(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 532 508**

(51) Int. CI.:

C12P 7/06 (2006.01)
C12N 1/19 (2006.01)
C12N 9/02 (2006.01)
C12N 9/04 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **18.11.2011 E 11788234 (0)**

(97) Fecha y número de publicación de la concesión europea: **17.12.2014 EP 2663645**

(54) Título: **Cepas de levaduras manipuladas para producir etanol a partir de glicerol**

(30) Prioridad:

18.11.2010 EP 10191736
18.11.2010 US 415054 P
05.04.2011 US 201161471836 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
27.03.2015

(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 y
TEUNISSEN, ALOYSIUS WILHELMUS RUDOLPHUS HUBERTUS

(74) Agente/Representante:

LEHMANN NOVO, María Isabel

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 levaduras manipuladas para producir etanol a partir de glicerol

Campo de la invención

La presente invención se refiere a la manipulación metabólica en microorganismos tales como levaduras. En particular, la invención se refiere a cepas de levadura que han sido manipuladas para producir etanol a partir de glicerol. Estas cepas han conservado su capacidad natural de producir etanol a partir de hexosas (glucosa, fructosa, galactosa, etc.) y comprenden una capacidad de manipulación para producir etanol a partir de pentosas tales como xilosa. La invención se refiere, además, a los procesos en los que las cepas manipuladas de la invención producen etanol a partir de glicerol, ya sea como material de alimentación principal de la fermentación, o junto con una o más hexosas y pentosas.

Antecedentes de la invención

Tanto la formación como la degradación de glicerol en la levadura *Saccharomyces cerevisiae* son procesos importantes que han sido estudiados durante más de un siglo. Durante las últimas décadas, se ha hecho hincapié en (1) la supresión de la formación de glicerol durante la producción de etanol. A comienzos del siglo pasado (2) se ha estudiado la optimización de la producción de glicerol a partir de azúcares. Ambos aspectos tienen profundas repercusiones económicas. Más recientemente, se ha introducido un aspecto totalmente nuevo en este campo del glicerol. Se trata de (3) la fermentación anaerobia de glicerol en etanol.

(1) Supresión de la formación de glicerol durante la producción de etanol

El bioetanol es producido por *Saccharomyces cerevisiae* a partir de una variedad de sustratos, incluidos hidrolizados lignocelulósicos de materiales de alimentación no alimentarios (p. ej., cultivos energéticos y residuos agrícolas). Uno de los problemas de la producción de etanol basada en levaduras es que durante la fermentación etanólica anaerobia de materiales de alimentación de azúcar, se forman invariablemente cantidades sustanciales de glicerol como un subproducto. La disimilación de azúcar durante el crecimiento anaerobio de *S. cerevisiae* se produce a través de la fermentación alcohólica. En este proceso, el NADH formado en la glucólisis se re-oxida a través de una alcohol deshidrogenasa NAD^+-dependiente que convierte acetaldehído (formado por descarboxilación de piruvato) en etanol. Esta vía disimilatoria es redox neutra y, por lo tanto, no puede compensar una reducción neta de NAD^+ en NADH que se produce en otros lugares en el metabolismo. Una reducción neta de este tipo de NAD^+ en NADH se produce en asimilación cuando la biomasa de levadura se sintetiza en condiciones anaerobias a partir de azúcares y, p. ej., amoníaco. Bajo condiciones anaerobias, la re-oxidación de NADH en *S. cerevisiae* para compensar la formación de NADH asimilatoria-asociada depende principalmente de la reducción de al menos parte de la fuente de carbono de azúcar en glicerol, resultando un rendimiento menor de etanol. Para abordar este problema, se han explorado varias estrategias de manipulación metabólica para reducir o eliminar la producción de glicerol en *S. cerevisiae* desarrollada en condiciones anaerobias. Dos ejemplos son:

Nissen et al. (2000) describen una estrategia en la que se eliminó la enzima glutamato deshidrogenasa NADPH-dependiente. Glutamina sintetasa y glutamato sintasa se sobre-expresaron. La cepa resultante ya no sintetizaba glutamato a partir de amonio y 2-oxoglutarato a través de una vía que requiere NADPH, sino más bien a través de una vía que requiere NADH y ATP. La cepa resultante tenía un rendimiento un 10% superior en etanol y un rendimiento un 38% inferior en glicerol que la cepa de tipo salvaje.

Guadalupe Medina et al. (2009, Appl. Environ. Microbiol, 76: 190-195) describen una cepa de *S. cerevisiae* en la que la producción del sub-producto glicerol se elimina por la interrupción de los genes glicerol 3-fosfato deshidrogenasa NAD-dependiente endógenos (*GPD1* y *GPD2*). La expresión del gen mhp^F de *E. coli* que codifica la acetaldehído deshidrogenasa NAD-dependiente acetilante restauró la capacidad de la cepa disgregada con GPD de desarrollarse en condiciones anaerobias. Sin embargo, la cepa disgregada con GPD sólo podría desarrollarse en condiciones anaerobias si el medio se suplementa con ácido acético.

(2) Optimización de la producción de glicerol a partir de azúcares

Producción de glicerol a partir de azúcares, en contraposición a minimizar su producción durante la formación de etanol, ha sido un proceso económicamente importante durante la 1ª Guerra Mundial en Alemania. En los últimos años se observa un renovado interés en este proceso como lo demuestra, por ejemplo, el trabajo de Overkamp et al (2002, Appl. Environ. Microbiol. 68: 2814-21).

(3) Fermentación de glicerol en etanol

La fermentación de glicerol en etanol no es factible en condiciones normales de *S. cerevisiae* y bajo condiciones de cultivo completamente anóxicas debido a un equilibrio desequilibrado de NADH. El glicerol se reduce más que el etanol y, por tanto, el organismo no puede deshacerse de su exceso de NADH por re-oxidación en condiciones

anóxicas normales. Algún trabajo inicial se ha hecho en el campo de la fermentación de glicerol en etanol: Yu et al. (2010, Bioresour.Technol. 101(11): 4157-61. Epub 9 de febrero de 2010) describen cepas de *S. cerevisiae* metabólicamente manipuladas para mejorar la producción de etanol a partir de glicerol mediante la sobre-expresión simultánea de glicerol deshidrogenasa (*GCY*), dihidroxiacetona quinasa (*DAK*) y la proteína de absorción de glicerol (*GUP1*). En una comunicación posterior, Yu et al (2010, J. Biotechnol doi: 10.1016 / j.jbiotec.2010.09.932) describen una optimización de su cepa para la producción de etanol a partir de glicerol mediante la supresión de dos genes de producción de glicerol, *FPS1* y *GPD2*. Se demostró que en sus cepas es posible la producción de etanol a partir de glicerol. Sin embargo, el aumento en el rendimiento dependía de las condiciones microaerobias, que puede explicarse por el requisito para la oxidación de NADH. El exceso de NADH producido en la vía a partir de glicerol a etanol al parecer sólo podría reoxidarse a través de una reacción dependiente de oxígeno.

Waks y Silver (2009, Appl. Environ. Microbiol, 75:1867-1875) describen un sistema de organismo dual sintético para la producción de biohidrógeno. En una primera etapa formiato es producido por una cepa manipulada de *S. cerevisiae*, en la que se ha implementado una vía super-productora de formiato. En una segunda etapa, el formiato producido por la levadura manipulada es procesado para formar hidrógeno por *Escherichia coli*. La cepa de *S. cerevisiae* se manipuló para producir formiato expresando la enzima piruvato formiato liasa (PFL) anaerobia de *E. coli*. La producción de formiato se incrementó aún más introduciendo también la expresión de una enzima de aguas abajo, la AdhE de *E. coli*, la enzima bifuncional que reduce acetil-CoA generada por PFL en etanol.

El documento WO 2010/019882 describe levaduras manipuladas genéticamente que son capaces de absorber glicerol y de convertir el glicerol en etanol.

Won-Kyung HONG et al., biotechnology letters vol. 32, nº 8, 2010, páginas 1077-1082 describen la producción mejorada de etanol a partir de glicerol por parte de *Hansenula polymorpha* manipulada que expresa genes piruvato descarboxilasa y aldehído deshidrogenasa de *Zymomonas mobilis*.

Es un objeto de la presente invención proporcionar para levaduras que sean capaces de producir etanol a partir de glicerol al tiempo que conserven sus capacidades de fermentación de hexosas (glucosa, fructosa, galactosa, etc.), así como pentosas tales como xilosa, así como procesos en los que estas cepas se utilizan para la producción de etanol y/u otros productos de la fermentación.

Descripción de la invención

Definiciones

La identidad de secuencia se define en esta memoria como una relación entre dos o más secuencias de aminoácidos (polipéptido o proteína) o dos o más secuencias de ácidos nucleicos (polinucleótidos), según se determina comparando las secuencias. En la técnica, "identidad" también significa el grado de relación de secuencia entre secuencias de aminoácidos o de ácidos nucleicos, según sea el caso, tal como se determina por la coincidencia entre cadenas de tales secuencias. La "similitud" entre dos secuencias de aminoácidos se determina comparando 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" se pueden calcular fácilmente por métodos conocidos. Las expresiones "identidad de secuencia" o "similitud de secuencia" significan que dos secuencias de (poli)péptidos o dos secuencias de nucleótidos, cuando se alinean óptimamente, de preferencia en toda su longitud (de al menos la secuencia más corta en la comparación) y maximizando el número de coincidencias y minimizando el número de huecos tal como mediante los programas ClustalW (1.83), GAP o BESTFIT utilizando parámetros por defecto, comparten al menos un determinado porcentaje de identidad de secuencia según se define en otra parte en esta memoria. GAP utiliza el algoritmo de alineación global de Needleman y Wunsch para alinear dos secuencias en toda su longitud, maximizando el número de coincidencias y minimizando el número de huecos. Generalmente, se utilizan los parámetros por defecto de GAP, con una penalización por creación de hueco = 50 (nucleótidos) / 8 (proteínas) y con una penalización por extensión de hueco = 3 (nucleótidos) / 2 (proteínas). Para nucleótidos, la matriz de puntuación por defecto utilizada es nwsgapdna y para proteínas la matriz de puntuación por defecto es Blosum62 (Henikoff y Henikoff, 1992, PNAS 89, 915-919). Un programa de alineamiento múltiple preferido para la alineación de secuencias de proteínas de la invención es ClustalW (1.83) utilizando una matriz blosum y configuraciones por defecto (penalización por apertura de hueco: 10; penalización por extensión de hueco: 0,05). Está claro que cuando se dice que las secuencias de ARN son esencialmente similares o que tienen un determinado grado de identidad de secuencia con secuencias de ADN, timina (T) en la secuencia de ADN se considera igual a uracilo (U) en la secuencia de ARN. Alineamientos de secuencias y puntuaciones de identidad de secuencia porcentual se pueden determinar utilizando programas de ordenador, como el Paquete GCG Wisconsin, versión 10.3, disponible de Accelrys Inc., 9685 Scranton Road, San Diego, CA 92121-3752 EE.UU. o el software de código abierto Emboss para Windows (versión actual 2.10.0-0.8). Alternativamente, el porcentaje de similitud o identidad se puede determinar mediante la búsqueda en bases de datos tales como FASTA, BLAST, etc.

Una variante de una secuencia de nucleótidos o de aminoácidos descrita en esta memoria también se puede definir como una secuencia de nucleótidos o de aminoácidos que tiene una o varias sustituciones, inserciones y/o

deleciones en comparación con la secuencia de nucleótidos o de aminoácidos específicamente descrita en este memoria (p. ej., en el listado de secuencias).

Opcionalmente, en la determinación del grado de similitud de aminoácidos, la persona experta puede tener en cuenta también las denominadas sustituciones de aminoácidos "conservativas", como resultará evidente para la persona experta. Sustituciones de aminoácidos conservativas se refieren a la capacidad de intercambio de residuos 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 alifáticas-hidroxilo 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 de carácter básico es lisina, arginina e histidina; y un grupo de aminoácidos que tienen cadenas laterales que contienen azufre es cisteína y metionina. Grupos de sustitución conservativa de aminoácidos preferidos son: valina-leucina-isoleucina, fenilalanina-tirosina, lisina-arginina, alanina-valina y asparagina-glutamina. Variantes de sustitución de la secuencia de aminoácidos descrita en esta memoria son aquellas en las que al menos un residuo en las secuencias descritas ha sido eliminado y en su lugar se ha insertado un residuo diferente. Preferiblemente, el cambio de aminoácido es conservativo. Sustituciones conservativas preferidas para cada uno de los aminoácidos que se producen de forma natural son como sigue: Ala a ser; Arg a lys; Asn a gln o his; Asp a glu; Cys a ser o ala; Gin a asn; Glu a asp; Gly a pro; His a asn o gln; Ile a leu o val; Leu a ile o val; Lys a arg; gln o glu; Met a leu o ile; Phe a met, leu o tyr; Ser a thr; Thr a ser; Trp a tyr; Tyr a trp o phe; y Val a ile o leu.

Las secuencias de nucleótidos de la invención también pueden definirse por su capacidad para hibridarse con partes de secuencias de nucleótidos específicas descritas en esta memoria, respectivamente, bajo condiciones de hibridación moderadas o, preferiblemente, rigurosas. Condiciones de hibridación rigurosas se definen aquí como condiciones que permiten que una secuencia de ácido nucleico de al menos aproximadamente 25, de preferencia aproximadamente 50 nucleótidos, 75 ó 100 y lo más preferiblemente de aproximadamente 200 o más nucleótidos se hibride a una temperatura de aproximadamente 65°C en una disolución que comprende aproximadamente sal 1 M, preferiblemente 6 x SSC o cualquier otra disolución que tenga una fuerza iónica equiparable, y lavado a 65°C en una disolución que comprende sal aproximadamente 0,1 M, o menos, preferiblemente 0,2 x SSC o cualquier otra disolución con una fuerza iónica equiparable. Preferiblemente, la hibridación se realiza durante la noche, es decir, al menos durante 10 horas, y preferiblemente el lavado se lleva a cabo durante al menos una hora con al menos dos cambios de la disolución de lavado. Estas condiciones permitirán habitualmente la hibridación específica de secuencias que tienen aproximadamente el 90% o más de identidad de secuencia.

Condiciones moderadas se definen en esta memoria como condiciones que permiten que una secuencia de ácido nucleico de al menos 50 nucleótidos, preferiblemente de aproximadamente 200 o más nucleótidos, se hibride a una temperatura de aproximadamente 45°C en una disolución que comprende sal aproximadamente 1 M, preferiblemente 6 x SSC o cualquier otra disolución con una fuerza iónica equiparable, y el lavado a temperatura ambiente en una disolución que comprende sal aproximadamente 1 M, preferiblemente 6 x SSC o cualquier otra disolución con una fuerza iónica equiparable. Preferiblemente, la hibridación se realiza durante una noche, es decir, al menos durante 10 horas, y preferiblemente el lavado se realiza durante al menos una hora con al menos dos cambios de la disolución de lavado. Estas condiciones permitirán habitualmente la hibridación específica de secuencias que tienen hasta el 50% de identidad de secuencia. La persona experta en la técnica será capaz de modificar estas condiciones de hibridación con el fin de identificar específicamente las secuencias que varían en identidad entre 50% y 90%.

Una "construcción de ácido nucleico" o "vector de ácido nucleico" se entiende en esta memoria que significa una molécula de ácido nucleico hecha por el hombre que resulta del uso de la tecnología de ADN recombinante. La expresión "construcción de ácido nucleico", por lo tanto, no incluye moléculas de ácido nucleico que se producen de forma natural, aunque una construcción de ácido nucleico puede comprender (partes de) moléculas de ácido nucleico que se producen de forma natural. Las expresiones "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 huéspedes o en organismos huéspedes compatibles con estas secuencias. Estos vectores de expresión incluyen típicamente al menos secuencias reguladoras de la transcripción adecuadas y, opcionalmente, señales de terminación de la transcripción 3’. También pueden estar presentes factores adicionales necesarios o útiles para efectuar la expresión tales como elementos potenciadores de la expresión. El vector de expresión se introducirá en una célula huésped adecuada y será capaz de efectuar la expresión de la secuencia codificadora en un cultivo celular in vitro de la célula huésped. El vector de expresión será adecuado para la replicación en la célula u organismo huésped de la invención.

Tal como se utiliza en esta memoria, el término "promotor" o la expresión "secuencia reguladora de la transcripción" se refiere a un fragmento de ácido nucleico que funciona para controlar la transcripción de una o más secuencias codificadoras, y está situado aguas arriba con respecto a la dirección de la transcripción del sitio de iniciación de la transcripción de la secuencia codificadora, y se identifica estructuralmente por la presencia de un sitio de unión para

la ARN polimerasa ADN-dependiente, sitios de iniciación de la transcripción y cualesquiera otras secuencias de ADN, que incluyen, pero no se limitan a sitios de unión del factor de transcripción, sitios de unión del represor y activador de la proteína, y cualesquiera otras secuencias de nucleótidos conocidas para un experto en la técnica que actúan directa o indirectamente para regular la cantidad de transcripción a partir del promotor. Un promotor "constitutivo" es un promotor que es activo en la mayoría de los tejidos bajo la mayoría de condiciones fisiológicas y de desarrollo. Un promotor "inducible" es un promotor que está regulado fisiológicamente o por el desarrollo, p. ej., por la aplicación de un inductor químico.

La expresión "marcador seleccionable" es un término familiar para un experto ordinario en la técnica y se utiliza en esta memoria para describir cualquier entidad genética que, cuando se expresa, se puede utilizar para seleccionar una célula o células que contienen el marcador seleccionable. El término "informador" se puede utilizar de manera indistinta con marcador, a pesar de que se utiliza principalmente para hacer referencia a marcadores visibles tales como la proteína verde fluorescente (GFP). Marcadores seleccionables pueden ser dominantes o recesivos o bidireccionales.

Tal como se utiliza en esta memoria, la expresión "enlazado operativamente" se refiere a una unión de elementos de polinucleótidos en una relación funcional. Un ácido nucleico está "enlazado operativamente" cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, una secuencia reguladora de la transcripción está enlazada operativamente a una secuencia codificadora si afecta a la transcripción de la secuencia codificadora. Enlazada operativamente significa que las secuencias de ADN que se enlazan son típicamente contiguas y, cuando es necesario unir dos regiones codificadoras de proteínas, son contiguas y están en marco de lectura.

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

"Hongos" (singular hongo) se entienden en esta memoria como microorganismos eucariotas heterótrofos que digieren su comida externamente, absorbiendo moléculas de nutrientes en sus células. Los hongos son un reino independiente de organismos eucariotas e incluyen levaduras, mohos y setas. Los términos hongos, hongo y fúngico, tal como se utilizan en esta memoria, incluyen, por lo tanto, expresamente levaduras, así como hongos filamentosos.

El término "gen" significa un fragmento de ADN que comprende una región (región transcrita), que se transcribe en una molécula de ARN (p. ej. un ARNm) en una célula, enlazado operativamente a regiones reguladoras adecuadas (p. ej., un promotor). Un gen comprenderá habitualmente varios fragmentos enlazados operativamente tales como un promotor, una secuencia conductora 5', una región codificadora y una secuencia de no traducida 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 enlazada operativamente a regiones reguladoras apropiadas, particularmente un promotor, se transcribe en un ARN, que es biológicamente activo, es decir, que es capaz de ser traducido a una proteína o péptido biológicamente activo.

El término "homólogo", cuando se utiliza para indicar la relación entre una molécula (recombinante) de ácido nucleico o molécula de polipéptido dada y un organismo huésped o célula huésped dado, se entiende que significa en la naturaleza la molécula de ácido nucleico o polipéptido es producida por una célula u organismos huéspedes de la misma especie, preferentemente de la misma variedad o cepa. Si es homóloga a una célula huésped, una secuencia de ácido nucleico que codifica un polipéptido estará típicamente (pero no necesariamente) enlazada operativamente a otra secuencia promotora (heteróloga) y, en su caso, otra secuencia señal secretora (heteróloga) y/o secuencia de terminador que en su entorno natural. Se entiende que las secuencias reguladoras, secuencias señal, secuencias de terminador, etc. también pueden ser homólogas a la célula huésped. En este contexto, el uso de sólo elementos de secuencia "homólogos" permite la construcción de organismos manipulados genéticamente (GMOs - siglas en inglés) "auto-clonados" (la auto-clonación se define aquí como en la Directiva Europea 98/81/EC Anexo II). Cuando se utiliza para indicar la relación de dos secuencias de ácidos nucleicos, el término "homólogo" significa que una secuencia de ácido nucleico monocatenario puede hibridar con una secuencia de ácido nucleico de una sola hebra complementaria. El grado de hibridación puede depender de un cierto número de factores, que incluyen la magnitud de identidad entre las secuencias y las condiciones de hibridación tales como temperatura y concentración de sal como se discute más adelante.

Los términos "heterólogos" y "exógenos", cuando se utilizan 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 el que está presente, o que se encuentra en una célula o lugar o lugares en el genoma o secuencia de ADN o ARN que difieren de aquel en el que se encuentra en la naturaleza. Ácidos nucleicos o proteínas heterólogos y exógenos no son endógenos a la célula en la que se introducen, pero se han obtenido de otra célula o se han producido de forma sintética o recombinante. Generalmente, aunque no necesariamente, tales ácidos nucleicos codifican proteínas, es decir, proteínas exógenas, que no se producen normalmente por la célula en la que se transcribe o se expresa el ADN. De manera similar, ARN exógeno codifica

proteínas que no se expresan normalmente en la célula en la que está presente el ARN exógeno. A ácidos nucleicos y proteínas exógenos/heterólogos también se les puede aludir como ácidos nucleicos o proteínas extraños. Cualquier ácido nucleico o proteína que un experto en la técnica reconocería como extraño a la célula en la que se expresa queda abarcado en esta memoria por la expresión ácido nucleico o proteína heterólogo o exógeno. Los términos heterólogo y exógeno también se aplican a combinaciones no naturales de secuencias de ácidos nucleicos o aminoácidos, es decir, combinaciones en donde al menos dos de las secuencias combinadas son extrañas una con respecto a la otra.

La "actividad específica" de una enzima se entiende en esta memoria que significa la cantidad de actividad de una enzima particular por cantidad de proteína total de la célula huésped, habitualmente expresada en unidades de actividad de la enzima por mg de proteína total de la célula huésped. En el contexto de la presente invención, la actividad específica de una enzima particular puede ser aumentada o disminuida en comparación con la actividad específica de esa enzima en una célula huésped (por lo demás idéntica) de tipo salvaje.

"Condiciones anaerobias" o un proceso de fermentación anaerobia se define en esta memoria como condiciones o una operación de proceso de fermentación en ausencia de oxígeno o en el que sustancialmente no se consume oxígeno, preferiblemente se consume menos de 5, 2,5 ó 1 mmol/L/h, más preferiblemente 0 mmol/L/h (es decir, el consumo de oxígeno no es detectable), y en el que las moléculas orgánicas sirven tanto como donantes de electrones como aceptores de electrones.

Descripción detallada de la invención

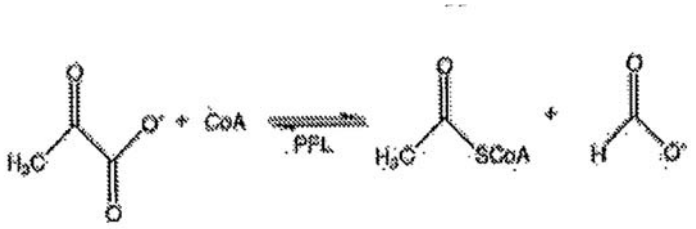
El biodiesel está teniendo y continuará teniendo un impacto sustancial en el sector de los biocombustibles. Se puede producir a partir de una diversidad de fuentes, incluyendo algas, colza y palma. Una amplia lista de fuentes de materiales de alimentación potenciales está disponible de "Feedstock and Biodiesel Characteristics Report," Renewable Energy Group, Inc., www.regfuel.com (2009)". Los aceites vegetales a partir de los cuales se produce el biodiesel mediante interesterificación resultan inevitablemente en grandes corrientes de subproductos de glicerol. Además, el material vegetal del que se extrae el aceite dará lugar a un subproducto lignocelulósico. Este último subproducto contendrá celulosa y hemicelulosas en diferentes concentraciones relativas. Estos azúcares poliméricos se pueden hidrolizar en glucosa y pentosas tales como xilosa, respectivamente, por procedimientos conocidos que todavía están optimizados. Por lo tanto, pueden resultar dos corrientes de subproductos de la producción de biodiesel. La presente invención describe un método para la fermentación concomitante tanto de glicerol como de los azúcares en etanol. Este proceso se puede aplicar para diversos materiales de alimentación. Como un ejemplo, a continuación se presenta con cierto detalle la situación en la industria del aceite de palma.

La industria del aceite de palma produce grandes cantidades de subproductos. El aceite consiste en sólo el 10% de la biomasa total producida en la plantación. El resto se compone de materiales lignocelulósicos tales como hojas de la palma de aceite, troncos y racimos vacíos (EFB - siglas en inglés) (véase, p. ej., http://umpir.ump Edu.my/697/l/Kamarul_Azlan_Abd_Hamid.pdf). Estos materiales lignocelulósicos se pueden utilizar para la producción de productos de valor añadido tal como etanol que se puede utilizar como biocombustible. Los materiales lignocelulósicos se pueden hidrolizar por métodos conocidos en corrientes que contienen hexosas y pentosas tales como glucosa y xilosa, que luego pueden ser fermentados a etanol por las cepas de levadura manipuladas tal como se describe, p. ej., por Kuypers et al (2005, FEMS Yeast Res. 5:399-409). Sin embargo, las corrientes que contienen hexosa y pentosa a partir de la hidrólisis de este tipo de materiales lignocelulósicos son relativamente diluidas y, por lo tanto, producirán concentraciones relativamente bajas de etanol de no más de aproximadamente 5% (v/v).

Al mismo tiempo, se generan grandes cantidades de glicerol concentrado como un subproducto en la producción de biodiesel a partir de reacciones de transesterificación utilizando aceites vegetales tales como aceite de palma y alcoholes. Se prevé que la disponibilidad de glicerol bruto aumente en los próximos años como consecuencia del crecimiento de la producción de biodiesel en todo el mundo. Por consiguiente, estarán disponibles grandes cantidades de glicerol concentrado a bajo costo cerca de las plantaciones de aceite de palma. Esto ofrece la posibilidad de aumentar la concentración de carbono de corrientes diluidas obtenidas en la hidrólisis de materiales lignocelulósicos anteriormente descrita, al mezclar en el glicerol altamente concentrado disponible. La utilización de tales corrientes mixtas requeriría cepas de levaduras manipuladas que pueden fermentar no sólo hexosas y pentosa, sino también glicerol para dar etanol. Sin embargo, el consumo anaerobio de gran cantidad de glicerol produciría un desequilibrio redox en la levadura. La presente invención aborda este problema mediante la manipulación de la levadura para producir ácido fórmico, además de etanol. Por lo tanto, la presente invención proporciona cepas de levadura manipuladas para producir etanol y ácido fórmico a partir de fuentes de carbono que contiene uno o más de glicerol, hexosa y pentosa, así como procesos en los que estas cepas se utilizan para producir etanol y ácido fórmico a partir de estas fuentes de carbono.

En un primer aspecto, la invención se refiere a una célula huésped fúngica que comprende un gen exógeno que codifica una enzima con la capacidad de convertir piruvato y coenzima-A en formiato y acetil-CoA. Una enzima con la capacidad de convertir el piruvato y la coenzima-A en formiato y acetil-CoA cataliza la reacción (CE 2.3.1.54):

piruvato + coenzima A (CoA) ↔ acetil-CoA + formiato.



Una enzima de este tipo se entiende en esta memoria como una enzima con actividad de piruvato formiato liasa y se la alude como una piruvato formiato liasa (PFL) o formiato C-acetiltransferasa.

Un gen exógeno adecuado que codifica una enzima que tiene actividad de piruvato formiato liasa es, p. ej., una piruvato formiato liasa procariota tal como la piruvato formiato liasa de *E. coli*. La piruvato formiato liasa de *E. coli*. es un dímero de PflB (codificado por *pfl*B), cuya maduración requiere la activación de la enzima PflAE (codificada por *pfl*A), S-adenosilmetionina radical, y un único donante de electrones, que en el caso de *E. coli* es flavodoxina (Buis y Broderick, 2005, Arch. Biochem. Biophys. 433:288-296; Sawers y Watson, 1998, Mol. Microbiol. 29: 945-954). Sin embargo, Waks y Silver (supra) han demostrado que para la activación de la piruvato formiato liasa en levaduras sólo se requiere la coexpresión de una enzima activante, pero la expresión de flavodoxina no es necesaria.

