaagafca	300
gffgggfgca cagagfggff acafcfgaacf ggaacfcaac agcfgfaaga fcfffgagag	360
ffffcgcfc gaagaacf ffcaafgaf gaccafffff aaagffcfgc fatgfggcf	420
ggaaffafcc cgaaffgacg cgggcaaga gcaacfcfga cgcfcfaca afcaffcfca	480
gaafgacf fgffgafacf caccagfcac agaaaagcaf cfafcfgaaf gcafgaacf	540
aagagaafaa fgcafggcf ccataaccaf gafgfaaac afcfggcca afcaffffcf	600
gacaacfafc ggaagaccga aggagcfaac ccfaffffff cacaacfagf ggaacfafg	660

ES 2 648 865 T3

aactgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaaacg acgagcgtga 720  
caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact 780  
tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc 840  
acttctgcgc tgggcccttc cggctggctg gtttattgct gataaatctg gagccgggtga 900  
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 960  
agttatctac acgacgggca gtcaggcaac tatggatgaa cgaatatagac agatcgctga 1020  
gataggtgcc tcaactgatta agcattggta actgtcagac caagtttact catatatact 1080  
ttagattgat ttaaaaacttc atttttaatt taaaaggatc taggtgaaga tcctttttga 1140  
taatctcatg accaaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgct 1200  
agaaaagatc aaaggatcct cttgagatcc tttttttctg cgcgtaatct gctgcttgca 1260  
aacaacaaaaa ccaccgctac cagcgggtgt ttgtttgccg gatcaagagc taccaactct 1320  
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgtg 1380  
gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 1440  
aatcctgtta ccagtggtct ctgccagtgg cgataagtct tgtcttaccg ggttggtactc 1500  
aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggt cgtgcacaca 1560  
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agcattgaga 1620  
aagcgcacag cttcccgaag ggagaaagcc ggacaggtat ccggtaaagc gcagggtcgg 1680  
aacaggagag cgcacgaggg agcttccagg ggggaacgcc tggatcttt atagtctgt 1740  
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggccgag 1800  
cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 1860  
tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 1920  
tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcagat cagtgagcga 1980  
ggaagcggaa gagcgcctaa tacgcaaac gcctctcccc gcgcggttggc cgattcatta 2040  
atgcagctgg cacgacaggt ttcccactg gaaagcgggc agtgagcga acgcaattaa 2100  
tgtgagttac ctcaactcatt aggcacccca ggctttacac tttatgcttc cggtcctat 2160  
gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 2220  
cgccaagctc ggaattaacc ctcaactaaag ggaacaaaag ctgggtaccg ggccccct 2280  
cgaggtcgac ggtatcgata agcttgatat cgaattcctg cagcccgggg gatccactag 2340  
tctcgagctc ttcaactcaa gacgcacaga tattataaca tctgcataat aggcatttgc 2400  
aagaattact cgtgagtaag gaaagagtga ggaactatcg catacctgca tttaaagatg 2460  
ccgatttggg cgcgaatcct ttattttggc ttcacctca tactattatc agggccagaa 2520  
aaaggaagtg tttccctcct tcttgaattg atgttaccct cataaagcac gtggcctctt 2580

ES 2 648 865 T3

atcgagaaag aaattaccgt cgctcgtgat ttgtttgcaa aaagaacaaa actgaaaaaa 2640  
 cccagacacg ctcgacttcc tgtcttcccta ttgattgcag cttccaattt cgtcacacaa 2700  
 caaggctcta gcgacggctc acagggtttg taacaagcaa tcgaaggttc tggaatggcg 2760  
 ggaaagggtt tagtaccaca tgctatgatg cccactgtga tctccagagc aaagttcgtt 2820  
 cgatcgtact gttactctct ctctttcaaa cagaattgtc cgaatcgtgt gacaacaaca 2880  
 gcctgttctc acacactctt ttcttctaac caagggggtg gtttagttta gtagaacctc 2940  
 gtgaaactta catttacata tatataaact tgcataaatt ggtcaatgca agaaatacat 3000  
 atttggcttt ttctaattcg tagtttttca agttcttaga tgctttcttt ttctcttttt 3060  
 tacagatcat caaggaagta attatctact ttttacaaca aatatactta gaaaatgtct 3120  
 acccaacaaa ccatgaccgt tgacgaacac atcaaccaat tagttagaaa ggctcaagtt 3180  
 gctttgaagg aatacttgaa gccagaatac acccaagaaa agatcgacta catcgttaag 3240  
 aaggctctcg ttgctgcttt ggaccaacac tgtgctttgg ctgctgctgc tgttgaagaa 3300  
 accggtagag gtatcttcca agacaaggct accaagaaca tcttcgcttg tgaacacgtt 3360  
 acccacgaaa tgagacacgc taagaccggt ggtatcatca acgttgacct attgtacggt 3420  
 atcaccgaaa tcgctgaacc agttgggtgt gtttgtggtg ttaccccagt taccaacca 3480  
 acctctaccg ctatcttcaa gtctttgatc tctatcaaga ccagaaacct aatcgttttc 3540  
 tctttccacc catctgcttt gaagtgttct attatggctg ctaaaatcgt tagagacgct 3600  
 gctatcgtcg ctgggtgctc agaaaactgt atccaatgga tcgaattcgg tggtatcgaa 3660  
 gcttctaaca agttgatgaa ccaccagggt gttgctacta tcttggttac cggtggtaac 3720  
 gctatggtta aggctgcata ctcttctggt aagccagctt tgggtgttgg tgctggtaac 3780  
 gttccaacct acatogaaaa gacctgtaac atcaagcaag ctgctaacga cgttgttatg 3840  
 tctaagtctt tcgacaacgg tatgatctgt gcttctgaac aagctgctat catcgacaag 3900  
 gaaatctacg accaagttgt tgaagaaatg aagaccttgg tgcttactt catcaacgaa 3960  
 gaagaaaagg ctaagttgga aaagttcatg ttcgggtgta acgcttactc tgctgacgtt 4020  
 aacaacgcta gattgaacct aaagtgcca ggtatgtctc cacaatgggt cgctgaacaa 4080  
 gttggtatca aggtaccaga agactgtaac atcatctgtg ctgtttghaa ggaagttggt 4140  
 ccaaacgaac cattgaccag agaaaagttg tctccagttt tggctatctt gaaagctgaa 4200  
 aacaccaag acggtatcga caaggctgaa gctatggttg aatttaacgg tagaggtcac 4260  
 totgctgcta tccactctaa cgacaaggct gttgttgaaa agtacgcttt gaccatgaag 4320  
 gctttagtaa tcttgacaaa cccccatct tctcaagggt gtatcggttc tatctacaac 4380  
 tacatctggc catctttcac cttgggttgt ggttcttacg gtggtaacctc tgtttctgct 4440

ES 2 648 865 T3

aacgttacct accacaactt gttgaacatc aagagattgg ctgacagaag aaacaacttg 4500  
caatggttca gagttccacc aaagatcttc ttogaaccac actctatcag atacttggct 4560  
gaattgaagg aattgtctaa gatcttcacg gtttctgaca gaatgatgta caagttgggt 4620  
tacgttgaca gagttatgga cgttttgaag agaagateta acgaagttga aatcgaaatc 4680  
ttcatcgacg ttgaaccaga cccatctatc caaaccttc aaaagggttt ggctgttatg 4740  
aacaccttcg gtccagacaa catcatcgct atcgggtggtg gttctgctat ggacgctgct 4800  
aagatcatgt ggttggtgta cgaacaccca gaagctgact tcttcgctat gaagcaaaaag 4860  
ttcatcgact tgagaaaagag agctttcaag ttccaacca tgggtaagaa ggctagattg 4920  
atctgtatcc caaccacctc tggtagcgtt tctgaagtta cccattcgc tgttatctct 4980  
gaccacgaaa ccggttaagaa gtaccattg gctgactact ctttgacccc atctgttgct 5040  
atcgttgacc caatgttcac catgtctttg ccaaagagag ctatcgctga caccggtttg 5100  
gacgttttg ttcacgctac cgaagcttac gtttctgta tggctaacga atacaccgac 5160  
ggtttggtta gagaagctgt taagttggtt tttgaaaact tgttgaagtc ttacaacggt 5220  
gacttgaag ctagagaaaa gatgcacaac gctgctacca tcgctggtat ggctttcgct 5280  
tctgctttct tgggtatgga ccaactctatg gctcacaagg ttggtgctgc tttccacttg 5340  
ccacacggta gatgtgttgc tgttttggtt ccacacgta tcagatacaa cggtaaaaag 5400  
ccaagaaagt tggctatgtg gccaaagtac aacttctaca aggctgacca aagatacatg 5460  
gaattggctc aaatggttgg tttgaagtgt aacaccccag ctgaaggtgt tgaagcttc 5520  
gctaaggctt gtgaagaatt gatgaaggct accgaaacca tcaccggttt caagaaggct 5580  
aacatcgacg aagctgcttg gatgtctaag gttccagaaa tggctttggt ggctttcgaa 5640  
gaccaatggt ctccagctaa cccaagagt ccaatggtta aggacatgga aaagatcttg 5700  
aaggctgctt actaccaat cgcttagctt aagcgcgcga atttcttatg atttatgatt 5760  
tttattatta aataagttat aaaaaaata agtgatatac aattttaaag tgactcttag 5820  
gttttaaac gaaaattctt attcttgagt aactcttcc tgtaggtcag gttgctttct 5880  
caggatagc atgaggtcgc tcttattgac cacacctta ccggcatgcc gagcaaatgc 5940  
ctgcaaatcg ctccccattt cacccaattg tagatatgct aactccagca atgagttgat 6000  
gaatctcggg gtgtatttta tgtcctcaga ggacaacacc tgttgtaatc gttcttccac 6060  
acgtacgaag ctaaaagggc gaattctgca gatatccatc aactggcgg ccgcatgcta 6120  
gctccggatt atcgatgata agctgtcaaa gatgagaatt aattccacgg actatagact 6180  
atactagata ctccgtctac tgtacgatac acttccgctc aggtccttgt cctttaacga 6240  
ggccttacca ctcttttggt actctattga tccagctcag caaaggcagt gtgatctaag 6300  
attctatctt cgcgatgtag taaaactagc tagaccgaga aagagactag aaatgcaaaa 6360

ES 2 648 865 T3

ggcacttcta caatggctgc catcattatt atccgatgtg acgctgcagc ttctcaatga 6420  
 tattcgaata cgctttgagg agatacagcc taatatccga caaactgttt tacagattta 6480  
 cgatcgtact tgttaccat cattgaatth tgaacatccg aacctgggag ttttccctga 6540  
 aacagatagt atatthgaac ctgtataata atatatagtc tagcgtttha cggaagacaa 6600  
 tgtatgtatt tccgttcctg gagaaactat tgcattctatt gcataggtaa tcttgcacgt 6660  
 cgcattcccc gttcattttc tgcgtttcca tcttgcactt caatagcata tctttgthaa 6720  
 cgaagcatct gtgcttcatt ttgtagaaca aaaatgcaac gcgagagcgc taatthttca 6780  
 aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgaaag cgctatthta 6840  
 ccaacgaaga atctgtgctt cattthttgta aaacaaaaat gcaacgcgac gagagcgccta 6900  
 atthttcaaa caaagaatct gagctgcatt tttacagaac agaaatgcaa cgcgagagcg 6960  
 ctatthttacc aacaaagaat ctatacttct tttttgttct acaaaaatgc atcccgcgag 7020  
 cgctatthtt ctaacaaagc atcttagatt actthttttc tctttgtgc gctctataat 7080  
 gcagctctct gataacttht tgcactgtag gtccgttaag gttagaagaa ggctactttg 7140  
 gtgtctatth tctcttccat aaaaaagcc tgactccact tcccgcgttt actgattact 7200  
 agcgaagctg cgggtgcatt tttcaagat aaaggcatcc ccgattatat tctataccga 7260  
 tgtggattgc gcatacttht tgaacagaaa gtgatagcgt tgatgattct tcattggtca 7320  
 gaaaattatg aacggthtct tctatthttgt ctctatatac tacgtatagg aaatgthtac 7380  
 atthttcgtat tgtthttcgt tcaactctatg aatagthtct actacaatth ttttgtctaa 7440  
 agagtaatac tagagataaa cataaaaaat gttagagtcg agthttagatg caagthtcaag 7500  
 gagcgaagg tggatgggta ggttatatag ggatatagca cagagatata tagcaaaag 7560  
 atactthttg gcaatgthtt tgggaagcgtt attcgcaatg ccggctthtc ccgtcaagct 7620  
 ctaaactcgg ggctccctth agggthtccga thtagtgctt tacggcacct cgaccccaaa 7680  
 aaacttgatt aggttgatg thcacgtagt gggccatcgc cctgatagac ggtthttcgc 7740  
 cctthtgact tggagthcac gthctthtaat agthgactct tgtthcaaac tggacaaca 7800  
 ctcaacccta tctcggctca thctthttgat thataaggga thttgccgat thcggcctat 7860  
 tggthaaaa atgagctgat thaacaaaa thtaacgcga atthtaacaa aatathaacg 7920  
 thtaacaatth cctgatgcgg ththttctcc thacgcact gtgcggtatt thcacaccgca 7980  
 tatgatccgt cgagthtcaag agaaaaaaa agaaaaagca aaaagaaaa aggaaagcgc 8040  
 gctcgttca gaatgacacg tatagaatga tgcattacct tgtcatctc agtatcatac 8100  
 tgttcgtata catacttact gacattcata ggtatacata tatacacatg tatatatatc 8160  
 gtatcgtgca gctthaaata atcgggtgca ctacataaga acacctthgg tggagggaac 8220

ES 2 648 865 T3

atcgttggtta ccattgggcg aggtggcttc tcttatggca accgcaagag ccttgaacgc 8280  
actctcacta cggatgatgat cattcttgcc tcgcagacaa tcaacgtgga gggtaattct 8340  
gctagcctct gcaaagcttt caagaaaatg cgggatcatc tcgcaagaga gatctcctac 8400  
tttctccctt tgcaaaccaa gttcgacaac tgcgtacggc ctgttcgaaa gatctaccac 8460  
cgctctggaa agtgcctcat ccaaaggcgc aaatcctgat ccaaaccctt ttactccacg 8520  
cgccagtagg gcctctttaa aagcttgacc gagagcaatc ccgcagtctt cagtgggtgtg 8580  
atggtcgtct atgtgtaagt caccaatgca ctcaacgatt agcgaccagc cggaatgctt 8640  
ggccagagca tgtatcatat ggtccagaaa ccctatacct gtgtggacgt taatcacttg 8700  
cgattgtgtg gcctgttctg ctactgcttc tgcctctttt tctgggaaga tcgagtgtctc 8760  
tategctagg ggaccaccct ttaaagagat cgcaatctga atcttggttt catttgtaat 8820  
acgctttact agggctttct gctctgtcat ctttgccttc gtttatcttg cctgctcatt 8880  
ttttagtata ttcttcgaag aaatcacatt actttatata atgtataatt cattatgtga 8940  
taatgccaat cgctaagaaa aaaaaagagt catccgctag gtggaaaaaa aaaaatgaaa 9000  
atcattaccg aggcataaaa aaatatagag tgtactagag gaggccaaga gtaatagaaa 9060  
aagaaaattg cgggaaagga ctgtgttatg acttccctga ctaatgcoct gttcaaacga 9120  
tacctggcag tgactcctag cgctcaccaa gctcttaaaa cgggaattatg gtgcactctc 9180  
agtacaatct gctctgatgc cgcatagtta agccagcccc gacacccgcc aacacccgct 9240  
gacgcgccct gacgggcttg tctgctcccg gcatccgctt acagacaagc tgtgaccgtc 9300  
tccgggagct gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc ga 9352

<210> 32

<211> 29

5 <212> ADN

<213> Artificial

<220>

<223> Cebador GPD2uf

10

<400> 32

ggtaccagat ctttgcggc gaggtgccg 29

<210> 33

15 <211> 42

<212> ADN

<213> Artificial

<220>

<223> Cebador GPD2ur

<400> 33

5 tctagactta aggaatgtgt atcttgtaa tcttctgaca gc 42

<210> 34

<211> 28

<212> ADN

10 <213> Artificial

<220>

<223> Cebador GPD2df

15 <400> 34

ctcgagatag tctacaacaa cgtccgca 28

<210> 35

<211> 39

20 <212> ADN

<213> Artificial

<220>

<223> Cebador GPD2dr

25

<400> 35

ccatggagat ctgcagtgaa aaagctcgaa gaaacagct 39

<210> 36

30 <211> 4397

<212> ADN

<213> Artificial

<220>

35 <223> pRN594

<400> 36

ES 2 648 865 T3

gacgacggcg acctggcggg ctctgtggtc gtctcgtact ccggctggaa ccgcccggctg 60  
accgtcgagg acatcgaggt cgccccggag caccgggggc acggggctcg gcgcgcttg 120  
atggggctcg cgacggagtt cgccccggag cggggcgccg ggcacctctg gctggaggtc 180  
accaacgtca acgcaccggc gatccacgcg taccggcgga tggggttcac cctctgcggc 240  
ctggacaccg ccctgtacga cggcaccgcc tcggacggcg agcaggcgct ctacatgagc 300  
atgcctgcc cctagtactg acaataaaaa gattcttgtt ttcaagaact tgtcatttgt 360  
atagtttttt tatattgtag ttgttctatt ttaatcaaat gttagcgtga tttatatattt 420  
ttttgcctc gacatcatct gccagatgc gaagttaagt gcgcagaaag taatatcatg 480  
cgtcaatcgt atgtgaatgc tggtcgctat actgctgtcg attcgatact aacgcgcga 540  
tccagtgtcg acggatccta ggtgtacagg gcccaaaagg gcgaattctg cagatatcca 600  
tcacactggc ggccgctcga gatagtctac aacaacgtcc gcatggaaga cctaccggag 660  
atgattgaag agctagacat cgatgacgaa tagacactct cccccccct cccctctga 720  
tctttcctgt tgcctctttt tcccccaacc aatttatcat tatacacaag ttctacaact 780  
actactagta acattactac agttattata attttctatt ctctttttct ttaagaatct 840  
atcattaacg ttaatttcta tatatacata actaccatta tacacgetat tatcgtttac 900

ES 2 648 865 T3

atatcacatc accgttaatg aaagatacga caccctgtac actaacacaa ttaataatc 960  
 gccataacct tttctgttat ctatagccct taaagctgtt tcttcgagct ttttcaactgc 1020  
 agatctccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg 1080  
 attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggttac 1140  
 ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg tgctttacgg 1200  
 tatcgccgct cccgattcgc agcgcacgc cttctatcgc cttcttgacg agttcttctg 1260  
 aattgaaaaa ggaagagtat gagtattcaa catttcctgt tcgcccttat tccctttttt 1320  
 gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct 1380  
 gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc 1440  
 cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta 1500  
 tgtggcggcg tattatcccg tattgacgcc gggcaagagc aactcggctc ccgcatacac 1560  
 tattctcaga atgacttggg tgagtactca ccagtcacag aaaagcatct tacggatggc 1620  
 atgacagtaa gagaattatg cagtgtgcc ataaccatga gtgataacac tgcggccaac 1680  
 ttacttctga caacgatcgg aggaccgaag gagctaaccg cttttttgca caacatgggg 1740  
 gatcatgtaa ctgccttga tcggtgggaa ccggagctga atgaagccat accaaacgac 1800  
 gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc 1860  
 gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt 1920  
 gcaggaccac ttctgcgctc ggcccttcg gctggctggg ttattgctga taaatctgga 1980  
 gccggtgagc gtgggtctcg cggtatcatt gcagcactgg gccagatgg taagccctcc 2040  
 cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 2100  
 atcgtcgaga taggtgcctc actgattaag cattggtaac tgtcagacca agtttactca 2160  
 tatatacttt agattgattt aaaacttcat ttttaattta aaaggatcta ggtgaagatc 2220  
 ctttttgata atctcatgac caaaatccct taacgtgagt tttcgttcca ctgagcgtca 2280  
 gaccccgtag aaaagatcaa aggatcttct tgagatcctt tttttctgcg cgtaatctgc 2340  
 tgcttgcaaa caaaaaaacc accgctacca gcggtggttt gtttgccgga tcaagagcta 2400  
 ccaactcttt ttccgaaggt aactggcttc agcagagcgc agataccaaa tactgttctt 2460  
 ctagtgtagc cgtagttagg ccaccacttc aagaactctg tagcaccgcc tacatactc 2520  
 gctctgctaa tctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg 2580  
 ttggaactca gacgatagtt accggataag gcgcagcggg cgggctgaa ggggggttcg 2640  
 tgcacacagc ccagcttggg gcgaacgacc tacaccgaac tgagatacct acagcgtgag 2700  
 ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc 2760  
 agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat 2820

ES 2 648 865 T3

agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg 2880  
 gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 2940  
 tggccttttg ctcacatggt ctttctgctg ttatcccctg attctgtgga taaccgtatt 3000  
 accgcctttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca 3060  
 gtgagcgagg aagcgggaaga gcgcccataa cgcaaaccgc ctctcccgcg gcgttggccg 3120  
 attcattaat gcagctggca cgacaggttt cccgactgga aagcgggagc tgagcgcaac 3180  
 gcaattaatg tgagtttagct cactcattag gcaccccagc ctttacctt tatgcttccg 3240  
 gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac 3300  
 catgattacg ccaagctatt taggtgacac tatagaatac tcaagctatg catcaagctt 3360  
 ggtaccagat cttttgcggc gaggtgccga tgggttgctg aggggaagag tgttttagctt 3420  
 acggacctat tgccattggt attccgatta atctattggt cagcagctct tctctacctt 3480  
 gtcattctag tatttttttt tttttttttt ggttttactt ttttttcttc ttgccttttt 3540  
 ttcttgttac tttttttcta gttttttttc cttccactaa gctttttcct tgatttatcc 3600  
 ttgggttctt ctttctactc ctttagattt tttttttata tattaatttt taagtttatg 3660  
 tatttttgga gattcaattc tctttccctt tccttttctt tcgctcccct tccttatcaa 3720  
 tgcttgctgt cagaagatta acaagataca cttccttaa ggcctcgtcc ccgcccgggc 3780  
 acccggccag cgacatggag gcccagaata ccctccttga cagtcttgac gtgcccagct 3840  
 caggggcatg atgtgactgt cgcccgtaca tttagcccat acatcccctt gtataatcat 3900  
 ttgcatccat acattttgat ggccgcacgg cgccaagcaa aaattacggc tcctcgtctc 3960  
 agacctgcga gcagggaaac gctcccctca cagacgcgtt gaattgtccc cacgccgcgc 4020  
 ccctgtagag aaatataaaa ggtaggatt tgccactgag gttcttcttt catatacttc 4080  
 cttttaaaat cttgctagga tacagttctc acatcacatc cgaacataaa caaccatgta 4140  
 aaatgaccac tcttgacgac acggcttacc ggtaccgcac cagtgtcccg ggggacgccg 4200  
 aggccatcga ggcactggat gggtccttca ccaccgacac cgtcttccgc gtcaccgcca 4260  
 ccggggacgg cttcaccctg cgggaggtgc cgggtggacc gccccctgacc aaggtgttcc 4320  
 ccgacgacga atcggacgac gaatcggacg ccggggagga cggcgaccgc gactcccgga 4380  
 cgttcgtcgc gtacggg 4397

<210> 37

<211> 14200

5

<212> ADN

<213> Artificial

<220>

<223> pRN957

10

ES 2 648 865 T3

<400> 37

gacgaaaggg cctcgtgata cgcctatfff tataggftaa tgtcatgata ataagacgtc 60  
aggtggcact ttfcggggaa atgtgcgcgg aaccctatt tgtttatfff tctaaataca 120  
ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 180  
aaggaagagt atgagtattc aacatttccg tgtcgcctt attcctfff ttgcggcatt 240  
ttgccttctt gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 300  
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tctttgagag 360  
ttttcgcccc gaagaacgtt ttccaatgat gagcactfff aaagtctgc tatgtggcgc 420  
ggtattatcc cgtattgacg ccgggcaaga gcaactcggc cgcgcatac actattctca 480  
gaatgacttg gttgagtact caccagtcaac agaaaagcat cttacggatg gcatgacagt 540  
aagagaatta tgcagtgctg ccataacat gatgataaac actgcggcca acttacttct 600  
gacaacgatc ggaggaccga aggagctaac cgtttttttt cacaacatgg gggatcatgt 660  
aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga 720  
caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact 780  
tactctagct tcccggaac aattaataga ctggatggag gcgataaag ttgcaggacc 840  
acttctgcgc tccggccttc cggctggctg gtttattgct gataaatctg gagccggtga 900  
gcgtgggtct cgcggtatca ttgcagcact ggggccagat ggtaagcctt cccgtatcgt 960  
agttatctac acgacgggca gtcaggcaac tatggatgaa cgaaatagac agatcgtctga 1020  
gataggtgcc tcaactgatta agcattggta actgtcagac caagtttact catatatact 1080  
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tcttttttga 1140  
taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccggt 1200  
agaaaagatc aaaggatctt cttgagatcc tttttttctg cgcgtaatct gctgcttgca 1260  
aacaaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 1320  
ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgtga 1380  
gccgtagtta gccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 1440  
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttgactc 1500  
aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca 1560  
gccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agcattgaga 1620  
aagcgcacg cttcccgaag ggagaaaggc ggacaggtat ccggtaaagc gcaggggtcgg 1680  
aacaggagag cgcacgaggg agcttccagg ggggaacgcc tggatatctt atagtcctgt 1740  
cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggccgag 1800  
cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctfff gctggccttt 1860

ES 2 648 865 T3

tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 1920  
 tgagtgagct gataccgctc gccgcagccg aacgaccgag cgcagcaggt cagtgagcga 1980  
 ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 2040  
 atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 2100  
 tgtgagttac ctcaactcatt aggcacccca ggctttacac tttatgcttc cggctcctat 2160  
 gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 2220  
 cgccaagctc ggaattaacc ctactaaag ggaacaaaag ctgggtaccg ggccccccct 2280  
 cgaggtcgag ctaagacgc gtttcttctt cagattccct catggagaaa gtgcggcaga 2340  
 tgtatatgac agagtcgcca gtttccaaga gactttattc aggcacttcc atgataggca 2400  
 agagagaaga cccagagatg ttgttgcct agttacacat ggtatttatt ccagagtatt 2460  
 cctgatgaaa tggtttagat ggacatacga agagttttaa tcgtttacca atgttcctaa 2520  
 cgggagcgtg atggtgatgg aactggacga atccatcaat agatacgtcc tgaggaccgt 2580  
 gctacccaaa tggactgatt gtgagggaga cctaactaca tagtgtttaa agattacgga 2640  
 tatttaactt acttagaata atgccatfff tttgagttat aataatccta cgttagtgtg 2700  
 agcgggattt aaactgtgag gaccttaata cattcagaca cttctgcggt atcaccctac 2760  
 ttattccctt cgagattata tctaggaacc catcaggttg gtggaagatt acccgttcta 2820  
 agacttttca gcttcctcta ttgatgttac acctggacac cccttttctg gcatccagtt 2880  
 tttaatcttc agtggcatgt gagattctcc gaaattaatt aaagcaatca cacaattctc 2940  
 toggatacca cctcggttga aactgacagg tggtttgta cgcagctaa tgcaaaggag 3000  
 cctatatacc tttggctcgg ctgctgtaac agggaatata aagggcagca taatttagga 3060  
 gtttagtgaa ctgcaacat ttactatfff cccttcttac gtaaatafff ttctttttaa 3120  
 ttctaaatca atctttttca attttttgtt tgtattcttt tcttgcttaa atctataact 3180  
 acaaaaaaca catacataa tctagaaaaa tgtctcaatt cttcttcaac cagagaaccc 3240  
 acttggtttc tgacgttatc gacggtgcta tcatcgcttc accatggaac aatttggtta 3300  
 gattggaatc tgaccagct atcagaatcg ttgttagaag agacttgaac aagaacaacg 3360  
 ttgctgttat ctctgggtgt ggttctggtc acgaaccagc tcacgttggg tcatcggtta 3420  
 agggatggt gaccgctgct gtttgggtg acgttttctc ttctccatct gttgacgctg 3480  
 ttttgactgc tatccaagct gttaccggtg aagctgggtg tttgtgatc gttagaact 3540  
 acaccggtga cagattgaac ttcggtttgg ctgctgaaaa ggctagaaga ttgggttaca 3600  
 acgttgaat gttgatcgtt ggtgacgaca tctctttgcc agacaacaag cacccaagag 3660  
 gtatcgctgg taccatcttg gttcacaaga tcgctgggta cttcgctgaa agaggttaca 3720

ES 2 648 865 T3

acttagctac cgttttgaga gaagctcaat acgctgcttc taacaccttc tctttgggtg 3780  
 ttgctttgtc ttcttgcac ttgccacaag aaaccgacgc tgcaccaaga caccacccag 3840  
 gtcacgctga attgggtatg ggtatccacg gtgaaccagg tcttctgtt atcgacacce 3900  
 aaaactctgc tcaagttgtt aacttgatgg ttgacaagtt gttggctgct ttgccagaaa 3960  
 ccggtagatt ggctgttatg atcaacaact tgggtggtgt ttctggtgct gaaatggcta 4020  
 tcatcaccag agaattggct tcttctccat tgcactcaag aatcgactgg ttgatcggtc 4080  
 cagcttcttt ggtaaccgct ttggacatga agggtttctc tttgaccgct atcgttttgg 4140  
 aagaatctat cgaaaaggct ttggtgaccg aagttgaaac ctctaactgg ccaaccccag 4200  
 ttccaccaag agaaatcacc tgtgttgttt cttctcacgc ttctgctaga gttgaattcc 4260  
 aaccatctgc taacgctttg gttgctggta tcggtgaatt ggttaccgct accttgtctg 4320  
 acttggaaac ccactggaac gctttggacg ctaagggtgg tgacgggtgac accggttcta 4380  
 ccttcgctgc tgctgctaga gaaatcgctt ctttgttgca cagacaacaa ttgccattga 4440  
 acaacttggc taccttgttc gctttgatcg gtgaaagatt gaccggtgtt atgggtggtt 4500  
 cttctggtgt tttgatgtct atcttcttca ccgctgctgg tcaaaaagttg gaacaaggtg 4560  
 ctaacgttgt tgaagctttg aacaccggtt tggctcaaat gaagttctac ggtggtgctg 4620  
 acgaagggtg cagaaccatg atcgacgctt tgcaaccagc tttgacctct ttgttggctc 4680  
 aaccaaaaga cttgcaagct gctttcgacg ctgctcaagc tgggtgctgaa agaacctgtt 4740  
 tgtctcttaa ggctaaccgct ggttagagctt cttacttgtc ttctgaatct ttgttgggtg 4800  
 acatggaccc aggtgctcaa agattggcta tggttttcaa ggctttggct gaatctgaat 4860  
 tgggttaata aggtcgagac aaatcgctct taaatatata cctaaagaac attaaagcta 4920  
 tattataagc aaagatacgt aaattttgc t tatattatta tacacatata atatttctat 4980  
 atttttaaga tttggttata taatgtacgt aatgcaaagg aaataaattt tatacattat 5040  
 tgaacagcgt ccaagtaact acattatgtg cactaatagt ttagcgtcgt gaagacttta 5100  
 ttgtgtcgcg aaaagtaaaa attttaaaaa ttagagcacc ttgaacttgc gaaaaaggtt 5160  
 ctcatcaact gtttaaaacg tacgcctagg aagccttcga gcgtcccaa accttctcaa 5220  
 gcaaggtttt cagtataatg ttacatgcgt acacgcgttt gtacagaaaa aaaagaaaa 5280  
 tttgaaatat aaataacggt ctttaacta acataactat taaaaaaat aaatagggac 5340  
 ctagaactca ggttgtctaa ctocctcctt ttcggttaga gcggatgtgg gaggagggcg 5400  
 tgaatgtaag cgtgacataa ctaattacat gatatcgaca aaggaaaaag gccgcgcctt 5460  
 attcccattc ttgtaggaat ctttgaccgt attggtcagc aaccaacaaa gcagcgtaaa 5520  
 cttggtctgg ggtagacca cctggcatgt tgtggatggt ttcacctca gcacaagcag 5580  
 cttcagcaac gattctcatc ttagctggaa cgtcttcctt gatgtccaat tgagccaagg 5640