El gen exógeno que codifica una enzima con actividad de piruvato formiato liasa comprende preferiblemente una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% de identidad de la secuencia de aminoácidos con SEQ ID NO: 1. Ejemplos adecuados de organismos que comprenden una enzima con actividad de piruvato formiato liasa se proporcionan en la Tabla 1. Ejemplos adicionales de organismos de este tipo se enumeran por Lehtiö y Goldman (2004, Prot. Engin. Design & Selection 17:545-552). Las secuencias de aminoácidos de estas enzimas están disponibles en bases de datos públicas y pueden ser utilizadas por la persona experta para diseñar secuencias de nucleótidos optimizadas en los codones que codifican la enzima correspondiente con actividad de piruvato formiato liasa (véase, p. ej., SEQ ID NO: 2). El gen exógeno que codifica una enzima con actividad de piruvato formiato liasa también puede comprender una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con la secuencia de aminoácidos de SEQ ID NO: 1. Preferiblemente, la secuencia de aminoácidos tiene no más de 420, 380, 300, 250, 200, 150, 100, 75, 50, 40, 30, 20, 10 ó 5 sustituciones, inserciones y/o deleciones de aminoácidos en comparación con SEQ ID NO: 1.

Tabla 1: Enzimas con actividad de piruvato formiato liasa relacionada con pflB de *E. coli*

Organismo	**Identidad de aminoácidos (%)**
Escherichia coli cepa K12 subcepa MG1655	100%
Shigella boydii	100%
Escherichi alberti TW07627	99%
Salmonella enteric	97%
Citrobacter rodentium ICC168	97%
Klebsiella pneumoniae NTUH-K2044	96%
Yersinia aldovae ATCC 35236	91%
Proteus mirabilis HI4320	87%
Haemophilus influenza Rd KW20	86%
Actinobacillus succiongenes 130Z	83%
Piromyces sp. E2	57%

La célula huésped de la invención comprende además, preferiblemente, un gen exógeno que codifica la enzima activante PflAE para la activación de la piruvato formiato liasa. La enzima activante piruvato formiato liasa se entiende en esta memoria como una enzima que cataliza la reacción:

S-adenosil-L-metionina + dihidroflavodoxina + [piruvato formiato liasa]-glicina

↕

5'-desoxiadenosina + L-metionina + flavodoxina semiquinona +

radical [formiato C-acetiltransferasa]-glicin-2-ilo

El gen exógeno que codifica la enzima activante piruvato formiato liasa comprende preferiblemente una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% de identidad de la secuencia de aminoácidos con SEQ ID NO: 3. Ejemplos adecuados de organismos que

comprenden una enzima con actividad de piruvato formiato liasa se proporcionan en la Tabla 2. Las secuencias de aminoácidos de estas enzimas están disponibles en bases de datos públicas y pueden ser utilizadas por la persona experta para diseñar secuencias de nucleótidos optimizadas en los codones que codifican la enzima correspondiente con actividad de piruvato formiato liasa (véase, p. ej., SEQ ID NO: 4). El gen exógeno que codifica la enzima con actividad de piruvato formiato liasa también puede comprender una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con la secuencia de aminoácidos de SEQ ID NO: 3. Preferiblemente, la secuencia de aminoácidos tiene no más de 135, 125, 100, 75, 50, 25, 20, 15, 10, 8, 5 ó 2 sustituciones, inserciones y/o deleciones de aminoácidos en comparación con SEQ ID NO: 3.

Tabla 2: Enzimas con actividad de piruvato formiato liasa relacionada con pflA de *E. coli*

Organismo	**Identidad de aminoácidos (%)**
Escherichia coli cepa K12 subcepa MG1655	100%
Shigella boydii	100%
Escherichi alberti TW07627	99%
Salmonella enteric	98%
Citrobacter rodentium ICC168	98%
Klebsiella pneumoniae NTUH-K2044	97%
Yersinia rhodei ATCC 43380	89%
Proteus penneri ATCC 35198	85%
Haemophilus parasuis 29755	70%

En una célula huésped preferida de la invención, los genes exógenos que codifican la enzima con actividad de piruvato formiato liasa y la enzima activante piruvato formiato liasa son del mismo organismo donante, es decir, pueden ser homólogos entre sí. Sin embargo, los genes exógenos que codifican la enzima con actividad de piruvato formiato liasa y la enzima activante piruvato formiato liasa también pueden ser de diferentes organismos donantes, es decir, pueden ser heterólogos entre sí.

En un aspecto, la invención se refiere a métodos para la preparación o construcción de las células de levadura de la invención. Para este propósito se utilizan técnicas de biología genética y molecular estándares que son generalmente conocidas en la técnica y que se han descrito, p. ej., 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, comps., "Current protocols in molecular biology", Green Publishing y Wiley Interscience, Nueva York). Además, la construcción de cepas de levaduras huéspedes mutadas se lleva a cabo por cruces genéticos, esporulación de los diploides resultantes, disección tétrada de las esporas haploides que contienen los marcadores auxotróficos deseados y purificación de colonias de tales levaduras huéspedes haploides en el medio de selección apropiado. Todos estos métodos son métodos genéticos estándares de levaduras conocidos por las personas en la técnica. Véase, por ejemplo, Sherman et al, Methods Yeast Genetics, Cold Spring Harbor Laboratory, NY (1978) y Guthrie et al. (comps.) Guide to Yeast Genetics and Molecular Biology Vol. 194, Academic Press, San Diego (1991).

Los genes exógenos que codifican la enzima con actividad de piruvato formiato liasa y la enzima activante piruvato formiato liasa son, preferiblemente, construcciones de expresión que comprenden la secuencia de nucleótidos que codifica las enzimas enlazadas operativamente a regiones/secuencias reguladoras de la expresión adecuadas para asegurar la expresión de las enzimas tras la transformación de las construcciones de expresión en la célula huésped de la invención. Por lo tanto, el gen o la construcción de expresión comprenderá al menos un promotor que es funcional en la célula huésped enlazada operativamente a la secuencia codificadora. El gen o la construcción puede comprender, además, una secuencia conductora 5’ aguas arriba de la región codificadora y una secuencia 3'-no traducida (extremo 3') que comprende un sitio de poliadenilación y un sitio de terminación de la transcripción aguas abajo de la secuencia codificadora. Se entiende que las secuencias de nucleótidos que codifican la enzima con actividad de piruvato formiato liasa y la enzima activante piruvato formiato liasa pueden estar presentes juntas en una sola construcción de expresión, o cada una de las enzimas puede estar presente en una construcción de expresión separada.

Promotores adecuados para la expresión de las secuencias de nucleótidos que codifican la enzima con actividad de piruvato formiato liasa y la enzima activante piruvato formiato liasa (así como otras enzimas de la invención, véase más adelante) incluyen promotores que son preferiblemente insensibles a la represión de catabolitos (glucosa), que son activos bajo condiciones anaerobias y/o que preferiblemente no requieren xilosa o arabinosa para la inducción. Promotores que tienen estas características están ampliamente disponibles y son conocidos por la persona experta. Ejemplos adecuados de tales promotores incluyen, p. ej., promotores de genes glicolíticos tales como los promotores fosfofructoquinasa (*PPK*), triosa fosfato isomerasa (*TPI*), gliceraldehído-3-fosfato deshidrogenasa (*GPD*, *TDH3* o *GAPDH*), piruvato quinasa (*PYK*), fosfoglicerato quinasa (*PGK*), promotor de glucosa-6-fosfato isomerasa (*PGI1*) de levaduras. Más detalles acerca de tales promotores de levaduras se pueden encontrar en (documento WO

93/03159). Otros promotores útiles son promotores de genes que codifican proteínas ribosomales (*TEF1*), el promotor del gen lactasa (*LAC4*), promotores de alcohol deshidrogenasa (*ADH1*, *ADH4* y similares), el promotor de enolasa (*ENO*) y el promotor transportador de hexosa (glucosa) (*HXT7*). Alternativamente, las secuencias de nucleótidos que codifican la enzima que tiene actividad de piruvato formiato liasa y la enzima activante PflAE se expresan en condiciones anaerobias utilizando un promotor anóxico tal como, p. ej., el promotor *ANB1* de *S. cerevisiae* (SEQ ID NO: 24). Otros promotores, tanto constitutivos como inducibles, y potenciadores o secuencias de activación aguas arriba serán conocidos por los expertos en la técnica. Preferiblemente, el promotor que está enlazado operativamente a la secuencia de nucleótidos según se define arriba es homólogo a la célula huésped. Secuencias de terminación adecuadas se pueden obtener, p. ej., a partir del gen citocromo c1 (*CYC1*) o un gen alcohol deshidrogenasa (p. ej., *ADH1*).

Para aumentar la probabilidad de que la enzima con actividad de piruvato formiato liasa se expresa a niveles suficientes y en forma activa en las células huésped transformadas de la invención, la secuencia de nucleótidos que codifica estas enzimas, así como la enzima activante piruvato formiato liasa y otras enzimas de la invención (véase más adelante), se adaptan preferiblemente para optimizar su uso de codones al de la célula huésped en cuestión. La capacidad de adaptación de una secuencia de nucleótidos que codifica una enzima al uso de codones de una célula huésped se puede expresar como índice de adaptación de codones (CAI). El índice de adaptación de codones se define en esta memoria como una medida 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 huésped u organismo particular. La capacidad de adaptación relativa (w) de cada uno de los codones es la relación del uso de cada uno de los codones al del codón más abundante para el mismo aminoácido. El índice CAI se define como la media geométrica de estos valores de capacidad de adaptación relativa. Se excluyen codones no sinónimos y codones de terminación (que dependen del código genético). Los valores del CAI oscilan entre 0 y 1, indicando los valores más altos una mayor proporción de los codones más abundantes (véase Sharp y 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 tiene preferiblemente un CAI de al menos 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8 ó 0,9. Las más preferidas son las secuencias que han sido optimizadas en codones para la expresión en la célula huésped fúngica en cuestión tal como, p. ej., células de *S. cerevisiae*.

La célula huésped fúngica a transformar con una construcción de ácido nucleico que comprende una secuencia de nucleótidos que codifica una enzima con actividad de piruvato formiato liasa es preferiblemente una célula huésped de levadura. Preferiblemente, la célula huésped es una célula cultivada. La célula huésped de la invención, preferiblemente, es un huésped capaz transportar, de forma activa o pasiva, pentosa (xilosa y preferiblemente también arabinosa) a la célula. La célula huésped contiene preferiblemente glicólisis activa. La célula huésped puede contener además, preferiblemente, una ruta de pentosa fosfato endógena y puede contener actividad de xilulosa quinasa endógena, de manera que xilulosa isomerizada a partir de xilosa puede ser metabolizada en piruvato. El huésped contiene además, preferiblemente, enzimas para la conversión de una pentosa (preferiblemente a través de piruvato) a 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 isopreno. Una célula huésped particularmente preferida es una célula de levadura que de forma natural es capaz de fermentación alcohólica, preferiblemente, fermentación alcohólica anaerobia. La célula huésped de levadura tiene además, preferiblemente, una alta tolerancia a etanol, una alta tolerancia a pH bajo (es decir, capaz de crecer a un pH menor que 5, 4 ó 3) y a ácidos orgánicos tal como ácido láctico, ácido acético o ácido fórmico y productos de degradación de azúcares tales como furfural e hidroxi-metilfurfural, y una alta tolerancia a las temperaturas elevadas. Cualquiera de estas características o actividades de la célula huésped puede estar presente de forma natural en la célula huésped o puede ser introducida o modificada por modificación genética, preferiblemente mediante auto-clonación o por los métodos de la invención que se describen más adelante. Una célula adecuada es una célula cultivada, una célula que se pueden cultivar en proceso de fermentación, p. ej., en fermentación sumergida o en estado sólido. Células huésped particularmente adecuadas son microorganismo eucariotas tales como, p. ej., hongos, sin embargo, las más adecuadas para uso en la presente invención son las levaduras.

Levaduras se definen en esta memoria como microorganismos eucarióticos e incluyen todas las especies de la subdivisión Eumycotina (Yeasts: characterístics and identification, J.A. Barnett, R.W. Payne, D. Yarrow, 2000, 3ª ed, Cambridge University Press, Cambridge, Reino Unido, y, The yeasts, a taxonomic study, C.P. Kurtzman y J.W. Fell (comps.) 1998, 4ª ed., Elsevier Science Publ. BV, Amsterdam, Países Bajos), que crecen predominantemente en forma unicelular. Las levaduras pueden crecer, ya sea por injerto de un tallo unicelular o pueden crecer por la fisión del organismo. Células de levaduras preferidas para uso en la presente invención pertenecen a los géneros *Saccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Schizosaccharomyces*, *Hansenula*, *Kloeckera*, *Schwanniomyces* y *Yarrowia*. Preferiblemente, la levadura es capaz de fermentación anaerobia, más preferiblemente fermentación alcohólica anaerobia. A lo largo de los años se han hecho sugerencias para la introducción de diversos organismos para la producción de bio-etanol a partir de azúcares de los cultivos. En la práctica, sin embargo, todos los principales procesos de producción de bioetanol han continuado utilizando las levaduras del género *Saccharomyces* como productoras de etanol. Esto se debe a las muchas características atractivas de especies de Saccharomyces para procesos industriales, es decir, una alta tolerancia a los ácidos, etanol y osmo-tolerancia, a la capacidad de

crecimiento anaerobio y naturalmente a su gran capacidad fermentativa alcohólica. Especies de levaduras como células huéspedes preferidas incluyen *S. cerevisiae*, *S. exiguus*, *S. bayanus*, *K. lactis*, *K. marxianus* y *Schizosaccharomyces pombe*.

En una realización adicional, la célula huésped de la invención comprende, además, una modificación genética que reduce la actividad específica de formiato deshidrogenasa NAD^+-dependiente en la célula. Formiato deshidrogenasas NAD^+-dependientes (FDH; EC 1.2.1.2) catalizan la oxidación de formiato en bicarbonato donando los electrones a NAD^+. En las células de la invención, la actividad específica de formiato deshidrogenasa se reduce preferiblemente en al menos un factor de 0,8, 0,5, 0,3, 0,1, 0,05 ó 0,01 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que provoca la reducción en la actividad específica, preferiblemente en condiciones anaerobias. La actividad de formiato deshidrogenasa se puede determinar según se describe por Overkamp et al. (2002, Yeast 19:509-520).

Preferiblemente, la actividad de formiato deshidrogenasa se reduce en la célula huésped mediante una o más modificaciones genéticas que reducen la expresión de o inactivan un gen que codifica una formiato deshidrogenasa. Preferiblemente, las modificaciones genéticas reducen o inactivan la expresión de cada una de las copias endógenas del gen que codifica una formiato deshidrogenasa específica en el genoma de la célula. Una célula dada puede comprender múltiples copias del gen que codifica una formiato deshidrogenasa específica con una y la misma secuencia de aminoácidos como resultado de diploidía, poliploidía o aneuploidía. En tales casos, preferiblemente se reduce o inactiva la expresión de cada una de las copias del gen específico que codifica la formiato deshidrogenasa. Alternativamente, una célula puede contener varias (iso)enzimas diferentes con actividad formiato deshidrogenasa que difieren en la secuencia de aminoácidos y que son cada una codificada por un gen diferente. En tales casos, en algunas realizaciones de la invención puede preferirse que sólo se reduzcan o inactiven determinados tipos de las isoenzimas, mientras que otros tipos no se ven afectados. Preferiblemente, sin embargo, se reduce o inactiva la expresión de todas las copias de genes que codifican (iso)enzimas con actividad de formiato deshidrogenasa.

Preferiblemente, una actividad de formiato deshidrogenasa que codifica un gen es inactivada por la deleción de al menos una parte del gen o por la interrupción del gen, con lo que en este contexto el término gen también incluye cualquier secuencia no codificadora aguas arriba o aguas abajo de la secuencia codificadora, cuya deleción o inactivación (parcial) resulta en una reducción de la expresión de la actividad de formiato deshidrogenasa en la célula huésped.

Un gen preferido que codifica una formato deshidrogenasa cuya actividad se ha de reducir o inactivar en la célula de la invención es la *S. cerevisiae FDH1* según se describe por van den Berg y Steensma (1997, Yeast 13:551-559), que codifica la secuencia de aminoácidos de SEQ ID NO: 5 y ortólogos de la misma en otras especies. Por lo tanto, un gen que codifica una formiato deshidrogenasa cuya actividad ha de ser reducida o inactivada en la célula de la invención es preferiblemente un gen que codifica una formiato deshidrogenasa que tiene una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98 ó 99% de identidad de secuencia con SEQ ID NO: 5, o un gen que codifica una formiato deshidrogenasa que tiene una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con la secuencia de aminoácidos de SEQ ID NO: 5.

Sin embargo, en algunas cepas de *S. cerevisiae* un segundo gen que codifica una formiato deshidrogenasa es activo, es decir, el *FDH2*, véase, p. ej., Overkamp et al. (2002, *supra*). Otro gen preferido que codifica una formiato deshidrogenasa cuya actividad se ha de reducir o inactivar en la célula de la invención es, por lo tanto, un *FDH2* de *S. cerevisiae* según se describe por Overkamp et al. (2002, *supra*), que codifica la secuencia de aminoácidos de SEQ ID NO: 6 y ortólogos de la misma en otras especies. Por lo tanto, un gen que codifica una formiato deshidrogenasa cuya actividad ha de ser reducida o inactivada en la célula de la invención es preferiblemente un gen que codifica una formiato deshidrogenasa que tiene una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98 ó 99% de identidad de secuencia con SEQ ID NO: 6 o un gen que codifica una formiato deshidrogenasa que tiene una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con la secuencia de aminoácidos de SEQ ID NO: 6.

En una realización adicional, se reduce o inactiva la actividad de todos los genes en la célula huésped que codifica una formiato deshidrogenasa. En este tipo de células preferiblemente todas las copias de genes endógenos que codifican una formiato deshidrogenasa que tiene una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98 ó 99% de identidad de secuencia con SEQ ID NO: 5 o 6 (o que tiene una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con las secuencias de aminoácidos de SEQ ID NO: 5 o 6) son inactivadas o al menos reducidas en la expresión.

En una realización preferida, la célula huésped de la invención comprende, además, un gen exógeno que codifica una enzima con la capacidad de reducir acetilCoA en acetaldehído, gen que confiere a la célula la capacidad de convertir acetilCoA (y/o ácido acético) en etanol. Una enzima con la capacidad de reducir acetilCoA en acetaldehído se entiende en esta memoria como una enzima que cataliza la reacción (ACDH; EC 1.2.1.10):

acetaldehído + NAD^+ + Coenzima A ↔ acetil-Coenzima A + NADH + H^+.

Por lo tanto, la enzima cataliza la conversión de acetilCoA en acetaldehído (y *viceversa*) y también se la alude como una acetaldehído deshidrogenasa (acetilante NAD-dependiente) 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 conveniencia, se aludirá en esta memoria a una enzima que tiene al menos la capacidad de reducir acetilCoA en cualquiera de acetaldehído o etanol como una "acetaldehído deshidrogenasa". Además, se entiende en esta memoria que la célula huésped tiene actividades de alcohol deshidrogenasa endógenas que permiten a la célula, estando provista de actividad de acetaldehído deshidrogenasa, completar la conversión de acetil-CoA en etanol. Además, también se prefiere que la célula huésped tenga una acetil-CoA sintetasa endógena que permita a la célula, estando provista de actividad de acetaldehído deshidrogenasa, completar la conversión de ácido acético (a través de acetil-CoA) en etanol.

El gen exógeno puede codificar una enzima monofuncional que tiene solamente actividad acetaldehído deshidrogenasa (es decir, una enzima que sólo tiene la capacidad de reducir acetilCoA en acetaldehído) tal como, p. ej., 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 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% identidad de secuencia de aminoácidos con SEQ ID NO: 7. Ejemplos adecuados de procariotas que comprenden enzimas monofuncionales con actividad de acetaldehído deshidrogenasa se proporcionan en la Tabla 3. Las secuencias de aminoácidos de estas enzimas monofuncionales están disponibles en bases de datos públicas y pueden utilizarse por la persona experta para diseñar secuencias de nucleótidos optimizadas en codones que codifiquen la enzima monofuncional correspondiente (véase, p. ej. SEQ ID NO: 8). El gen exógeno que codifica la enzima monofuncional que tiene sólo actividad de acetaldehído deshidrogenasa también puede comprender una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 7.

Tabla 3: Enzimas con actividad de acetaldehído deshidrogenasa relacionada con mhpF de *E. coli*

Organismo	**Identidad de aminoácidos (%)**
Escherichia coli cepa K12 subcepa MG1655	100%
Shigella sonnei	100%
Escherichi coli IAI39	99%
Citrobacter youngae ATCC 29220	93%
*Citrobacter sp. 3*0_2	92%
Klebsiella pneumoniae 342)	87%
Klebsiella variicola	87%
Pseudomonas putida	81%
Ralstonia *eutropha* JMP134	82%
Burkolderia sp. H160	81%
Azotobacter.vinelandii DJ	79%
Ralstonia metallidurans CH34	70%
Xanthobacter autotrophicus Py2	67%
Burkolderia cenocepacia J2315	68%
Frankia sp. EAN1pec	67%
Polaromonas sp. JS666	68%
Burkholderia phytofirmans PsJN	70%
Rhodococcus opacus B4	64%

Preferiblemente, la célula huésped comprende un gen exógeno que codifica una enzima bifuncional con actividad de acetaldehído deshidrogenasa y de alcohol deshidrogenasa, gen que confiere a la célula la capacidad de convertir acetilCoA en etanol. La ventaja de utilizar 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 la canalización directa del producto intermedio entre las enzimas que catalizan reacciones consecutivas en una ruta y ofrece la posibilidad de medios eficientes, exclusivos y protegidas de suministro de metabolitos. La canalización del sustrato, por lo tanto, disminuye el tiempo de tránsito de los productos intermedios, evita la pérdida de los productos intermedios por difusión, protege productos intermedios lábiles de disolvente y previene la entrada de productos intermedios que compitan en las vías metabólicas. Por consiguiente, la enzima bifuncional permite una conversión más eficiente de acetilCoA en etanol, en comparación con las enzimas acetaldehído deshidrogenasa y alcohol deshidrogenasa por separado. Una ventaja adicional de utilizar la enzima bifuncional es que también se puede utilizar en células huésped que tienen poca o ninguna actividad de alcohol deshidrogenasa bajo la condición utilizada tal como, p. ej., condiciones anaerobias y/o condiciones de represión catabólica.

Enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y alcohol deshidrogenasa son conocidas en la técnica de procariotas y protozoos, incluyendo, p. ej., las enzimas bifuncionales codificadas por los genes adhE de *Escherichia coli* y ADH2 de Entamoeba histolytica (véase, p. ej., Bruchaus y Tannich, 1994, J. Biochem. 303: 743-748; Burdette y 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). Enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y actividad de alcohol deshidrogenasa son proteínas más grandes que consisten en alrededor de 900 aminoácidos y son bifuncionales, debido a que exhiben actividad tanto de acetaldehído deshidrogenasa (ACDH; EC 1.2.1.10) como de alcohol deshidrogenasa (ADH; EC 1.1.1.1). El adhE de *E. coli* y ADH2 de *Entamoeba histolytica* muestran una identidad de aminoácidos del 45%. Por lo tanto, en una realización de la invención, un gen exógeno adecuado que codifica una enzima bifuncional con actividad de acetaldehído deshidrogenasa y de alcohol deshidrogenasa comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% identidad de secuencia de aminoácidos con al menos una de SEQ ID NO: 9 y 11. Ejemplos adecuados de procariotas comprenden enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y actividad de alcohol deshidrogenasa se proporcionan en las Tablas 4 y 5. Las secuencias de aminoácidos de estas enzimas bifuncionales están disponibles en bases de datos públicas y pueden ser utilizadas por la persona experta para diseñar secuencias de nucleótidos optimizadas en codones que codifican la correspondiente enzima bifuncional (véase, p. ej., SEQ ID NO: 10 ó 12). El gen exógeno que codifica una enzima bifuncional con actividad de acetaldehído deshidrogenasa y de alcohol deshidrogenasa también puede comprender una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con al menos una de SEQ ID NO: 9 y 11.

Tabla 4: Enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y de alcohol deshidrogenasa relacionadas con adhE de *E. coli*

Organismo	**Identidad de aminoácidos (%)**
Escherichia coli O157:H7 cepa Sakai	100%
Shigella sonnei	100%
Shigella dysenteriae 1012	99%
Klebsiella pneumonia 342	97%
Enterobacter sp. 638	94%
Yersinia pestis biovar Microtus cepa 91001	90%
Serratia proteamaculans 568	90%
Pectobacterium carotovorum WPP14	90%
Sodalis glossinidius cepa 'morsitans'	87%
Erwinia tasmaniensis Et1/99	86%
Aeromonas hydrophila ATCC 7966	81%
Vibria vulnificus YJ016]	76%

Tabla 5: Enzimas bifuncionales con actividad de acetaldehído deshidrogenasa y alcohol deshidrogenasa relacionadas conADH2 de *Entamoeba histolytica*

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

Para la expresión de la secuencia de nucleótidos que codifica la enzima bifuncional que tiene actividades de acetaldehído deshidrogenasa y de alcohol deshidrogenasa, o la enzima que tiene actividad de acetaldehído deshidrogenasa, la secuencia de nucleótidos (a expresar) se coloca en una construcción de expresión en la que está operativamente enlazada a regiones/secuencias reguladoras de la expresión adecuadas para asegurar la expresión de la enzima después de la transformación de la construcción de expresión en la célula huésped de la invención (véase arriba). Promotores adecuados para la expresión de la secuencia de nucleótidos que codifica la enzima bifuncional que tiene actividades de acetaldehído deshidrogenasa y de alcohol deshidrogenasa, o la enzima que tiene actividad de acetaldehído deshidrogenasa incluyen promotores que son preferiblemente insensible a la represión de catabolitos (glucosa), que son activas en condiciones anaerobias y/o que preferiblemente no requieren xilosa o arabinosa para la inducción. Ejemplos de promotores de este tipo se han dado arriba.

Preferiblemente, la secuencia de nucleótidos que codifica la enzima bifuncional que tiene actividades de acetaldehído deshidrogenasa y de alcohol deshidrogenasa, o la enzima que tiene actividad de acetaldehído deshidrogenasa está adaptada para optimizar su uso de codones al de la célula huésped en cuestión (según se describe arriba).

La enzima que tiene actividades de acetaldehído deshidrogenasa y, opcionalmente, de alcohol deshidrogenasa se expresa preferiblemente en forma activa en la célula huésped transformada. Por lo tanto, la expresión de la secuencia de nucleótidos en la célula huésped produce una acetaldehído deshidrogenasa con una actividad específica de al menos 0,005, 0,010, 0,020, 0,050 ó 0,10 µmol min^{-1} (mg de proteína)$^{-1}$, determinada como la tasa de reducción de NADH, dependiente de acetil-CoA, en los extractos celulares de la célula huésped transformada a 30°C según se describe en los Ejemplos en esta memoria.

En una realización adicional, la célula huésped de la invención comprende, además, una modificación genética que aumenta al menos uno de: i) la actividad específica de glicerol deshidrogenasa; ii) la actividad específica de dihidroxiacetona quinasa; y iii) el transporte de glicerol a la célula.

Preferiblemente, la modificación genética que aumenta la actividad específica de al menos una de glicerol deshidrogenasa y de dihidroxiacetona quinasa es la sobre-expresión de una secuencia de nucleótidos que codifica al menos una de una glicerol deshidrogenasa y una dihidroxiacetona quinasa. Sin embargo, alternativamente, la actividad específica de la glicerol deshidrogenasa y/o dihidroxiacetona quinasa se puede aumentar mediante la expresión de una enzima que tiene actividad incrementada en comparación con la enzima de tipo salvaje endógena de la célula huésped, además de, o como un reemplazo de la enzima de tipo salvaje.

Una glicerol deshidrogenasa se entiende en esta memoria como una enzima que cataliza la reacción química (EC 1.1.1.6):

$$\text{glicerol} + NAD^+ \leftrightarrow \text{glicerona} + NADH + H^+$$

Otros nombres en uso común incluyen glicerina deshidrogenasa, glicerol deshidrogenasa enlazada a NAD^+ y glicerol:NAD^+ 2-oxidorreductasa. Preferiblemente, la modificación genética determina la sobre-expresión de una glicerol deshidrogenasa, p. ej., mediante la sobre-expresión de una secuencia de nucleótidos que codifica una glicerol deshidrogenasa. La secuencia de nucleótidos que codifica la glicerol deshidrogenasa pueden ser endógena a la célula o puede ser una glicerol deshidrogenasa que es heteróloga a la célula. Secuencias de nucleótidos que se pueden utilizar para la sobre-expresión de glicerol deshidrogenasa en las células de la invención son, p. ej., el gen glicerol deshidrogenasa de *S. cerevisiae* (*GCY1*) tal como, p. ej., se describe por Oechsner et al. (1988, FEBS Lett. 238: 123-128) o Voss et al. (1997, Yeast 13: 655-672). Preferiblemente, la secuencia de nucleótidos que codifica la glicerol deshidrogenasa comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% identidad de secuencia de aminoácidos con SEQ ID NO: 13, o una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 13. En una realización preferida, se sobre-expresa una secuencia de nucleótidos optimizada en codones (véase arriba) que codifica la glicerol deshidrogenasa tal como, p. ej., una secuencia de nucleótidos optimizada en codones que codifica la glicerol deshidrogenasa de SEQ ID NO: 13.