ES 2 648 865 T3

tgattggcaa accaacagcg tgagacaaag cagcaacggt ttcgatttct tcaactggag 5700  
 cgttttccaa aaccaattgg gtcaaggtac cgaaagcaac cttttcaccg tggtagtagt 5760  
 ggtgagcgtc tgggatagcg gtcaaaccgt tgtgaacagc gtgagcagca gccaaaccac 5820  
 cagattcgaa accaacacca gacaagtagg tgtagcttc gataactctt tccaaagctg 5880  
 gggtaacaac gtgttgttca gcagccaaca tagcctttc accttcttcc aacaaggtgt 5940  
 tgtaacacia ttcagccaaa gccaaagcag cttgggtaca cttaccacca gccatggtgg 6000  
 tagcaccaga tctagaacia gctctagctt cgaaccaggt agccaaagcg tcaccgatac 6060  
 cagctgccaa caatctagct ggagcaccag caacgatctt ggtgtcaacg ataaccatgt 6120  
 ttgggttgtt tggcaacaac aagtatctgt cgaattcacc ttcgtcgggtg tagataacag 6180  
 acaaagcaga acatggagcg tcggtagaag cgatggttgg agcgatagca actggaacac 6240  
 ccatgaagtg agccaaagcc ttagcgggtg ccaaggtctt accaccaccg atacccaaga 6300  
 tagcaccaca ttgagcgggt tcagcgatac ctctcaatct gtcgatttctg ttttgagaac 6360  
 attcaccacc gaatggagcg atttcaacia ccaaaccagc gtccttgaaa gacttttcaa 6420  
 cggtagattg agcgaaacc aaacgaact tgtcaccaac aaccaaccat ctttcagcca 6480  
 atggctcaa gtattcacc aatctgttga taacgtcagc accttggatg tacttacctg 6540  
 gagattgat gattctgtcc atttctgca gttaatcag taaatttctg atcttgggaa 6600  
 gaaaaagca gtaagcgtga aaaatctaaa agctgatgta gtagaagatc ctattcttta 6660  
 acaaagattg accttttctt tttcttcttg gtttgagtag aaaggggaag gaagaataca 6720  
 agagagagga aaaaaaggaa gataaaaaga gagcgtgata taaatgaata tatattaaac 6780  
 aagagagatt ggaaggaaa ggatcaaaca aacccaaaaa tatttcaaaa aggagagaga 6840  
 gaggcgagtt tggtttcaaa acggtttatt tatttatgca agaggacgtg gaagaaaaag 6900  
 aagaaggag aaaaaaattt gaaagaaaa aacgcgtggc gggtaagaa gaaaatggaa 6960  
 aatagaggcc gggtgacaga gaaatattga gggtaattg gaaaatatgt tagggtgagg 7020  
 catatgttt taagggttt gaggatccga taaggaagaa ttaggttaa atgttgtgca 7080  
 ttaattgctg tggcagctta cccgcttccc cacacattta ctagtctcga gctcttcaac 7140  
 tcaagacgca cagatattat aacatctgca taataggcat ttgcaagaat tactcgtgag 7200  
 taaggaaaga gtgaggaact atcgatacc tgcatttaa gatgccgatt tgggcgcgaa 7260  
 tcctttattt tggcttacc ctcatactat tatcagggcc agaaaagga agtgtttccc 7320  
 tccttcttga attgatgta ccctcataaa gcacgtggcc tcttatcgag aaagaaatta 7380  
 cgcctcctcg tgatttgttt gcaaaaagaa caaaactgaa aaaaccaga cacgctcgac 7440  
 ttctgtctt cctattgatt gcagcttcca atttctcac acaacaaggt cctagcagcg 7500

ES 2 648 865 T3

gctcacaggt tttgtaacaa gcaatcgaag gttctggaat ggcgggaaag ggtttagtac 7560  
cacatgctat gatgccact gtgatctcca gagcaaagt cgttcgatcg tactgttact 7620  
ctctctcttt caaacagaat tgtccgaatc gtgtgacaac aacagcctgt tctcacacac 7680  
tcttttcttc taaccaaggg ggtggtttag tttagtagaa cctcgtgaaa cttacattta 7740  
catatatata aacttgcata aattggtaa tgcaagaaat acatatttgg tcttttctaa 7800  
ttcgtagttt ttcaagtct tagatgctt cttttctct tttttacaga tcatcaagga 7860  
agtaattatc tactttttac aacaaatata tctagaaaat ggctgttacc aacgttgctg 7920  
aattgaacgc tttggttgaa agggttaaga aggotcaaag agaatacgt tctttcacc 7980  
aagaacaagt tgacaagatc tcagagctg ctgctttggc tgctgctgac gctagaatcc 8040  
cattggctaa gatggctgt gctgaatctg gtatgggtat cgttgaagac aaggttatca 8100  
agaaccactt cgcttctgaa tacatctaca acgcttaca ggacgaaaag acctgtggtg 8160  
tttgtcaga agacgacacc ttcggtacca tcaccatcgc tgaaccaatc ggtatcatct 8220  
gtggtatcgt tccaaccacc aaccaacct ctaccgctat cttcaagtct ttgatctctt 8280  
tgaagaccag aaacgctatc atcttctctc cacaccaag agctaaagac gctaccaaca 8340  
aggctgctga catcgtttt caagctgcta tcgctgctgg tgctccaaag gacttgatcg 8400  
gttgatcga ccaaccatct gttgaattgt ctaacgctt gatgcaccac ccagacatca 8460  
acttgatctt ggctaccggt ggtccaggta tggttaaggc tgcttactct tctggttaagc 8520  
cagctatcgg tgttggtgct ggtaacacc cagttgttat cgacgaaacc gctgacatca 8580  
agagagctgt tgcttctggt ttgatgtcta agacctcga caacggtggt atctgtgctt 8640  
ctgaacaatc tgttggtggt gttgactctg tttacgaagc tgttagagaa agattcgtca 8700  
cccacggtg ttacttgtt caaggtaagg aattgaaggc tgttcaagac gttatcttga 8760  
agaacggtgc tttgaaacgt gctatcgttg gtcaaccagc ttacaagatc gctgaattag 8820  
ctggtttctc tgttccagaa aacaccaaga tcttgatcgg tgaagttacc gttgttgacg 8880  
aatctgaacc attcgtcac gaaaagttgt ctccaacctt ggctatgtac agagctaagg 8940  
acttcaaga cgctgttgaa aaagctgaaa agttggtgct tatgggtggt attggtcaca 9000  
cctcttgtt gtacaccgac caagacaacc aaccagctag agtttcttac ttcggtcaaa 9060  
agatgaagac cgctagaatc ttgatcaaca cccagcttc tcaaggtggt atcggtgact 9120  
tgtacaactt caagttggct ccatcttga ccttgggttg tggttcttgg ggtggttaact 9180  
ctatctctga aaacgttgg ccaaacgact tgatcaacaa gaagaccgtt gctaagagag 9240  
ctgaaaacat gttgtggcac aagttgcaa aatctatcta cttcagaaga gttctttgc 9300  
caatcgctt ggacgaagt atcaccgac gtcaacaag agctttgatc gttaccgaca 9360  
gattcttgtt caacaacggt tacgctgacc aatcacctc tgttttgaag gctgctggtg 9420

ES 2 648 865 T3

ttgaaaccga agttttcttc gaagttgaag ctgacccaac cttgtctatc gttagaaagg 9480  
 gtgctgaatt ggctaactct ttcaagccag acgttatcat cgctttgggt ggtggttctc 9540  
 caatggacgc tgctaagatc atgtgggta tgtagaaca cccagaaacc cacttcgaag 9600  
 aattggcttt gagattcatg gacatcagaa agagaatcta caagttccca aagatgggtg 9660  
 ttaaggctaa gatgatcgct gttaccacca cctctggtac cggttctgaa gttaccccat 9720  
 tcgctgttgt taccgacgac gctaccggtc aaaagtacc attggetgac tacgctttga 9780  
 cccagacat ggctatcggt gacgctaact tggttatgga catgccaaag tctttgtgtg 9840  
 ctttcggtgg tttggacgct gttaccaccg ctatggaagc ttacgtttct gttttggctt 9900  
 ctgaattctc tgacgggtcaa gctttgcaag ctttgaagtt gttgaaggaa tacttgccag 9960  
 cttcttacca cgaaggttct aagaaccag ttgctagaga aagagttcac tctgctgcta 10020  
 ccatcgctgg tatcgctttc gctaacgctt tcttgggtgt ttgtcactct atggctcaca 10080  
 agttgggttc tcaattccac atcccacag gtttggctaa cgctttgttg atctgtaacg 10140  
 ttatcagata caacgctaac gacaacccaa ccaagcaaac cgctttctct caatacgaca 10200  
 gaccacaagc tagaagaaga tacgctgaaa tcgctgacca cttgggtttg tctgctccag 10260  
 gtgacagaac cgctgcaaag atcgaaaagt tgttggcttg gttggaacc ttgaaggctg 10320  
 aattgggtat cccaaagtct atcagagaag ctggtgttca agaagctgac ttcttggcta 10380  
 acgttgaaa gttgtctgaa gacgctttcg acgaccaatg taccggtgct aacccaagat 10440  
 accattgat ctctgaattg aagcaaatct tgttggacac ctactacggt agagactacg 10500  
 ttgaaggaga aaccgctgct aagaaggaag ctgctccagc taaggctgaa aagaaggcta 10560  
 agaagtctgc ttagcttaag cgcggaatt tcttatgatt tatgatTTTT attattaaat 10620  
 aagttataaa aaaaataagt gtatacaaat tttaaagtga ctcttaggtt ttaaaacgaa 10680  
 aattcttatt cttgagtaac tctttctgt aggtcaggtt gctttctcag gtatagcatg 10740  
 aggtcgctct tattgaccac acctctaccg gcatgccgag caaatgcctg caaatcgctc 10800  
 cccattcac ccaattgtag atatgctaac tccagcaatg agttgatgaa tctcgggtg 10860  
 tattttatgt cctcagagga caacacctgt tgtaatcgtt cttccacacg tacgaagctt 10920  
 aaaagggcga atctgcaga tatccatcac actggcgcc gcatgctagc tccggattat 10980  
 cgatgataag ctgtcaaaga tgagaattaa ttccacggac tatagactat actagatact 11040  
 ccgtctactg tacgatacac ttccgctcag gtcttctcc ttaacgagg ccttaccact 11100  
 cttttgttac tctattgatc cagctcagca aaggcagtgt gatctaagat tctatctctg 11160  
 cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg cacttctaca 11220  
 atggctgcca tcattattat ccgatgtgac gctgcagctt ctcaatgata ttogaatacg 11280

ES 2 648 865 T3

ctttgaggag atacagccta atatccgaca aactgtttta cagatttacg atcgtacttg 11340  
 ttacccatca ttgaattttg aacatccgaa cctgggagtt ttccctgaaa cagatagtat 11400  
 atttgaacct gtataataat atatagtcta gcgcttttacg gaagacaatg tatgtatttc 11460  
 ggttcctgga gaaactattg catctattgc ataggtaato ttgcacgtcg catccccggt 11520  
 tcattttctg cgtttccatc ttgcacttca atagcatatc tttgttaacg aagcatctgt 11580  
 gcttcatttt gtagaacaaa aatgcaacgc gagagcgcta atttttcaa caaagaatct 11640  
 gagctgcatt ttacagaac agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat 11700  
 ctgtgcttca tttttgtaa acaaaaatgc aacgcgacga gagcgcta at tttcaaaca 11760  
 aagaatctga gctgcatttt tacagaacag aaatgcaacg cgagagcgct attttaccaa 11820  
 caaagaatct atacttcttt tttgtttctac aaaaatgcat cccgagagcg ctatttttct 11880  
 aacaaagcat cttagattac tttttttctc ctttgtgcgc tctataatgc agtctcttga 11940  
 taactttttg cactgtaggt ccgttaaggt tagaagaagg ctacttttgt gtctattttc 12000  
 tcttcataa aaaaagcctg actccacttc ccgcgtttac tgattactag cgaagctgcg 12060  
 ggtgcatttt ttcaagataa aggcattccc gattatattc tataccgatg tggattgcgc 12120  
 atactttgtg aacagaaagt gatagcgttg atgattcttc attggtcaga aaattatgaa 12180  
 cggtttcttc tattttgtct ctatatacta cgtataggaa atgtttacat tttcgtattg 12240  
 ttttcgattc actctatgaa tagttcttac tacaattttt ttgtctaaag agtaatacta 12300  
 gagataaaca taaaaaatgt agaggtcgag tttagatgca agttcaagga gcgaaaggtg 12360  
 gatgggtagg ttatataggg atatagcaca gagatatata gcaaagagat acttttgagc 12420  
 aatgtttgtg gaagcggat tcgcaatgcc ggctttcccc gtcaagctct aaatcggggg 12480  
 ctcccttag ggttcogatt tagtgctta cggcacctcg accccaaaaa acttgattag 12540  
 ggtgatgggt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc tttgacgttg 12600  
 gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc 12660  
 tcggtctatt cttttgattt ataagggatt ttgocgattt cggcctattg gttaaaaaat 12720  
 gagctgattt aacaaaaatt taacgcgaat ttaacaaaa tattaacgtt tacaatttcc 12780  
 tgatcgggta ttttctcctt acgcatctgt gcggtatttc acaccgcata tgatccgtcg 12840  
 agttcaagag aaaaaaaaaag aaaaagcaaa aagaaaaaag gaaagcgcgc ctcgttcaga 12900  
 atgacacgta tagaatgatg cattaccttg tcatcttcag tatcatactg ttcgtataca 12960  
 tacttactga cattcatagg tatacatata tacacatgta tatatatcgt atgctgcagc 13020  
 tttaaataat cgggtgcact acataagaac acctttgggt gagggaacat cgttgggtacc 13080  
 attgggcgag gtggtctctc ttatggcaac cgcaagagcc ttgaacgcac tctcactacg 13140  
 gtgatgatca ttcttgccct gcagacaatc aacgtggagg gtaattctgc tagcctctgc 13200

ES 2 648 865 T3

aaagctttca agaaaatgcg ggatcatctc gcaagagaga tctcctactt tctccctttg 13260  
 caaaccaagt tcgacaactg cgtacggcct gttcgaaaga tctaccaccg ctctggaaag 13320  
 tgcctcatcc aaaggcgcaa atcctgatcc aaaccttttt actccacgcy ccagtagggc 13380  
 ctctttaaaa gcttgaccga gagcaatccc gcagtcttca gtggtgtgat ggtcgtctat 13440  
 gtgtaagtca ccaatgcact caacgattag cgaccagccg gaatgcttgg ccagagcatg 13500  
 tatcatatgg tccagaaacc ctatacctgt gtggacgtta atcacttggc attgtgtggc 13560  
 ctgttctgct actgcttctg cctctttttc tgggaagatc gagtgctcta tcgctagggg 13620  
 accacccttt aaagagatcg caatctgaat cttggtttca tttgtaatac gctttactag 13680  
 ggctttctgc tctgtcatct ttgccttcgt ttatcttggc tgctcatttt ttagtatatt 13740  
 cttogaagaa atcacattac tttatataat gtataattca ttatgtgata atgccaatcg 13800  
 ctaagaaaaa aaaagagtca tccgctaggt ggaaaaaaaa aatgaaaat cattaccgag 13860  
 gcataaaaaa atatagagtg tactagagga ggccaagagt aatagaaaaa gaaaattgcy 13920  
 ggaaaggact gtgttatgac ttccctgact aatgcccgtt tcaaacgata cctggcagtg 13980  
 actcctagcy ctcaccaagc tcttaaaacg gaattatggt gcactctcag tacaatctgc 14040  
 tctgatgccg catagttaag ccagccccga caccgccea caccgctga cgcgccctga 14100  
 cgggcttgtc tgctcccggc atccgcttac agacaagctg tgaccgtctc cgggagctgc 14160  
 atgtgtcaga ggttttcacc gtcacaccg aaacgcgcga 14200

<210> 38

<211> 3042

5

<212> ADN

<213> Artificial

<220>

<223> pRNDAK

10

<400> 38

ES 2 648 865 T3

ggatccacta gtaacggccg ccagtgtgct ggaattgcc ettctcgagc ttaagacgcg 60  
tttcttcttc agattccctc atggagaaag tgcggcagat gtatatgaca gagtcgccag 120  
tttccaagag actttatca ggcacttcca tgataggcaa gagagaagac ccagagatgt 180  
tgttgtccta gttacacatg gtatttattc cagagtattc ctgatgaaat ggtttagatg 240  
gacatacgaa gagtttgaat cgtttaccaa tgttcctaac gggagcgtaa tggatgatga 300  
actggacgaa tccatcaata gatacgtcct gaggaccgtg ctacccaaat ggactgattg 360  
tgaggagac ctaactacat agtgtttaa gattacggat atttaactta cttagaataa 420  
tgccatttt ttgagttata ataatcctac gttagtgtga gcgggattta aactgtgagg 480  
acctaatac attcagacac ttctgcggtg tcaccctact tattcccttc gagattatat 540

ES 2 648 865 T3

ctaggaaccc atcaggttgg tggaagatta cccggttetaa gacttttcag ctctctctat 600  
tgatgttaca cctggacacc cctttcttgg catccagttt ttaatcttca gtggcatgtg 660  
agattctccg aaattaatta aagcaatcac acaattctct cggataccac ctcggttgaa 720  
actgacaggt ggtttgttac gcatgctaat gcaaaggagc ctatatacct ttggctcggc 780  
tgctgtaaca gggaatataa agggcagcat aatntagtag tttagtgaac ttgcaacatt 840  
tactattttc ccttcttacg taaatatttt tctttttaat tctaaatcaa tctttttcaa 900  
tttttggtt gtattctttt ctgcttaaa tctataacta caaaaaacac atacataaat 960  
ctagaaaatg tccgctaaat cgtttgaagt cacagatcca gtcaattcaa gtctcaaagg 1020  
gtttgccctt gctaaccctt ccattacgct ggtccctgaa gaaaaaatc tcttcagaaa 1080  
gaccgattcc gacaagatcg cattaatttc tgggtgtgtg agtggacatg aacctacaca 1140  
cgccggtttc attggttaagg gtatgttgag tggcgccgtg gttggcgaaa tttttgcatc 1200  
cccttcaaca aaacagattt taaatgcaat ccgtttagtc aatgaaaatg cgtctggcgt 1260  
tttattgatt gtgaagaact acacaggtga tgttttgcac tttggtctgt ccgctgagag 1320  
agcaagagcc ttgggtatta actgccgcgt tgctgtcata ggtgatgatg ttgcagttgg 1380  
cagagaaaag ggtggtatgg ttggtagaag agcattggca ggtaccgtt tggttcataa 1440  
gattgtaggt gccttcgcag aagaatattc tagtaagtat ggcttagacg gtacagctaa 1500  
agtggctaaa attatcaacg acaatttggg gaccattgga tcttctttag accattgtaa 1560  
agttcctggc aggaaattcg aaagtgaatt aaacgaaaaa caaatggaat tgggtatggg 1620  
tattcataac gaacctggtg tgaaagtttt agacctatt ccttctaccg aagacttgat 1680  
ctccaagtat atgctaccaa aactattgga tccaaacgat aaggatagag cttttgtaaa 1740  
gtttgatgaa gatgatgaag ttgtcttgtt agttaacaat ctcggcggtg tttctaattt 1800  
tgttattagt tctatcactt ccaaaactac ggatttctta aaggaaaatt acaacataac 1860  
cccgttcaa acaattgctg gcacattgat gacctcttc aatggtaatg ggttcagtat 1920  
cacattacta aacgccacta aggctacaaa ggctttgcaa tctgattttg aggagatcaa 1980  
atcagtacta gacttgttga acgcatttac gaacgcaccg ggctggccaa ttgcagattt 2040  
tgaaaagact tctgccccat ctgttaacga tgacttgta cataatgaag taacagcaaa 2100  
ggccgtcggg acctatgact ttgacaagtt tgctgagtgg atgaagagt gtgctgaaca 2160  
agttatcaag agcgaaccgc acattacgga actagacaat caagttggtg atggtgattg 2220  
tggttacact ttagtggcag gagttaaagg catcaccgaa aacctgaca agctgtcgaa 2280  
ggactcatta tctcagggcg ttgccc aaat ttcagatttc attgaaggct caatggggag 2340  
tacttctggg ggtttatatt ctattctttt gtcgggtttt tcacacggat taattcaggt 2400  
ttgtaaatca aaggatgaac ccgtcactaa ggaaattgtg gctaagtcac tcggaattgc 2460

ES 2 648 865 T3

```
attggatact ttatacaaat atacaaaggc aaggaaggga tcatccacca tgattgatgc 2520
tttagaacca ttcgttaaag aatttactgc atctaaggat ttcaataagg cggtaaaagc 2580
tgcagaggaa ggtgctaaat ccaactgctac attcgaggcc aaatttgca gagcttcgta 2640
tgtcggcgat tcatctcaag tagaagatcc tgggtgcagta ggcctatgtg agtttttgaa 2700
gggggttcaa agcgccttgt aagtcgagac aaatcgctct taaatatata cctaaagaac 2760
attaaagcta tattataagc aaagatacgt aaattttgc t tatattatta tacacatatc 2820
atatttctat atttttaaga tttggttata taatgtacgt aatgcaaagg aaataaattt 2880
tatacattat tgaacagcgt ccaagtaact acattatgtg cactaatagt ttagcgtcgt 2940
gaagacttta ttgtgtcgcg aaaagtaaaa attttaaaaa ttagagcacc ttgaacttgc 3000
gaaaaagggtt ctcatcaact gtttaaaacg tacgaagctt aa 3042
```

<210> 39

<211> 14291

5

<212> ADN

<213> Artificial

<220>

<223> pRN977

10

<400> 39

ES 2 648 865 T3

gacgaaaggg cctcgtgata cgcctatfff tataggfmaa tgfcatgata ataagacgfc	60
aggtggcact fttcggggaa atgtgcgcgg aaccctatt tgfthatttt tctaaataca	120
ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa	180
aaggaagagt atgagtattc aacatftccg tgtcgcctt attccctfff ttgoggcatt	240
ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca	300
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag	360
tttcgcccc gaagaacgtt ttccaatgat gagcactfff aaagttctgc tatgtggcgc	420
ggtattatcc cgtattgacg ccgggcaaga gcaactcggf cggccatac actattctca	480
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt	540
aagagaatta tgcagtgctg ccataacct gatgataac actgcggcca acttacttct	600
gacaacgatc ggaggaccga aggagctaac cgctfttttt cacaacatgg gggatcatgt	660
aactcgcctt gatcgttggg aaccggagct gaatgaagcc ataccaaacg acgagcgtga	720
caccacgatg cctgtagcaa tggcaacaac gttgcgcaaa ctattaactg gcgaactact	780
tactctagct tcccggcaac aattaataga ctggatggag gcggataaag ttgcaggacc	840
acttctgcgc tcggcccttc cggctggctg gtttattgct gataaatctg gagccggtga	900
gcgtgggtct cggggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt	960

ES 2 648 865 T3

agttatctac acgacgggca gtcaggcaac tatggatgaa cgaatatagac agatcgctga 1020  
 gataggtgcc tcaactgatta agcatttqgta actgtcagac caagtttact catatatact 1080  
 ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tcctttttga 1140  
 taatctcatg accaaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 1200  
 agaaaaagatc aaaggatcct cttgagatcc ttttttctg cgcgtaatct gctgcttgca 1260  
 aacaaaaaaa ccaccgctac cagcgggtgt ttgtttgccg gatcaagagc taccaactct 1320  
 ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgt 1380  
 gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 1440  
 aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggtactc 1500  
 aagacgatag ttaccggata aggcgcagcg gtcgggctga acggggggtt cgtgcacaca 1560  
 gccagcctt gagcgaacga cctacaccga actgagatac ctacagcgtg agcattgaga 1620  
 aagcgcacg cttcccgaag ggagaaaggc ggacaggtat ccggaagcg gcagggtcgg 1680  
 aacaggagag cgcacgaggg agcttcacag gggaacgcc tggatcttt atagtcctgt 1740  
 cgggtttcgc cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggccgag 1800  
 cctatggaaa aacgccagca acgcggcctt tttacggttc ctggcctttt gctggccttt 1860  
 tgctcacatg ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt 1920  
 tgagttagct gataccgctc gccgcagccg aacgaccgag cgcagcaggt cagttagcga 1980  
 ggaagcggaa gagcgcocaa tacgcaaacc gcctctcccc gcgcggttggc cgattcatta 2040  
 atgcagctgg cacgacaggt ttcccgaact gaaagcgggc agtgagcga acgcaattaa 2100  
 tgtgagttac ctcaactcatt aggcacccca ggctttacac tttatgcttc cggtcctat 2160  
 gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 2220  
 cgccaagctc ggaattaacc ctcaactaaag ggaacaaaag ctgggtaccg ggccccccct 2280  
 cgaggtcgag cttaagacgc gtttctctt cagattccct catggagaaa gtgcggcaga 2340  
 tgtatatgac agagtcgcca gtttccaaga gactttatc aggcacttcc atgataggca 2400  
 agagagaaga cccagagatg ttgttgcct agttacacat ggtatttatt ccagagtatt 2460  
 cctgatgaaa tggtttagat ggacatacga agagtttgaa tcgtttacca atgttcctaa 2520  
 cgggagcgtg atggtgatgg aactggacga atccatcaat agatacgtcc tgaggaccgt 2580  
 gctacccaaa tggactgatt gtgagggaga cctaactaca tagtgttta agattacgga 2640  
 tatttaactt acttagaata atgccatfff tttgagttat aataatccta cgttagtgtg 2700  
 agcgggattt aaactgtgag gaccttaata cattcagaca cttctgcggt atcacccctac 2760  
 ttatccctt cgagattata tctaggaacc catcaggttg gtggaagatt acccgttcta 2820  
 agacttttca gcttctcta ttgatgttac acctggacac cccttttctg gcatccagtt 2880

ES 2 648 865 T3

tttaatcttc agtggcatgt gagattctcc gaaattaatt aaagcaatca cacaattctc 2940  
 tccgatacca cctcggttga aactgacagg tggtttgta cgcattgtaa tgcaaaggag 3000  
 cctatatacc ttggctcgg ctgctgtaac agggaatata aagggcagca taatttagga 3060  
 gtttagtgaa cttgcaacat ttactatfff cccttcttac gtaaatatff ttctttttaa 3120  
 ttctaaatca atctttttca attttttggt tgtattcttt tcttgcttaa atctataact 3180  
 acaaaaaaca catacataa tctagaaaat gtccgctaaa tcgtttgaag tcacagatcc 3240  
 agtcaattca agtctcaaag ggtttgcct tgcatacccc tccattacgc tggccctga 3300  
 agaaaaaatt ctcttcagaa agaccgattc cgacaagatc gcattaatff ctgggtggtg 3360  
 tagtggacat gaacctacac acgccggttt cattggtaag ggtatggtga gtggcgccgt 3420  
 ggttggcgaa atttttgcat ccccttcaac aaaacagatt ttaaatgcaa tccgtttagt 3480  
 caatgaaaat gcgtctggcg ttttattgat tgtgaagaac tacacagggtg atgttttga 3540  
 ttttggctg tccgctgaga gagcaagagc cttgggtatt aactgccgcg ttgctgtcat 3600  
 aggtgatgat gttgcagttg gcagagaaaa ggggtgatg gttggtagaa gagcattggc 3660  
 aggtaccggt ttggttcata agattgtagg tgccttcgca gaagaatatt ctagtaagta 3720  
 tggcttagac ggtacagcta aagtggctaa aattatcaac gacaatttgg tgaccattgg 3780  
 atcttcttta gaccattgta aagttcctgg caggaaatc gaaagtgaat taaacgaaaa 3840  
 acaaatggaa ttgggtatgg gtattcataa cgaacctggt gtgaaagttt tagaccctat 3900  
 tccttctacc gaagacttga tctccaagta tatgctacca aactatttgg atccaaacga 3960  
 taaggataga gcttttgtaa agtttgatga agatgatgaa gttgtcttgt tagttaacaa 4020  
 tctcggcggg ttttctaatt ttgttattag ttctatcact tccaaaacta cggatttctt 4080  
 aaaggaaaat tacaacataa ccccgttca aacaattgct ggcacattga tgacctctt 4140  
 caatggtaat gggttcagta tcacattact aaacgccact aaggctacaa aggctttgca 4200  
 atctgatfff gaggatgca aatcagtact agacttgtt gacgcattta cgaacgcacc 4260  
 gggctggcca attgcagatt ttgaaaagac ttctgcccc aactgttaacg atgacttgtt 4320  
 acataatgaa gtaacagcaa aggccgtcgg tacctatgac tttgacaagt ttgctgagtg 4380  
 gatgaagagt ggtgctgaac aagttatcaa gagcgaaccg cacattacgg aactagacaa 4440  
 tcaagttggt gatggtgatt gtggttacac tttagtggca ggagttaaag gcatcaccga 4500  
 aaacctgac aagctgtcga aggactcatt atctcaggcg gttgccccaa tttcagatff 4560  
 cattgaaggc tcaatgggag gtacttctgg tggtttatat tctattcttt tgtcgggttt 4620  
 ttcacacgga ttaattcagg tttgtaaatc aaaggatgaa cccgtcacta aggaaattgt 4680  
 ggctaagtca ctccgaattg cattggatac tttatacaaa tatacaaagg caaggaaggg 4740