Sin embargo, la glicerol deshidrogenasa codificada por el gen *GCY1* de levadura parece ser específica para el cofactor $NADP^+$ (EC 1.1.1.72) en oposición a NAD^+ (EC 1.1.1.6). Las levaduras tales como *S. cerevisiae* parecen carecer de actividad de glicerol deshidrogenasa dependiente de NAD^+ (EC 1.1.1.6) (véase, p. ej., la ruta KEGG 00561). Secuencias de nucleótidos preferida para la sobreexpresión de una deshidrogenasa de glicerol heteróloga en las células de la invención son, por tanto, por ejemplo, secuencias que codifican deshidrogenasas glicerol bacterianas que utilizan NAD^+ como factor de Co (EC 1.1.1.6), tales como por ejemplo, el gen gldA de E. coli descrito por Truniger y Boos (1994, J Bacteriol 176 (6):. 1796- 1800)., la expresión de que en la levadura ya ha sido reportado (Lee y Dasilva, 2006, Metab Ing 8 (I): 58-65). Preferiblemente, la secuencia de nucleótidos que codifica una deshidrogenasa de glicerol heteróloga comprende una secuencia de nucleótidos que codifica una secuencia de 176(6):1796-1800), cuya expresión en levaduras ya ha sido reseñada (Lee y Dasilva, 2006, Metab Eng. 8(1):58-65). Preferiblemente, la secuencia de nucleótidos que codifica una glicerol deshidrogenasa heteróloga comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% de identidad de secuencia de aminoácidos con SEQ ID NO: 49 o una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 49. En una realización preferida, se sobre-expresa una secuencia de nucleótidos optimizada en codones (véase arriba) que codifica la glicerol deshidrogenasa heteróloga tal como, p. ej., una secuencia de nucleótidos optimizada en codones que codifica la secuencia de aminoácidos de la glicerol deshidrogenasa de SEQ ID NO: 49. Una de estas secuencias de nucleótidos optimizadas en codones se proporciona, p. ej., en SEQ ID NO: 50 (posiciones 10-1113; CAI = 0,976).

Para la sobre-expresión de la secuencia de nucleótidos que codifica la glicerol deshidrogenasa, la secuencia de nucleótidos (a ser sobre-expresada) se puede colocar en una construcción de expresión en la que está enlazada

operativamente a regiones/secuencias reguladoras de la expresión adecuadas para asegurar la sobre-expresión de la enzima glicerol deshidrogenasa después de la transformación de la construcción de expresión en la célula huésped de la invención (véase arriba). Promotores adecuados para la (sobre)expresión de la secuencia de nucleótidos que codifica la enzima que tiene actividad de glicerol deshidrogenasa incluyen promotores que son preferiblemente insensibles a la represión de catabolitos (glucosa),que son activos en condiciones anaerobias y/o que preferiblemente no requieren xilosa o arabinosa para la inducción. Ejemplos de tales promotores se dan arriba. En las células de la invención, una glicerol deshidrogenasa a ser sobre-expresada se sobre-expresa preferiblemente en al menos un factor de 1,1, 1,2, 1,5, 2, 5, 10 ó 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que determina la sobre-expresión. Preferiblemente, la glicerol deshidrogenasa se sobre-expresa en condiciones anaerobias en al menos un factor de 1,1, 1,2, 1,5, 2, 5, 10 ó 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que determina la sobre-expresión. Ha de entenderse que estos niveles de sobre-expresión pueden aplicarse al nivel de estado estacionario de la actividad de la enzima (actividad específica en la célula), al nivel de estado estacionario de la proteína de la enzima, así como al nivel de estado estacionario de la codificación del transcrito para la enzima en la célula. La sobre-expresión de la secuencia de nucleótidos en la célula huésped produce una actividad de glicerol deshidrogenasa específica de al menos 0,2, 0,5, 1,0, 2,0 ó 5,0 U min^{-1} (mg de proteína)$^{-1}$, determinado en extractos celulares de las células huésped la transformadas a 30°C según se describe en los Ejemplos de esta memoria.

Una dihidroxiacetona quinasa se entiende en esta memoria como una enzima que cataliza la reacción química ((EC 2.7.1.29):

ATP + glicerona ↔ ADP + fosfato de glicerona

Otros nombres en uso común incluyen glicerona quinasa, ATP:glicerona fosfotransferasa y acetol quinasa (fosforilante). Se entiende que glicerona y dihidroxiacetona son la misma molécula. Preferiblemente, la modificación genética determina la sobre-expresión de una dihidroxiacetona quinasa, p. ej., mediante la sobre-expresión de una secuencia de nucleótidos que codifica una dihidroxiacetona quinasa. La secuencia de nucleótidos que codifica la dihidroxiacetona quinasa puede ser endógena a la célula o puede ser una dihidroxiacetona quinasa que es heteróloga a la célula. Las secuencias de nucleótidos que se pueden utilizar para la sobre-expresión de dihidroxiacetona quinasa en las células de la invención son, p. ej. los genes dihidroxiacetona quinasa de *S. cerevisiae* (*DAK1*) y (*DAK2*) tal como, p. ej., se describe por Molin et al. (2003, J. Biol. Chem. 278:1415-1423). Preferiblemente, la secuencia de nucleótidos que codifica la dihidroxiacetona quinasa comprende una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% identidad de secuencia de aminoácidos con al menos una de SEQ ID NOs: 14 y 15, o una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con al menos una de SEQ ID NOs: 14 y 15. En una realización preferida, se sobre-expresa una secuencia de nucleótidos optimizada en codones (véase arriba) que codifica la dihidroxiacetona quinasa tal como, p, ej,, una secuencia de nucleótidos optimizada en codones que codifica la dihidroxiacetona quinasa de SEQ ID NO: 14 o una secuencia de nucleótidos optimizada en codones que codifica la dihidroxiacetona quinasa de SEQ ID NO: 15. Una secuencia de nucleótidos preferida para la sobre-expresión de un dihidroxiacetona quinasa es una secuencia de nucleótidos que codifica una dihidroxiacetona quinasa que comprende una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% de identidad de secuencia de aminoácidos con SEQ ID NO: 14 (*S. cerevisiae* (*DAK1*), que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 14.

Secuencias de nucleótidos que se pueden utilizar para la sobre-expresión de una dihidroxiacetona quinasa heterólogo en las células de la invención son, p. ej., secuencias que codifican dihidroxiacetona quinasas bacterianas tales como el gen *dhaK* de *Citrobacter freundii*, p. ej., descrito por Daniel et al. (1995, J. Bacteriol. 177:4392-4401). Preferiblemente, la secuencia de nucleótidos que codifica una dihidroxiacetona quinasa heteróloga comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% de identidad de secuencia de aminoácidos con SEQ ID NO: 52 o una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 52. En una realización preferida se sobre-expresa una secuencia de nucleótidos optimizada en codones (véase arriba) que codifica la dihidroxiacetona quinasa heteróloga tal como, p. ej., una secuencia de nucleótidos optimizada en codones que codifica la secuencia de aminoácidos de la dihidroxiacetona quinasa de SEQ ID NO: 52. Una secuencia de nucleótidos optimizada en codones de este tipo se proporciona, p. ej., en SEQ ID NO: 53 (posiciones 10 - 1668).

Para la sobreexpresión de la secuencia de nucleótidos que codifica la dihidroxiacetona quinasa, la secuencia de nucleótidos (a ser sobre-expresada) se puede colocar en una construcción de expresión en la que está operativamente enlazada a regiones/secuencias reguladoras de la expresión adecuadas para asegurar la sobre-expresión de la enzima dihidroxiacetona quinasa después de la transformación de la construcción de expresión en la célula huésped de la invención (véase arriba). Promotores adecuados para la (sobre)expresión de la secuencia de nucleótidos que codifica la enzima que tiene actividad de dihidroxiacetona quinasa incluyen promotores que son preferiblemente insensibles a la represión del catabolito (glucosa), que son activos en condiciones anaerobias y/o

que preferiblemente no requieren xilosa o arabinosa para la inducción. Ejemplos de tales promotores se dan arriba. En las células de la invención, un dihidroxiacetona quinasa a ser sobre-expresada se sobre-expresa preferiblemente en al menos un factor de 1,1, 1,2, 1,5, 2, 5, 10 ó 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que determina la sobreexpresión. Preferiblemente, la dihidroxiacetona quinasa se sobre-expresa en condiciones anaerobias por al menos un factor de 1,1, 1,2, 1,5, 2, 5, 10 ó 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que determina la sobre-expresión. Ha de entenderse que estos niveles de sobre-expresión pueden aplicarse al nivel de estado estacionario de la actividad de la enzima (actividad específica en la célula), al nivel de estado estacionario de la proteína de la enzima, así como al nivel de estado estacionario de la codificación del transcrito que codifica la enzima en la célula. La sobre-expresión de la secuencia de nucleótidos en la célula huésped produce una actividad de dihidroxiacetona quinasa específica de al menos 0,002, 0,005, 0,01, 0,02 ó 0,05 U min^{-1} (mg de proteína)$^{-1}$, determinado en extractos celulares de las células huéspedes transformadas a 30°C según se describe en los Ejemplos en esta memoria.

Preferiblemente, la modificación genética que aumenta el transporte de glicerol a la célula es preferiblemente una modificación genética que determina la sobre-expresión de una secuencia de nucleótidos que codifica al menos uno de una proteína de absorción de glicerol y un canal de glicerol.

Una proteína de absorción de glicerol se entiende en esta memoria como una proteína con múltiples segmentos transmembranales que pertenece a la incluida en la superfamilia de las O-aciltransferasas ligadas a la membrana (MBOAT) que incluyen, p. ej., las proteínas de absorción de glicerol de *S. cerevisiae* codificadas por los genes *GUP1* y *GUP2*. Preferiblemente, la modificación genética determina la sobre-expresión de una proteína de absorción de glicerol, p. ej., mediante la sobre-expresión de una secuencia de nucleótidos que codifica una proteína de absorción de glicerol. La secuencia de nucleótidos que codifica la proteína de absorción de glicerol puede ser endógena a la célula o puede ser una proteína de absorción de glicerol que es heteróloga a la célula. Secuencias de nucleótidos que se pueden utilizar para la sobre-expresión de la proteína de absorción de glicerol en las células de la invención son, p. ej., los genes de proteínas de absorción de glicerol de *S. cerevisiae* (*GUP1*) y (GUP2) y ortólogos de los mismos según se describe, p. ej., por Neves et al. (2004, FEMS Yeast Res. 5:51-62). Preferiblemente, la secuencia de nucleótidos que codifica la proteína de absorción de glicerol comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 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 NOs: 16 (*Gup1p*) y 17 (*Gup2p*) o una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con al menos una de SEQ ID NOs : 16 y 17. En una realización preferida, se sobre-expresa una secuencia de nucleótidos optimizada en codones (véase arriba) que codifica la proteína de absorción de glicerol tal como, p. ej., una secuencia de nucleótidos optimizada en codones que codifica la proteína de absorción de glicerol de SEQ ID NO: 16 o una secuencia de nucleótidos optimizada en codones que codifica la proteína de absorción de glicerol de SEQ ID NO: 17. Aunque la naturaleza exacta de la influencia de GUP1 sobre el transporte de glicerol aún no está clara, Yu et al. (2010, *supra*) han demostrado que la sobre-expresión de GUP1 en *S. cerevisiae* mejora la producción de etanol en células de crecimiento en glicerol. Por lo tanto, una secuencia de nucleótidos preferida para la sobre-expresión de una proteína de absorción de glicerol es una secuencia de nucleótidos que codifica una proteína de absorción de glicerol que es capaz de rescatar el fenotipo, asociado al estrés de una sal, de un mutante *gup1Δ* de *S. cerevisiae* mediante complementación según se describe por Neves et al. (2004, *supra*). Tales ortólogos de complementación de *GUP1* de *S. cerevisiae* incluyen secuencias de nucleótidos que codifican secuencias de aminoácidos que tienen al menos 60, 68, 72, 75, 80, 85, 90, 95, 98, 99% de identidad con la secuencia de aminoácidos de SEQ ID NO: 16 y pueden obtenerse a partir de especies de levaduras que pertenecen a los géneros de *Saccharomyces*, *Zygosaccharomyces*, *Kluyveromyces*, *Candida*, *Pichia*, *Hansenula*, *Kloeckera*, *Schwanniomyces* y *Yarrowia*.

Un canal de glicerol en esta memoria se entiende como un miembro de la familia MIP de proteínas de los canales revisado por Reizer et al. (1993, CRC Crit. Rev. Biochem. Mol. Biol., 28: 235-257), proteínas de los canales que comprender un dominio transmembrana de 250 -280 aminoácidos que consiste en seis dominios con segmentos membranales y que tienen al menos 30, 35, 40, 45, 50, 60, 70, 80, 90, 95, 98 o 99% de identidad de aminoácidos, o al menos 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: 18, el *FPS1* de acuagliceroporina de *S. cerevisiae.* Alternativamente, la proteína de los canales comprende un dominio transmembranal de 250 -280 aminoácidos que consiste en seis dominios que abarcan la membrana, y que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con la secuencia de aminoácidos entre los aminoácidos 250 y 530 de SEQ ID NO: 18.

Las secuencias de nucleótidos que se pueden utilizar para la sobreexpresión de un canal de glicerol en las células de la invención incluyen secuencias de nucleótidos que codifican el gen *FPS1* de acuagliceroporina de levadura, p. ej. *S. cerevisiae* (Van Aelst et al, 1991, EMBO J. 10: 2095-2104) y ortólogos del mismo de otras levaduras, incluyendo *Kluyveromyces lactis*, *Kluyveromyces marxianus* y *Zygosaccharomyces rouxii* tal como se describe, p. ej., por Neves et al. (2004, *supra*). Sin embargo, no se excluye el uso de canales de glicerol bacterianos o vegetales, ya que, p. ej., Luyten et al. (1995, EMBO J. 14: 1360-1371) han demostrado que el facilitador de glicerol de *E. coli*, que tiene solamente un 30% de identidad de secuencia con la secuencia de aminoácidos entre los aminoácidos 250 y

530 de la *FPS1* de acuagliceroporina de S. cerevisiae, puede complementar la absorción 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, se sobre-expresa una secuencia de nucleótidos optimizada en codones (véase arriba) que codifica el canal de glicerol tal como, p. ej., una secuencia de nucleótidos optimizada en codones que codifica la acuagliceroporina de SEQ ID NO: 18.

Para la sobre-expresión de la secuencia de nucleótidos que codifica la proteína de absorción de glicerol y/o la proteína del canal de glicerol, la secuencia de nucleótidos (a ser sobre-expresada) se puede colocar en una construcción de expresión en la que está enlazada operativamente a regiones/secuencias reguladoras de la expresión adecuadas para asegurar la sobre-expresión de la proteína de absorción de glicerol y/o la proteína del canal de glicerol después de la transformación de la construcción de expresión en la célula huésped de la invención (véase arriba). Promotores adecuados para la (sobre)expresión de la secuencia de nucleótidos que codifica la proteína de absorción de glicerol y/o la proteína del canal de glicerol incluyen promotores que son preferiblemente insensible a la represión de catabolitos (glucosa), que son activos en condiciones anaerobias y/o que preferiblemente no requieren xilosa o arabinosa para la inducción. Ejemplos de promotores de este tipo se dan arriba. En las células de la invención, una proteína de absorción de glicerol y/o una proteína del canal de glicerol a ser sobre-expresada se sobre-expresa preferiblemente en al menos un factor de 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 determina la sobre-expresión. Preferiblemente, la proteína de absorción de glicerol y/o la proteína del canal de glicerol se sobre-expresa en condiciones anaerobias en al menos un factor de 1,1, 1,2, 1,5, 2, 5, 10 ó 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que determina la sobre-expresión. Ha de entenderse que estos niveles de sobre-expresión pueden aplicarse al nivel de estado estacionario de la actividad de la enzima (actividad específica en la célula), al nivel de estado estacionario de la proteína de la enzima, así como al nivel de estado estacionario de la codificación del transcrito para la enzima en la célula.

En una realización adicional, la célula huésped de la invención comprende, además, una modificación genética que aumenta la actividad específica de acetil-CoA sintetasa en la célula, preferiblemente bajo 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 esta memoria como una enzima que cataliza la formación de un nuevo enlace químico entre acetato y coenzima A (CoA). Preferiblemente, la modificación genética determina la sobre-expresión de una acetil-CoA sintetasa, p. ej., mediante la sobre-expresión 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 se pueden utilizar para la sobre-expresión de acetil-CoA sintetasa en las células de la invención son, p. ej., los genes de acetil CoA sintetasa de S. cerevisiae (*ACS1* y *ACS2*) tal como se describe, p. ej., por de Jong-Gubbels et al. (1998, FEMS Microbiol Lett. 165: 15-20). Preferiblemente, la secuencia de nucleótidos que codifica la acetil-CoA sintetasa comprende una secuencia de aminoácidos con al menos 45, 50, 60, 65, 70, 75, 80, 85, 90, 95, 98, 99% de identidad de secuencia de aminoácidos con al menos una de las SEQ ID NOs: 19 y 20, o una secuencia de nucleótidos que codifica una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con al menos una de SEQ ID NOs: 19 y 20.

En una realización, la secuencia de nucleótidos que se sobre-expresa codifica una acetil CoA sintetasa con una alta afinidad para acetato. Se prefiere el uso de una acetil-CoA sintetasa con una alta afinidad para acetato de condiciones en las que hay una concentración relativamente baja de ácido acético en el medio de cultivo, p. ej., no más de 2 g de ácido acético/L de medio de cultivo. Una acetil-CoA sintetasa con una alta afinidad para acetato se define en esta memoria como un acetil-CoA sintetasa con una afinidad más alta para acetato que la acetil-CoA sintetasa codificada por *S. cerevisiae ACS2* (SEQ ID NO: 20). Preferiblemente, una acetil-CoA sintetasa con una alta afinidad para acetato tiene una Km para acetato de no más de 10, 5, 2, 1, 0,5, 0,2 ó 0,1 mM tal como, p. ej., la acetil-CoA sintetasa codificada por el gen *ACS1* de *S. cerevisiae*. Más preferiblemente, se sobre-expresa una secuencia de nucleótidos optimizada en codones (véase arriba) que codifica la secuencia de aminoácidos de SEQ ID NO: 19.

En otra realización, la secuencia de nucleótidos que es sobre-expresada codifica una acetil-CoA sintetasa con una tasa máxima ($V_{máx}$). Se prefiere el uso de una acetil-CoA sintetasa con una alta tasa máxima para la condición bajo la cual hay una concentración relativamente alta de ácido acético en el medio de cultivo, p. ej., al menos 2, 3, 4 ó 5 g de ácido acético/L de medio de cultivo. Una acetil-CoA sintetasa con una alta tasa máxima se define en esta memoria como una acetil-CoA sintetasa con una tasa máxima más alta que la acetil-CoA sintetasa codificada por *ACS1* de *S. cerevisiae*. Preferiblemente, la acetil-CoA sintetasa con una alta tasa máxima es la acetil-CoA sintetasa codificada por el gen *ACS2* de S. cerevisiae. Más preferiblemente, se sobre-expresa una secuencia de nucleótidos optimizada en codones (véase arriba) que codifica la secuencia de aminoácidos de SEQ ID NO: 20.

Para la sobre-expresión de la secuencia de nucleótidos que codifica la acetil-CoA sintetasa (a ser sobre-expresada) se puede colocar en una construcción de expresión en la que está enlazada operativamente a regiones/secuencias reguladoras de la expresión adecuadas para asegurar la sobre-expresión de la enzima acetil-CoA sintetasa después de la transformación de la construcción de expresión en la célula huésped de la invención (véase arriba).

Promotores adecuados para la (sobre)expresión de la secuencia de nucleótidos que codifica la enzima que tiene actividad de acetil CoA sintetasa incluyen promotores que son preferiblemente insensibles a la represión de catabolitos (glucosa),que son activos en condiciones anaerobias y/o que preferiblemente no requieren xilosa o arabinosa para la inducción. Ejemplos de tales promotores se dan arriba. En las células de la invención, una acetil CoA sintetasa a ser sobre-expresada se sobre-expresa en al menos un factor de 1,1, 1,2, 1,5, 2, 5, 10 ó 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que determina la sobre-expresión. Preferiblemente, la acetil CoA sintetasa se sobre-expresa en condiciones anaerobias en al menos un factor de 2, 5, 10, 20, 50 ó 100, en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que determina la sobre-expresión. Ha de entenderse que estos niveles de sobre-expresión pueden aplicarse al nivel de estado estacionario de la actividad de la enzima (actividad específica), al nivel de estado estacionario de la proteína de la enzima, así como al nivel de estado estacionario de la codificación del transcrito para la enzima.

En una realización adicional, la célula huésped 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 la dihidroxiacetona fosfato en sn-glicerol 3-fosfato, mientras que NADH se oxida en NAD^+. En las células de la invención, la actividad de glicerolfosfato deshidrogenasa específica se reduce preferiblemente en al menos un factor de 0,8, 0,5, 0,3, 0,1, 0,05 ó 0,01 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que determina la sobre-expresión, preferiblemente en condiciones anaerobias.

Preferiblemente, la actividad de glicerolfosfato deshidrogenasa se reduce en la célula huésped en una o más modificaciones genéticas que reducen la expresión de o inactivan un gen que codifica una glicerolfosfato deshidrogenasa. Preferiblemente, las modificaciones genéticas reducen o inactivan la expresión de cada una de las copias endógenas 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 una y la misma secuencia de aminoácidos como resultado de diploidía, poliploidía o aneuploidía. En tales casos, preferiblemente la expresión de cada una de las copias 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 difieren en la secuencia de aminoácidos y que son cada una codificada por un gen diferente. En tales casos, en algunas realizaciones de la invención, se prefiere que sólo determinados tipos de las isoenzimas se reduzcan o inactiven mientras que otros tipos no se ven afectados (véase más adelante). Preferiblemente, el gen es inactivado por deleción de al menos una parte del gen o por la interrupción del gen, con lo que en este contexto el término gen también incluye cualquier secuencia no codificante aguas arriba o aguas abajo de la secuencia codificadora, cuya deleción o inactivación (parcial) resulta en una reducción de la expresión de la actividad glicerolfosfato deshidrogenasa en la célula huésped.

Un gen preferido que codifica una glicerolfosfato deshidrogenasa, cuya actividad ha de ser reducida o inactivada en la célula de la invención es el gen *GPD2* de *S. cerevisiae* según se describe por Eriksson et al. (1995, Mol. Microbiol. 17: 95-107), que codifica la secuencia de aminoácidos de SEQ ID NO: 21 y ortólogos de la misma en otras especies. Por lo tanto, un gen que codifica una glicerolfosfato deshidrogenasa, cuya actividad ha de ser reducida o inactivada en la célula de la invención es preferiblemente un gen que codifica una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos 70, 75, 80, 85, 90, 95, 98 o 99% de identidad de secuencia con SEQ ID NO: 21, o una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 21.

En una realización preferida de la invención, la célula huésped de la invención comprende una ruta de respuesta de glicerol de alta osmolaridad funcional. Preferiblemente, por lo tanto, sólo se reduce o inactiva la actividad del o de los genes que codifican una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos 70% de identidad de secuencia con SEQ ID NO: 21, mientras que al menos es funcional un gen endógeno que codifica una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con la secuencia de al menos 70, 75, 80, 85, 90, 95, 98 ó 99% de identidad con SEQ ID NO: 22. SEQ ID NO: 22 representa la secuencia de aminoácidos codificada por el gen *GPD1* de S. cerevisiae según se describe por Albertyn et al. (1994, Mol. Cell. Biol. 14: 4135-4144), que tiene un 69% de identidad de aminoácidos con la glicerolfosfato deshidrogenasa *S. cerevisiae GPD2*. El gen *GPD1* de *S. cerevisiae* es la glicerolfosfato deshidrogenasa inducida por estrés de *S. cerevisiae*, que es importante para el crecimiento bajo estrés osmótico como puede ocurrir bajo condiciones de fermentaciones industriales. Su expresión es regulada, entre otros, por la ruta de respuesta de glicerol de alta osmolaridad. Por lo tanto, es ventajoso que una célula huésped 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 70, 75, 80, 85, 90, 95, 98 ó 99% de identidad de secuencia con SEQ ID NO: 22 o una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con SEQ ID NO: 22.

En una realización adicional, se reduce o inactiva la actividad de todos los genes en la célula huésped que codifica una glicerolfosfato deshidrogenasa. En tales células, preferiblemente se inactivan o al menos se reducen en la

expresión todas las copias de genes endógenos que codifican una glicerolfosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos 70, 75, 80, 85, 90, 95, 98 ó 99% de identidad de secuencia con SEQ ID NO: 21 ó 22 (o que tienen una secuencia de aminoácidos que tiene una o varias sustituciones, inserciones y/o deleciones en comparación con al menos una de SEQ ID NO de: 21 y 22).

En una realización preferida adicional, la célula huésped de la invención tiene al menos uno 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 tiene preferiblemente 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 preferiblemente genes exógenos funcionales que codifican una L-arabinosa isomerasa, una L-ribuloquinasa y una L-ribulosa-5-fosfato 4-epimerasa, genes que juntos confieren a la célula la capacidad de convertir por isomerización L-arabinosa en D-xilulosa 5-fosfato.

Células huésped fúngicas que tienen la capacidad de isomerizar xilosa en xilulosa tal como se describe, p. ej., en el documento WO 03/0624430 y en el documento WO 06/009434. La capacidad de isomerizar xilosa a xilulosa es conferida preferiblemente a la célula mediante transformación con una construcción de ácido nucleico que comprende una secuencia de nucleótidos que codifica una xilosa isomerasa. Por lo tanto, preferiblemente la célula adquiere la capacidad de isomerizar directamente xilosa en xilulosa. Por lo tanto, más preferiblemente, la célula adquiere la capacidad de crecer en condiciones aerobias y/o anaerobias en xilosa como única fuente de energía y/o carbono, aunque la isomerización directa de xilosa en xilulosa (y el posterior metabolismo de xilulosa). Es en esta memoria se entiende que la isomerización directa de xilosa en xilulosa se produce en una única reacción catalizada por una xilosa isomerasa, en oposición a la conversión en dos etapas de xilosa en xilulosa a través de un compuesto intermedio de xilitol catalizada por la xilosa reductasa y xilitol deshidrogenasa, respectivamente.

Varias xilosas isomerasas (y sus secuencias de aminoácidos y codificadoras de nucleótidos) que se pueden utilizar con éxito para conferir a la célula de la invención la capacidad de isomerizar directamente xilosa en xilulosa se han descrito en la técnica. 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 *Bacteroides* bacteriano, incluyendo, p. ej., *B. thetaiotaomicron* (documento WO 06/009434), *B. fragilis* y *B. uniformis* (documento WO 09/109633), la xilosa isomerasa del *Clostridium phytofermentans* de bacteria anaerobia (Brat et al., 2009, Appl. Environ. Microbiol. 75: 2304-2311), y las xilosa isomerasas de *Clostridium difficile*, *Ciona intestinales* y *Fusobacterium mortiferum* (documento WO 10/074577).

Células huéspedes fúngicas que tienen la capacidad de convertir L-arabinosa en D-xilulosa 5- fosfato tal como se describe, p. ej., en Wisselink et al. (2007, AEM Acepts, publicado en línea previo a la impresión 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 preferiblemente a la célula mediante transformación con uno o más constructos de ácido nucleico que comprenden secuencias de nucleótidos que codifican a) una arabinosa isomerasa; b) una ribuloquinasa, preferiblemente una L-ribuloquinasa una xilosa isomerasa; y c) una ribulosa-5-P-4 epimerasa, preferiblemente una L-ribulosa-5-P-4-epimerasa. Preferiblemente, 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 a través de las reacciones subsiguientes de 1) isomerización de arabinosa en ribulosa ; 2) fosforilación de ribulosa a ribulosa 5-fosfato; y, 3) epimerización de ribulosa 5-fosfato en D-xilulosa 5-fosfato. Secuencias de nucleótidos adecuadas que codifican arabinosa isomerasas, un ribuloquinasa y ribulosa-5-P-4-epimerasas pueden obtenerse de *Bacillus subtilis*, *Escherichia coli* (véase, p. ej. el documento EP 1 499 708), *Lactobacilli*, p. ej., *Lactobacillus plantarum* (véase, p. ej. Wisselink et al. *supra*), o especies de *Clavibacter*, *Arthrobacter* y *Gramella*, de los cuales preferiblemente *Clavibacter michiganensis*, *Arthrobacter aurescens* y *Gramella forsetii* (véase el documento WO2009/011591).