ES 2 648 865 T3

atcatccacc atgattgatg ctttagaacc attcgttaaa gaatttactg catctaagga 4800  
 tttcaataag gcggtaaaag ctgcagagga aggtgctaaa tccactgcta cattcgaggc 4860  
 caaatTTGGC agagcttCGT atgtcggcga ttcatctcaa gtagaagatc ctggtgcagt 4920  
 aggccatagt gagtttttga agggggttca aagcgccttg taagtcgaga caaatcgctc 4980  
 ttaaataat acctaaagaa cattaagact atattataag caagatacg taaatTTTGC 5040  
 ttataattatt atacacatat catatttcta tatttttaag atttggttat ataatgtacg 5100  
 taatgcaaag gaaataaatt ttatacatta ttgaacagcg tccaagtaac tacattatgt 5160  
 gcactaatag ttttagcgtcg tgaagacttt attgtgtcgc gaaaagtaaa aattttaaaa 5220  
 attagagcac cttgaacttg cgaaaaaggt tctcatcaac tgtttaaaaac gtacgcctag 5280  
 gaagccttcg agcgtcccaa aaccttctca agcaaggttt tcagtataat gttacatgcg 5340  
 tacacgcggt tgtacagaaa aaaaagaaaa atttgaaata taaataacgt tcttaatact 5400  
 aacataacta ttaaaaaaaa taaataggga cctagacttc agttgtcta actccttctc 5460  
 tttcggttag agcggatgtg ggaggagggc gtgaatgtaa gcgtgacata actaattaca 5520  
 tgatatcgac aaaggaaaag gggcgcgcct tattccatt cttgtaggaa tctttgaccg 5580  
 tattggtcag caaccaacaa agcagcgtaa acttggctcg ggttagcacc acctggcatg 5640  
 ttgtggatg tttcaccttc agcacaagca gcttcagcaa cgattctcat cttagctgga 5700  
 acgtcttctc tgatgtccaa ttgagccaag gtgattggca aaccaacagc gtgagacaaa 5760  
 gcagcaacgg tttcgatttc ttcaactgga gcgttttcca aaaccaattg ggtcaaggta 5820  
 ccgaaagcaa ccttttcacc gtggtagtag tgggtgagcgt ctgggatagc ggtcaaaccg 5880  
 ttgtgaacag cgtgagcagc agccaaacca ccagattoga aaccaacacc agacaagtag 5940  
 gtgttagctt cgataactct ttocaaagct ggggtaacaa cgtgttcttc agcagccaac 6000  
 atagcctttt caccttcttc caacaagggt ttgtaacaca attcagcca agccaaagca 6060  
 gcttgggtac acttaccacc agccatggtg gtagcaccag atctagaaca agctctagct 6120  
 tcgaaccag tagccaaagc gtcaccgata ccagctgcca acaatctagc tggagcacca 6180  
 gcaacgatct tgggtgtcaac gataaccatg tttgggttgt ttggcaacaa caagtatctg 6240  
 tcgaattcac cttcgtcggg ttagataaca gacaaagcag aacatggagc gtcggtagaa 6300  
 gcgatggttg gagcgatagc aactggaaca cccatgaagt gagccaaagc cttagcggtg 6360  
 tccaaggtct taccaccacc gataccaag atagcaccac attgagcggg ttcagcgata 6420  
 cctctcaatc tgtcgatttc gttttgagaa cattcaccac cgaatggagc gatttcaaca 6480  
 accaaaaccg cgtccttgaa agacttttca acggtagatt gagcgaacc caaacgaac 6540  
 ttgtcaccaa caaccaacca tctttcagcc aatggcttca agtattcacc caatctgttg 6600  
 ataacgtcag caccttggat gtacttaact ggagattgga tgattctgtc cattttctgc 6660

ES 2 648 865 T3

agttaattca gtaaattttc gatcttggga agaaaaaagc agtaagcgtg aaaaatctaa 6720  
 aagctgatgt agtagaagat cctattcttt aacaaagatt gaccttttct ttttcttctt 6780  
 ggtttgagta gaaaggggaa ggaagaatac aagagagagg aaaaaaagga agataaaaag 6840  
 agagcgtgat ataaatgaat atatattaaa caagagagat tgggaaggaa aggatcaaac 6900  
 aaacccaaaa atatttcaaa aaggagagag agaggcgagt ttggtttcaa aacggtttat 6960  
 ttatttatgc aagaggacgt ggaagaaaaa gaagaaggaa gaaaaaaatt tgaaagaaaa 7020  
 aaacgcgtgg cgggtaaaga agaaaatgga aatagaggc cgggtgacag agaaatattg 7080  
 agggttaatt gaaaaatag ttagggtag gcatatgtt ttaagggtt tgaggatccg 7140  
 ataaggaaga atgtaggta aatgttgtgc attaattgct gtggcagctt acccgcttcc 7200  
 ccacacattt actagtctcg agctcttcaa ctcaagacgc acagatatta taacatctgc 7260  
 ataataggca ttgcaagaa ttactcgtga gtaaggaaag agtgaggaac tatcgcatac 7320  
 ctgcatttaa agatgccgat ttgggcgga atcctttatt ttggcttcac cctcatacta 7380  
 ttatcagggc cagaaaaag aagtgttcc ctcttcttg aattgatgtt accctcataa 7440  
 agcacgtggc ctcttatcga gaaagaaatt accgtcgcctc gtgatttgtt tgcaaaaaga 7500  
 acaaaactga aaaaaccag acacgcctga ctctctgtct tcctattgat tgcagcttcc 7560  
 aatttcgtca cacaacaagg tcctagcgc ggcacacag ttttgtaaca agcaatcgaa 7620  
 ggttctggaa tggcgggaaa gggtttagta ccacatgcta tgatgccac tgtgatctcc 7680  
 agagcaaagt tcgttcgatc gtactgttac tctctctct tcaaacagaa ttgtccgaat 7740  
 cgtgtgacaa caacagcctg ttctcacaca ctcttttct ctaaccaagg ggggtggtta 7800  
 gtttagtaga acctcgtgaa acttacattt acatataat aaacttgcac aaattggtca 7860  
 atgcaagaaa tacatatttg gtcttttcta attcgtagt tttcaagttc ttagatgctt 7920  
 tcttttctc ttttttacag atcatcaagg aagtaattat ctacttttta caacaaatat 7980  
 atctagaaaa tggctgttac caacgttgc gaattgaacg ctttggttga aagggttaag 8040  
 aaggctcaaa gagaatacgc ttctttcacc caagaacaag ttgacaagat cttcagagct 8100  
 gctgctttgg ctgctgctga cgctagaatc ccattggcta agatggctgt tgctgaatct 8160  
 ggtatgggta tcgttgaaga caaggttatc aagaaccact tcgcttctga atacatctac 8220  
 aacgcttaca aggacgaaaa gacctgtggt gttttgtcag aagacgacac cttcgggtacc 8280  
 atcaccatcg ctgaaccaat cggatcatc tgtggtatcg ttccaaccac caaccaacc 8340  
 totaccgcta tcttcaagtc tttgatctct ttgaagacca gaaacgctat catcttctct 8400  
 ccacacccaa gagctaaaga cgctaccaac aaggctgctg acatcgtttt gcaagctgct 8460  
 atcgtctgctg gtgctccaaa ggacttgatc ggttggatcg accaaccatc tgttgaattg 8520

ES 2 648 865 T3

tctaacgctt tgatgcacca cccagacatc aacttgatct tggctaccgg tggccaggt 8580  
atggtaag ctgcttactc ttctgtaag ccagctatcg gtgttggtgc tgtaacacc 8640  
ccagttgta tcgacgaaac cgctgacatc aagagagctg ttgcttctgt ttgatgtct 8700  
aagacctcg acaacggtgt tatctgtgct tetgaacaat ctgttggtgt tgttgactct 8760  
gtttacgacg ctgtagaga aagattcgct acccaggtg gttacttgtt gcaaggtaag 8820  
gaattgaag ctgttcaaga cgttatcttg aagaacggtg ctttgaacgc tgctatcgtt 8880  
ggccaaccag cttacaagat cgctgaatta gctggtttct ctgttccaga aaacaccaag 8940  
atcttgatcg gtgaagttac cgttgttgac gaatctgaac cattcgctca cgaaggtt 9000  
tctccaacct tggctatgta cagagetaag gacttcgaag acgctgttga aaaagtgaa 9060  
aagttggtt ctatgggtgg tattggtcac acctcttgtt tgtacaccga ccaagacaac 9120  
caaccagcta gagtttctta cttcggtaa aagatgaaga ccgctagaat cttgatcaac 9180  
acccagctt ctcaaggtgg tatcgggtgac ttgtacaact tcaagttggc tccatcttg 9240  
acctgggtt gtggttcttg ggggtgtaac tctatctctg aaaacgttgg tccaaagcac 9300  
ttgatcaaca agaagaccgt tgctaagaga gctgaaaaca tgttggtgca caagtgcca 9360  
aaatctatct acttcagaag aggttctttg ccaatcgctt tggacgaagt tatcaccgac 9420  
ggtcacaaga gagctttgat cgttaccgac agattcttgt tcaacaacgg ttacgctgac 9480  
caaatcacct ctgttttgaa ggctgctggt gttgaaaccg aagttttctt cgaagttgaa 9540  
gctgacccaa ccttgtctat cgttagaaag ggtgctgaat tggctaactc tttcaagcca 9600  
gacgttatca tcgctttggg tgggtggttct ccaatggacg ctgctaagat catgtgggtt 9660  
atgtacgaac acccagaaac ccacttcgaa gaattggctt tgagattcat ggacatcaga 9720  
aagagaatct acaagttocc aaagatgggt gtttaaggcta agatgatcgc tgttaccacc 9780  
acctctggtc cgggttctga agttacccca ttcgctgttg ttaccgacga cgctaccggt 9840  
caaaagtacc cattggtgta ctacgcttg accccagaca tggctatcgt tgacgctaac 9900  
ttggttatgg acatgccaaa gtctttgtgt gctttcgggt gtttggacgc tgttaccac 9960  
gctatggaag cttacgtttc tgttttggt tctgaattct ctgacggtca agctttgcaa 10020  
gctttgaagt tgttgaagga atacttgcca gcttcttacc acgaaggttc taagaacca 10080  
gttgctagag aaagagttca ctctgctgct accatcgctg gtatcgcttt cgctaacgct 10140  
ttcttgggtg tttgtcactc tatggtcac aagttgggtt ctcaattcca catcccacac 10200  
ggtttggtc aogctttggt gatctgtaac gttatcagat acaacgctaa cgacaacca 10260  
accaagcaaa ccgctttctc tcaatacgac agaccacaag ctagaagaag atacgctgaa 10320  
atcgctgacc acttgggttt gtctgctcca ggtgacagaa ccgctgcaa gatcgaaaag 10380  
ttgttggtt gggtggaac cttgaaggct gaattgggtt tcccaaagtc tatcagagaa 10440

ES 2 648 865 T3

gctggtgttc aagaagctga cttcttggtc aacggtgaca agttgtctga agacgctttc 10500  
gacgaccaat gtaccgggtc taaccaaga taccattga tctctgaatt gaagcaaadc 10560  
ttgttgaca cctactacgy tagagactac gttgaaggty aaaccgctgc taagaaggaa 10620  
gctgctccag ctaaggctga aaagaaggct aagaagtctg cttagcttaa gcgcgcaat 10680  
ttcttatgat ttatgatttt tattattaaa taagttataa aaaaaataag tgtatacaaa 10740  
ttttaaagtg actcttaggt tttaaaacga aaattcttat tcttgagtaa ctctttcctg 10800  
taggtcaggt tgctttctca ggtatagcat gaggtcgtc ttattgacca cacctctacc 10860  
ggcatgccga gcaaatgcct gcaaatcgtc cccatttca cccaattgta gatatgctaa 10920  
ctccagcaat gagttgatga atctcgggtg gtattttatg tcctcagagg acaacacctg 10980  
ttgtaatcgt tctccacac gtacgaagct taaaagggcg aattctgcag atatccatca 11040  
cactggcggc cgcagcttag ctccgatta tcgatgataa gctgtcaaag atgagaatta 11100  
attccacgga ctatagacta tactagatac tccgtctact gtacgataca cttccgctca 11160  
ggccttgctc ctttaacgag gccttaccac tcttttgta ctctattgat ccagctcagc 11220  
aaaggcagty tgatctaaga ttctatcttc gcgatgtagt aaaactagct agaccgagaa 11280  
agagactaga aatgcaaaag gcacttctac aatggctgcc atcattatta tccgatgtga 11340  
cgctgcagct tctcaatgat attcgaatac gctttgagga gatacagcct aatatccgac 11400  
aaactgtttt acagatttac gatcgtactt gttaccatc attgaatttt gaacatccga 11460  
acctgggagt tttccctgaa acagatagta tatttgaacc tgtataataa tatatagtct 11520  
agcgttttac ggaagacaat gtatgtattt cggttcctgy agaaactatt gcactctattg 11580  
cataggtaat cttgcacgtc gcacccccg ttcattttct gcgtttccat cttgcacttc 11640  
aatagcatat ctttgtaac gaagcatctg tgcttcattt tgtagaacaa aaatgcaacg 11700  
cgagagcgtc aatttttcaa acaaagaatc tgagctgcat ttttacagaa cagaaatgca 11760  
acgcgaaagc gctattttac caacgaagaa tctgtgcttc atttttgtaa aacaaaaatg 11820  
caacgcgacg agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca 11880  
gaaatgcaac gcgagagcgc tattttacca acaaagaatc tatacttctt ttttgttcta 11940  
caaaaatgca tcccgagagc gctatttttc taacaaagca tcttagatta ctttttttct 12000  
cctttgtgcy ctctataatg cagtctcttg ataacttttt gcactgtagg tccgttaaggy 12060  
ttagaagaag gctacttttg tgtctatttt ctcttcata aaaaaagcct gactccactt 12120  
ccgcggttta ctgattacta gcgaagctgc ggggtcattt tttcaagata aaggcatccc 12180  
cgattatatt ctataccgat gtggattgcy catactttgt gaacagaaag tgatagcgtt 12240  
gatgattctt cattggctcag aaaattatga acggtttctt ctattttgc tctatatact 12300

ES 2 648 865 T3

acgtatagga aatgtttaca ttttcgtatt gttttcgatt cactctatga atagttctta 12360  
 ctacaatfff tttgtctaaa gagtaatact agagataaac ataaaaaatg tagaggtcga 12420  
 gtttagatgc aagttcaagg agcgaaaagg ggatgggtag gttatatagg gatatagcac 12480  
 agagatatat agcaaaagaga tacttttgag caatgtttgt ggaagcggta ttcgcaatgc 12540  
 cggcctttccc cgtcaagctc taaatcgggg gctcccttta gggttccgat ttagtgcttt 12600  
 acggcacctc gacccccaaa aacttgatta gggatgatgg tccacgtatg gcccatcgcc 12660  
 ctgatagacg gtttttcgcc ctttgacgtt ggagtcacg tctttaata gtggactctt 12720  
 gttccaaact ggaacaacac tcaaccctat ctccggtctat tcttttgatt tataagggat 12780  
 tttgccgatt tcggcctatt ggtaaaaaa tgagctgatt taacaaaaat ttaacgcgaa 12840  
 ttttaacaaa atattaacgt ttacaatttc ctgatgcggg attttctcct tacgcatctg 12900  
 tggcgtatft cacaccgcat atgatccgtc gaggttcaaga gaaaaaaaaa gaaaaagcaa 12960  
 aaagaaaaaa ggaaagcgcg cctcgttcag aatgacacgt atagaatgat gcattacctt 13020  
 gtcactttca gtatcatact gttcgtatac atacttactg acattcatag gtatacatat 13080  
 atacacatgt atatatatcg tatgctgcag ctttaaataa tcgggtgtcac tacataagaa 13140  
 cacctttggg ggaggaaca tcggtgttac cattgggcga ggtggcttct cttatggcaa 13200  
 ccgcaagagc cttgaaacga ctctcactac ggtgatgac attcttgccct cgcagacaat 13260  
 caacgtggag ggtaattctg ctagecctcg caaagctttc aagaaaatgc gggatcatct 13320  
 cgcaagagag atctcctact ttctcccttt gcaaaccaag ttcgacaact gcgtacggcc 13380  
 tgttcgaaag atctaccacc gctctgaaa gtgcctcatc caaaggcgcg aatcctgatc 13440  
 caaacctfff tactccaagc gccagtaggg cctctttaa agcttgaccg agagcaatcc 13500  
 cgcagtcttc agtgggtgta tggctgtcta tgtgtaagtc accaatgcac tcaacgatta 13560  
 gcgaccagcc ggaatgcttg gccagagcat gtatcatatg gtccagaaac cctatacctg 13620  
 tgtggaagtt aatcacttgc gattgtgtgg cctgttctgc tactgcttct gccctttttt 13680  
 ctgggaagat cgagtgtctc atcgctaggg gaccaccctt taaagagatc gcaatctgaa 13740  
 tcttggtttc atttgtaata cgctttacta gggctttctg ctctgtcatc tttgccttcg 13800  
 tttatcttgc ctgctcattt tttagtatat tcttcgaaga aatcacatta ctttatataa 13860  
 tgtataattc attatgtgat aatgccaatc gctaagaaaa aaaaagagtc atccgctagg 13920  
 tggaaaaaaaa aaaatgaaaa tcattaccga ggcataaaaa aatatagagt gtactagagg 13980  
 aggccaagag taatagaaaa agaaaattgc gggaaaggac tgtgttatga ctccctgac 14040  
 taatgccgtg ttcaaacgat acctggcagt gactcctagc gctcaccaag ctcttaaac 14100  
 ggaattatgg tgcaetctca gtacaatctg ctctgatgcc gcatagttaa gccagccccg 14160  
 acaccggcca acaccgctg acgcgccctg acgggcttgt ctgctcccgg catccgctta 14220  
 cagacaagct gtgaccgtct ccgggagctg catgtgtcag aggttttcac cgctcatcacc 14280  
 gaaacgcgcg a 14291

<210> 40

5 <211> 4891

<212> ADN

ES 2 648 865 T3

<213> Artificial

<220>

<223> pRN593

5

<400> 40

```

gttttaaatc aatctaaagt atatatgagt aaacttggtc tgacagttac caatgcttaa      60
tcagtgaggc acctatctca gcgatctgtc tatttcgttc atccatagtt gcctgactcc      120
ccgtcgtgta gataactacg atacgggagg gcttaccatc tggccccagt gctgcaatga      180
taccgcgaga cccacgctca ccggtccag atttatcagc aataaaccag ccagccggaa      240
gggccgagcg cagaagtggc cctgcaactt tatccgcctc catccagtct attaattggt      300
gccgggaagc tagagtaagt agttcgccag ttaatagttt gcgcaacggt gttgccattg      360
ctacaggcat cgtgggtgca cgctcgtcgt ttggatggc ttcattcagc tccggttccc      420
aacgatcaag gcgagttaca tgatcccca tgttggtgcaa aaaagcgggt agctccttcg      480
gtcctccgat cgttgtcaga agtaagttgg ccgcagtggt atcactcatg gttatggcag      540
cactgcataa ttctcttact gtcatgccat ccgtaagatg cttttctgtg actggtgagt      600
actcaaccaa gtcattctga gaatagtgta tgcggcgacc gagttgctct tgccccgcgt      660
caatacggga taataccgcg ccacatagca gaactttaa agtgctcatc attggaaaac      720
gttcttcggg gogaaaactc tcaaggatct taccgctgtt gagatccagt tcgatgtaac      780
ccactcgtgc acccaactga tcttcagcat cttttacttt caccagcgtt tctgggtgag      840
caaaaacagg aaggcaaaat gccgcaaaa agggaataag ggcgacacgg aaatggtgaa      900
tactcatact cttccttttt caattcagaa gaactcgtca agaaggcgat agaaggcgat      960
gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg aagcggtcag ccatttcgcc     1020
gccaaactct tcagcaatat cacggtagc caacgctatg tcctgatagc ggtccgccac     1080
accagccgg ccacagtcga tgaatccaga aaagcggcca tttccacca tgatattcgg     1140
caagcaggca tcgccatggt accgagtggg gttgtaacca ccctcaaaa atgcggaaga     1200
ggtgtacagc taatattata atgaaaaatc cttttgctct ataggatfff ttttttaaaa     1260
tttaaatata catatttatt tatatttatt tggttactaa attttgtatt aatttaataa     1320
taataagtaa tttctaactg gataaagtag ttatgagaaa tgacataatg ctaatttatg     1380
aatatgatat agaagagcct cgaaaaaagt gggggaaagt atgatatggt atctttctcc     1440

```

ES 2 648 865 T3

aataaatcta atcttcatgt agatctaatt cttcaatcat gtccggcagg ttcttcattg 1500  
ggtagttggt gtaaaccgatt tggatatacgg cttcaaataa tgggaagtct tcgacagagc 1560  
cacatgtttc caaccattcg tgaacttctt tgcaggtaat taaaccttga gcggattggc 1620  
cattcaacaa ctcgtaacacc taggatccgt cgacactgga tggcggcgtt agtatcgaat 1680  
cgacagcagt atagcgacca gcattcacat acgattgacg catgatatta ctttctgcgc 1740  
acttaacttc gcatctgggc agatgatgtc gaggcgaaaa aaaatataaa tcacgctaac 1800  
atgtgattaa aatagaacaa ctacaatata aaaaaactat acaaatgaca agttcttgaa 1860  
aacaagaatc tttttattgt cagtaactgat tattcctttg ccctcggacg agtgctgggg 1920  
cgtcgggttc cactatcggc gagtaactct acacagccat cggtcacagc ggccgcgctt 1980  
ctgcgggcga tttgtgtacg cccgacagtc ccggctccgg atcggacgat tgcgtcgcac 2040  
cgacctgcg cccaagctgc atcatcgaaa ttgccgtcaa ccaagctctg atagagttgg 2100  
tcaagaccaa tgcggagcat atacgcccgg agccgcggcg atcctgcaag ctccggatgc 2160  
ctccgctcga agtagcgcgt ctgctgctcc atacaagcca accacggcct ccagaagaag 2220  
atgttgccga cctcgtattg ggaatccccg aacatcgctt cgctccagtc aatgaccgct 2280  
gttatgcggc cattgtccgt caggacattg ttggagccga aatccgcgtg cacgaggtgc 2340  
cggacttcgg ggcagtcctc ggccaaaagc atcagctcat cgagagcctg cgcgacggac 2400  
gcactgacgg tgtcgtccat cacagtttgc cagtataca catggggatc agcaatcgcg 2460  
catatgaaat cacgccatgt agtgtattga ccgattcctt gcggtccgaa tgggcccgaac 2520  
ccgctcgtct ggctaagatc ggccgcagcg atcgcatcca tggcctccgc gaccggctgc 2580  
agaacagcgg gcagttcggg ttcaggcagg tcttgcaacg tgacaccctg tgcacggcgg 2640  
gagatgcaat aggtcaggct ctocgtgaat tccccaatgt caagcacttc cggaatcggg 2700  
agcgcggccg atgcaaagtg ccgataaaca taacgatctt tgtagaaacc atcggcgcag 2760  
ctatttaccg gcaggacata tccacgcctt cctacatcga agctgaaagc acgagattct 2820  
tcgccctccg agagctgcat caggctcggag acgctgtcga acttttcgat cagaaacttc 2880  
tcgacagacg tcgcggtgag ttcaggcttt ttacccatgg ttgtttatgt tcggatgtga 2940  
tgtgagaact gtatcctagc aagattttaa aaggaagtat atgaaagaag aacctcagtg 3000  
gcaaatccta accttttata tttctctaca ggggcgcggc gtggggacaa ttcacgcgtc 3060  
tgtgagggga gcgtttccct gctcgcaggt ctgcagcgag gagccgtaat ttttgcttcg 3120  
cgccgtgcgg ccatcaaaat gtatggatgc aatgattat acatggggat gtatgggcta 3180  
aatgtacggg cgacagtcac atcatgcccc tgagctgcgc acgtcaagac tgtcaaggag 3240  
ggtattctgg gctccatgt cgctggccgg gtgacccggc ggggacgagg cettaagttc 3300  
gaacgtacga gctccgatt aagggcgaat tccagcacac tggcggccgt tactagacca 3360

ES 2 648 865 T3

gcattcaagt ggccggaagt taagtttaat ctatcagcag cagcagacat ctttatatta 3420  
tcaatatttg tgtttgtgga gggggggggt gtacaatata caattgtttc ttgatgtcct 3480  
atgtaggaga gtaagataaa agaaaattaa agggagaaaag aatttaagggt aacgtaacag 3540  
tttcaactcat ggtagtatat atataactac gaaactgcc aaccaagcc agtctacgtg 3600  
cgaattaggt tggatagatt tgatgaacct attctgatac ttggtgtgcg gaaaaagga 3660  
caaatataggg aaaactggaa aaagaaaagg aaggggagag aagcaaggcg ggtaccaagc 3720  
ttgatgcata gcttgagtat tctatagtgt cacctaaata gcttggcgta atcatggtca 3780  
tagctgtttc ctgtgtgaaa ttgttatccg ctcaacaattc cacacaacat acgagccgga 3840  
agcataaagt gtaaagcctg ggtgcctaa tgagtgagct aactcacatt aattgcggtg 3900  
cgctcaactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc 3960  
caacgcgcgg ggagagcggg ttgctgtatt gggcgtctt ccgcttcctc gctcaactgac 4020  
tcgctgcgct cggctcgtcg gctgcggcga gcggtatcag ctcaactcaa gccggtaata 4080  
cggttatcca cagaatcagg ggataacgca ggaaagaaca tgtgagcaaa aggccagcaa 4140  
aagccagga accgtaaaaa gccgcggtg ctggcgtttt tccataggct ccgccccct 4200  
gacgagcatc acaaaaatcg acgctcaagt cagaggtggc gaaaccgcac aggactataa 4260  
agataccagg cgtttcccc tggaagctcc ctgctgcgct ctccctgttc gaccctgccc 4320  
cttaccggat acctgtccgc cttctccct tcgggaagcg tggcgctttc tcatagctca 4380  
cgctgtaggt atctcagttc ggtgtaggtc gttcgtcca agctgggctg tgtgcacgaa 4440  
cccccgttc agcccgaccg ctgcgcctta tccggtaact atcgtcttga gtccaaccg 4500  
gtaagacacg acttatcgcc actggcagca gccactggta acaggattag cagagcgagg 4560  
tatgtaggag gtgctacaga gttcttgaag tgggtggccta actacggcta cactagaaga 4620  
acagtatttg gtatctgcgc tctgctgaag ccagttacct tcggaaaaag agttggtagc 4680  
tcttgatccg gcaaaaaaac caccgctggg agcgggtggt tttttgttg caagcagcag 4740  
attacgcgca gaaaaaagg atctcaagaa gatcctttga tctttctac ggggtctgac 4800  
gctcagtgga acgaaaactc acgttaaggg attttggtca tgagattatc aaaaggatc 4860  
ttcacctaga tccttttaaa ttaaaaatga a 4891

<210> 41

<211> 30

5 <212> ADN

<213> Artificial

<220>

<223> Cebador gpd1uf

10

<400> 41

ES 2 648 865 T3

aagcttgga cccgcctgc ttctctccc 30

<210> 42  
<211> 26  
5 <212> ADN  
<213> Artificial

<220>  
<223> Cebador gpd1ur  
10

<400> 42  
tctagaccag cattcaagtg gccgga 26

<210> 43  
15 <211> 29  
<212> ADN  
<213> Artificial

<220>  
20 <223> Cebador gpd1df

<400> 43  
cgtacgagtt gttgaatggc caatccgct 29

25 <210> 44  
<211> 31  
<212> ADN  
<213> Artificial

30 <220>  
<223> Cebador gpd1dr

<400> 44  
35 ccatgtacc gagtgggtt gtaaccacc t 31

<210> 45  
<211> 20

<212> ADN  
<213> Artificial

<220>

5 <223> Cebador gpd1cf

<400> 45  
accaatacgt aaacggggcg 20

10 <210> 46  
<211> 23  
<212> ADN  
<213> Artificial

15 <220>

<223> Cebador gpd1cr

<400> 46  
aatacaccca tacatacga cgc 23

20 <210> 47  
<211> 28  
<212> ADN  
<213> Artificial

25 <220>

<223> Cebador acs1f

<400> 47

30 ttaagcttaa aatgtcgccc tctgccgt 28

<210> 48  
<211> 50  
<212> ADN

35 <213> Artificial

<220>

<223> Cebador acs1r

<400> 48

aagcgcgcta caactgacc gaatcaatta gatgtctaac aatgccaggg 50

5

<210> 49

<211> 49

<212> ADN

<213> Artificial

10

<220>

<223> Cebador acs2f

<400> 49

15

aactgcagaa aatgacaatc aaggaacata aagtagttaa tgaagctca 49

<210> 50

<211> 49

<212> ADN

20

<213> Artificial

<220>

<223> Cebador acs2r

25

<400> 50

acgctgacta ttcttttt tgagagaaaa attggttctc tacagcaga 49

<210> 51

<211> 13018

30

<212> ADN

<213> Artificial

<220>

<223> pRN753

35

<400> 51

ES 2 648 865 T3

gacgaaaggg cctcgtgata cgcctatfff tataggfmaa tgfcatgata ataatggfff 60  
cttagacgfc aggtggcact fffcgggaa atgtgcgcgg aaccctatt fgtffatfff 120  
tctaaataca ffcaaatag fatccgctca fgagacaata accctgataa atgctfcaat 180  
aatatfgaaa aaggaagagt atgagtattc aacattffcgg fgtcgcctt attccctfff 240  
ffgcggcatt ffccctcct gfffffgtc acccagaaac gctggfgaaa gtaaaagatg 300  
ctgaagatca gffggfgca cggatggff acatcgaact ggatctcaac agcggtaaga 360  
fcctfgagag ffffcgccc gaagaacgff ffccaatgat gagcactfff aaagffctgc 420  
fatgffggcgc ggtattatcc cgtattgacg cggggcaaga gcaactcggf cggcgcatac 480  
actatfcfca gaatgactfg gffgagfacc caccagfcac agaaaagcat cffacggatg 540  
gcatgacagf aagagaatta fgcagfgctg ccataacct gatfgataac actcgggcca 600  
actfaccff cacaacgafc gaggagccga aggagfcaac cgcfffffff cacaacatgg 660  
gggatcatgt aactcgcctf gatcgtfggg aaccggagct gaatgaagcc ataccaaacg 720  
acgagcgtga caccacgatg cctgtagcaa fggaacaac gffgcgcaa cffaaactg 780  
gcgaactact factctagct fcccgcaac aattaataga cffgatggag gcgataaag 840  
ffgcaggacc actfcfgcgc fcggccctfc cggctggctg gfffatfgct gataaatctg 900  
gagccggfga cgcfggffct cggcgfatca ffgcagcact ggggccagat gfgaacccf 960  
cccgatcgt agffatctac acgacgggca gfcaggcaac fatggatgaa cgaatagac 1020  
agatcgcfga gatagfgcc fcactgatta agcattgfga actgfcagac caagffact 1080  
catatatact ffagatfgat ftaaaactfc atffffaatt faaaagfat tagfgaaga 1140  
fcctffffga faatctcatg accaaaatcc cffaacfga gffffcgtfc cactgagcgt 1200  
cagaccccgf agaaaagatc aaagfatctf cffgagatcc ffffffcfg cgcgfaatc 1260  
gctgctfga acaaaaaaa ccaccgctac cagcggfggff ffgfffcgg gatcaagagc 1320  
faccactct ffffcggaag gfaactggct fcagcagagc gcagatacca aatactgfc 1380  
ffctagfga gccgtagffa ggcaccact fcaagaactc fgtagaccg cffacatacc 1440  
fcgctctgct aatcctgffa ccagfgctg cffccagfg cgaafagfc fgctffaccg 1500  
ggffggactc aagacgatag ffaccggata aggcgcagcg gfcgggctga accgggggff 1560  
cgtgcacaca gccagctfg gagcgaacga cffacaccga actgagatac cffacagcftg 1620  
agcattgaga aagcgcacg cffcccgaa ggaagaaaggc ggacagfat ccgfgaagc 1680  
gcagggfcgg aacaggagag cgcacgaggg agctfccag ggggaacgcc fgfatctff 1740  
atagfcctgt cgggfffcgc cactctgac ffgagcfcg atfffffgfa fgctcfcag 1800  
gggggcccag cctatgfaa aaccgcagca accggcctf ffacgfffc cffgcctfff 1860  
gctggcctff fgctcacatg ffctffctg cgttatccc fgatfcfgfg gataaccfga 1920  
ffaccgcctf fgagfgagct gataccctc gccgcagccg aaccaccgag cgcagcagf 1980  
cagfgagcga ggaagcggaa gagcgcctaa facgcaaac gcctctccc gcgcgfffgc 2040  
cgatfcatta atgcagcgg cacgacaggf ffcccgactg gaaagcgggc agfgagcga 2100