La célula huésped transformada de la invención comprende, además, preferiblemente la actividad de xilulosa quinasa, de modo que xilulosa isomerizada a partir de xilosa puede ser metabolizada a piruvato. Preferiblemente, la célula contiene actividad de xilulosa quinasa endógena. Más preferiblemente, una célula de la invención comprende una modificación genética que aumenta la actividad específica xilulosa quinasa. Preferiblemente, la modificación genética determina la sobre-expresión de una xilulosa quinasa, p. ej., mediante la sobre-expresión de una secuencia de nucleótidos que codifica una xilulosa quinasa. El gen que codifica la xilulosa quinasa puede ser endógeno a la célula o puede ser una xilulosa quinasa que es heteróloga a la célula. Una secuencia de nucleótidos que se puede utilizar para la sobre-expresión de xilulosa quinasa en las células de la invención es, p. ej., el gen de la xilulosa quinasa de *S. cerevisiae* (*XKS1*) según se describe por Deng y Ho (1990, Appl. Biochem. Biotechnol. 24-25: 193-199). Otra xilulosa quinasa preferida es una xilulosa quinasa que se relaciona con la xilulosa quinasa de *Piromyces* (*xylB*; véase el documento WO 03/0624430). Este xilulosa quinasa de *Piromyces* está en realidad más relacionada con quinasa procariota que con todas las quinasas eucarióticas conocidas tales como la levadura quinasa. Las xilulosa quinasas eucarióticas se han indicado como azúcar quinasas no específicas, que tienen una gama de sustratos amplia que incluye xilulosa. En contraposición, se ha indicado que las xilulosa quinasas procarióticas, con

las que la quinasa de *Piromyces* está más estrechamente relacionada, son quinasas más específicas para xilulosa, es decir, que tienen un intervalo de sustrato más estrecho. En las células de la invención, una xilulosa quinasa a ser sobre-expresada se sobre-expresa en al menos un factor de 1,1, 1,2, 1,5, 2, 5, 10 ó 20 en comparación con una cepa que es genéticamente idéntica, excepto por la modificación genética que determina la sobreexpresión. Ha de entenderse que estos niveles de sobre-expresión pueden aplicarse al nivel de estado estacionario de la actividad de la enzima, al nivel de estado estacionario de la proteína de la enzima, así como al nivel de estado estacionario de la codificación del transcrito que codifica la enzima.

Una célula de la invención comprende, además, preferiblemente una modificación genética que aumenta el flujo de la vía de la pentosa fosfato según se describe en el documento WO 06/009434. En particular, la modificación genética determina un aumento del flujo de la parte no oxidativa de la pentosa fosfato. Una modificación genética que determina un flujo incrementado de la parte no oxidativa de la vía de la pentosa fosfato se entiende en esta memoria que significa una modificación que aumenta el flujo en al menos un factor de 1,1, 1,2, 1,5, 2, 5, 10 ó 20 en comparación con el flujo en una cepa que es genéticamente idéntica, excepto por la modificación genética que determina el flujo incrementado. El flujo de la parte no oxidativa de la vía de las pentosa fosfato se puede medir como se describe en el documento WO 06/009434.

Modificaciones genéticas que aumentan el flujo de la vía de la pentosa fosfato se pueden introducir en las células de la invención de diversas maneras. Éstas incluyen, p. ej., conseguir mayores niveles de actividad de estado estacionario de xilulosa quinasa y/o una o más de las enzimas de la parte no oxidativa de la vía de pentosa fosfato y/o un nivel de estado estacionario reducido de actividad inespecífica de la aldosa reductasa. Estos cambios en los niveles de actividad de estado estacionario se puede efectuar mediante la selección de mutantes (espontáneos o inducidos por sustancias químicas o radiación) y/o por tecnología de ADN recombinante, p. ej., mediante la sobre-expresión o la 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 sobre-expresión de al menos una enzima de la (parte no oxidativa) de la vía de pentosa fosfato. Preferiblemente, la enzima se selecciona del grupo que consiste en las enzimas que codifican ribulosa-5-fosfato isomerasa, ribulosa-5-fosfato 3-epimerasa, transcetolasa y transaldolasa. Se pueden sobre-expresar diversas combinaciones de enzimas de la (parte no oxidativa) de la vía de pentosa fosfato. P. ej., las enzimas que se sobre-expresan 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, 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 sobre-expresa en la célula de la invención. Se prefiere una célula en la que la modificación genética comprende al menos la sobre-expresión de la enzima transaldolasa. Más preferida es una célula en la que la modificación genética comprende al menos la sobre-expresión tanto de las enzimas transcetolasa como transaldolasa, ya que una célula huésped de este tipo ya es capaz de crecimiento anaerobio en xilosa. De hecho, bajo algunas condiciones, los autores de la invención han encontrado que las células que sobre-expresan únicamente la transcetolasa y la transaldolasa ya tienen la misma tasa de crecimiento anaerobio en xilosa que las células que sobre-expresan las cuatro de las enzimas, es decir, la ribulosa-5-fosfato isomerasa, ribulosa-5 fosfato 3-epimerasa, transcetolasa y transaldolasa. Además, se prefieren células de la invención que sobre-expresan tanto las enzimas ribulosa-5-fosfato isomerasa como ribulosa-5-fosfato 3-epimerasa frente a células que sobre-expresan únicamente la isomerasa o sólo la 3-epimerasa, dado que la sobre-expresión de sólo una de estas enzimas puede producir desequilibrios metabólicos.

Existen diversos medios disponibles en la técnica para la sobre-expresión de enzimas en las células huéspedes de la invención. En particular, una enzima puede ser sobre-expresada aumentando el número de copias del gen que codifica la enzima en la célula, p. ej., mediante la integración de copias adicionales del gen en el genoma de la célula, mediante la expresión del gen a partir de un vector de expresión multicopia episomal o mediante la introducción de un vector de expresión episomal que comprende múltiples copias del gen. La secuencia codificadora utilizada para la sobre-expresión de las enzimas es preferiblemente homóloga a la célula huésped de la invención. Sin embargo, igualmente se pueden emplear secuencias codificadoras que son heterólogas a la célula huésped de la invención. Alternativamente, la sobre-expresión de enzimas en las células de la invención se puede lograr mediante el uso de un promotor que no es nativo a la secuencia codificadora para la enzima a ser sobre-expresada, es decir, un promotor que es heterólogo a la secuencia codificadora a la que está enlazado operativamente. Aunque el promotor es preferiblemente heterólogo a la secuencia codificadora a la que está enlazado operativamente, también se prefiere que el promotor sea homólogo, es decir, endógeno a la célula de la invención. Preferiblemente, el promotor heterólogo es capaz de producir un nivel de estado estacionario más alto del transcrito que comprende

la secuencia codificadora (o es capaz de producir más moléculas de transcrito, es decir, moléculas de ARNm, por unidad de tiempo) que el promotor que es nativo de la secuencia codificadora, preferiblemente bajo condiciones en las que una o más de xilosa, arabinosa y glucosa están disponibles como fuentes de carbono, más preferiblemente como fuentes de carbono principales (es decir, más de 50% de la fuente de carbono disponible consiste en una o más de xilosa, arabinosa y glucosa), lo más preferiblemente como únicas fuentes de carbono.

Una célula preferida adicional de la invención comprende una modificación genética que reduce la actividad inespecífica de aldosa reductasa en la célula. Preferiblemente, la actividad inespecífica de la aldosa reductasa se reduce en la célula huésped mediante una o más modificaciones genéticas que reducen la expresión de o neutralizan un gen que codifica una aldosa reductasa inespecífica. Preferiblemente, las modificaciones genéticas reducen o inactivan la expresión de cada una de las copias endógenas de un gen que codifica una aldosa reductasa inespecífica que es capaz de reducir una aldopentosa, incluyendo xilosa, xilulosa y arabinosa, en el genoma de la célula. Una célula dada puede comprender múltiples copias de genes que codifican aldosa reductasas inespecíficas como resultado de diploidía, poliploidía o aneuploidía, y/o una célula puede contener varias (iso)enzimas diferentes con actividad de aldosa reductasa que difieren en la secuencia de aminoácidos y que son cada una codificada por un gen diferente. También en tales casos, preferiblemente se reduce o inactiva la expresión de cada uno de los genes que codifica una aldosa reductasa inespecífica. Preferiblemente, el gen se inactiva por deleción de al menos una parte del gen o por la interrupción del gen, con lo que en este contexto, el término gen también incluye cualquier secuencia no codificadora aguas arriba o aguas abajo de la secuencia codificadora, cuya deleción o inactivación (parcial) resulta en una reducción de la expresión de la actividad inespecífica de la aldosa reductasa en la célula huésped. Una secuencia de nucleótidos que codifica una aldosa reductasa cuya actividad se ha de reducir en la célula de la invención y secuencias de aminoácidos de dichas aldosa reductasas se describen en el documento WO 06/009434 e incluyen, p. ej., los genes (inespecíficos) de la aldosa reductasa del gen *GRE3* de *S. cerevisiae* (Traff et al, 2001, Appl. Environm. Microbiol. 67: 5668-5674) y ortólogos de los mismos en otras especies.

Una célula huésped transformada aún más preferida de acuerdo con la invención puede comprender modificaciones genéticas adicionales que resultan en una o más de las características seleccionadas del grupo que consiste en (a) el transporte incrementado de xilosa y/o arabinosa en la célula; (b) una sensibilidad disminuida a la represión catabólica; (c) una tolerancia incrementada al etanol, osmolaridad o ácidos orgánicos; y (d) una producción reducida de sub-productos. Sub-productos se entienden que significan moléculas con contenido en carbono que no sean el producto de fermentación deseado, e incluyen, p. ej., xilitol, arabinitol, glicerol y/o ácido acético. Cualquier modificación genética descrita en esta memoria puede introducirse mediante mutagénesis clásica y rastreo y/o selección para el mutante deseado, o simplemente por rastreo y/o selección de los mutantes espontáneos con las características deseadas. Alternativamente, las modificaciones genéticas pueden consistir en la sobre-expresión de genes endógenos y/o la inactivación de genes endógenos. Genes cuya sobre-expresión se desea para el transporte incrementado de arabinosa y/o xilosa en la célula se eligen preferiblemente de genes que codifican un transportador de hexosa o pentosa. En *S. cerevisiae* y otras levaduras estos genes incluyen *HXT1*, *HXT2*, *HXT3*, *HXT4*, *HXT5*, *HXT7* y *GAL2*, de los cuales los más preferidos son *HXT7*, *HXT5* y *GAL2* (véase Sedlack y Ho, Yeast 2004; 21: 671-684). Otro transportador preferido para la expresión en levaduras es el transportador de glucosa codificado por el gen *SUT1* de *P. stipitis* (Katahira et al, 2008, Enzyme Microb. Technol. 43: 115-119). De manera similar, se pueden sobre-expresar ortólogos de estos genes transportadores en otras especies. Otros genes que pueden ser sobre-expresados en las células de la invención incluyen genes que codifican enzimas glucolíticas y/o enzimas etanologénicas tales como alcohol deshidrogenasas. Genes endógenos preferidos para la inactivación incluyen genes de hexosa quinasa, p. ej., el gen *HXK2* de *S. cerevisiae* (véase Diderich et al, 2001, Appl. Environ. Microbiol. 67: 1587-93); los genes *MIG1* o *MIG2* de *S. cerevisiae*; genes que codifican enzimas implicadas en el metabolismo de glicerol tales como los genes de *S. cerevisiae* glicerol fosfato deshidrogenasa 1 y/o 2, u ortólogos (hibridantes) de estos genes en otras especies. Otras modificaciones adicionales preferidos de células huéspedes para la fermentación de xilosa se describen en van Maris et al. (2006, Antonie van Leeuwenhoek 90:391-418), documentos WO2006/009434, WO2005/023998, WO2005/111214 y WO2005/091733. Cualquiera de las modificaciones genéticas de las células de la invención según se describe en esta memoria, en la medida de lo posible, son introducidas o modificadas preferiblemente por modificación genética de auto-clonación.

Una célula huésped preferida de acuerdo con la invención tiene la capacidad de crecer en al menos una de xilosa y arabinosa como fuente de carbono/energía, preferiblemente como única fuente de carbono/energía, y preferiblemente bajo condiciones anaerobias, es decir, condiciones según se definen en esta memoria más adelante para el proceso de fermentación anaerobia. Preferiblemente, cuando se cultivan en xilosa como fuente de carbono/energía, la célula huésped no produce esencialmente xilitol, p. ej., el xilitol producido está por debajo del límite de detección o, p. ej., menos de 5, 2, 1, 0,5 ó 0,3% del carbono consumido sobre una base molar. Preferiblemente, cuando se cultivan en arabinosa como fuente de carbono/energía, la célula no produce esencialmente arabinitol, p. ej., el arabinitol producido está por debajo del límite de detección o, p. ej., menos de 5, 2, 1, 0,5 ó 0,3% del carbono consumido sobre una base molar.

Una célula huésped preferida de la invención tiene la capacidad de crecer en al menos una de una hexosa, una pentosa, glicerol, ácido acético y combinaciones de los mismos a una tasa de al menos 0,01, 0,02, 0,05, 0,1, 0,2,

0,25 ó 0,3 h^{-1} en condiciones aerobias o, más preferiblemente, a una tasa de al menos 0,005, 0,01, 0,02, 0,05, 0,08, 0,1, 0,12, 0,15 ó 0,2 h^{-1} en condiciones anaerobias. Por lo tanto, preferiblemente la célula huésped tiene la capacidad de crecer en al menos una de xilosa y arabinosa como única fuente de carbono/energía a una tasa de al menos 0,01, 0,02, 0,05, 0,1, 0,2, 0,25 ó 0,3 h^{-1} en condiciones aerobias o, más preferiblemente, a una tasa de al menos 0,005, 0,01, 0,02, 0,05, 0,08, 0,1, 0,12, 0,15 ó 0,2 h^{-1} bajo condiciones anaerobias. Más preferiblemente, la célula huésped tiene la capacidad de crecer en una mezcla de una hexosa (p. ej., glucosa) y al menos una de xilosa y arabinosa (en una relación ponderal de 1: 1) como única fuente de carbono/energía a una tasa de al menos 0,01, 0,02, 0,05, 0,1, 0,2, 0,25 ó 0,3 h^{-1} en condiciones aerobias o, más preferiblemente, a una tasa de al menos 0,005, 0,01, 0,02, 0,05, 0,08, 0,1, 0,12 , 0.15 ó 0.2 h^{-1} bajo condiciones anaerobias. Más preferiblemente, la célula huésped tiene la capacidad de crecer en una mezcla de una hexosa (p. ej., glucosa), al menos una de xilosa y arabinosa y glicerol (en una relación ponderal de 1: 1: 1) como única fuente de carbono/energía a una tasa de al menos 0,01, 0,02, 0,05, 0,1, 0,2, 0,25 ó 0,3 h^{-1} en condiciones aerobias o, más preferiblemente, a una tasa de al menos 0,005, 0,01, 0,02, 0,05, 0,08 , 0,1, 0,12, 0,15 ó 0,2 h^{-1} bajo condiciones anaerobias.

En un aspecto, la invención se refiere al uso de una célula de levadura de acuerdo con la invención para la preparación de un producto de fermentación seleccionado del grupo que consiste en etanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, 1,3-propano-diol, butanoles y productos derivados de isoprenoides.

En otro aspecto, la invención se refiere a un procedimiento para producir un producto de fermentación seleccionado del grupo que consiste en etanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, 1,3-propano-diol, butanoles (1-butanol, 2 butanol, isobutanol) y productos derivados de isoprenoides. El procedimiento puede ser un procedimiento en el que se produce formiato además del producto de fermentación. El procedimiento comprende preferiblemente la etapa de: a) fermentar un medio con una célula de levadura, con lo que el medio contiene o es alimentado con: a) una fuente de al menos una de una hexosa, una pentosa y glicerol y en donde la célula de levadura fermenta la al menos una de una hexosa, pentosa y glicerol al producto de fermentación y, opcionalmente, a formiato. Preferiblemente, la célula de levadura es una célula (huésped) según se define arriba en esta memoria. El procedimiento comprende preferiblemente, además, una o más etapas adicionales en las que se recuperan el producto de fermentación y/o el formiato. El procedimiento puede ser un procedimiento discontinuo, un procedimiento de alimentación por lotes o un procedimiento continuo como son bien conocidos en la técnica.

En un procedimiento preferido, la fuente de al menos una de una hexosa, una pentosa y glicerol comprende o consiste en: hexosa y pentosa; hexosa y glicerol; pentosa y glicerol; hexosa, pentosa y glicerol. En un procedimiento preferido adicional, la fuente de al menos una de una hexosa, una pentosa y glicerol comprende, además, ácido acético.

En un procedimiento preferido la fuente de hexosa comprende o consiste en glucosa. Preferiblemente, la fuente de pentosa comprende o consiste en al menos una de xilosa y arabinosa, de las cuales se prefiere xilosa. Preferiblemente, el medio fermentado por las células de la invención comprende o es alimentado con (fracciones de) biomasa hidrolizada que comprende al menos una de al menos una hexosa y una pentosa tal como glucosa, xilosa y/o arabinosa. La (fracciones de) biomasa hidrolizada que comprenden las hexosas y las pentosas comprenderán habitualmente también ácido acético (o una sal del mismo). Un ejemplo de biomasa hidrolizada a fermentar en los procedimientos de la invención es, p. ej., biomasa lignocelulósica hidrolizada. Biomasa lignocelulósica se entiende aquí como biomasa vegetal que se compone de celulosa, hemicelulosa y lignina. Los polímeros de hidratos de carbono (celulosa y hemicelulosa) están estrechamente vinculados a la lignina. Ejemplos de biomasa lignocelulósica a ser hidrolizada para uso en la presente invención incluyen los residuos agrícolas (incluyendo, p. ej., racimos vacíos (EFB) de la palma de aceite, rastrojo de maíz y bagazo de caña de azúcar), residuos de madera (incluyendo desechos de aserraderos y fábricas de celulosa) y residuos de papel (municipales). Métodos para la hidrólisis de biomasa tal como lignocelulosa son conocidos en la técnica per se, e incluyen, p. ej., ácidos tal como ácido sulfúrico y enzimas tales como celulasas y hemicelulasas.

En el procedimiento 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 oligómero o polímero de hidratos de carbono que comprende unidades de xilosa, glucosa y/o arabinosa tales como, p. ej., lignocelulosa, arabinanos, xilanos, celulosa, almidón y similares. Para la liberación de unidades de xilosa, glucosa y/o arabinosa a partir de este tipo de hidratos de carbono, carbohidrasas apropiadas (tales como arabinasas, xilanasas, glucanasas, amilasas, celulasas, glucanasas y similares) pueden añadirse al medio de fermentación o pueden ser producidas por la célula huésped modificada. En este último caso, la célula huésped puede ser modificada por ingeniería genética para producir y excretar estas carbohidrasas. Una ventaja adicional del uso de fuentes oligoméricas o poliméricas de glucosa es que permite mantener una concentración baja (más baja) de glucosa libre durante la fermentación, p. ej., mediante el uso de cantidades limitantes de la tasa de las carbohidrasas preferiblemente durante la fermentación. Esto, a su vez, evitará la represión de los sistemas necesarios para el metabolismo y transporte de azúcares no glucosa tales como xilosa y arabinosa. En un procedimiento preferido, la célula huésped modificada fermenta tanto la glucosa y al menos una de xilosa y arabinosa, preferiblemente de forma simultánea, en cuyo caso preferiblemente se utiliza una célula huésped modificada que es insensible a la represión

de glucosa para prevenir el crecimiento diáuxico. Además de una fuente de al menos una de una hexosa, una pentosa y glicerol como fuente de carbono, el medio de fermentación comprenderá, además, los ingredientes apropiados requeridos para el crecimiento de la célula huésped modificada. Composiciones de medios de fermentación para el crecimiento de microorganismos eucariotas tales como levaduras son bien conocidas en la técnica.

En el procedimiento de la invención, el medio comprende además preferiblemente y/o es alimentado con una fuente de glicerol. El glicerol para su uso en el procedimiento de la presente invención puede ser ventajosamente glicerol que se genera como un subproducto en la producción de biodiesel a partir de reacciones de transesterificación utilizando aceites vegetales tales como aceite de palma, o grasas animales y un alcohol.

El proceso de fermentación puede ser un proceso de fermentación aerobio o anaerobio. Un proceso de fermentación anaerobio se define aquí como una operación de proceso de fermentación en ausencia de oxígeno o en la que sustancialmente no se consume oxígeno, preferiblemente se consume menos de 5, 2,5 ó 1 mmol/L/h, más preferiblemente 0 mmol/L/h (es decir, el consumo de oxígeno no es detectable), y en la que las moléculas orgánicas sirven tanto como donantes de electrones y aceptores de electrones. En ausencia de oxígeno, NADH producido en la glucólisis y la formación de biomasa no puede ser oxidado mediante fosforilación oxidativa. Para resolver este problema muchos microorganismos utilizan piruvato o uno de sus derivados como un aceptor de electrones y de hidrógeno, regenerando con ello NAD^+. Así, en un proceso de fermentación anaerobia preferido se utiliza piruvato como un electrón (y aceptor de hidrógeno) y se reduce a productos de fermentación tales como etanol, así como productos de fermentación no etanol tales como ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, 1,3-propano-diol, butanoles (1-butanol, 2-butanol, isobutanol) y productos derivados de isoprenoides, preferiblemente bajo la producción concomitante de formiato. Se prefieren los procesos anaerobios de la invención frente a los procesos aerobios, ya que los procesos anaerobios no requieren inversiones ni energía para la aireación y, además, los procesos anaerobios producen mayores rendimientos de producto que los procesos aerobios. Alternativamente, el proceso de fermentación de la invención puede ser ejecutado bajo condiciones limitadas de oxígeno aerobias. Preferiblemente, en un proceso aerobio bajo condiciones limitadas de oxígeno, la tasa de consumo de oxígeno es al menos 5,5, más preferiblemente al menos 6 e incluso más preferiblemente al menos 7 mmol/L/h.

El proceso de fermentación se ejecuta preferiblemente a una temperatura que es óptima para las células modificadas de la invención. Por lo tanto, para la mayoría de las células de levadura, el proceso de fermentación se lleva a cabo a una temperatura que es inferior a 42°C, preferiblemente inferior a 38°C. Para las células de levadura, el proceso de fermentación se lleva a cabo preferiblemente a una temperatura que es inferior a 35, 33, 30 ó 28°C y a una temperatura que es superior a 20, 22 ó 25°C.

Debido a que el ácido fórmico no disociado es más tóxico para las células huésped de la invención en comparación con el ion ácido disociado, el proceso de fermentación se ejecuta preferiblemente a un pH que es mayor que el pKa del ácido fórmico, que es 3,75. Por consiguiente, el proceso de fermentación se ejecuta preferiblemente a un pH de al menos 3,8, 4,0, 4,5, 5,0, 5,5, 6,0, 6,5, 7,0, 7,5 u 8,0. Preferiblemente, el proceso de fermentación se lleva a cabo a al menos los pHs antes mencionados durante toda la duración del proceso. Alternativamente, el pH puede ser regulado a fin de mantener la concentración de ácido fórmico no disociado por debajo de una concentración dada. Preferiblemente, el pH del medio se regula durante el proceso de fermentación para mantener una concentración de ácido fórmico no disociado que no es mayor que 50,0, 30,0, 20,0, 18,1, 15,0, 10,0, 5,0, 2,0 ó 1,0 mM.

Un proceso de fermentación preferido de acuerdo con la invención es un procedimiento para la producción de etanol y, opcionalmente, formiato, en donde el proceso comprende la etapa de fermentar un medio con una célula de levadura, con lo que el medio contiene o es alimentado con una fuente de al menos una de una hexosa, una pentosa y glicerol y en el que la célula de levadura fermenta la al menos una de una hexosa, pentosa y glicerol a etanol y, opcionalmente, formiato. Opcionalmente, el proceso comprende la etapa de recuperación de al menos uno de etanol y, opcionalmente, formiato. La fermentación puede además realizarse según se describe arriba. En el proceso la productividad volumétrica de etanol es preferiblemente al menos 0,5, 1,0, 1,5, 2,0, 2,5, 3,0, 5,0 ó 10,0 g de etanol por litro por hora. El rendimiento de etanol en hexosa y/o pentosa y/o glicerol (y/o acetato) en el proceso es preferiblemente al menos 50, 60, 70, 80, 90, 95 ó 98%. El rendimiento de etanol se define aquí como un porcentaje del rendimiento máximo teórico que, para xilosa, glucosa y arabinosa, es 0,51 g de etanol por g de hexosa o pentosa. Para glicerol el rendimiento máximo teórico es 0,50 g de etanol por g de glicerol, y para ácido acético el rendimiento máximo teórico es de 0,77 g de etanol por g de ácido acético.

En este documento y en sus reivindicaciones, el verbo "comprender" y sus conjugaciones se utilizan en su sentido no limitativo en el sentido de que están incluidos los artículos siguientes de la palabra, pero no se excluyen los artículos que no se mencionan específicamente. Además, la referencia a un elemento por el artículo indefinido "un" o "una" no excluye la posibilidad de que esté presente más de uno de los elementos, a menos que el contexto requiera claramente que haya uno y sólo uno de los elementos. El artículo indefinido "un" o "una", por lo tanto, significa habitualmente "al menos uno".

Todas las referencias a patentes y bibliografía citadas en la presente memoria se incorporan como referencia en su totalidad.

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

Ejemplos

1. Ensayos de actividad enzimática

Extractos celulares para ensayos de actividad se prepararon a partir de cultivos por lotes aerobios o anaerobios en crecimiento exponencial y se analizaron en cuanto al contenido de proteína según se describe por Abbot et al, (2009, Appl. Environ. Microbiol. 75:2320-2325).

La actividad de acetaldehído deshidrogenasa NAD^+-dependiente (EC 1.2.1.10) se midió a 30°C monitorizando la oxidación de NADH a 340 nm. La mezcla de reacción (volumen total 1 ml) contenía tampón fosfato potásico 50 mM (pH 7,5), NADH 0,15 mM y extracto celular. La reacción se inició mediante la adición de acetil-coenzima A 0,5 mM.

Para determinar la actividad de glicerol 3-fosfato deshidrogenasa (EC 1.1.1.8), se prepararon extractos celulares según se describe arriba, excepto que el tampón fosfato fue reemplazado por tampón trietanolamina (10 mM, pH 5). Las actividades de glicerol-3-fosfato deshidrogenasa se ensayaron en extractos celulares a 30°C según se describió previamente (Blomberg y Adler, 1989, J. Bacteriol. 171:1087-1092,9). Las velocidades de reacción eran proporcionales a las cantidades de extracto de células añadidas.

La actividad de acetil-CoA sintasa (EC 6.2.1.1) se midió según 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 medida se realiza espectrofotométricamente cuando la acetil-CoA producida se acopla con reacciones de citrato sintasa y malato deshidrogenasa. El sistema de ensayo contenía Tris-Cl 100 mM (pH 7,6), $MgCl_2$ 10 mM, ATP 6 mM, malato 5 mM, NAD^+ 1 mM, NADH 0,1 mM, ditiotreitol 2,5 mM o 2-mercaptoetanol, 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 la tasa se midió a 340 nm y se calculó a partir del coeficiente de extinción de NADH (6,22 x 10^6 cm^2/mol).

La actividad de la glicerol deshidrogenasa y de la dihidroxiacetona quinasa se miden a 30°C en extractos celulares, esencialmente como se describe anteriormente (González et al, 2008, Metab. Eng. 10, 234-245). Las actividades enzimáticas de glicerol deshidrogenasa y de dihidroxiacetona quinasa se reseñan como µmoles de sustrato/min/mg de proteína celular.

2. Construcción de la cepa

Todas las modificaciones se inician con la cepa RN1041 fermentadora de xilosa. RN1041 es una cepa basada en CEN.PK con el siguiente genotipo:

Mat a, ura3-52, leu2-112, his3::loxP, gre3 ::loxP, loxP-Ptpi::TAL1, loxP-Ptpi::RK11, loxP-Ptpi-TKL1, loxP-Ptpi-RPE1, delta::Padh1XKSITcyc1-LEU2, delta::URA3-Ptpi-xylA-TcicI

Mat a = tipo de apareamiento a

Mutaciones *ura3-52, leu2-112, his3::loxP* en los genes ura3, leu2 y his3, el ura3 se complementa con la construcción de sobre-expresión xylA de *Piromyces*, leu2 se complementa con la construcción de sobre-expresión XKS1. his3 podría ser utilizado para la selección de plásmidos adicionales, RN1041 necesita histidina en el medio de crecimiento.

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

loxP-Ptpi ….= sobre-expresión de la vía pentosas fosfato het, el sitio loxP aguas arriba del promotor constitutivo se deja después de la eliminación del marcador.

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

Construcciones para la expresión de genes pflA y pflB de *E. coli* en levadura

Los genes pflA y pflB de *E. coli* se obtuvieron como genes sintéticos optimizados en codones. La construcción de expresión de pflA ligando un fragmento promotor de TEF1 (cortado con las enzimas de restricción *Sal*I y *Hind*III, de la colección de plásmidos Royal Nedalco), un ORF (marco de lectura abierto) sintético de pflA (cortado con *Hind*III y

*Bss*HII) y un fragmento terminador de ADH1 (cortado con *Bss*HII y *Bsi*WI; de la colección de plásmidos Royal Nedalco) junto con pCRII romo (Invitrogen) para producir pRN613.

La construcción de expresión de pflB ligando un fragmento de promotor de PGK1 (cortado con las enzimas de restricción *Spe*I y *Pst*I; de la colección de plásmidos Royal Nedalco), un ORF sintético de pflB (cortado con *Pst*I y *Sal*I) y un fragmento terminador de PGI1 (cortado con *Xho*I y *Bsi*WI; de la colección de plásmidos Royal Nedalco) juntos en pCRII romo (Invitrogen) para producir pRN614.