ES 2 648 865 T3

acgcaattaa tgtgagttac ctcactcatt aggcacccca ggctttacac tttatgcttc 2160  
 cggctcctat gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg 2220  
 accatgatta cgccaagctc ggaattaacc ctcactaaag ggaacaaaag ctgggtaccg 2280  
 ggccccccct cgaccggctc cacacacccat agcttcaaaa tgtttctact ccttttttac 2340  
 tcttccagat tttctcggac tccgcgcac cccgtaccac ttcaaaacac ccaagcacag 2400  
 cataactaat ttcccctctt tcttctccta ggggtgctgtt aattaccctg actaaaggtt 2460  
 tggaaaagaa aaaagagacc gcctcgtttc tttttcttcg tcgaaaagg caataaaaat 2520  
 ttttatcaog tttctttttc ttgaaaattt ttttttttga tttttttctc tttcogatgac 2580  
 ctcccattga tatttaagtt aataaacggt cttcaatttc tcaagtttca gtttcatttt 2640  
 tcttgttcta ttacaacttt ttttacttct tgctcattag aaagaagca tagcaatcta 2700  
 atctaagctg cagaagctta aaatgtcgc ctctgccgta caatcatcaa aactagaaga 2760  
 acagtcaagt gaaattgaca agttgaaagc aaaaatgtcc cagtctgccg ccaactgcga 2820  
 gcagaagaag gaacatgagt atgaacattt gacttcggtc aagatcgtgc cacaacggcc 2880  
 catctcagat agactgcagc ccgcaattgc taccactat tctccacact tggacgggtt 2940  
 gcaggactat cagcgttgc acaaggagtc tattgaagac cctgctaagt tcttcggttc 3000  
 taaagctacc caatttttaa actggtctaa gccattcgat aaggtgttca tcccagaccc 3060  
 taaaacgggc aggccctcct tccagaacaa tgcatggttc ctcaacggcc aattaaacgc 3120  
 ctgttacaac tgtgttgaca gacatgcctt gaagactcct aacaagaaag ccattatttt 3180  
 cgaaggtgac gagcctggcc aaggctattc cattacctac aaggaactac ttgaagaagt 3240  
 ttgtcaagtg gcacaagtgc tgacttactc tatgggcgtt cgcaagggcg atactgttgc 3300  
 cgtgtacatg cctatggtcc cagaagcaat cataaccttg ttggccattt cccgtatcgg 3360  
 tgccattcac tccgtagtct ttgccgggtt ttcttccaac tccttgagag atcgtatcaa 3420  
 cgatggggac tctaaagttg tcatcactac agatgaatcc aacagaggtg gtaaagtcac 3480  
 tgagactaaa agaattggtg atgacgcgct aagagagacc ccaggcgtga gacacgtctt 3540  
 ggtttataga aagaccaaca atccatctgt tgctttccat gccccagag atttggattg 3600  
 ggcaacagaa aagaagaaat acaagacctt ctatccatgc acaccggtg attctgagga 3660  
 tccattattc ttgttgata cgtctggttc tactggtgcc cccaaggggtg ttcaacattc 3720  
 taccgcaggt tacttgctgg gagctttggt gaccatgccc tacacttttg acactcacca 3780  
 agaagacggt ttcttcacag ctggagacat tggttgatt acaggccaca cttatgtggt 3840  
 ttatggtccc ttactatag gttgtgccac tttggtcttt gaagggactc ctgcgtaccc 3900  
 aaattactcc cgttattggg atattattga tgaacacaaa gtcacccaat tttatgttgc 3960

ES 2 648 865 T3

gccaaactgct ttgcgtttgt tgaaaagagc tggtgattcc tacatcgaaa atcattcctt 4020  
 aaaatcctttg cgttgcttgg gttcggtcgg tgagccaatt gctgctgaag tttgggagtg 4080  
 gtactctgaa aaaataggta aaaatgaaat ccccatgtga gacacctact ggcaaacaga 4140  
 atctggttcg catctggtca ccccgctggc tggtggtggt acaccaatga aaccgggttc 4200  
 tgcctcattc cccttcttgc gtattgatgc agttgttctt gaccctaaca ctggtgaaga 4260  
 acttaacacc agccacgcag aggggtgtcct tgccgtcaaa gctgcatggc catcatttgc 4320  
 aagaactatt tggaaaaatc atgataggta tctagacact tatttgaacc cttaccctgg 4380  
 ctactatctt actggtgatg gtgctgcaaa ggataaggat ggttatatct ggattttggg 4440  
 tcgtgtagac gatgtggtga acgtctctgg tcaccgtctg tctaccgtg aaattgaggc 4500  
 tgctattatc gaagatccaa ttgtggccga gtgtgctggt gtcggattca acgatgactt 4560  
 gactggtcaa gcagttgctg catttgtggt gttgaaaaac aaatctagtt ggtccaccgc 4620  
 aacagatgat gaattacaag atatcaagaa gcatttgtgc tttactgtta gaaaagacat 4680  
 cgggccattt gccgcaccaa aattgatcat tttagtggat gacttgccca agacaagatc 4740  
 cggcaaaaatt atgagacgta ttttaagaaa aatcctagca ggagaaagt accaactagg 4800  
 cgacgtttct acattgtcaa accctggcat tgttagacat ctaattgatt cggccaagtt 4860  
 gtaacttaag cgcgcgaatt tcttatgatt tatgattttt attattaaat aagttataaa 4920  
 aaaaataagt gtatacaaat tttaaagtga ctcttaggtt ttaaaacgaa aattcttatt 4980  
 cttgagtaac tctttcctgt aggtcaggtt gctttctcag gtatagcatg aggtcgctct 5040  
 tattgaccac acctctaccg gcattgccag caaatgcctg caaatcgctc cccatttcac 5100  
 ccaattgtag atatgctaac tccagcaatg agttgatgaa tctcgggtgtg tattttatgt 5160  
 cctcagagga caacacctgt tgtaatcgtt cttccacacg tacgttttaa acagttgatg 5220  
 agaacctttt tcgcaagtcc aagggtgctct aatttttaa atttttactt ttcgcgacac 5280  
 aataaagtct tcacgacgct aaactattag tgcacataat gtagttactt ggacgctgtt 5340  
 caataatgta taaaatttat ttoccttgca ttacgtacat tatataacca aatcttaaaa 5400  
 atatagaaat atgatatgtg tataataata taagcaaaat ttacgtatct ttgcttataa 5460  
 tatagcttta atgttcttta ggtatatatt taagagcgat ttgtctcgac tatttctttt 5520  
 tttgagagaa aaattggttc tctacagcag aatgatggc aggtacaact tctgggttgg 5580  
 ccaaagtagt taggtcacct agctgttcgg cttcgttaga agcaaccttt cttagaactc 5640  
 ttctcataat ctttctgac cttgttcttg gtagatctct aactagaata atggtttttg 5700  
 gtgaggcgaa aggaccaatc tcaccctaa cttgtaagat caattctcta cgtaaattat 5760  
 ctggtgtgat gtgttctgca tcacctcag tagcgttgtt ttgtagataa ccatctttaa 5820  
 gggaaacata tgcaacgacg gtttgaccgg tcaattcatc tggaataccg acaacagcag 5880

ES 2 648 865 T3

ctccgagac gttttcgtga tttgagatag atgcttcaat ttctgatgtg gataatctat 5940  
 gaccggaac atttacaacg tcgtcaactc taccctgat ccagtagtaa ccatcatgat 6000  
 ctctaccagc accatcacct gtgaaatagt gaccaggata aggtttcaag taagtatcca 6060  
 tgtaacggtc gtggtggttc caaacagatc tagccattga tggccatggt gatttaacgg 6120  
 caaggacacc ttcgacatca ttaccttcta attccacacc tgtaacaggg tcaatgatac 6180  
 aagcgtaat accaaagaat ggcacggtag cagaaccagg ttttgttggg acagcacctg 6240  
 ccaaaggagc aattaaatga gaaccagact ctgtttgcca catagtgtca caaatgacac 6300  
 agtttttgtt acccactttt tcatgatacc attoccataa gtctggagag attggttcac 6360  
 cgacggaacc caagacacgt aatgaggaag tgtcatattt ggcaatttcg gcttcaccta 6420  
 cacgtttgat taatcttaaa gcagttggag ccacatagaa atgggtagcc ttgtgacgtt 6480  
 ggataattct ccaatatcta ccataatctg ggtaggcagg agtggattcg aaaattattg 6540  
 aggcgttacc caaggttaat ggaccatata gagcataggt gtgaccctg atccagccga 6600  
 cgtcaccggc agtgaagaga acatcttctg ggtgaatata aaaaacgtat ctagtgttta 6660  
 aagcggcacc taataataaa ccacctgtag tgtgaacgac accctttgga gaaccagtgg 6720  
 aaccggaagt gtataataaa aatagaggat cttcagcgtc acatgaaaca ggaggtaggt 6780  
 aagttctctg cttagcggcc tcctcatgcc accagtaate tctaccggcc ttcattggaa 6840  
 taccttcagt accagttctt tggaaaacca agatacggga aaccaaatcg actccgttca 6900  
 aaccttcgtc aacaattttt ttagtgttga tggctctacc acctctttta ccttcatcac 6960  
 aagtgatgac cacttttagaa ttagcgtcaa cgacacgatc tttcaacgaa ccagcggaga 7020  
 accagcaaa gacaacagag tgaatagcac caatacgagc cacagccaac atagcaatga 7080  
 ccgcttctgg aatcattggc aaatagatag ccaactgtgc acctttotta acgccccagc 7140  
 tttttaagac accagcgatt tgggaaactt ttctgagtaa ttcaccaa atgtgatgatt 7200  
 tgttgcgga ttcgtcatca gcttcataga tcaaagctgg cttgtcggga ttagcaaagg 7260  
 catgtctgtc aacacaattg tatgatgcat tcaatttacc gttcaaaaac catgcaacat 7320  
 caccattggt caatgaacca gattgaactt tgggtgatgg agcatcccaa tgcaagtatt 7380  
 ccttagccat ctatcaaaag aatttttctg gctcattgat agattgttga tacatttctt 7440  
 gataatgttg catatcagta acgtaaccct tgccgggttg gctgttgtaa aatgttgag 7500  
 gagcctaag agcctttacg ttgtgagctt cataaactac tttatgttcc ttgattgtca 7560  
 ttttctgcag tctagatata tttgttgtaa aaagtagata attacttctt tgatgatctg 7620  
 taataaagag aaaaagaaag catctaagaa cttgaaaaac tacgaattag aaaagaccaa 7680  
 atatgtattt cttgcattga ccaatttatg caagtttata tataatgtaa tgtaagtttc 7740

ES 2 648 865 T3

acgaggttct actaaactaa accaccccct tggttagaag aaaagagtgt gtgagaacag 7800  
 gctgttgttg tcacacgatt cggacaattc tgtttgaaag agagagagta acagtacgat 7860  
 cgaacgaact ttgctctgga gatcacagtg ggcacatag catgtggtac taaacccttt 7920  
 cccgccattc cagaaccttc gattgcttgt tacaaaaact gtgagccgtc gctaggacct 7980  
 tgttgtgtga cgaaattgga agctgcaatc aataggaaga caggaagtgc agcgtgtctg 8040  
 ggttttttca gttttgttct ttttgcaaac aatcacgag cgacggtaat ttctttctcg 8100  
 ataagaggcc acgtgcttta tgagggtaac atcaattcaa gaaggaggga aacacttctc 8160  
 tttctgggcc ctgataatag tatgagggtg aagccaaaat aaaggattcg cgcccaaatc 8220  
 ggcactctta aatgcaggta tgcgatagtt cctcactctt tccttactca cgagtaattc 8280  
 ttgcaaatgc ctattatgca gatgttataa tatctgtgcg tcttgagttg aagagctcga 8340  
 gactagatgc atgctcgagc ggccgccagt gtgatggata tctgcagaat tcgccctttt 8400  
 gggccctgta cacctaggat ccgtcgacac tggatggcgg cgttagtata gaatcgacag 8460  
 cagtatagcg accagcattc acatacgatt gacgcatgat attactttct gcgcacttaa 8520  
 cttcgcactc gggcagatga tgtcgaggcg aaaaaaata taaatcacgc taacatttga 8580  
 ttaaaataga acaactacaa tataaaaaaa ctatacaaat gacaagttct tgaaaacaag 8640  
 aatcttttta ttgtcagtac tgattagaaa aactcatcga gcatcaaatg aaactgcaat 8700  
 ttattcatat caggattatc aataccatat ttttgaaaaa gccgtttctg taatgaagga 8760  
 gaaaaactac cgaggcagtt ccataggatg gcaagatcct ggtatcggtc tgcgattccg 8820  
 actcgtccaa catcaataca acctattaat ttcccctcgt caaaaataag gttatcaagt 8880  
 gagaaatcac catgagtgac gactgaatcc ggtgagaatg gcaaaagctt atgcatttct 8940  
 ttccagactt gttcaacagg ccagccatta cgctcgtcat caaaatcact cgcacaaacc 9000  
 aaaccgttat tcattcgtga ttgcccctga gcgagacgaa atacgcgatc gctgttaaaa 9060  
 ggacaattac aaacaggaat cgaatgcaac cggcgcagga acaactgccag cgcacaaaca 9120  
 atattttcac ctgaatcagg atattcttct aatacctgga atgctgtttt gccggggatc 9180  
 gcagtgggtga gtaaccatgc atcatcagga gtacggataa aatgcttgat ggtcgggaaga 9240  
 ggcataaatt ccgtcagcca gtttagtctg accatctcat ctgtaacatc attggcaacg 9300  
 ctacctttgc catgtttcag aaacaactct ggcgcacatc gcttcccata caatcgatag 9360  
 attgtcgcac ctgattgccc gacattatcg cgagcccatt tatacccata taaatcagca 9420  
 tccatgttgg aatttaatcg cggcctcgaa acgtgagtct tttccttacc catggttgtt 9480  
 tatgttcogga tgtgatgtga gaactgtatc ctagcaagat tttaaaagga agtatatgaa 9540  
 agaagaacct cagtggcaaa tctaactctt ttatatctct ctacaggggc gccgcgtggg 9600  
 gacaattcaa cgcgtctgtg aggggagcgt ttccctgctc gcaggtctgc agcgaggagc 9660

ES 2 648 865 T3

cgtaattttt gcttgcgcc gtgcggccat caaaatgtat ggatgcaaat gattatacat 9720  
 ggggatgtat gggctaaatg tacggggcag agtcacatca tgcccctgag ctgcgcacgt 9780  
 caagactgtc aaggagggtg ttctgggect ccattgctgc ggccgggtga cccggcgggg 9840  
 acgaggcctt aagttcgaac gtacgagctc cggattatcg atgataagct gtcaaagatg 9900  
 agaattaatt ccacggacta tagactatac tagatactcc gtctactgta cgatacactt 9960  
 ccgctcaggt ccttgtcctt taacgaggcc ttaccactct tttgttactc tattgatcca 10020  
 gctcagcaaa ggcagtgtga tctaagattc tatcttcgcg atgtagtaaa actagctaga 10080  
 ccgagaaaga gactagaaat gcaaaaaggca cttctacaat ggctgcoatc attattatcc 10140  
 gatgtgacgc tgcagcttct caatgatatt cgaatacgtc ttgaggagat acagcctaata 10200  
 atccgacaaa ctgttttaca gatttacgat cgtacttgtt acccatcatt gaattttgaa 10260  
 catccgaacc tgggagtttt ccctgaaaca gatagtatat ttgaacctgt ataataatat 10320  
 atagtctagc gctttacgga agacaatgta tgtatttcgg ttccctggaga aactattgca 10380  
 tctattgcat aggtaatctt gcacgtcgca tccccggctc attttctgcg tttccatctt 10440  
 gcacttcaat agcatatctt tgttaacgaa gcatctgtgc ttcattttgt agaacaaaaa 10500  
 tgcaacgcga gagcgctaata ttttcaaca aagaatctga gctgcatttt tacagaacag 10560  
 aatgcaacg cgaaagcgct attttaccac cgaagaatct gtgcttcatt tttgtaaaac 10620  
 aaaaatgcaa cgcgacgaga gcgctaattt ttcaacaaa gaatctgagc tgcattttta 10680  
 cagaacagaa atgcaacgcg agagcgctat tttaccaaca aagaatctat acttcttttt 10740  
 tgttctacaa aaatgcatcc cgagagcgct atttttctaa caaagcatct tagattactt 10800  
 tttttctcct ttgtgcgctc tataatgcag tctcttgata actttttgca ctgtaggtcc 10860  
 gttaagggtg gaagaaggct actttgggtg ctattttctc ttccataaaa aaagcctgac 10920  
 tccacttccc gcgtttactg attactagcg aagctgcggg tgcatttttt caagataaag 10980  
 gcatccccga ttatattcta taccgatgtg gattgcgcat actttgtgaa cagaaagtga 11040  
 tagcgttgat gattcttcat tggtcagaaa attatgaacg gtttcttcta ttttgtctct 11100  
 atatactacg tataggaaat gtttacattt tcgtattgtt ttcgattcac tctatgaata 11160  
 gttcttacta caattttttt gtctaaagag taatactaga gataaacata aaaaatgtag 11220  
 aggtcgagtt tagatgcaag ttcaaggagc gaaagggtga tgggtagggt atatagggat 11280  
 atagcacaga gatatatagc aaagagatac ttttgagcaa tgtttgtgga agcgggtattc 11340  
 gcaatgccgg ctttccccgt caagctctaa atcgggggct cccttaggg ttccgattta 11400  
 gtgctttacg gcacctgac cccaaaaaac ttgattaggg tgatggttca cgtagtgggc 11460  
 catcgccctg atagacggtt tttcgccctt tgacgttggg gtccacgctc tttaatagtg 11520

ES 2 648 865 T3

gactcttggt ccaaactgga acaacactca accctatctc ggtctattct tttgatttat 11580  
aagggatttt gccgatttcg gcctattggt taaaaaatga gctgatttaa caaaaattta 11640  
acgcgaattt taacaaaata ttaacgttta caatttcctg atgcggtatt ttctccttac 11700  
gcatctgtgc ggtatttcac accgcagggt aataactgat ataattaaat tgaagctcta 11760  
atltgtgagt ttagtataca tgcatttact tataatacag ttttttagtt ttgctggccg 11820  
catcttctca aatatgcttc ccagcctgct tttctgtaac gttcaccctc taccttagca 11880  
tcctttccct ttgcaaatag tcctcttcca acaataataa tgtcagatcc tgtagagacc 11940  
acatcatcca cggttctata ctgttgaccc aatgcgtctc ccttgtcatc taaaccaca 12000  
ccgggtgtca taatcaacca atcgtaacct tcactctctc cacccatgct tctttgagca 12060  
ataaagccga taacaaaatc tttgtcgtc ttcgcaatgt caacagtacc cttagtatat 12120  
tctccagtag atagggagcc cttgcatgac aattctgcta acatcaaag gcctctaggt 12180  
tcctttgta cttcttctgc cgctgcttc aaaccgctaa caatacctgg gccaccaca 12240  
ccgtgtgcat tcgtaatgtc tgccattct gctattctgt atacaccgcg agagtactgc 12300  
aatttgactg tattaccaat gtcagcaaat tttctgtctt cgaagagtaa aaaattgtac 12360  
ttggcggata atgcctttag cggcttaact gtgccctcca tggaaaaatc agtcaagata 12420  
tccacatgtg tttttagtaa acaaattttg ggacctaag cttcaactaa ctccagtaat 12480  
tccttggtag tacgaacatc caatgaagca cacaagtttg tttgcttttc gtgcatgata 12540  
ttaaatagct tggcagcaac aggactagga tgagtagcag cacgttcctt atatgtagct 12600  
ttcgacatga tttatcttcg tttcctgcag gtttttgctc tgtgcagttg ggtaagaat 12660  
actgggcaat ttcattgttc ttcaacacta catatgcgta tatataccaa tctaagtctg 12720  
tgctccttcc ttcgttcttc cttctgttcg gagattaccg aatcaaaaaa atttcaaaga 12780  
aaccgaaatc aaaaaaaga ataaaaaaa aatgatgaat tgaattgaaa agcgtgggtgc 12840  
actctcagta caatctgctc tgatgcgca tagttaagcc agccccgaca cccgccaca 12900  
cccgtgacg cgcctgacg ggcttgtctg ctcccggcat ccgcttacag acaagctgtg 12960  
accgtctccg ggagctgcat gtgtcagagg ttttcaccgt catcaccgaa acgcgcga 13018

## REIVINDICACIONES

1. Un proceso de producción de etanol, por el que el proceso comprende la etapa de fermentar un medio con una célula de levadura, por el que el medio contiene o se alimenta con:

- 5 a) una fuente de al menos una de una hexosa y una pentosa;  
 b) una fuente de ácido acético; y,  
 c) una fuente de glicerol,

10 por el que la célula de levadura fermenta ácido acético, glicerol y al menos una de la hexosa y pentosa en etanol, y opcionalmente, recuperación del etanol, por el que la célula de levadura comprende un gen exógeno que codifica una enzima con actividad de acetaldehído deshidrogenasa, gen que confiere a la célula la capacidad de convertir el ácido acético en etanol, y por el que la célula de levadura comprende un gen bacteriano que codifica una enzima con actividad de glicerol deshidrogenasa unida a NAD<sup>+</sup>, y en el que la célula de levadura comprende una modificación genética que aumenta la actividad específica de dihidroxiacetona cinasa, en el que la célula de levadura no es una célula de levadura que comprende una modificación genética que reduce la actividad de formiato deshidrogenasa dependiente de NAD<sup>+</sup> específica en la célula, en el que el aumento o reducción de una actividad específica de una enzima particular es un aumento o disminución en comparación con la actividad específica de esa enzima en una célula de levadura no mutada por lo demás idéntica.

2. Un proceso según la reivindicación 1, en el que el gen exógeno que codifica la enzima con actividad de acetaldehído deshidrogenasa comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos uno de:

- 20 i) al menos el 64 % de identidad de secuencia de aminoácidos con SEQ ID NO: 1,  
 ii) al menos el 76 % de identidad de secuencia de aminoácidos con SEQ ID NO: 3 y,  
 iii) al menos el 61 % de identidad de secuencia de aminoácidos con SEQ ID NO: 5; y,

25 en el que el gen bacteriano que codifica una enzima con actividad de glicerol deshidrogenasa unida a NAD<sup>+</sup> comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos el 45 % de identidad de secuencia de aminoácidos con SEQ ID NO: 7.

3. Un proceso según la reivindicación 1 o 2, en el que la célula de levadura comprende una modificación genética que reduce la actividad específica de glicerol 3-fosfato deshidrogenasa dependiente de NAD<sup>+</sup> en la célula.

30 4. Un proceso según la reivindicación 3, en el que la modificación genética que reduce la actividad específica de glicerol 3-fosfato deshidrogenasa dependiente de NAD<sup>+</sup> en la célula es una modificación genética que reduce o inactiva la expresión de un gen endógeno que codifica una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos el 70 % de identidad de secuencia con SEQ ID NO: 16.

35 5. Un proceso según la reivindicación 1 o 2, en el que la modificación genética que aumenta la actividad específica de dihidroxiacetona cinasa es expresión en exceso de una secuencia de nucleótidos que codifica una dihidroxiacetona cinasa, y en el que, preferentemente, la secuencia de nucleótidos que codifica la dihidroxiacetona cinasa comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos el 50 % de identidad de secuencia de aminoácidos con al menos una de SEQ ID NO: 8, 9 y 25.

6. Un proceso según una cualquiera de las reivindicaciones 2 - 5, en el que la célula comprende además una modificación genética que aumenta al menos uno de:

- 40 i) la actividad de acetil-CoA sintetasa específica, y en el que la modificación genética es la expresión en exceso de una secuencia de nucleótidos que codifica una acetil-CoA sintetasa; y,  
 ii) el transporte de glicerol en la célula mediante expresión en exceso de una secuencia de nucleótidos que codifica al menos una de una proteína de captación de glicerol y un canal de glicerol, en el que, preferentemente, la secuencia de nucleótidos que codifica la proteína de captación de glicerol comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos el 50 % de identidad de secuencia de aminoácidos con al menos una de SEQ ID NO: 10 y 11, y en el que, preferentemente, la secuencia de nucleótidos que codifica el canal de glicerol comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos el 30 % de identidad de secuencia de aminoácidos con la secuencia de aminoácidos entre los aminoácidos 250 y 530 de SEQ ID NO: 12.

50 7. Un proceso según la reivindicación 6, en el que la secuencia de nucleótidos que codifica una acetil-CoA sintetasa es la acetil-CoA sintetasa codificada por el gen ACS1 o ACS2 de *S. cerevisiae*.

8. Un proceso según la reivindicación 6 o 7, en el que la modificación genética aumenta la actividad de acetil-CoA sintetasa específica en condiciones anaerobias mediante la expresión en exceso de una secuencia de nucleótidos que codifica una acetil-CoA sintetasa, por el que la secuencia de nucleótidos está operativamente unida a un promotor que es activo en condiciones anaerobias.
- 5 9. Un proceso según una cualquiera de las reivindicaciones 2-8, en el que la célula de levadura comprende al menos uno de:
- i) un gen de xilosa isomerasa exógeno funcional, gen que confiere a la célula la capacidad de isomerizar xilosa en xilulosa; y,
  - 10 ii) genes exógenos funcionales que codifican una L-arabinosa isomerasa, una L-ribulocinasa y una L-ribulosa-5-fosfato 4-epimerasa, genes que juntos confieren a la célula la capacidad de convertir L-arabinosa en D-xilulosa 5-fosfato, y en el que, preferentemente, la célula de levadura comprende al menos una modificación genética adicional que produce una característica seleccionada del grupo que consiste en:
    - a) elevada actividad de xilulosa cinasa específica;
    - b) elevado flujo de la vía de pentosa fosfato;
    - 15 c) reducida actividad específica de aldosa reductasa no específica; y,
    - d) elevado transporte de al menos una de xilosa y arabinosa en la célula hospedadora.
10. Un proceso según una cualquiera de las reivindicaciones precedentes, en el que la célula de levadura es de un género seleccionado del grupo que consiste en *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces* y *Yarrowia*.
- 20 11. Un proceso según la reivindicación 10, en el que la célula de levadura pertenece a una especie seleccionada del grupo que consiste en *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis*, *K. marxianus* y *Schizosaccharomyces pombe*.
12. Un proceso según una cualquiera de las reivindicaciones precedentes, en el que el medio contiene o se alimenta con un hidrolizado lignocelulósico.
- 25 13. Un proceso según una cualquiera de las reivindicaciones precedentes, en el que la célula de levadura fermenta en condiciones anaerobias.
14. Una célula de levadura que comprende:
- a) un gen exógeno que codifica una enzima con actividad de acetaldehído deshidrogenasa, gen que confiere a la célula la capacidad de convertir ácido acético en etanol; y,
  - b) un gen bacteriano que codifica una enzima con actividad de glicerol deshidrogenasa unida a NAD<sup>+</sup>,
  - 30 en la que la célula de levadura no es una célula de levadura que comprende una modificación genética que reduce la actividad de formiato deshidrogenasa dependiente de NAD<sup>+</sup> específica en la célula, en la que la célula de levadura comprende una modificación genética que aumenta la actividad específica de dihidroxiacetona cinasa, en la que el aumento o reducción de una actividad específica de una enzima particular es un aumento o disminución en comparación con la actividad específica de esa enzima en una célula de
  - 35 levadura no mutada por lo demás idéntica.
15. Una célula de levadura según la reivindicación 14, en la que la célula de levadura es además como se define en una cualquiera de las reivindicaciones 2 - 11.

Fig. 1

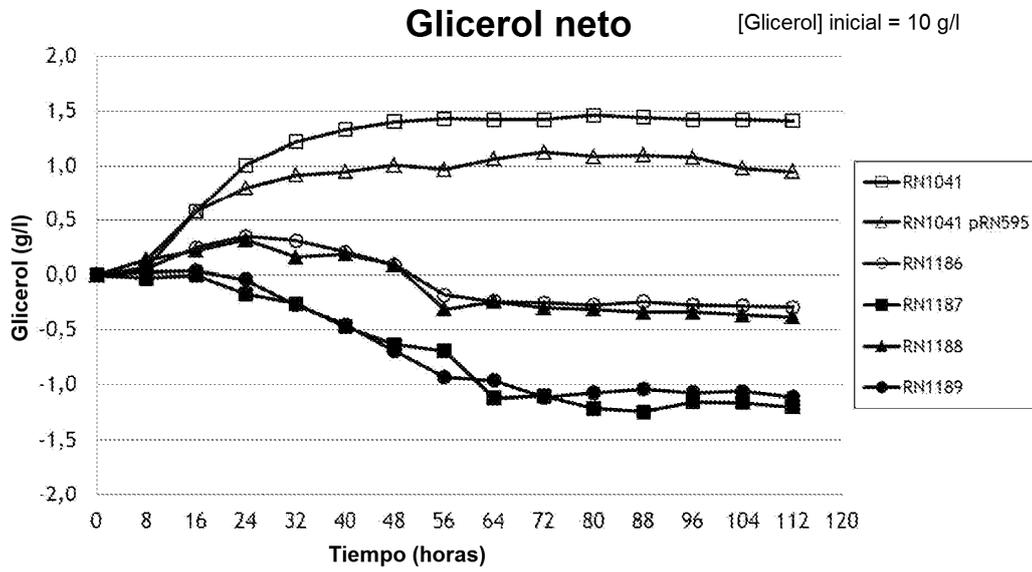


Fig. 2