El marcador LEU2 del plásmido pRS305 (Sikorski y Hieter, 1989, Genetics 122:19-27) se intercambia para el marcador ZeoMX (*zeo*MX = resistencia a fleomicina debido a la expresión del gen Tn5 ble con una secuencia de promotor de TEF1 y de terminador de *Ashbya gossypii*). Los transformantes de levadura se seleccionan en medios que contienen zeomicina, pRS305 se corta con *BSt*BI y *Bsr*GI y un fragmento de ZeoMX (de la colección de plásmidos Royal Nedalco) se corta con las mismas enzimas de restricción. El ligamiento proporciona el plásmido pRS30z.

Un ORI de 2µ se introduce cortando pRS30z con *Afl*II en *Mfe*I y cortando un fragmento de 2µ (de la colección de plásmidos Royal Nedalco) con *Afl*II y *Eco*RI. Los fragmentos se ligan para producir pRN615 (6341 pb).

Para la construcción final pRN615 se corta con *Xho*I y *Spe*I, este vector se combina con los insertos de pRN613 cortados con *Sal*I y *Bsi*WI (1510 pb) y pRN614 cortado con *Spe*I y *Bsi*WI (3381 pb) para producir pRN616 (SEQ ID NO: 31).

Para la expresión de sólo el gen pflA de *E. coli*, pRN613 se corta con *Sal*I y *Bsi*WI y este fragmento se liga en pRN615 cortado con *Acc*65I y *Xho*I para producir pRN619 (SEQ ID NO: 32).

Para la expresión de sólo el gen pflB de *E. coli*, pRN614 se corta con XhoI y *Bsi*WI y este fragmento se liga en pRN615 cortado con *Acc*65I y *Xho*I para producir pRN620 (SEQ ID NO: 33).

Los plásmidos pRN616, pRN619 y pRN620 se utilizan para transformar la cepa de levadura RN1041 junto con los otros plásmidos como se indica en la Tabla 7 para generar las cepas como se indica en la Tabla 7.

Construcciones para la sobre-expresión de GCY1, gldA, DAK1 dhaK y GUP1

GCY1

La PCR se lleva a cabo en ADN genómico de *S. cerevisiae* con cebadores introduciendo un sitio *Pst*I 5' de la ATG y un sitio *Afl*II 3' de la TAA para producir el fragmento de SEQ ID NO: 25. Un fragmento de ADN que comprende el promotor Actin1 de *S. cerevisiae* se liga aguas arriba del ORF de GCY1 y el fragmento de ADN que comprende el fragmento de terminador ADH1 de *S. cerevisiae* se ligó aguas abajo del ORF de GCY1.

gldA

La construcción para la expresión en la levadura del *gld*A de *E. coli* se hizo mediante ligamiento de un fragmento de promotor ACT1 de levadura (cortado con las enzimas de restricción *SpeI* y *PstI*), un ORF sintético (SEQ ID NO: 50), que codifica el *gldA* de E. coli (cortado con *PstI* en *Bss*HII) y un fragmento de terminador de CYC1 de levadura (cortado con *Bss*HII y *Bsi*WI) juntos en pCRII romo (Invitrogen) para producir pRNgldA (SEQ ID NO: 51).

DAK1

La PCR se lleva a cabo en ADN genómico de S. cerevisiae con cebadores introduciendo un sitio *Xba*I 5' de la ATG y un sitio *Sal*I 3' de la TAA para producir el fragmento que figura en SEQ ID NO: 26. Un fragmento de ADN que comprende el promotor TPI1 de *S. cerevisiae* se liga aguas arriba del ORF de DAK1 y el fragmento de ADN que comprende el fragmento de terminador PGI1 de *S. cerevisiae* se ligó aguas abajo del ORF de DAK1 para producir la construcción de expresión de SEQ ID NO: 26.

dhaK

La construcción para la expresión en la levadura de *dhaK* de *Citrobacter freundii* se hizo ligando el fragmento de promotor TPI1 de levadura (cortado con las enzimas de restricción *Xho*I y *Xba*I), un ORF sintético (SEQ ID NO: 53) que codifica *dhaK* de *C. freundii* (cortado con *Xba*I y *Sal*I) y un fragmento de terminador PGI1 levadura (cortado con *Xho*I y *Bsi*WI) juntos en pCRII romo (Invitrogen) para producir pRNdhaK (SEQ ID NO: 54).

GUP1

La PCR se lleva a cabo en ADN genómico de *S. cerevisiae* con cebadores introduciendo un sitio *Hind*III 5' de la ATG y un sitio *Bam*HI 3' de la TAA para producir el fragmento que figura en SEQ ID NO: 27. Un fragmento de ADN que comprende el promotor TDH3 de *S. cerevisiae* se liga aguas arriba del ORF de GUP1 y el fragmento de ADN que

comprende el fragmento terminador CYC1 de *S. cerevisiae* se liga aguas abajo del ORF de GUP1 para producir la construcción de expresión de SEQ ID NO: 27.

Cepas que sobre-expresan GCY1, DAK1 y GUP1

Las construcciones de expresión arriba descritas de GCY1, DAK1 y GUP1 se combinan en un plásmido centroméricо a base de pRS315 que tiene el gen de resistencia a higromicina como marcador seleccionable para producir pRN605 (SEQ ID NO: 29). El plásmido pRN605 se utiliza para transformar la cepa de levadura RN1041 junto con los otros plásmidos según se indica en la Tabla 7 para generar las cepas según se indica en la Tabla 7.

Las construcciones de expresión arriba descritas de GCY1 y DAK1 se combinan en un plásmido centromérico a base de pRS315 que tiene el gen de resistencia a higromicina como marcador seleccionable para producir pRN608 (SEQ ID NO: 30) se utilizan para transformar la cepa de levadura RN1041 junto con los otros plásmidos según se indica en la Tabla 7 para generar la cepa RN1098, según se indica en la Tabla 7.

Cepas que expresan gldA de *E. coli* con o sin la expresión de dhaK de *C. freundii*

La construcción de expresión gldA de *E.coli* se corta del plásmido pRNgldA con las enzimas de restricción *Spe*I y *Bsi*WI. La construcción de expresión dhaK de *C.freundii* se corta a partir del plásmido pRNdhaK con las enzimas de restricción *Xho*I y *Bsi*WI. Estos fragmentos se ligan en el plásmido pRN595 cortado con las enzimas de restricción *Spe*I y *Sal*I para producir pRN957 (SEQ ID NO: 55). El plásmido pRN958 (SEQ ID NO: 56) se hace posteriormente a partir del plásmido pRN957 mediante la supresión de la construcción de expresión dhaK de *C.freundii*. Para ello, el plásmido pRN957 se corta con las enzimas de restricción *Psp*OMI y *Avr*II, las partes de cadena sencilla se rellenan con la ADN polimerasa phusion (Finnzymes) y el plásmido se vuelve a ligar.

Las construcciones de expresión de DAK1 y GUP1 se combinan en un plásmido centromérico a base de pRS315 que tiene el gen de resistencia a higromicina como marcador seleccionable para producir pRN607 (SEQ ID NO: 57).

Los plásmidos pRN957, pRN958 y pRN607 utilizados para transformar la levadura junto con los otros plásmidos indicados en la Tabla 7 para producir la cepa RN1194, RN1195 y RN1196.

Construcción para la sobreexpresión de FPS1

La PCR se lleva a cabo en ADN genómico de *S. cerevisiae* con cebadores introduciendo un sitio *Nsi*I 5' de la ATG y un sitio *Bam*HI 3' de la TAA para producir el fragmento que figura en SEQ ID NO: 28. Un fragmento de ADN que comprende el promotor ADH1 (medio) de *S. cerevisiae* se liga aguas arriba del ORF de FSP1 y el fragmento de ADN que comprende el fragmento de terminador CYC1 de *S. cerevisiae* se liga aguas abajo del ORF de FSP1 para producir la construcción de expresión de SEQ ID NO: 28. El ligamiento de estos fragmentos en pCRII romo produce pRN617. El inserto de pRN617 se corta con *Bsi*WI y *Xho*I y se clona en pRN616 cortado con *Bsr*GI y *Sal*I para producir pRN618 (SEQ ID NO: 34). El plásmido pRN618 se utiliza para transformar la cepa de levadura RN1041 junto con los otros plásmidos según se indica en la Tabla 7 para generar las cepas según se indica en la Tabla 7.

Construcciones para la deleción de FDH1 y FDH2

Los cebadores FDHuf, FDHur, FDHdf y FDHdr se utilizan para la amplificación de fragmentos de secuencias genómicas aguas arriba y aguas abajo de los dos genes FDH1 y FDH2 para su inactivación. Los mismos cebadores se utilizan para los dos alelos FDH1 y FDH2.

Para la inactivación del alelo FDH1, un fragmento de PCR de 423 pb aguas arriba con un sitio *Bsp*EI en el extremo 3' para el ligamiento al marcador de MX, y un sitio *Bbs*I en el extremo 5' para el aislamiento de la construcción de deleción se amplifica utilizando FDHuf y FDHur y se clona en pCRII (romo) utilizando la clonación topo (Invitrogen), que resulta en un sitio *Bst*BI adicional en el extremo 5'.

FDHuf: TCGAAGACTCCGAATGAAAAAGACATGCCAG (SEQ ID NO: 35)

FDHur: TCCGGATACCAAGTTCATTTTCAATACACCCCA (SEQ ID NO: 36)

Un fragmento de PCR de 378 pb aguas abajo con un sitio *Nsi*I y un sitio *Sph*I en el extremo 5' para el ligamiento al marcador de MX, y un sitio *Bbs*I en el extremo 3' para el aislamiento de la construcción de deleción se amplifica utilizando FDHdf y FDHdr.

FDHdf: ATGCATGCAGAATGGTTCTTATGCCAC (SEQ ID NO: 37)

FDHdr: GAAGACAGTTCTGTTATTAACGACGAGCCA (SEQ ID NO: 38)

Para la construcción final el fragmento aguas abajo cortado con *Xho*I y *Nsi*I se liga al marcador patMX (colección Royal Nedalco) cortado con *Bsp*EI y *Nsi*I en el plásmido que contiene el fragmento aguas arriba cortado con *Bsp*EI y

*Xho*I. El plásmido final, pRN621 (SEQ ID NO: 39), se corta con *Bbs*I antes de la transformación de la levadura. Se seleccionan transformantes para la resistencia a fosfinotricina. La integración correcta del fragmento de interrupción se verifica mediante PCR con cebadores específicos.

FDH2 se encuentra como un pseudogen en algunas cepas (p. ej. S288C, YPL275W e YPL276W). Dos mutaciones restaurar la homología de ORF con FDH1 (T436C y delA476). El análisis de la secuencia muestra que en el fondo CEN.PK (RN1041) FDH2 es un ORF homólogo a FDH1.

Para la construcción final el fragmento aguas abajo cortado con *Xho*I y *Sph*I se liga al marcador kanMX (colección Royal Nedalco) cortado con *Bsp*EI y *Sph*I en el plásmido que contiene el fragmento aguas arriba cortado con BspEI y *Xho*I. El plásmido final, pRN622 (SEQ ID NO: 40), se corta con *Bbs*I antes de la transformación de la levadura. Los transformantes se seleccionan para resistencia a G418. La integración correcta del fragmento de interrupción se verifica mediante PCR con cebadores específicos.

Fragmentos de interrupción de pRN621 y pRN622 se utilizan para transformar la cepa de levadura RN1041 junto con los otros plásmidos según se indica en la Tabla 7 para generar las cepas según se indica en la Tabla 7.

Construcciones para la expresión de los genes mhpF, adhE de E. coli o adh2 de Entamoebe histolytica en levaduras

Para la expresión del gen mhpF de *E. coli*, un fragmento de promotor de PGK1 de levadura (*Spe*I-*Pst*I) y un fragmento de terminador de ADH1 (AflII-NotI) (ambos de la colección de plásmidos Nedalco) se ligaron en el fragmento sintético optimizado en codones que codifica el mhpF de *E. coli* (SEQ ID NO: 8).

pRS 303 con 2µ ori (= pRN347, colección de plásmidos Royal Nedalco) se cortó con *Spe*I y *Not*I y la construcción de expresión de mhpF se clonó en este vector para producir pRN558 (SEQ ID NO: 41). pRN558 se utiliza para transformar la cepa de levadura RN1041 (selección en medio que carece de histidina) junto con los otros plásmidos según se indica en la Tabla 7 para generar las cepas según se indica en la Tabla 7. La sobre-expresión de mhpF se verifica mediante qPCR.

Para la expresión del gen adhE de *E. coli*, un fragmento sintético optimizado en codones que codifica el adhE de *E. coli* (SEQ ID NO: 10) se corta con *Xba*I y *Afl*II y se liga en pRN558 cortado con *Xba*I y *Afl*II (reemplazando el gen mhpF de *E. coli* en pRN558) para producir pRN595 (SEQ ID NO: 42). pRN595 se utiliza para transformar la cepa de levadura RN1041 (selección en medio que carece de histidina), junto con los otros plásmidos según se indica en la Tabla 7 para generar las cepas según se indica en la Tabla 7. La sobre-expresión de adhE se verifica mediante qPCR.

Para la expresión del adh2 de Entamoebe histolytica, un fragmento sintético optimizado en codones que codifica el adh2 de *E. coli* (SEQ ID NO: 12) se corta con *Xba*I y *Afl*II y se liga en pRN558 cortado con *Xba*I y *Afl*II (reemplazando el gen mhpF de *E. coli* en pRN558) para producir pRN596 (SEQ ID NO: 43). pRN596 se utiliza para transformar la cepa de levadura RN1041 (selección en medio que carece de histidina), junto con los otros plásmidos según se indica en la Tabla 7 para generar las cepas según se indica en la Tabla 7. La sobre-expresión de adh2 se verifica mediante qPCR.

Construcciones para la deleción de GPD2

Los cebadores GPD2uf, GPD2ur, GPD2df y GPD2dr se utilizan para la amplificación de fragmentos de secuencias genómicas aguas arriba y aguas abajo del gen GPD2 para su inactivación. Un fragmento de PCR de 407bp aguas arriba con un sitio *Afl*II en el extremo 3' (derivado de la secuencia de GPD2) y un sitio *Bgl*II en el extremo 5' (para el aislamiento de la construcción de deleción) se amplifica utilizando GPD2uf, GPD2ur y se clona en pCR2.1 (topo T / A, Invitrogen).

GPD2uf: GGTACCAGATCTTTTGCGGCGAGGTGCCG (SEQ ID NO: 44)

GPD2ur: TCTAGACTTAAGGAATGTGTATCTTGTTAATCTTCTGACAGC (SEQ ID NO: 45)

Un fragmento de PCR de 417 pb aguas abajo con un sitio *Xho*I en el extremo 5' y un sitio *Bgl*II en el extremo 3' se amplifica utilizando GPD2df y GPD2dr.

GPD2df: CTCGAGATAGTCTACAACAACGTCCGCA (SEQ ID NO: 46)

GPD2dr: CCATGGAGATCTGCAGTGAAAAAGCTCGAAGAAACAGCT (SEQ ID NO: 47)

Para la construcción final, el plásmido que contiene el fragmento aguas arriba se corta con *Afl*II y *Kpn*, el fragmento aguas abajo se corta con *Xho*I en *Nco*I y el marcador natMX (colección de plásmidos Royal Nedalco) se corta con *Afl*II en *Xho*I y los fragmentos se ligan para producir el plásmido pRN594 (SEQ ID NO: 48). pRN594 se corta con *Bgl*II antes de la transformación de la levadura. pRN594 se utiliza para transformar la cepa de levadura RN1041 (selección para la resistencia a nourseotricina) junto con los otros plásmidos según se indica en la Tabla 7 para

generar las cepas según se indica en la Tabla 7. La integración correcta se verifica mediante PCR utilizando cebadores específicos.

3. Fermentaciones con las cepas construidas

3.1 Ajuste experimental

Las cepas RN1041, RN1088, RN1089, RN1090, RN1091, RN1092, RN1093, RN1094, RN1095, RN1096, RN1097, RN1098, RN1099, RN1194, RN1195 y RN1196 se someten todas a ensayo en cuanto a su capacidad para crecer tanto en condiciones anóxicas (sin oxígeno presente) a expensas de glicerol como fuente de carbono. Al mismo tiempo, se monitoriza la formación de ácido fórmico y etanol. En varios casos se confirma mediante experimentos que las cepas consumidoras de glicerol también conservan su capacidad de fermentar glucosa y xilosa. Precultivos de las cepas se preparan mediante la inoculación de un cultivo patrón de glicerol congelado de la levadura en un medio YP (extracto de levadura - peptona) con la adición del azúcar glucosa (2,5% p/v) a 32°C y pH 5,5. Después de 24 h de incubación, este cultivo se utiliza para inocular los cultivos del fermentador. Las células se recogen mediante centrifugación y se lavan con dH_2O fría. La inoculación de la levadura utilizada es 2 gramos de levadura en materia seca por litro de medio de fermentación.

Las fermentaciones se llevan a cabo como cultivo en quimiostato a 32°C en fermentadores de 1 litro con un volumen de trabajo de 800 ml. La tasa de dilución se fija en 0,1 h^{-1}. Estas fermentaciones discontinuas anóxicas se realizan en medio YP al que se añade glicerol en combinación con xilosa y/o glucosa. El pH del cultivo se mantiene en pH = 7,0 mediante la adición automática de KOH 2M. Los cultivos se agitan a 300 rpm.

Muestras para análisis de producción de glucosa, xilosa, etanol, glicerol y ácido fórmico se toman en situaciones de estado estacionario que se obtienen después de 5 cambios de volumen del recipiente de fermentación. Las concentraciones de etanol, glicerol y ácido fórmico se controlan mediante análisis por HPLC. La glucosa y la xilosa se determinan mediante análisis de HPAEC (Dionex).

3.2 Resultados

La cepa RN1041 (Tablas 6 y 7) es una cepa que se deriva de la cepa RN1001 (RN1001 his3::loxP). La cepa RN1001 se ha descrito previamente. Si se añade glucosa (25 mM) o xilosa (25 mM) al medio YP, entonces las fermentaciones anóxicas en las condiciones arriba descritas conducen a un consumo total de glucosa o xilosa por parte de esta cepa. Las concentraciones de etanol en los medios de glucosa y xilosa, respectivamente, son 40 mM y 33 mM en estado estable. Se forman pequeñas cantidades de glicerol tanto para glucosa como para xilosa (1,8 y 1,5 mM, respectivamente). No se observa la formación de ácido fórmico. En ausencia de azúcares en el medio YP, no se forma etanol ni glicerol o ácido fórmico. Si se añade glicerol al medio a una concentración de 50 mM, entonces el organismo no produce etanol adicional ni ácido fórmico en comparación con las situaciones con glucosa o xilosa sola.

La cepa RN1088 se derivó de la cepa RN1041 por

1. supra-regulación de DAK1, GCY1 y GUP1 permitiendo una absorción y consumo incrementados de glicerol en condiciones anóxicas

2. expresión heteróloga de pflA + pflB, lo que permite la conversión de piruvato en ácido fórmico y acetilCoA

3. expresión heteróloga de adhE2, lo que permite la conversión de acetil-CoA en etanol

4. eliminación de FDH1 y FDH2 evitando la oxidación de ácido fórmico por parte de esta cepa.

El organismo se somete a ensayo en cuanto al crecimiento y la formación de producto de la misma manera que RN1041 según se describe arriba. Con cualquiera de glucosa o xilosa se obtienen resultados similares a RN1041 en términos de producción de etanol. Sin embargo, la formación de glicerol fue ligeramente inferior y se observan trazas de ácido fórmico. Si se añade glicerol al medio con contenido en xilosa o glucosa a una concentración de 50 mM, entonces el organismo no consume más glicerol ni produce etanol adicional y, además, no se produce ácido fórmico adicional. Las cepas RN1089 y RN1090 tienen una base genética similar a RN1088. La única diferencia es que adhE no se expresa en estos organismos. En su lugar se expresan adh2 o mhpF. Las características de estas cepas durante las fermentaciones de glicerol son similares a las características de la cepa RN1088. Ambas cepas mostraron características similares a RN1088 para la fermentación tanto de glucosa como de xilosa como cepa RN1088.

Las cepas RN1091, RN1092 y RN1093 tienen el mismo fondo genético que la cepa RN1088 con la excepción de deshidrogenasas de ácido fórmico. Estas cepas tienen 1 ó 2 genes todavía presentes que codifican una formiato deshidrogenasa. Ninguno de los 3 organismos es capaz de realizar buenas fermentaciones en glicerol. Sólo se

observaron trazas de cualquiera de etanol y ácido fórmico durante los experimentos de fermentación. Los organismos hacen fermentar tanto glucosa como xilosa a un ritmo equiparable a la cepa RN1088.

Las cepas RN1094, RN1095 y RN1096 tienen el mismo fondo genético que la cepa RN1088, con la excepción de la expresión de cualquiera de pflA o pflB. Estas cepas carecen de cualquiera de estos dos genes o de cualquiera de los genes. Ninguno de los 3 organismos es capaz de realizar buenas fermentaciones en glicerol. No se observa etanol o ácido fórmico adicional basado en la adición de glicerol 50 mM a un medio con contenido en xilosa (25 mM) durante los experimentos de fermentación. Los organismos hacen fermentar tanto la glucosa o xilosa a tasas equiparables a la cepa RN1088.

La cepa RN1097 se deriva de la cepa RN1088 suprimiendo GPD2. Al hacerlo, se podría evitar una operación parcial de un ciclo fútil en el nivel de consumo/producción de glicerol. Con glucosa o xilosa, la concentración de etanol en estados estacionarios son 38 mM y 31 mM, respectivamente. El glicerol se acumula en 0,9 mM y 0,5 mM, respectivamente, y ácido fórmico se había acumulado en 1,0 mM y 0,8 mM, respectivamente. Si se añade glicerol al medio con contenido en xilosa o glucosa a una concentración de 50 mM, entonces el organismo no consume más glicerol ni produce etanol adicional y, además, no se produce ácido fórmico adicional.

Las cepas RN1098 y RN1099 se asemejan a la cepa RN1088, pero difieren genéticamente en el nivel del o de los transportadores de glicerol. La cepa RN1098 tiene FPS1 sobre-expresado en lugar de GUP1, mientras que la cepa RN1099 tiene los dos genes transportadores expresados. Los organismos se someten a ensayo en cuanto al crecimiento y la formación de producto de la misma manera que RN1041 según se describe arriba. Con cualquiera de glucosa o xilosa se obtienen resultados similares a RN1041 en términos de producción de etanol. Sin embargo, la formación de glicerol era ligeramente inferior y se observaron trazas de ácido fórmico. Si se añadió glicerol ya sea al medio con contenido en xilosa o glucosa a una concentración de 50 mM, entonces los organismos RN1098 y RN1099 no consumían más glicerol ni producían etanol adicional y además no se produjo ácido fórmico adicional alguno.

RN1194 se somete a ensayo en cuanto al crecimiento y la formación de producto de la misma manera que RN1041 según se describe arriba. Con cualquiera de glucosa o xilosa, se obtienen resultados similares a RN1041 en términos de producción de etanol. Sin embargo, la formación de glicerol es ligeramente inferior y se observaron trazas de ácido fórmico. Si se añade glicerol ya sea al medio con contenido en xilosa o glucosa a una concentración de 50 mM, entonces el organismo produce etanol adicional y, además, se produce ácido fórmico. La concentración de etanol en estado estacionario en el medio de glicerol/glucosa es 75 mM y la concentración de ácido fórmico es 32 mM. El nivel de glicerol restante es de 7 mM. La concentración de etanol en estado estacionario en el medio de glicerol/xilosa es 66 mM y la concentración de ácido fórmico es 30 mM. El nivel de glicerol restante es de 8 mM. No se hacen intentos para someter a ensayo cualquier otro producto de fermentación que podrían haber sido producidos sobre la base de glicerol. RN1196 produce resultados similares a los de RN1194, aunque el consumo de glicerol y la producción de etanol y ácido fórmico se reducen en cierta medida en comparación con RN1194.

La cepa RN1195 se deriva de la cepa RN1194 suprimiendo GPD2. Al hacerlo, se podría evitar una operación parcial de un ciclo fútil en el nivel de consumo/producción de glicerol. Con glucosa o xilosa, la concentración de etanol en estados estacionarios son 38 mM y 31 mM, respectivamente. El glicerol se acumula en 0,9 mM y 0,5 mM, respectivamente, y ácido fórmico se acumula en 1,0 mM y 0,8 mM, respectivamente. Si se añade glicerol al medio con contenido en xilosa (25 mM) a una concentración de 50 mM, entonces el organismo produce tanto etanol como ácido fórmico. La concentración final de etanol en estado estacionario era de 72 mM, para glicerol 2 mM, y para ácido fórmico 37 mM. No se hicieron intentos para someter a ensayo cualesquiera otros productos de fermentación que podrían haber sido producidos sobre la base de glicerol. Estos resultados indican claramente que la deleción del gen GPD2 tiene un efecto beneficioso sobre la fermentación de glicerol en la cepa RN1195 en comparación con la cepa RN1194.

Tabla 6

Gen	GPD2	adhE	ade2	mphF	DAK1	GCY1	GUP1	FPS1	FDH1	FDH2	pflA+pflB	E.c gldA	C.f. dhaK
Cepa / Marcardor	natMX	HIS3	HIS3	HIS3	hphMX	hphMX	hphMX	zeoMX	patMX	kanMX	zeo MX	HIS3	HIS3
RN1041	ts	ausente	ausente	ausente	ts	ts	ts	ts	ts	ts	ausente/ausente	ausente	ausente
RN1088	ts	expr	ausente	ausente	asc	asc	asc	ts	del	del	expr/expr	ausente	ausente
RN1089	ts	ausente	expr	ausente	asc	asc	asc	ts	del	del	expr/expr	ausente	ausente
RN1090	ts	ausente	ausente	expr	asc	asc	asc	ts	del	del	expr/expr	ausente	ausente
RN1091	ts	expr	ausente	ausente	asc	asc	asc	ts	ts	ts	expr/expr	ausente	ausente
RN1092	ts	expr	ausente	ausente	asc	asc	asc	ts	del	ts	expr/expr	ausente	ausente
RN1093	ts	expr	ausente	ausente	asc	asc	asc	ts	ts	del	expr/expr	ausente	ausente
RN1094	ts	expr	ausente	ausente	asc	asc	asc	ts	del	del	ausente/ausente	ausente	ausente
RN1095	ts	expr	ausente	ausente	asc	asc	asc	ts	del	del	expr/ausente	ausente	ausente
RN1096	ts	expr	ausente	ausente	asc	asc	asc	ts	del	del	ausente/expr	ausente	ausente
RN1097	del	expr	ausente	ausente	asc	asc	asc	ts	del	del	expr/expr	ausente	ausente
RN1098	ts	expr	ausente	ausente	asc	asc	ts	asc	del	del	expr/expr	ausente	ausente
RN1099	ts	expr	ausente	ausente	asc	asc	asc	asc	del	del	expr/expr	ausente	ausente
RN1194	ts	expr	ausente	ausente	ts	ts	ts	ts	del	del	expr/expr	expr	expr
RN1195	del	expr	ausente	ausente	ts	ts	ts	ts	del	del	expr/expr	expr	expr
RN1196	ts	expr	ausente	ausente	asc	ts	asc	ts	del	del	expr/expr	expr	ausente

Tabla 7

	plásmidos					
selección	His3	hph	loxPNatMXloxP	nat	kan	zeo
Cepa						
RN1041	pRN347					
RN1088	pRN595	pRN605		pRN621	pRN622	pRN616
RN1089	pRN596	pRN605		pRN621	pRN622	pRN616
RN1090	pRN558	pRN605		pRN621	pRN622	pRN616
RN1091	pRN595	pRN605				pRN616
RN1092	pRN595	pRN605		pRN621		pRN616
RN1093	pRN595	pRN605			pRN622	pRN616
RN1094	pRN595	pRN605		pRN621	pRN622	
RN1095	pRN595	pRN605		pRN621	pRN622	pRN619
RN1096	pRN595	pRN605		pRN621	pRN622	pRN620
RN1097	pRN595	pRN605	pRN594	pRN621	pRN622	pRN616
RN1098	pRN595	pRN608		pRN621	pRN622	pRN618
RN1099	pRN595	pRN605		pRN621	pRN622	pRN618
RN1194	pRN597			pRN621	pRN622	pRN616
RN1195	pRN957		pRN594	pRN621	pRN622	pRN616
RN1196	pRN958	pRN607		pRN621	pRN622	pRN616

pRN347 = pRS303 + 2µORI

LISTADO DE SECUENCIAS

<110> Royal Nedalco s. v.
C5 Yeast Company s. v.

<120> Cepas de levaduras manipuladas para producer etanol y formiato

<130> P6031691PCT

<150> EP10191736.7
<151> 2010-11-18

<150> US61/415.054
<151> 2010-11-18

<150> US61/471.836
<151> 2011-04-05

<160> 57

<170> PatentIn version 3.3

<210> 1
<211> 760
<212> PRT
<213> Escherichia coli

<400> 1

```
Met Ser Glu Leu Asn Glu Lys Leu Ala Thr Ala Trp Glu Gly Phe Thr
1               5                  10                  15

Lys Gly Asp Trp Gln Asn Glu Val Asn Val Arg Asp Phe Ile Gln Lys
            20                  25                  30

Asn Tyr Thr Pro Tyr Glu Gly Asp Glu Ser Phe Leu Ala Gly Ala Thr
        35                  40                  45

Glu Ala Thr Thr Thr Leu Trp Asp Lys Val Met Glu Gly Val Lys Leu
    50                  55                  60

Glu Asn Arg Thr His Ala Pro Val Asp Phe Asp Thr Ala Val Ala Ser
65                  70                  75                  80

Thr Ile Thr Ser His Asp Ala Gly Tyr Ile Asn Lys Gln Leu Glu Lys
                85                  90                  95

Ile Val Gly Leu Gln Thr Glu Ala Pro Leu Lys Arg Ala Leu Ile Pro
            100                 105                 110

Phe Gly Gly Ile Lys Met Ile Glu Gly Ser Cys Lys Ala Tyr Asn Arg
        115                 120                 125

Glu Leu Asp Pro Met Ile Lys Lys Ile Phe Thr Glu Tyr Arg Lys Thr
    130                 135                 140
```

```
His Asn Gln Gly Val Phe Asp Val Tyr Thr Pro Asp Ile Leu Arg Cys
145                 150                 155                 160

Arg Lys Ser Gly Val Leu Thr Gly Leu Pro Asp Ala Tyr Gly Arg Gly
                165                 170                 175

Arg Ile Ile Gly Asp Tyr Arg Arg Val Ala Leu Tyr Gly Ile Asp Tyr
            180                 185                 190

Leu Met Lys Asp Lys Leu Ala Gln Phe Thr Ser Leu Gln Ala Asp Leu
        195                 200                 205

Glu Asn Gly Val Asn Leu Glu Gln Thr Ile Arg Leu Arg Glu Glu Ile
    210                 215                 220

Ala Glu Gln His Arg Ala Leu Gly Gln Met Lys Glu Met Ala Ala Lys
225                 230                 235                 240

Tyr Gly Tyr Asp Ile Ser Gly Pro Ala Thr Asn Ala Gln Glu Ala Ile
                245                 250                 255

Gln Trp Thr Tyr Phe Gly Tyr Leu Ala Ala Val Lys Ser Gln Asn Gly
            260                 265                 270

Ala Ala Met Ser Phe Gly Arg Thr Ser Thr Phe Leu Asp Val Tyr Ile
        275                 280                 285

Glu Arg Asp Leu Lys Ala Gly Lys Ile Thr Glu Gln Glu Ala Gln Glu
    290                 295                 300

Met Val Asp His Leu Val Met Lys Leu Arg Met Val Arg Phe Leu Arg
305                 310                 315                 320

Thr Pro Glu Tyr Asp Glu Leu Phe Ser Gly Asp Pro Ile Trp Ala Thr
                325                 330                 335

Glu Ser Ile Gly Gly Met Gly Leu Asp Gly Arg Thr Leu Val Thr Lys
            340                 345                 350

Asn Ser Phe Arg Phe Leu Asn Thr Leu Tyr Thr Met Gly Pro Ser Pro
        355                 360                 365

Glu Pro Asn Met Thr Ile Leu Trp Ser Glu Lys Leu Pro Leu Asn Phe
    370                 375                 380

Lys Lys Phe Ala Ala Lys Val Ser Ile Asp Thr Ser Ser Leu Gln Tyr
```

385 390 395 400

Glu Asn Asp Asp Leu Met Arg Pro Asp Phe Asn Asn Asp Asp Tyr Ala
405 410 415

Ile Ala Cys Cys Val Ser Pro Met Ile Val Gly Lys Gln Met Gln Phe
420 425 430

Phe Gly Ala Arg Ala Asn Leu Ala Lys Thr Met Leu Tyr Ala Ile Asn
435 440 445

Gly Gly Val Asp Glu Lys Leu Lys Met Gln Val Gly Pro Lys Ser Glu
450 455 460

Pro Ile Lys Gly Asp Val Leu Asn Tyr Asp Glu Val Met Glu Arg Met
465 470 475 480

Asp His Phe Met Asp Trp Leu Ala Lys Gln Tyr Ile Thr Ala Leu Asn
485 490 495

Ile Ile His Tyr Met His Asp Lys Tyr Ser Tyr Glu Ala Ser Leu Met
500 505 510

Ala Leu His Asp Arg Asp Val Ile Arg Thr Met Ala Cys Gly Ile Ala
515 520 525

Gly Leu Ser Val Ala Ala Asp Ser Leu Ser Ala Ile Lys Tyr Ala Lys
530 535 540

Val Lys Pro Ile Arg Asp Glu Asp Gly Leu Ala Ile Asp Phe Glu Ile
545 550 555 560

Glu Gly Glu Tyr Pro Gln Phe Gly Asn Asn Asp Pro Arg Val Asp Asp
565 570 575

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

His Thr Tyr Arg Asp Ala Ile Pro Thr Gln Ser Val Leu Thr Ile Thr
595 600 605

Ser Asn Val Val Tyr Gly Lys Lys Thr Gly Asn Thr Pro Asp Gly Arg
610 615 620

Arg Ala Gly Ala Pro Phe Gly Pro Gly Ala Asn Pro Met His Gly Arg
625 630 635 640

Asp Gln Lys Gly Ala Val Ala Ser Leu Thr Ser Val Ala Lys Leu Pro
645 650 655

Phe Ala Tyr Ala Lys Asp Gly Ile Ser Tyr Thr Phe Ser Ile Val Pro
660 665 670

Asn Ala Leu Gly Lys Asp Asp Glu Val Arg Lys Thr Asn Leu Ala Gly
675 680 685

Leu Met Asp Gly Tyr Phe His His Glu Ala Ser Ile Glu Gly Gly Gln
690 695 700

His Leu Asn Val Asn Val Met Asn Arg Glu Met Leu Leu Asp Ala Met
705 710 715 720

Glu Asn Pro Glu Lys Tyr Pro Gln Leu Thr Ile Arg Val Ser Gly Tyr
725 730 735

Ala Val Arg Phe Asn Ser Leu Thr Lys Glu Gln Gln Gln Asp Val Ile
740 745 750

Thr Arg Thr Phe Thr Gln Ser Met
755 760

<210> 2
<211> 2283
<212> ADN
<213> Artificial

<220>
<223> pflB de E. coli optimizado en codones sintético

<400> 2
atgtctgaat tgaacgagaa gttggctacc gcttgggaag gtttcaccaa gggtgactgg 60

caaaacgaag ttaacgttag agacttcatc caaaagaact acaccccata cgaaggtgac 120

gaatctttct tggctggtgc taccgaagct accaccacct tgtgggacaa ggttatggaa 180

ggtgttaagt tggaaaacag aacccacgct ccagttgact tcgacaccgc tgttgcttct 240

accatcacct ctcacgacgc tggttacatc aacaagcaat tggaaaagat cgttggttta 300

caaaccgaag ctccattgaa gagagctttg atcccattcg gtggtatcaa gatgatcgaa 360

ggttcttgta aggcttacaa cagagaattg gacccaatga tcaagaagat tttcaccgaa 420

tacagaaaga cccacaacca aggtgttttc gacgtttaca ctccagacat cttgagatgt 480

agaaagtctg gtgttttgac tggtttgcca gacgcttacg gtagaggtag aatcatcggt 540

gactacagaa gagttgcttt gtacggtatc gactacttga tgaaggacaa gttggctcaa 600

ttcacctctt tgcaagctga cttggaaaac ggtgttaact tggaacaaac catcagattg 660

agagaagaaa tcgctgaaca acacagagct ttgggtcaaa tgaaggaaat ggctgctaag 720

tacggttacg acatctctgg tccagctacc aacgctcaag aagctatcca atggacctac 780

ttcggttact tggctgctgt taagtctcaa aacggtgctg ctatgtcttt cggtaggacc 840

tctaccttct tggacgttta catcgaaaga gacttgaagg ctggtaagat caccgaacaa 900

gaagctcaag aaatggttga ccacttggtt atgaagttga gaatggttag attcttgaga 960

accccagaat acgacgaatt gttctctggt gacccaatct gggctaccga atctatcggt 1020

ggtatgggtt tggacggtag aaccttggtt accaagaact ctttcagatt cttgaacacc 1080

ttatacacca tgggtccatc tccagaacca aacatgacca tcttgtggtc tgaaaagtta 1140

ccattgaact tcaagaagtt cgctgctaag gtttctatcg acacctcttc tttgcaatac 1200

gaaaacgacg acttgatgag accagacttc aacaacgacg actacgctat cgcttgttgt 1260

gtttctccaa tgatcgttgg taagcaaatg caattcttcg gtgctagagc taacttggct 1320

aagaccatgt tgtacgctat caacggtggt gttgacgaaa agttgaagat gcaagttggt 1380

ccaaagtctg aaccaatcaa gggtgacgtt ttgaactacg acgaagttat ggaaagaatg 1440

gaccacttca tggactggtt ggctaagcaa tacatcaccg ctttgaacat catccactac 1500

atgcacgaca agtactctta cgaagcatca ttgatggctt tgcacgacag agacgtaatc 1560

agaaccatgg cttgtggtat cgctggtttg tctgttgctg ctgactcttt gtctgctatc 1620

aagtacgcta aggttaagcc aatcagagac gaagacggtt tggctatcga cttcgaaatc 1680

gaaggtgaat accctcaatt cggtaacaac gacccaagag ttgacgactt ggctgttgac 1740

ttggttgaaa gatttatgaa gaagatccaa aagttgcaca cctacagaga cgctatccca 1800

acccaatctg ttttgactat cacatctaac gttgtttacg gtaagaagac tggtaacacc 1860

ccagacggta gaagagctgg tgctccattc ggtccaggtg ctaacccaat gcacggtaga 1920

gaccaaaagg gtgctgtagc atctttgacc tctgttgcta agttgccatt cgcttacgct 1980

aaggacggta tctcttacac cttctctatc gttccaaacg ctttgggtaa ggacgatgaa 2040

gttagaaaga ccaacttggc tggtttgatg gacggttact tccaccacga agcatctatc 2100

gaaggtggtc aacacttgaa cgtaaatgtt atgaacagag aaatgttgtt ggacgctatg 2160

gaaaacccag aaaagtaccc acaattgacc atcagagttt ctggttacgc tgttagattc 2220

aactctttga ccaaggaaca acaacaagac gttatcacca gaaccttcac ccaatctatg 2280

taa 2283

<210> 3
<211> 246
<212> PRT
<213> Escherichia coli

<400> 3

```
Met Ser Val Ile Gly Arg Ile His Ser Phe Glu Ser Cys Gly Thr Val
1               5                  10                  15

Asp Gly Pro Gly Ile Arg Phe Ile Thr Phe Phe Gln Gly Cys Leu Met
            20                  25                  30

Arg Cys Leu Tyr Cys His Asn Arg Asp Thr Trp Asp Thr His Gly Gly
        35                  40                  45

Lys Glu Val Thr Val Glu Asp Leu Met Lys Glu Val Val Thr Tyr Arg
    50                  55                  60

His Phe Met Asn Ala Ser Gly Gly Gly Val Thr Ala Ser Gly Gly Glu
65                  70                  75                  80

Ala Ile Leu Gln Ala Glu Phe Val Arg Asp Trp Phe Arg Ala Cys Lys
                85                  90                  95

Lys Glu Gly Ile His Thr Cys Leu Asp Thr Asn Gly Phe Val Arg Arg
            100                 105                 110

Tyr Asp Pro Val Ile Asp Glu Leu Leu Glu Val Thr Asp Leu Val Met
        115                 120                 125

Leu Asp Leu Lys Gln Met Asn Asp Glu Ile His Gln Asn Leu Val Gly
    130                 135                 140

Val Ser Asn His Arg Thr Leu Glu Phe Ala Lys Tyr Leu Ala Asn Lys
145                 150                 155                 160

Asn Val Lys Val Trp Ile Arg Tyr Val Val Val Pro Gly Trp Ser Asp
                165                 170                 175

Asp Asp Asp Ser Ala His Arg Leu Gly Glu Phe Thr Arg Asp Met Gly
            180                 185                 190

Asn Val Glu Lys Ile Glu Leu Leu Pro Tyr His Glu Leu Gly Lys His
        195                 200                 205

Lys Trp Val Ala Met Gly Glu Glu Tyr Lys Leu Asp Gly Val Lys Pro
    210                 215                 220

Pro Lys Lys Glu Thr Met Glu Arg Val Lys Gly Ile Leu Glu Gln Tyr
225                 230                 235                 240

Gly His Lys Val Met Phe
                245
```

<210> 4
<211> 741
<212> ADN
<213> Artificial

<220>
<223> pflA de E. coli optimizado en codones sintético

<400> 4

atgtctgtta tcggtagaat ccactctttc gaatcttgtg gtactgttga cggtccaggt 60

atcagattca tcaccttctt ccaaggttgt ttgatgagat gtttgtactg tcacaacaga 120

gacacctggg acacccacgg tggtaaggaa gttactgttg aagacttgat gaaggaagtt 180

gttacctaca gacactttat gaacgcttca ggaggtggtg ttaccgcttc tggtggtgaa 240

gctatcttgc aagctgaatt tgttagagac tggttcagag cttgtaagaa ggaaggtatc 300

cacacctgtt tggacaccaa cggtttcgtt agaagatacg acccagttat cgacgaattg 360

ttggaagtta ccgacttggt tatgttggac ttgaagcaaa tgaacgacga aatccaccaa 420

aacttggttg gtgtttctaa ccacagaacc ttggaatttg ctaagtactt ggctaacaag 480

aacgttaagg tttggatcag atacgttgtt gttccaggtt ggtctgacga cgacgactct 540

gctcacagat tgggtgagtt caccagagac atgggtaacg ttgaaaagat cgaattgttg 600

ccataccacg aattgggtaa gcacaagtgg gttgctatgg gtgaagaata caagttggac 660

ggtgttaagc caccaaagaa ggaaaccatg gaaagagtta agggtatctt ggaacaatac 720

ggtcacaagg ttatgttcta a 741

<210> 5
<211> 376
<212> PRT
<213> Saccharomyces cerevisiae

<400> 5

Met Ser Lys Gly Lys Val Leu Leu Val Leu Tyr Glu Gly Gly Lys His
1 5 10 15

Ala Glu Glu Gln Glu Lys Leu Leu Gly Cys Ile Glu Asn Glu Leu Gly
20 25 30

Ile Arg Asn Phe Ile Glu Glu Gln Gly Tyr Glu Leu Val Thr Thr Ile
35 40 45

Asp Lys Asp Pro Glu Pro Thr Ser Thr Val Asp Arg Glu Leu Lys Asp
50 55 60

Ala Glu Ile Val Ile Thr Thr Pro Phe Phe Pro Ala Tyr Ile Ser Arg
65 70 75 80

```
Asn Arg Ile Ala Glu Ala Pro Asn Leu Lys Leu Cys Val Thr Ala Gly
                85                  90                  95

Val Gly Ser Asp His Val Asp Leu Glu Ala Ala Asn Glu Arg Lys Ile
            100                 105                 110

Thr Val Thr Glu Val Thr Gly Ser Asn Val Val Ser Val Ala Glu His
        115                 120                 125

Val Met Ala Thr Ile Leu Val Leu Ile Arg Asn Tyr Asn Gly Gly His
    130                 135                 140

Gln Gln Ala Ile Asn Gly Glu Trp Asp Ile Ala Gly Val Ala Lys Asn
145                 150                 155                 160

Glu Tyr Asp Leu Glu Asp Lys Ile Ile Ser Thr Val Gly Ala Gly Arg
                165                 170                 175

Ile Gly Tyr Arg Val Leu Glu Arg Leu Val Ala Phe Asn Pro Lys Lys
            180                 185                 190

Leu Leu Tyr Tyr Asp Tyr Gln Glu Leu Pro Ala Glu Ala Ile Asn Arg
        195                 200                 205

Leu Asn Glu Ala Ser Lys Leu Phe Asn Gly Arg Gly Asp Ile Val Gln
    210                 215                 220

Arg Val Glu Lys Leu Glu Asp Met Val Ala Gln Ser Asp Val Val Thr
225                 230                 235                 240

Ile Asn Cys Pro Leu His Lys Asp Ser Arg Gly Leu Phe Asn Lys Lys
                245                 250                 255

Leu Ile Ser His Met Lys Asp Gly Ala Tyr Leu Val Asn Thr Ala Arg
            260                 265                 270

Gly Ala Ile Cys Val Ala Glu Asp Val Ala Glu Ala Val Lys Ser Gly
        275                 280                 285

Lys Leu Ala Gly Tyr Gly Gly Asp Val Trp Asp Lys Gln Pro Ala Pro
    290                 295                 300

Lys Asp His Pro Trp Arg Thr Met Asp Asn Lys Asp His Val Gly Asn
305                 310                 315                 320

Ala Met Thr Val His Ile Ser Gly Thr Ser Leu Asp Ala Gln Lys Arg
```

325 330 335

Tyr Ala Gln Gly Val Lys Asn Ile Leu Asn Ser Tyr Phe Ser Lys Lys
340 345 350

Phe Asp Tyr Arg Pro Gln Asp Ile Ile Val Gln Asn Gly Ser Tyr Ala
355 360 365

Thr Arg Ala Tyr Gly Gln Lys Lys
370 375

<210> 6
<211> 376
<212> PRT
<213> Saccharomyces cerevisiae

<400> 6
Met Ser Lys Gly Lys Val Leu Leu Val Leu Tyr Glu Gly Gly Lys His
1 5 10 15

Ala Glu Glu Gln Glu Lys Leu Leu Gly Cys Ile Glu Asn Glu Leu Gly
20 25 30

Ile Arg Asn Phe Ile Glu Glu Gln Gly Tyr Glu Leu Val Thr Thr Ile
35 40 45

Asp Lys Asp Pro Glu Pro Thr Ser Thr Val Asp Arg Glu Leu Lys Asp
50 55 60

Ala Glu Ile Val Ile Thr Thr Pro Phe Phe Pro Ala Tyr Ile Ser Arg
65 70 75 80

Asn Arg Ile Ala Glu Ala Pro Asn Leu Lys Leu Cys Val Thr Ala Gly
85 90 95

Val Gly Ser Asp His Val Asp Leu Glu Ala Ala Asn Glu Arg Lys Ile
100 105 110

Thr Val Thr Glu Val Thr Gly Ser Asn Val Val Ser Val Ala Glu His
115 120 125

Val Met Ala Thr Ile Leu Val Leu Ile Arg Asn Tyr Asn Gly Gly His
130 135 140

Gln Gln Ala Ile Asn Gly Glu Trp Asp Ile Ala Gly Val Ala Lys Asn
145 150 155 160

Glu Tyr Asp Leu Glu Asp Lys Ile Ile Ser Thr Val Gly Ala Gly Arg

```
                165                 170                 175

Ile Gly Tyr Arg Val Leu Glu Arg Leu Val Ala Phe Asn Pro Lys Lys
            180                 185                 190

Leu Leu Tyr Tyr Asp Tyr Gln Glu Leu Pro Ala Glu Ala Ile Asn Arg
        195                 200                 205

Leu Asn Glu Ala Ser Lys Leu Phe Asn Gly Arg Gly Asp Ile Val Gln
    210                 215                 220

Arg Val Glu Lys Leu Glu Asp Met Val Ala Gln Ser Asp Val Val Thr
225                 230                 235                 240

Ile Asn Cys Pro Leu His Lys Asp Ser Arg Gly Leu Phe Asn Lys Lys
                245                 250                 255

Leu Ile Ser His Met Lys Asp Gly Ala Tyr Leu Val Asn Thr Ala Arg
            260                 265                 270

Gly Ala Ile Cys Val Ala Glu Asp Val Ala Glu Ala Val Lys Ser Gly
        275                 280                 285

Lys Leu Ala Gly Tyr Gly Gly Asp Val Trp Asp Lys Gln Pro Ala Pro
    290                 295                 300

Lys Asp His Pro Trp Arg Thr Met Asp Asn Lys Asp His Val Gly Asn
305                 310                 315                 320

Ala Met Thr Val His Ile Ser Gly Thr Ser Leu Asp Ala Gln Lys Arg
                325                 330                 335

Tyr Ala Gln Gly Val Lys Asn Ile Leu Asn Ser Tyr Phe Ser Lys Lys
            340                 345                 350

Phe Asp Tyr Arg Pro Gln Asp Ile Ile Val Gln Asn Gly Ser Tyr Ala
        355                 360                 365

Thr Arg Ala Tyr Gly Gln Lys Lys
    370                 375
```

<210> 7
<211> 316
<212> PRT
<213> Escherichia coli

<400> 7

```
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
145                 150                 155                 160

Lys Ser Ala Gly Pro Gly Thr Arg Ala Asn Ile Asp Glu Phe Thr Glu
                165                 170                 175

Thr Thr Ser Arg Ala Ile Glu Val Val Gly Gly Ala Ala Lys Gly Lys
            180                 185                 190

Ala Ile Ile Val Leu Asn Pro Ala Glu Pro Pro Leu Met Met Arg Asp
        195                 200                 205

Thr Val Tyr Val Leu Ser Asp Glu Ala Ser Gln Asp Asp Ile Glu Ala
    210                 215                 220

Ser Ile Asn Glu Met Ala Glu Ala Val Gln Ala Tyr Val Pro Gly Tyr
225                 230                 235                 240

Arg Leu Lys Gln Arg Val Gln Phe Glu Val Ile Pro Gln Asp Lys Pro
                245                 250                 255
```

Val Asn Leu Pro Gly Val Gly Gln Phe Ser Gly Leu Lys Thr Ala Val
260 265 270

Trp Leu Glu Val Glu Gly Ala Ala His Tyr Leu Pro Ala Tyr Ala Gly
275 280 285

Asn Leu Asp Ile Met Thr Ser Ser Ala Leu Ala Thr Ala Glu Lys Met
290 295 300

Ala Gln Ser Leu Ala Arg Lys Ala Gly Glu Ala Ala
305 310 315

<210> 8
<211> 968
<212> ADN
<213> Artificial

<220>
<223> mhpF de E.coli optimizado en codones sintético

<400> 8
ctgcagtcta gatgtctaag agaaaggttg ctatcatcgg ttctggtaac atcggtactg 60

acttgatgat caagatccta agacacggtc aacacttgga aatggctgtt atggttggta 120

tcgacccaca atctgacggt ttggctagag ctagaagaat gggtgttgct accacccacg 180

aaggtgttat cggtttgatg aacatgccag aattcgctga catcgacatc gttttcgacg 240

ctacctctgc tggtgctcac gttaagaacg acgctgcttt gagagaagct aagccagaca 300

tcagattgat cgacttgacc ccagctgcta tcggtccata ctgtgttcca gttgttaact 360

tggaagctaa cgttgaccaa ttaaacgtta acatggttac ctgtggtggt caagctacca 420

tcccaatggt tgctgctgtt tcaagagttg ctagagttca ctacgctgaa atcatcgctt 480

ctatcgcttc taagtctgct ggtccaggta ccagagctaa catcgacgaa ttcaccgaaa 540

ccacctctag ggctatcgaa gttgttggtg gtgctgctaa gggtaaggct atcatcgttt 600

tgaacccagc tgaaccacca ttgatgatga gagacaccgt ttacgttttg tctgacgaag 660

catctcaaga cgacatcgaa gcttcaatca acgaaatggc tgaagctgtt caagcatacg 720

ttccaggtta cagattgaag caaagagttc aattcgaagt tatcccacaa gacaagccag 780

ttaacttgcc aggtgttggt caattctctg gtttgaagac cgctgtttgg ttggaagttg 840

aaggtgctgc tcactacttg ccagcttacg ctggtaactt ggacattatg acctcttctg 900

ctttggctac cgctgaaaag atggctcaat ctttggctag aaaggctggt gaagctgctt 960

aagcgcgc 968

<210> 9
<211> 891
<212> PRT

<213> Escherichia coli

<400> 9

```
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

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

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

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> 10
<211> 2690
<212> ADN
<213> Artificial

<220>
<223> adhE de E. coli optimizado en codones sintético

<400> 10
tctagaaaat ggctgttacc aacgttgctg aattgaacgc tttggttgaa agggttaaga 60

aggctcaaag agaatacgct tctttcaccc aagaacaagt tgacaagatc ttcagagctg 120

ctgctttggc tgctgctgac gctagaatcc cattggctaa gatggctgtt gctgaatctg 180

gtatgggtat cgttgaagac aaggttatca agaaccactt cgcttctgaa tacatctaca 240

acgcttacaa ggacgaaaag acctgtggtg ttttgtcaga agacgacacc ttcggtacca 300

tcaccatcgc tgaaccaatc ggtatcatct gtggtatcgt tccaaccacc aacccaacct 360

ctaccgctat cttcaagtct ttgatctctt tgaagaccag aaacgctatc atcttctctc 420

cacacccaag agctaaagac gctaccaaca aggctgctga catcgttttg caagctgcta 480

tcgctgctgg tgctccaaag gacttgatcg gttggatcga ccaaccatct gttgaattgt 540

ctaacgcttt gatgcaccac ccagacatca acttgatctt ggctaccggt ggtccaggta 600

tggttaaggc tgcttactct tctggtaagc cagctatcgg tgttggtgct ggtaacaccc 660

cagttgttat cgacgaaacc gctgacatca agagagctgt tgcttctgtt ttgatgtcta 720

agaccttcga caacggtgtt atctgtgctt ctgaacaatc tgttgttgtt gttgactctg 780

tttacgacgc tgttagagaa agattcgcta cccacggtgg ttacttgttg caaggtaagg 840

aattgaaggc tgttcaagac gttatcttga agaacggtgc tttgaacgct gctatcgttg 900

gtcaaccagc ttacaagatc gctgaattag ctggtttctc tgttccagaa aacaccaaga 960

tcttgatcgg tgaagttacc gttgttgacg aatctgaacc attcgctcac gaaaagttgt 1020

ctccaacctt ggctatgtac agagctaagg acttcgaaga cgctgttgaa aaagctgaaa 1080

agttggttgc tatgggtggt attggtcaca cctcttgttt gtacaccgac caagacaacc 1140

aaccagctag agtttcttac ttcggtcaaa agatgaagac cgctagaatc ttgatcaaca 1200

ccccagcttc tcaaggtggt atcggtgact tgtacaactt caagttggct ccatctttga 1260

ccttgggttg tggttcttgg ggtggtaact ctatctctga aaacgttggt ccaaagcact 1320

tgatcaacaa gaagaccgtt gctaagagag ctgaaaacat gttgtggcac aagttgccaa 1380

aatctatcta cttcagaaga ggttctttgc caatcgcttt ggacgaagtt atcaccgacg 1440

gtcacaagag agctttgatc gttaccgaca gattcttgtt caacaacggt tacgctgacc 1500

aaatcacctc tgttttgaag gctgctggtg ttgaaaccga agttttcttc gaagttgaag 1560

ctgacccaac cttgtctatc gttagaaagg gtgctgaatt ggctaactct ttcaagccag 1620

acgttatcat cgctttgggt ggtggttctc caatggacgc tgctaagatc atgtgggtta 1680

tgtacgaaca cccagaaacc cacttcgaag aattggcttt gagattcatg gacatcagaa 1740

agagaatcta caagttccca aagatgggtg ttaaggctaa gatgatcgct gttaccacca 1800

cctctggtac cggttctgaa gttaccccat tcgctgttgt taccgacgac gctaccggtc 1860

aaaagtaccc attggctgac tacgctttga ccccagacat ggctatcgtt gacgctaact 1920

tggttatgga catgccaaag tctttgtgtg ctttcggtgg tttggacgct gttacccacg 1980

ctatggaagc ttacgtttct gttttggctt ctgaattctc tgacggtcaa gctttgcaag 2040

ctttgaagtt gttgaaggaa tacttgccag cttcttacca cgaaggttct aagaacccag 2100

ttgctagaga aagagttcac tctgctgcta ccatcgctgg tatcgctttc gctaacgctt 2160

tcttgggtgt ttgtcactct atggctcaca agttgggttc tcaattccac atcccacacg 2220

gtttggctaa cgctttgttg atctgtaacg ttatcagata caacgctaac gacaacccaa 2280

ccaagcaaac cgctttctct caatacgaca gaccacaagc tagaagaaga tacgctgaaa 2340

tcgctgacca cttgggtttg tctgctccag gtgacagaac cgctgcaaag atcgaaaagt 2400

tgttggcttg gttggaaacc ttgaaggctg aattgggtat cccaaagtct atcagagaag 2460

ctggtgttca agaagctgac ttcttggcta acgttgacaa gttgtctgaa gacgctttcg 2520

acgaccaatg taccggtgct aacccaagat acccattgat ctctgaattg aagcaaatct 2580

tgttggacac ctactacggt agagactacg ttgaaggtga aaccgctgct aagaaggaag 2640

ctgctccagc taaggctgaa aagaaggcta agaagtctgc ttagcttaag 2690

<210> 11
<211> 870
<212> RPT
<213> Entamoeba histolítica

<400> 11

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

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

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                 400

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

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
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> 12
<211> 2627
<212> ADN
<213> Artificial

<220>
<223> ADH2 de E. histolítica optimizado en codones sintético

<400> 12
tctagaaaat gtctacccaa caaaccatga ccgttgacga acacatcaac caattagtta 60

gaaaggctca agttgctttg aaggaatact tgaagccaga atacacccaa gaaaagatcg 120

actacatcgt taagaaggct tctgttgctg ctttggacca acactgtgct ttggctgctg 180

ctgctgttga agaaaccggt agaggtatct tcgaagacaa ggctaccaag aacatcttcg 240

cttgtgaaca cgttacccac gaaatgagac acgctaagac cgttggtatc atcaacgttg 300

acccattgta cggtatcacc gaaatcgctg aaccagttgg tgttgtttgt ggtgttaccc 360

cagttaccaa cccaacctct accgctatct tcaagtcttt gatctctatc aagaccagaa 420

acccaatcgt tttctctttc cacccatctg ctttgaagtg ttctattatg gctgctaaaa 480

tcgttagaga cgctgctatc gctgctggtg ctccagaaaa ctgtatccaa tggatcgaat 540

tcggtggtat cgaagcttct aacaagttga tgaaccaccc aggtgttgct actatcttgg 600

ctaccggtgg taacgctatg gttaaggctg catactcttc tggtaagcca gctttgggtg 660

ttggtgctgg taacgttcca acctacatcg aaaagacctg taacatcaag caagctgcta 720

acgacgttgt tatgtctaag 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 atcaaggtac cagaagactg taacatcatc tgtgctgttt 1020

gtaaggaagt tggtccaaac gaaccattga ccagagaaaa gttgtctcca gttttggcta 1080

tcttgaaagc tgaaaacacc caagacggta tcgacaaggc tgaagctatg gttgaattta 1140

acggtagagg tcactctgct gctatccact ctaacgacaa ggctgttgtt gaaaagtacg 1200

ctttgaccat gaaggcttgt agaatcttgc acaacacccc atcttctcaa ggtggtatcg 1260

gttctatcta caactacatc tggccatctt tcaccttggg ttgtggttct tacggtggta 1320

actctgtttc tgctaacgtt acctaccaca acttgttgaa catcaagaga ttggctgaca 1380

gaagaaacaa cttgcaatgg ttcagagttc caccaaagat cttcttcgaa ccacactcta 1440

tcagatactt ggctgaattg aaggaattgt ctaagatctt catcgtttct gacagaatga 1500

tgtacaagtt gggttacgtt gacagagtta tggacgtttt gaagagaaga tctaacgaag 1560

ttgaaatcga aatcttcatc gacgttgaac cagacccatc tatccaaacc gttcaaaagg 1620

gtttggctgt tatgaacacc ttcggtccag acaacatcat cgctatcggt ggtggttctg 1680

ctatggacgc tgctaagatc atgtggttgt tgtacgaaca cccagaagct gacttcttcg 1740

ctatgaagca aaagttcatc gacttgagaa agagagcttt caagttccca accatgggta 1800

agaaggctag attgatctgt atcccaacca cctctggtac cggttctgaa gttaccccat 1860

tcgctgttat ctctgaccac gaaaccggta agaagtaccc attggctgac tactctttga 1920

ccccatctgt tgctatcgtt gacccaatgt tcaccatgtc tttgccaaag agagctatcg 1980

```
ctgacaccgg tttggacgtt ttggttcacg ctaccgaagc ttacgtttct gttatggcta   2040

acgaatacac cgacggtttg gctagagaag ctgttaagtt ggtttttgaa aacttgttga   2100

agtcttacaa cggtgacttg gaagctagag aaaagatgca caacgctgct accatcgctg   2160

gtatggcttt cgcttctgct ttcttgggta tggaccactc tatggctcac aaggttggtg   2220

ctgctttcca cttgccacac ggtagatgtg ttgctgtttt gttgccacac gttatcagat   2280

acaacggtca aaagccaaga aagttggcta tgtggccaaa gtacaacttc tacaaggctg   2340

accaaagata catggaattg gctcaaatgg ttggtttgaa gtgtaacacc ccagctgaag   2400

gtgttgaagc tttcgctaag gcttgtgaag aattgatgaa ggctaccgaa accatcaccg   2460

gtttcaagaa ggctaacatc gacgaagctg cttggatgtc taaggttcca gaaatggctt   2520

tgttggcttt cgaagaccaa tgttctccag ctaacccaag agttccaatg gttaaggaca   2580

tggaaaagat cttgaaggct gcttactacc caatcgctta gcttaag                 2627
```

<210> 13
<211> 312
<212> PRT
<213> Saccharomyces cerevisiae

<400> 13

```
Met Pro Ala Thr Leu His Asp Ser Thr Lys Ile Leu Ser Leu Asn Thr
1               5                  10                  15

Gly Ala Gln Ile Pro Gln Ile Gly Leu Gly Thr Trp Gln Ser Lys Glu
            20                  25                  30

Asn Asp Ala Tyr Lys Ala Val Leu Thr Ala Leu Lys Asp Gly Tyr Arg
        35                  40                  45

His Ile Asp Thr Ala Ala Ile Tyr Arg Asn Glu Asp Gln Val Gly Gln
    50                  55                  60

Ala Ile Lys Asp Ser Gly Val Pro Arg Glu Glu Ile Phe Val Thr Thr
65                  70                  75                  80

Lys Leu Trp Cys Thr Gln His His Glu Pro Glu Val Ala Leu Asp Gln
                85                  90                  95

Ser Leu Lys Arg Leu Gly Leu Asp Tyr Val Asp Leu Tyr Leu Met His
            100                 105                 110

Trp Pro Ala Arg Leu Asp Pro Ala Tyr Ile Lys Asn Glu Asp Ile Leu
        115                 120                 125

Ser Val Pro Thr Lys Lys Asp Gly Ser Arg Ala Val Asp Ile Thr Asn
```

```
    130                 135                 140

Trp Asn Phe Ile Lys Thr Trp Glu Leu Met Gln Glu Leu Pro Lys Thr
145                 150                 155                 160

Gly Lys Thr Lys Ala Val Gly Val Ser Asn Phe Ser Ile Asn Asn Leu
                165                 170                 175

Lys Asp Leu Leu Ala Ser Gln Gly Asn Lys Leu Thr Pro Ala Ala Asn
            180                 185                 190

Gln Val Glu Ile His Pro Leu Leu Pro Gln Asp Glu Leu Ile Asn Phe
        195                 200                 205

Cys Lys Ser Lys Gly Ile Val Val Glu Ala Tyr Ser Pro Leu Gly Ser
    210                 215                 220

Thr Asp Ala Pro Leu Leu Lys Glu Pro Val Ile Leu Glu Ile Ala Lys
225                 230                 235                 240

Lys Asn Asn Val Gln Pro Gly His Val Val Ile Ser Trp His Val Gln
                245                 250                 255

Arg Gly Tyr Val Val Leu Pro Lys Ser Val Asn Pro Asp Arg Ile Lys
            260                 265                 270

Thr Asn Arg Lys Ile Phe Thr Leu Ser Thr Glu Asp Phe Glu Ala Ile
        275                 280                 285

Asn Asn Ile Ser Lys Glu Lys Gly Glu Lys Arg Val Val His Pro Asn
    290                 295                 300

Trp Ser Pro Phe Glu Val Phe Lys
305                 310
```

<210> 14
<211> 584
<212> PRT
<213> Saccharomyces cerevisiae

<400> 14

```
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
    130                 135                 140

Val Gly Arg Arg Ala Leu Ala Gly Thr Val Leu Val His Lys Ile Val
145                 150                 155                 160

Gly Ala Phe Ala Glu Glu Tyr Ser Ser Lys Tyr Gly Leu Asp Gly Thr
                165                 170                 175

Ala Lys Val Ala Lys Ile Ile Asn Asp Asn Leu Val Thr Ile Gly Ser
            180                 185                 190

Ser Leu Asp His Cys Lys Val Pro Gly Arg Lys Phe Glu Ser Glu Leu
        195                 200                 205

Asn Glu Lys Gln Met Glu Leu Gly Met Gly Ile His Asn Glu Pro Gly
    210                 215                 220

Val Lys Val Leu Asp Pro Ile Pro Ser Thr Glu Asp Leu Ile Ser Lys
225                 230                 235                 240

Tyr Met Leu Pro Lys Leu Leu Asp Pro Asn Asp Lys Asp Arg Ala Phe
                245                 250                 255

Val Lys Phe Asp Glu Asp Asp Glu Val Val Leu Leu Val Asn Asn Leu
            260                 265                 270

Gly Gly Val Ser Asn Phe Val Ile Ser Ser Ile Thr Ser Lys Thr Thr
        275                 280                 285
```

```
Asp Phe Leu Lys Glu Asn Tyr Asn Ile Thr Pro Val Gln Thr Ile Ala
    290                 295                 300

Gly Thr Leu Met Thr Ser Phe Asn Gly Asn Gly Phe Ser Ile Thr Leu
305                 310                 315                 320

Leu Asn Ala Thr Lys Ala Thr Lys Ala Leu Gln Ser Asp Phe Glu Glu
                325                 330                 335

Ile Lys Ser Val Leu Asp Leu Leu Asn Ala Phe Thr Asn Ala Pro Gly
            340                 345                 350

Trp Pro Ile Ala Asp Phe Glu Lys Thr Ser Ala Pro Ser Val Asn Asp
        355                 360                 365

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

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> 15
<211> 591
<212> PRT
<213> Saccharomyces cerevisiae

<400> 15

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

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 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

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

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> 16
<211> 560
<212> PRT
<213> Saccharomyces cerevisiae

<400> 16

```
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

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
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> 17
<211> 609
<212> PRT
<213> Saccharomyces cerevisiae

<400> 17

```
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
65                  70                  75                  80

Phe Thr Val Val Val Pro Phe Met Ile Lys Ala Ala Leu Ala Thr Ser
                85                  90                  95

Ser Glu Ser Asn Pro Asn Tyr Tyr Lys Phe Ser Gly Leu Leu Ala His
            100                 105                 110

Gly Trp Ile Leu Gly Arg Lys Val Asp Asn Ser Asp Pro Gln Tyr Arg
        115                 120                 125

Phe Phe Arg Ser Asn Phe Phe Leu Leu Ala Ile Leu Ile Leu Leu Gln
    130                 135                 140

Ile Ile Leu Lys Lys Val Phe Val Lys Phe Ser Lys Ile Pro Lys Thr
145                 150                 155                 160

Lys Phe Asp Phe Ala Cys Gly Leu Val Phe Val Cys Phe Met Tyr Gly
                165                 170                 175

Ile Asn Ser Val Lys Leu Phe Thr His Ala Phe Ile Phe Phe Thr Leu
            180                 185                 190

Ala His Ser Leu Lys Arg Lys Arg Leu Ile Ala Ala Phe Ala Ile Trp
        195                 200                 205

Ser Tyr Gly Ile Phe Thr Leu Phe Ile Asn Gln Lys Met Lys Asn Leu
    210                 215                 220
```

```
Pro Phe Asn Asn Ile Ala Ile Ile Leu Ser Pro Met Asp Gln Trp Tyr
225                 230                 235                 240

Lys Gly Ile Val Pro Arg Trp Asp Phe Phe Phe Asn Phe Thr Leu Leu
                245                 250                 255

Arg Leu Leu Ser Tyr Ser Met Asp Phe Leu Glu Arg Trp His Glu Gln
            260                 265                 270

Leu Ser Arg Gln Pro Ser Ile Asp Tyr Asp Asp Arg Arg Pro Glu Phe
        275                 280                 285

Arg Lys Ser Leu Ser Gly Ser Thr Leu Gln Thr Ile Tyr Glu Ser Gly
    290                 295                 300

Lys Asn Val Leu Glu Glu Lys Glu Arg Leu Val Ala Glu His His Ile
305                 310                 315                 320

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
```

<210> 18
<211> 669
<212> PRT
<213> Saccharomyces cerevisiae

<400> 18

```
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 Asn Asn Asn Asn Asn Asn Asn Asn Asn Asn 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

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

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> 19
<211> 713
<212> PRT
<213> Saccharomyces cerevisiae

<400> 19
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
225                 230                 235                 240

Ile Thr Thr Asp Glu Ser Asn Arg Gly Gly Lys Val Ile Glu Thr Lys
                245                 250                 255

Arg Ile Val Asp Asp Ala Leu Arg Glu Thr Pro Gly Val Arg His Val
            260                 265                 270

Leu Val Tyr Arg Lys Thr Asn Asn Pro Ser Val Ala Phe His Ala Pro
        275                 280                 285

Arg Asp Leu Asp Trp Ala Thr Glu Lys Lys Lys Tyr Lys Thr Tyr Tyr
    290                 295                 300

Pro Cys Thr Pro Val Asp Ser Glu Asp Pro Leu Phe Leu Leu Tyr Thr
305                 310                 315                 320

Ser Gly Ser Thr Gly Ala Pro Lys Gly Val Gln His Ser Thr Ala Gly
                325                 330                 335

Tyr Leu Leu Gly Ala Leu Leu Thr Met Arg Tyr Thr Phe Asp Thr His
            340                 345                 350

Gln Glu Asp Val Phe Phe Thr Ala Gly Asp Ile Gly Trp Ile Thr Gly
        355                 360                 365

His Thr Tyr Val Val Tyr Gly Pro Leu Leu Tyr Gly Cys Ala Thr Leu
    370                 375                 380
```

```
Val Phe Glu Gly Thr Pro Ala Tyr Pro Asn Tyr Ser Arg Tyr Trp Asp
385                 390                 395                 400

Ile Ile Asp Glu His Lys Val Thr Gln Phe Tyr Val Ala Pro Thr Ala
                405                 410                 415

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

Leu Lys Ser Leu Arg Cys Leu Gly Ser Val Gly Glu Pro Ile Ala Ala
        435                 440                 445

Glu Val Trp Glu Trp Tyr Ser Glu Lys Ile Gly Lys Asn Glu Ile Pro
    450                 455                 460

Ile Val Asp Thr Tyr Trp Gln Thr Glu Ser Gly Ser His Leu Val Thr
465                 470                 475                 480

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> 20
<211> 683
<212> PRT
<213> Saccharomyces cerevisiae

<400> 20
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

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

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> 21
<211> 440
<212> PRT
<213> Saccharomyces cerevisiae

<400> 21
Met Leu Ala Val Arg Arg Leu Thr Arg Tyr Thr Phe Leu Lys Arg Thr
1 5 10 15

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

Ser Thr Phe Leu Arg Arg Ser Leu Leu Gln Thr Gln Leu His Ser Lys
35 40 45

Met Thr Ala His Thr Asn Ile Lys Gln His Lys His Cys His Glu Asp
50 55 60

His Pro Ile Arg Arg Ser Asp Ser Ala Val Ser Ile Val His Leu Lys
65 70 75 80

Arg Ala Pro Phe Lys Val Thr Val Ile Gly Ser Gly Asn Trp Gly Thr
85 90 95

Thr Ile Ala Lys Val Ile Ala Glu Asn Thr Glu Leu His Ser His Ile
100 105 110

Phe Glu Pro Glu Val Arg Met Trp Val Phe Asp Glu Lys Ile Gly Asp
115 120 125

Glu Asn Leu Thr Asp Ile Ile Asn Thr Arg His Gln Asn Val Lys Tyr
130 135 140

Leu Pro Asn Ile Asp Leu Pro His Asn Leu Val Ala Asp Pro Asp Leu
145 150 155 160

Leu His Ser Ile Lys Gly Ala Asp Ile Leu Val Phe Asn Ile Pro His
165 170 175

Gln Phe Leu Pro Asn Ile Val Lys Gln Leu Gln Gly His Val Ala Pro

```
            180                 185                 190

His Val Arg Ala Ile Ser Cys Leu Lys Gly Phe Glu Leu Gly Ser Lys
        195                 200                 205

Gly Val Gln Leu Leu Ser Ser Tyr Val Thr Asp Glu Leu Gly Ile Gln
    210                 215                 220

Cys Gly Ala Leu Ser Gly Ala Asn Leu Ala Pro Glu Val Ala Lys Glu
225                 230                 235                 240

His Trp Ser Glu Thr Thr Val Ala Tyr Gln Leu Pro Lys Asp Tyr Gln
                245                 250                 255

Gly Asp Gly Lys Asp Val Asp His Lys Ile Leu Lys Leu Leu Phe His
            260                 265                 270

Arg Pro Tyr Phe His Val Asn Val Ile Asp Asp Val Ala Gly Ile Ser
        275                 280                 285

Ile Ala Gly Ala Leu Lys Asn Val Val Ala Leu Ala Cys Gly Phe Val
    290                 295                 300

Glu Gly Met Gly Trp Gly Asn Asn Ala Ser Ala Ala Ile Gln Arg Leu
305                 310                 315                 320

Gly Leu Gly Glu Ile Ile Lys Phe Gly Arg Met Phe Phe Pro Glu Ser
                325                 330                 335

Lys Val Glu Thr Tyr Tyr Gln Glu Ser Ala Gly Val Ala Asp Leu Ile
            340                 345                 350

Thr Thr Cys Ser Gly Gly Arg Asn Val Lys Val Ala Thr Tyr Met Ala
        355                 360                 365

Lys Thr Gly Lys Ser Ala Leu Glu Ala Glu Lys Glu Leu Leu Asn Gly
    370                 375                 380

Gln Ser Ala Gln Gly Ile Ile Thr Cys Arg Glu Val His Glu Trp Leu
385                 390                 395                 400

Gln Thr Cys Glu Leu Thr Gln Glu Phe Pro Leu Phe Glu Ala Val Tyr
                405                 410                 415

Gln Ile Val Tyr Asn Asn Val Arg Met Glu Asp Leu Pro Glu Met Ile
            420                 425                 430

Glu Glu Leu Asp Ile Asp Asp Glu
        435                 440
```

<210> 22
<211> 391
<212> PRT
<213> Saccharomyces cerevisiae

<400> 22

```
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
        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
```

<210> 23
<211> 2058
<212> ADN
<213> Artificial

<220>
<223> interrupción de GPD2

<400> 23

```
agatcttttg cggcgaggtg ccgatgggtt gctgagggga agagtgttta gcttacggac     60

ctattgccat tgttattccg attaatctat tgttcagcag ctcttctcta ccctgtcatt    120

ctagtatttt ttttttttt ttttggtttt acttttttt cttcttgcct ttttttcttg    180
```

ttacttttt tctagttttt tttccttcca ctaagctttt tccttgattt atccttgggt 240

tcttctttct actcctttag attttttttt tatatattaa tttttaagtt tatgtatttt 300

ggtagattca attctctttc cctttccttt tccttcgctc cccttcctta tcaatgcttg 360

ctgtcagaag attaacaaga tacacattcc ttaaggcctc gtccccgccg ggtcacccgg 420

ccagcgacat ggaggcccag aataccctcc ttgacagtct tgacgtgcgc agctcagggg 480

catgatgtga ctgtcgcccg tacatttagc ccatacatcc ccatgtataa tcatttgcat 540

ccatacattt tgatggccgc acggcgcgaa gcaaaaatta cggctcctcg ctgcagacct 600

gcgagcaggg aaacgctccc ctcacagacg cgttgaattg tccccacgcc gcgcccctgt 660

agagaaatat aaaaggttag gatttgccac tgaggttctt ctttcatata cttcctttta 720

aaatcttgct aggatacagt tctcacatca catccgaaca taaacaacca tgtaaaatga 780

ccactcttga cgacacggct taccggtacc gcaccagtgt cccgggggac gccgaggcca 840

tcgaggcact ggatgggtcc ttcaccaccg acaccgtctt ccgcgtcacc gccaccgggg 900

acggcttcac cctgcgggag gtgccggtgg acccgccccc gaccaaggtg ttccccgacg 960

acgaatcgga cgacgaatcg gacgccgggg aggacggcga cccggactcc cggacgttcg 1020

tcgcgtacgg ggacgacggc gacctggcgg gcttcgtggt cgtctcgtac tccggctgga 1080

accgccggct gaccgtcgag gacatcgagg tcgccccgga gcaccggggg cacggggtcg 1140

ggcgcgcgtt gatggggctc gcgacggagt tcgcccgcga gcggggcgcc gggcacctct 1200

ggctggaggt caccaacgtc aacgcaccgg cgatccacgc gtaccggcgg atggggttca 1260

ccctctgcgg cctggacacc gccctgtacg acggcaccgc ctcggacggc gagcaggcgc 1320

tctacatgag catgccctgc ccctagtact gacaataaaa agattcttgt tttcaagaac 1380

ttgtcatttg tatagttttt ttatattgta gttgttctat tttaatcaaa tgttagcgtg 1440

atttatattt tttttcgcct cgacatcatc tgcccagatg cgaagttaag tgcgcagaaa 1500

gtaatatcat gcgtcaatcg tatgtgaatg ctggtcgcta tactgctgtc gattcgatac 1560

taacgccgcc atccagtgtc gacggatcct aggtgtacag ggcccaaaag ggcgaattct 1620

gcagatatcc atcacactgg cggccgctcg aggatagtct acaacaacgt ccgcatggaa 1680

gacctaccgg agatgattga agagctagac atcgatgacg aatagacact ctccccccc 1740

ctccccctct gatctttcct gttgcctctt tttcccccaa ccaatttatc attatacaca 1800

agttctacaa ctactactag taacattact acagttatta taattttcta ttctcttttt 1860

ctttaagaat ctatcattaa cgttaatttc tatatataca taactaccat tatacacgct 1920

attatcgttt acatatcaca tcaccgttaa tgaaagatac gacaccctgt acactaacac 1980

aattaaataa tcgccataac cttttctgtt atctatagcc cttaaagctg tttcttcgag 2040

ctttttcact gcagatct 2058

<210> 24
<211> 816
<212> ADN
<213> Saccharomyces cerevisiae

<400> 24

ccatatgatc atgtgtcgtc gcacacatat atatatgcct gtatgtgtca gcactaaagt 60

tgcctggcca tccacgctat atatacacgc ctggcggatc tgctcgagga ttgcctacgc 120

gtgggcttga tccaccaacc aacgctcgcc aaatgaactg gcgctttggt cttctgccat 180

cgtccgtaaa ccccggccaa agagaccgga aagatcggtg aaaacatctt gatcttgctc 240

ccgggaattt tagattcagg taggaaattg attacatcaa tactgttacc ctgaatcata 300

ttcgacgatg tcgtctcaca cggaaatata attcatttct tggttttcca aaaaaatttt 360

catttttttt cacttttttg tttcgtcctc cttttttttt ttttttttgt attttttttc 420

ctgtgttcac cttttttttt ttcagttgac atctttctgc attcttttct gtgttttttt 480

ttttttttgt cgtttttcca ttgttcgttc gttgcctgtt ttttcgccct attgttctcg 540

agcctaaaaa ttttttcctt tcctgctttc ctttcttcgt tcaaagtttc ctattccatt 600

gttctctttg gtaaactcat tgttgtcgga actcagatat attcaggtca atttactgta 660

cttcaattga ctttttttctt gaaatttcaa cttgcctttt caacttgttc ttctttttta 720

atcttattct acactttagt tcccttacct tgttcctaat tattgtctag caaaaagaaa 780

acatacacct atttcattca cacactgcag aaaatg 816

<210> 25
<211> 954
<212> ADN
<213> Artificial

<220>
<223> fragmento de PCR de GCY1

<400> 25
ctgcagaaaa tgcctgctac tttacatgat tctacgaaaa tcctttctct aaatactgga 60

gcccaaatcc ctcaaatagg tttaggtacg tggcagtcga aagagaacga tgcttataag 120

gctgttttaa ccgctttgaa agatggctac cgacacattg atactgctgc tatttaccgt 180

aatgaagacc aagtcggtca agccatcaag gattcaggtg ttcctcggga agaaatcttt 240

gttactacaa agttatggtg tacacaacac cacgaacctg aagtagcgct ggatcaatca 300

ctaaagaggt taggattgga ctacgtagac ttatatttga tgcattggcc tgccagatta 360

gatccagcct acatcaaaaa tgaagacatc ttgagtgtgc caacaaagaa ggatggttct 420

cgtgcagtgg atatcaccaa ttggaatttc atcaaaacct gggaattaat gcaggaacta 480

ccaaagactg gtaaaactaa ggccgttgga gtctccaact tttctataaa taacctgaaa 540

gatctattag catctcaagg taataagctt acgccagctg ctaaccaagt cgaaatacat 600

ccattactac ctcaagacga attgattaat ttttgtaaaa gtaaaggcat tgtggttgaa 660

gcttattctc cgttaggtag taccgatgct ccactattga aggaaccggt tatccttgaa 720

attgcgaaga aaaataacgt tcaacccgga cacgttgtta ttagctggca cgtccaaaga 780

ggttatgttg tcttgccaaa atctgtgaat cccgatcgaa tcaaaacgaa caggaaaata 840

tttactttgt ctactgagga ctttgaagct atcaataaca tatcgaagga aaagggcgaa 900

aaaagggttg tacatccaaa ttggtctcct ttcgaagtat tcaagtaact taag 954

<210> 26

<211> 2986
<212> ADN
<213> Artificial

<220>
<223> fragmento de PCR de DAK1

<400> 26

ctcgagtacc ggtttcttct tcagattccc tcatggagaa agtgcggcag atgtatatga 60

cagagtcgcc agtttccaag agactttatt caggcacttc catgataggc aagagagaag 120

acccagagat gttgttgtcc tagttacaca tggtatttat tccagagtat tcctgatgaa 180

atggtttaga tggacatacg aagagtttga atcgtttacc aatgttccta acgggagcgt 240

aatggtgatg gaactggacg aatccatcaa tagatacgtc ctgaggaccg tgctacccaa 300

atggactgat tgtgagggag acctaactac atagtgttta aagattacgg atatttaact 360

tacttagaat aatgccattt ttttgagtta taataatcct acgttagtgt gagcgggatt 420

taaactgtga ggaccttaat acattcagac acttctgcgg tatcacccta cttattccct 480

tcgagattat atctaggaac ccatcaggtt ggtggaagat tacccgttct aagacttttc 540

agcttcctct attgatgtta cacctggaca ccccttttct ggcatccagt ttttaatctt 600

cagtggcatg tgagattctc cgaaattaat taaagcaatc acacaattct ctcggatacc 660

acctcggttg aaactgacag gtggtttgtt acgcatgcta atgcaaagga gcctatatac 720

ctttggctcg gctgctgtaa cagggaatat aaagggcagc ataatttagg agtttagtga 780

acttgcaaca tttactattt tcccttctta cgtaaatatt tttctttttа attctaaatc 840

aatctttttc aatttttttgt ttgtattctt ttcttgctta aatctataac tacaaaaaac 900

acatacataa atctagaaaa tgtccgctaa atcgtttgaa gtcacagatc cagtcaattc 960

aagtctcaaa gggtttgccc ttgctaaccc ctccattacg ctggtccctg aagaaaaaat 1020

tctcttcaga aagaccgatt ccgacaagat cgcattaatt tctggtggtg gtagtggaca 1080

tgaacctaca cacgccggtt tcattggtaa gggtatgttg agtggcgccg tggttggcga 1140

aatttttgca tcccccttcaa caaaacagat tttaaatgca atccgtttag tcaatgaaaa 1200

tgcgtctggc gttttattga ttgtgaagaa ctacacaggt gatgttttgc attttggtct 1260

gtccgctgag agagcaagag ccttgggtat taactgccgc gttgctgtca taggtgatga 1320

tgttgcagtt ggcagagaaa agggtggtat ggttggtaga agagcattgg caggtaccgt 1380

tttggttcat aagattgtag gtgccttcgc agaagaatat tctagtaagt atggcttaga 1440

cggtacagct aaagtggcta aaattatcaa cgacaatttg gtgaccattg gatcttcttt 1500

agaccattgt aaagttcctg gcaggaaatt cgaaagtgaa ttaaacgaaa aacaaatgga 1560

attgggtatg ggtattcata acgaacctgg tgtgaaagtt ttagacccta ttccttctac 1620

cgaagacttg atctccaagt atatgctacc aaaactattg gatccaaacg ataaggatag 1680

agcttttgta aagtttgatg aagatgatga agttgtcttg ttagttaaca atctcggcgg 1740

tgtttctaat tttgttatta gttctatcac ttccaaaact acggatttct taaaggaaaa 1800

ttacaacata accccggttc aaacaattgc tggcacattg atgacctcct tcaatggtaa 1860

tgggttcagt atcacattac taaacgccac taaggctaca aaggctttgc aatctgattt 1920

tgaggagatc aaatcagtac tagacttgtt gaacgcattt acgaacgcac cgggctggcc 1980

aattgcagat tttgaaaaga cttctgcccc atctgttaac gatgacttgt tacataatga 2040

agtaacagca aaggccgtcg gtacctatga ctttgacaag tttgctgagt ggatgaagag 2100

tggtgctgaa caagttatca agagcgaacc gcacattacg gaactagaca atcaagttgg 2160

tgatggtgat tgtggttaca ctttagtggc aggagttaaa ggcatcaccg aaaaccttga 2220

caagctgtcg aaggactcat tatctcaggc ggttgcccaa atttcagatt tcattgaagg 2280

ctcaatggga ggtacttctg gtggtttata ttctattctt ttgtcgggtt tttcacacgg 2340

attaattcag gtttgtaaat caaaggatga acccgtcact aaggaaattg tggctaagtc 2400

actcggaatt gcattggata ctttatacaa atatacaaag gcaaggaagg gatcatccac 2460

catgattgat gctttagaac cattcgttaa agaatttact gcatctaagg atttcaataa 2520

ggcggtaaaa gctgcagagg aaggtgctaa atccactgct acattcgagg ccaaatttgg 2580

cagagcttcg tatgtcggcg attcatctca agtagaagat cctggtgcag taggcctatg 2640

tgagtttttg aaggggttc aaagcgcctt gtaagtcgag acaaatcgct cttaaatata 2700

tacctaaaga acattaaagc tatattataa gcaaagatac gtaaattttg cttatattat 2760

tatacacata tcatatttct atatttttaa gatttggtta tataatgtac gtaatgcaaa 2820

ggaaataaat tttatacatt attgaacagc gtccaagtaa ctacattatg tgcactaata 2880

gtttagcgtc gtgaagactt tattgtgtcg cgaaaagtaa aaattttaaa aattagagca 2940

ccttgaactt gcgaaaaagg ttctcatcaa ctgtttaaaa cgtacg 2986

<210> 27
<211> 2651
<212> ADN
<213> Artificial

<220>
<223> fragmento de GUP1 de PCR

<400> 27

```
ggtacctagg accggtttat cattatcaat actgccattt caaagaatac gtaaataatt     60
aatagtagtg attttcctaa ctttatttag tcaaaaaatt agccttttaa ttctgctgta    120
acccgtacat gcccaaaata gggggcgggt tacacagaat atataacatc gtaggtgtct    180
gggtgaacag tttattcctg gcatccacta aatataatgg agcccgcttt ttaagctggc    240
atccagaaaa aaaaagaatc ccagcaccaa aatattgttt tcttcaccaa ccatcagttc    300
ataggtccat tctcttagcg caactacaga gaacaggggc acaaacaggc aaaaaacggg    360
cacaacctca atggagtgat gcaacctgcc tggagtaaat gatgacacaa ggcaattgac    420
ccacgcatgt atctatctca ttttcttaca ccttctatta ccttctgctc tctctgattt    480
ggaaaaagct gaaaaaaaag gttgaaacca gttccctgaa attattcccc tacttgacta    540
ataagtatat aaagacggta ggtattgatt gtaattctgt aaatctattt cttaaacttc    600
ttaaattcta cttttatagt tagtcttttt tttagtttta aaacaccaag aacttagttt    660
cgaataaaca cacataaaga attcgaaaat gtcgctgatc agcatcctgt ctcccctaat    720
tacttccgag ggcttagatt caagaatcaa accttcacca aaaaaggatg cctctactac    780
cactaagcca tcactatgga aaactactga gttcaaattc tactacattg catttctggt    840
cgtggttccc ttgatgttct atgctgggtt acaagctagt tcgcccgaaa atccaaacta    900
tgcaagatac gaacgtctcc tatctcaagg ttggttattt ggcagaaaag tagacaatag    960
tgattctcaa tataggtttt tcagggacaa ttttgcgcta ttgtcagttt taatgctagt   1020
ccacacttct ataaaacgca ttgtacttta ttcaacaaat atcactaaat tgaggtttga   1080
tctgatattt ggtttgatct ttttagtggc cgctcatggt gtcaattcga taagaatttt   1140
agcccatatg ctaattttat atgccatcgc ccatgtacta aagaacttta gaagaatagc   1200
caccatcagc atttggattt atggtatttc tacgcttttt attaacgaca acttcagagc   1260
atatccattt ggtaatattt gctctttttt aagcccattg gaccattggt atagaggtat   1320
cattccaaga tgggatgtct ttttcaattt tactcttttg agagtcttaa gttacaactt   1380
ggacttctta gagaggtggg agaatttaca aaagaagaaa agtccatcct atgaatcaaa   1440
agaagctaaa tcagccattt tgctcaatga acgtgctaga ttaactgctg cacaccccat   1500
acaggactac agcttaatga attatattgc atatgttact tacacgccac ttttcattgc   1560
cggccccatt ataacattca atgattatgt ttaccaatcg aaacatacct tgccatcaat   1620
```

aaatttcaaa ttcatttttt actatgcggt gagattcgtt attgctctct tatctatgga 1680

gttcatttta cactttctcc acgttgtggc aatctcaaaa accaaagcgt gggaaaatga 1740

cacacctttc cagatttcca tgattggctt atttaatttg aatattattt ggctaaaact 1800

actgattccg tggaggctgt ttaggctgtg ggctttgcta gacggaatcg atacacctga 1860

aaatatgatc aggtgtgttg ataacaatta cagttcacta gcattctgga gagcttggca 1920

tagaagctac aataagtggg ttgtccgtta catatatatt cctctaggtg gttcaaaaaa 1980

tagagttttg acatcactag cagtcttttc cttcgtagct atatggcatg acatcgaact 2040

aaagttatta ttatggggtt ggctaatagt tttgttcctc ttaccagaaa tttttgctac 2100

ccaaattttc tctcattata ccgacgcagt ctggtacaga cacgtttgcg ctgtcggtgc 2160

tgttttcaac atatgggtta tgatgatcgc taatcttttt ggattctgct tgggctctga 2220

cggtactaaa aaattactaa gcgatatgtt ctgtaccgta tctggtttca aatttgtaat 2280

tttggcaagc gttagtttat tcatcgcagt acaaataatg tttgaaatca gagaagaaga 2340

aaagaggcac ggaatttacc taaaatgctg aggatcccct tttcctttgt cgatatcatg 2400

taattagtta tgtcacgctt acattcacgc cctcctccca catccgctct aaccgaaaag 2460

gaaggagtta gacaacctga agtctaggtc cctatttatt ttttttaata gttatgttag 2520

tattaagaac gttatttata tttcaaattt ttctttttttt tctgtacaaa cgcgtgtacg 2580

catgtaacat tatactgaaa accttgcttg agaaggtttt gggacgctcg aaggcttcct 2640

aggctcgagt t 2651

<210> 28
<211> 3037
<212> ADN
<213> Artificial

<220>
<223> fragmento de FPS1 de PCR

<400> 28
aagagctccg gactagtcgt acgaattcta tccttttgtt gtttccgggt gtacaatatg 60

gacttcctct tttctggcaa ccaaacccat acatcgggat tcctataata ccttcgttgg 120

tctccctaac atgtaggtgg cggaggggag atatacaata gaacagatac cagacaagac 180

ataatgggct aaacaagact acaccaatta cactgcctca ttgatggtgg tacataacga 240

actaatactg tagccctaga cttgatagcc atcatcatat cgaagtttca ctacccttttt 300

tccatttgcc atctattgaa gtaataatag gcgcatgcaa cttcttttct ttttttttct 360

tttctctctc ccccgttgtt gtctcaccat atccgcaatg acaaaaaaat gatggaagac 420

actaaaggaa aaaattaacg acaaagacag caccaacaga tgtcgttgtt ccagagctga 480

tgaggggtat ctcgaagcac acgaaacttt ttccttcctt cattcacgca cactactctc 540

taatgagcaa cggtatacgg ccttccttcc agttacttga atttgaaata aaaaaagttt 600

gctgtcttgc tatcaagtat aaatagacct gcaattatta atcttttgtt tcctcgtcat 660

tgttctcgtt ccctttcttc cttgtttctt tttctgcaca atatttcaag ctataccaag 720

catacaatca actccagctg cattaaaatg agtaatcctc aaaaagctct aaacgacttt 780

ctgtccagtg aatctgttca tacacatgat agttctagga aacaatctaa taagcagtca 840

tccgacgaag gacgctcttc atcacaacct tcacatcatc actctggtgg tactaacaac 900

aataataaca ataataataa taataataac agtaacaaca acaacaacgg caacgatggg 960

ggaaatgatg acgactatga ttatgaaatg caagattata gaccttctcc gcaaagtgcg 1020

cggcctactc ccacgtatgt tccacaatat tctgtagaaa gtgggactgc tttcccgatt 1080

caagaggtta ttcctagcgc 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 ggtgcccaaa ctaatatgga aagcaatgaa 1380

tcaccacgta acgtccccat tatggtgaag ccaaagacat tataccagaa ccctcaaaca 1440

cctacagtct tgccctccac ataccatcca attaataaat ggtcttccgt caaaaacact 1500

tatttgaagg aatttttagc cgagtttatg ggaacaatgg ttatgattat tttcggtagt 1560

gctgttgttt gtcaggtcaa tgttgctggg aaaatacagc aggacaattt caacgtggct 1620

ttggataacc ttaacgttac cgggtcttct gcagaaacga tagacgctat gaagagttta 1680

acatccttgg tttcatccgt tgcgggcggt acctttgatg atgtggcatt gggctgggct 1740

gctgccgtgg tgatgggcta tttctgcgct ggtggtagtg ccatctcagg tgctcatttg 1800

aatccgtcta ttacattagc caatttggtg tatagaggtt ttcccctgaa gaaagttcct 1860

tattactttg ctggacaatt gatcggtgcc ttcacaggcg ctttgatctt gtttatttgg 1920

tacaaaaggg tgttacaaga ggcatatagc gattggtgga tgaatgaaag tgttgcggga 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 tcttgcacta 2220

tatgcagttg gatttgatca taaaatgctt tgggtgcatc atcatcattt cttttgggtt 2280

cccatggtag gcccatttat tggtgcgtta atgggggggt tggtttacga tgtctgtatt 2340

tatcagggtc atgaatctcc agtcaactgg tctttaccag tttataagga aatgattatg 2400

agagcctggt ttagaaggcc tggttggaag aagagaaata gagcaagaag aacatcggac 2460

ctgagtgact tctcatacaa taacgatgat gatgaggaat ttggagaaag aatggctctt 2520

caaaagacaa agaccaagtc atctatttca gacaacgaaa atgaagcagg agaaaagaaa 2580

gtgcaattta aatctgttca gcgcggcaaa agaacgtttg gtggtatacc aacaattctt 2640

gaagaagaag attccattga aactgcttcg ctaggtgcga cgacgactga ttctattggg 2700

ttatccgaca catcatcaga agattcgcat tatggtaatg ctaagaaggt aacatgagga 2760

tccccttttc ctttgtcgat atcatgtaat tagttatgtc acgcttacat tcacgccctc 2820

ctcccacatc cgctctaacc gaaaaggaag gagttagaca acctgaagtc taggtcccta 2880

tttatttttt ttaatagtta tgttagtatt aagaacgtta tttatatttc aaatttttct 2940

ttttttttctg tacaaacgcg tgtacgcatg taacattata ctgaaaacct tgcttgagaa 3000

ggttttggga cgctcgaagg cttcctaggc tcgagtt 3037

<210> 29
<211> 12954
<212> ADN
<213> Artificial

<220>
<223> plásmido pRN605 para la sobre-expresión de GCY1, DAK1 y GUP1

<400> 29
tacgccaagc tcggaattaa ccctcactaa agggaacaaa agctgggtac cgggcccccc 60

ctcgagccta ggaagccttc gagcgtccca aaaccttctc aagcaaggtt ttcagtataa 120

tgttacatgc gtacacgcgt ttgtacagaa aaaaaagaaa aatttgaaat ataaataacg 180

ttcttaatac taacataact attaaaaaaa ataaataggg acctagactt caggttgtct 240

aactccttcc ttttcggtta gagcggatgt gggaggaggg cgtgaatgta agcgtgacat 300

aactaattac atgatatcga caaaggaaaa ggggatcctc agcattttag gtaaattccg 360

tgcctctttt cttcttctct gatttcaaac attatttgta ctgcgatgaa taaactaacg 420

cttgccaaaa ttacaaattt gaaaccagat acggtacaga acatatcgct tagtaatttt 480

ttagtaccgt cagagcccaa gcagaatcca aaaagattag cgatcatcat aacccatatg 540

ttgaaaacag caccgacagc gcaaacgtgt ctgtaccaga ctgcgtcggt ataatgagag 600

aaaatttggg tagcaaaaat ttctggtaag aggaacaaaa ctattagcca accccataat 660

aataacttta gttcgatgtc atgccatata gctacgaagg aaaagactgc tagtgatgtc 720

aaaactctat tttttgaacc acctagagga atatatatgt aacggacaac ccacttattg 780

tagcttctat gccaagctct ccagaatgct agtgaactgt aattgttatc aacacacctg 840

atcatatttt caggtgtatc gattccgtct agcaaagccc acagcctaaa cagcctccac 900

ggaatcagta gttttagcca aataatattc aaattaaata agccaatcat ggaaatctgg 960

aaaggtgtgt cattttccca cgctttggtt tttgagattg ccacaacgtg gagaaagtgt 1020

aaaatgaact ccatagataa gagagcaata acgaatctca ccgcatagta aaaaatgaat 1080

ttgaaattta ttgatggcaa ggtatgtttc gattggtaaa cataatcatt gaatgttata 1140

atggggccgg caatgaaaag tggcgtgtaa gtaacatatg caatataatt cattaagctg 1200

tagtcctgta tggggtgtgc agcagttaat ctagcacgtt cattgagcaa aatggctgat 1260

ttagcttctt ttgattcata ggatggactt ttcttctttt gtaaattctc ccacctctct 1320

aagaagtcca agttgtaact taagactctc aaaagagtaa aattgaaaaa gacatcccat 1380

cttggaatga tacctctata ccaatggtcc aatgggctta aaaaagagca aatattacca 1440

aatggatatg ctctgaagtt gtcgttaata aaaagcgtag aaataccata aatccaaatg 1500

ctgatggtgg ctattcttct aaagttcttt agtacatggg cgatggcata taaaattagc 1560

atatgggcta aaattcttat cgaattgaca ccatgagcgg ccactaaaaa gatcaaacca 1620

aatatcagat caaacctcaa tttagtgata tttgttgaat aaagtacaat gcgttttata 1680

gaagtgtgga ctagcattaa aactgacaat agcgcaaaat tgtccctgaa aaacctatat 1740

tgagaatcac tattgtctac ttttctgcca aataaccaac cttgagatag gagacgttcg 1800

tatcttgcat agtttggatt ttcgggcgaa ctagcttgta acccagcata gaacatcaag 1860

ggaaccacga ccagaaatgc aatgtagtag aatttgaact cagtagtttt ccatagtgat 1920

ggcttagtgg tagtagaggc atccttttt ggtgaaggtt tgattcttga atctaagccc 1980

tcggaagtaa ttaggggaga caggatgctg atcagcgaca ttttgaattc tttatgtgtg 2040

tttattcgaa actaagttct tggtgtttta aaactaaaaa aaagactaac tataaaagta 2100

gaatttaaga agtttaagaa atagatttac agaattacaa tcaataccta ccgtctttat 2160

atacttatta gtcaagtagg ggaataattt cagggaactg gtttcaacct ttttttcag 2220

ctttttccaa atcagagaga gcagaaggta atagaaggtg taagaaaatg agatagatac 2280

atgcgtgggt caattgcctt gtgtcatcat ttactccagg caggttgcat cactccattg 2340

aggttgtgcc cgtttttttgc ctgtttgtgc ccctgttctc tgtagttgcg ctaagagaat 2400

ggacctatga actgatggtt ggtgaagaaa acaatatttt ggtgctggga ttctttttt 2460

ttctggatgc cagcttaaaa agcgggctcc attatattta gtggatgcca ggaataaact 2520

gttcacccag acacctacga tgttatatat tctgtgtaac ccgcccccta ttttgggcat 2580

gtacgggtta cagcagaatt aaaaggctaa ttttttgact aaataaagtt aggaaaatca 2640

ctactattaa ttatttacgt attctttgaa atggcagtat tgataatgat aaaccggttt 2700

cttcttcaga ttccctcatg gagaaagtgc ggcagatgta tatgacagag tcgccagttt 2760

ccaagagact ttattcaggc acttccatga taggcaagag agaagaccca gagatgttgt 2820

tgtcctagtt acacatggta tttattccag agtattcctg atgaaatggt ttagatggac 2880

atacgaagag tttgaatcgt ttaccaatgt tcctaacggg agcgtaatgg tgatggaact 2940

ggacgaatcc atcaatagat acgtcctgag gaccgtgcta cccaaatgga ctgattgtga 3000

gggagaccta actacatagt gtttaaagat tacggatatt taacttactt agaataatgc 3060

catttttttg agttataata atcctacgtt agtgtgagcg ggatttaaac tgtgaggacc 3120

ttaatacatt cagacacttc tgcggtatca ccctacttat tcccttcgag attatatcta 3180

ggaacccatc aggttggtgg aagattaccc gttctaagac ttttcagctt cctctattga 3240

tgttacacct ggacacccct tttctggcat ccagttttta atcttcagtg gcatgtgaga 3300

ttctccgaaa ttaattaaag caatcacaca attctctcgg ataccacctc ggttgaaact 3360

gacaggtggt ttgttacgca tgctaatgca aaggagccta tatacctttg gctcggctgc 3420

tgtaacaggg aatataaagg gcagcataat ttaggagttt agtgaacttg caacatttac 3480

tattttccct tcttacgtaa atatttttct ttttaattct aaatcaatct ttttcaattt 3540

tttgtttgta ttcttttctt gcttaaatct ataactacaa aaaacacata cataaatcta 3600

gaaaatgtcc gctaaatcgt ttgaagtcac agatccagtc aattcaagtc tcaaagggtt 3660

tgcccttgct aacccctcca ttacgctggt ccctgaagaa aaaattctct tcagaaagac 3720

cgattccgac aagatcgcat taatttctgg tggtggtagt ggacatgaac ctacacacgc 3780

cggtttcatt ggtaagggta tgttgagtgg cgccgtggtt ggcgaaattt ttgcatcccc 3840

ttcaacaaaa cagattttaa atgcaatccg tttagtcaat gaaaatgcgt ctggcgtttt 3900

attgattgtg aagaactaca caggtgatgt tttgcatttt ggtctgtccg ctgagagagc 3960

aagagccttg ggtattaact gccgcgttgc tgtcataggt gatgatgttg cagttggcag 4020

agaaaagggt ggtatggttg gtagaagagc attggcaggt accgttttgg ttcataagat 4080

tgtaggtgcc ttcgcagaag aatattctag taagtatggc ttagacggta cagctaaagt 4140

ggctaaaatt atcaacgaca atttggtgac cattggatct tctttagacc attgtaaagt 4200

tcctggcagg aaattcgaaa gtgaattaaa cgaaaaacaa atggaattgg gtatgggtat 4260

tcataacgaa cctggtgtga aagttttaga ccctattcct tctaccgaag acttgatctc 4320

caagtatatg ctaccaaaac tattggatcc aaacgataag gatagagctt ttgtaaagtt 4380

tgatgaagat gatgaagttg tcttgttagt taacaatctc ggcggtgttt ctaattttgt 4440

tattagttct atcacttcca aaactacgga tttcttaaag gaaaattaca acataacccc 4500

ggttcaaaca attgctggca cattgatgac ctccttcaat ggtaatgggt tcagtatcac 4560

attactaaac gccactaagg ctacaaaggc tttgcaatct gattttgagg agatcaaatc 4620

agtactagac ttgttgaacg catttacgaa cgcaccgggc tggccaattg cagattttga 4680

aaagacttct gccccatctg ttaacgatga cttgttacat aatgaagtaa cagcaaaggc 4740

cgtcggtacc tatgactttg acaagtttgc tgagtggatg aagagtggtg ctgaacaagt 4800

tatcaagagc gaaccgcaca ttacggaact agacaatcaa gttggtgatg gtgattgtgg 4860

ttacacttta gtggcaggag ttaaaggcat caccgaaaac cttgacaagc tgtcgaagga 4920

ctcattatct caggcggttg cccaaatttc agatttcatt gaaggctcaa tgggaggtac 4980

ttctggtggt ttatattcta ttcttttgtc gggtttttca cacggattaa ttcaggtttg 5040

taaatcaaag gatgaacccg tcactaagga aattgtggct aagtcactcg gaattgcatt 5100

ggatacttta tacaaatata caaaggcaag gaagggatca tccaccatga ttgatgcttt 5160

agaaccattc gttaaagaat ttactgcatc taaggatttc aataaggcgg taaaagctgc 5220

agaggaaggt gctaaatcca ctgctacatt cgaggccaaa tttggcagag cttcgtatgt 5280

cggcgattca tctcaagtag aagatcctgg tgcagtaggc ctatgtgagt ttttgaaggg 5340

ggttcaaagc gccttgtaag tcgagacaaa tcgctcttaa atatatacct aaagaacatt 5400

aaagctatat tataagcaaa gatacgtaaa ttttgcttat attattatac acatatcata 5460

tttctatatt tttaagattt ggttatataa tgtacgtaat gcaaaggaaa taaattttat 5520

acattattga acagcgtcca agtaactaca ttatgtgcac taatagttta gcgtcgtgaa 5580

gactttattg tgtcgcgaaa agtaaaaatt ttaaaaatta gagcaccttg aacttgcgaa 5640

aaaggttctc atcaactgtt taaaacgtac gtgtggaaga acgattacaa caggtgttgt 5700

cctctgagga cataaaatac acaccgagat tcatcaactc attgctggag ttagcatatc 5760

tacaattggg tgaaatgggg agcgatttgc aggcatttgc tcggcatgcc ggtagaggtg 5820

tggtcaataa gagcgacctc atgctatacc tgagaaagca acctgaccta caggaaagag 5880

ttactcaaga ataagaattt tcgttttaaa acctaagagt cactttaaaa tttgtataca 5940

cttatttttt ttataactta tttaataata aaaatcataa atcataagaa attcgcgcgc 6000

ttacttgaat acttcgaaag gagaccaatt tggatgtaca acccttttt cgccctttc 6060

cttcgatatg ttattgatag cttcaaagtc ctcagtagac aaagtaaata ttttcctgtt 6120

cgttttgatt cgatcgggat tcacagattt tggcaagaca acataacctc tttggacgtg 6180

ccagctaata acaacgtgtc cgggttgaac gttatttttc ttcgcaattt caaggataac 6240

cggttccttc aatagtggag catcggtact acctaacgga gaataagctt caaccacaat 6300

gcctttactt ttacaaaaat taatcaattc gtcttgaggt agtaatggat gtatttcgac 6360

ttggttagca gctggcgtaa gcttattacc ttgagatgct aatagatctt tcaggttatt 6420

tatagaaaag ttggagactc caacggcctt agttttacca gtctttggta gttcctgcat 6480

taattcccag gttttgatga aattccaatt ggtgatatcc actgcacgag aaccatcctt 6540

ctttgttggc acactcaaga tgtcttcatt tttgatgtag gctggatcta atctggcagg 6600

ccaatgcatc aaatataagt ctacgtagtc caatcctaac ctctttagtg attgatccag 6660

cgctacttca ggttcgtggt gttgtgtaca ccataacttt gtagtaacaa agatttcttc 6720

ccgaggaaca cctgaatcct tgatggcttg accgacttgg tcttcattac ggtaaatagc 6780

agcagtatca atgtgtcggt agccatcttt caaagcggtt aaaacagcct tataagcatc 6840

gttctctttc gactgccacg tacctaaacc tatttgaggg atttgggctc cagtatttag 6900

agaaaggatt ttcgtagaat catgtaaagt agcaggcatt ttctgcagtt aattcagtaa 6960

attttcgatc ttgggaagaa aaaagcagta agcgtgaaaa atctaaaagc tgatgtagta 7020

gaagatccta ttctttaaca aagattgacc ttttcttttt cttcttggtt tgagtagaaa 7080

ggggaaggaa gaatacaaga gagaggaaaa aaaggaagat aaaaagagag cgtgatataa 7140

atgaatatat attaaacaag agagattggg aaggaaagga tcaaacaaac ccaaaaatat 7200

ttcaaaaagg agagagagag gcgagtttgg tttcaaaacg gtttatttat ttatgcaaga 7260

ggacgtggaa gaaaaagaag aaggaagaaa aaaatttgaa agaaaaaaac gcgtggcggg 7320

taaagaagaa aatggaaaat agaggccggg tgacagagaa atattgaggg ttaattggaa 7380

aatatgttag ggtgaggcat atgtttttaa gggttttgag gatccgataa ggaagaatgt 7440

aggttaaatg ttgtgcatta attgctgtgg cagcttaccc gcttccccac acatttacta 7500

gttctagagc ggccgccacc gcggtggagc tccaattcgc cctatagtga gtcgtattac 7560

aattcactgg ccgtcgtttt acaacgtcgt gactgggaaa accctggcgt tacccaactt 7620

aatcgccttg cagcacatcc cccttcgcc agctggcgta atagcgaaga ggcccgcacc 7680

gatcgccctt cccaacagtt gcgcagcctg aatggcgaat ggcgcgacgc gccctgtagc 7740

ggcgcattaa gcgcggcggg tgtggtggtt acgcgcagcg tgaccgctac acttgccagc 7800

gccctagcgc ccgctccttt cgctttcttc ccttcctttc tcgccacgtt cgccggtagt 7860

gttagacctg aacaaggttt actaaaaatc cgtaaagaac ttcaattgta cgccaactta 7920

aggcctcgtc cccgccgggt cacccggcca gcgacatgga ggcccagaat accctccttg 7980

acagtcttga cgtgcgcagc tcaggggcat gatgtgactg tcgcccgtac atttagccca 8040

tacatcccca tgtataatca tttgcatcca tacattttga tggccgcacg gcgcgaagca 8100

aaaattacgg ctcctcgctg cagacctgcg agcagggaaa cgctcccctc acagacgcgt 8160

gaattgtccc cacgccgcgc ccctgtagag aaatataaaa ggttaggatt tgccactgag 8220

gttcttcttt catatacttc cttttaaaat cttgctagga tacagttctc acatcacatc 8280

cgaacataaa caaccatggg taaaaagcct gaactcaccg cgacgtctgt cgagaagttt 8340

ctgatcgaaa agttcgacag cgtctccgac ctgatgcagc tctcggaggg cgaagaatct 8400

cgtgctttca gcttcgatgt aggagggcgt ggatatgtcc tgcgggtaaa tagctgcgcc 8460

gatggtttct acaaagatcg ttatgtttat cggcactttg catcggccgc gctcccgatt 8520

ccggaagtgc ttgacattgg ggaattcagc gagagcctga cctattgcat ctcccgccgt 8580

gcacagggtg tcacgttgca agacctgcct gaaaccgaac tgcccgctgt tctgcagccg 8640

gtcgcggagg ccatggatgc gatcgctgcg gccgatctta gccagacgag cgggttcggc 8700

ccattcggac cgcaaggaat cggtcaatac actacatggc gtgatttcat atgcgcgatt 8760

gctgatcccc atgtgtatca ctggcaaact gtgatggacg acaccgtcag tgcgtccgtc 8820

gcgcaggctc tcgatgagct gatgctttgg gccgaggact gccccgaagt ccggcacctc 8880

gtgcacgcgg atttcggctc caacaatgtc ctgacggaca atggccgcat aacagcggtc 8940

attgactgga gcgaggcgat gttcggggat tcccaatacg aggtcgccaa catcttcttc 9000

tggaggccgt ggttggcttg tatggagcag cagacgcgct acttcgagcg gaggcatccg 9060

gagcttgcag gatcgccgcg gctccgggcg tatatgctcc gcattggtct tgaccaactc 9120

tatcagagct tggttgacgg caatttcgat gatgcagctt gggcgcaggg tcgatgcgac 9180

gcaatcgtcc gatccggagc cgggactgtc gggcgtacac aaatcgcccg cagaagcgcg 9240

gccgtctgga ccgatggctg tgtagaagta ctcgccgata gtggaaaccg acgccccagc 9300

actcgtccga gggcaaagga ataatcagta ctgacaataa aaagattctt gttttcaaga 9360

acttgtcatt tgtatagttt ttttatattg tagttgttct attttaatca aatgttagcg 9420

tgatttatat ttttttcgc ctcgacatca tctgcccaga tgcgaagtta agtgcgcaga 9480

aagtaatatc atgcgtcaat cgtatgtgaa tgctggtcgc tatactgctg tcgattcgat 9540

actaacgccg ccatccagtg tcgacggatc ctaggtgtac ataaacttta taaatgaaat 9600

tcataataga aacgacacga aattacaaaa tggaatatgt tcatagggta gacgaaacta 9660

tatacgcaat ctacatacat ttatcaagaa ggagaaaaag gaggatagta aaggaataca 9720

ggtaagcaaa ttgatactaa tggctcaacg tgataaggaa aaagaattgc actttaacat 9780

taatattgac aaggaggagg gcaccacaca aaaagttagg tgtaacagaa aatcatgaaa 9840

ctacgattcc taatttgata ttggaggatt ttctctaaaa aaaaaaaaat acaacaaata 9900

aaaaacactc aatgacctga ccatttgatg gagtttaagt caataccttc ttgaaccatt 9960

tcccataatg gtgaaagttc cctcaagaat tttactctgt cagaaacggc cttacgacgt 10020

agtcgatatg gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagcccc 10080

gacacccgcc aacacccgct gacgcgccct gacgggcttg tctgctcccg gcatccgctt 10140

acagacaagc tgtgaccgtc tccgggagct gcatgtgtca gaggttttca ccgtcatcac 10200

cgaaacgcgc gagacgaaag ggcctcgtga tacgcctatt tttataggtt aatgtcatga 10260

taataatggt ttcttaggac ggatcgcttg cctgtaactt acacgcgcct cgtatctttt 10320

aatgatggaa taatttggga atttactctg tgtttattta tttttatgtt ttgtatttgg 10380

attttagaaa gtaaataaag aaggtagaag agttacggaa tgaagaaaaa aaaataaaca 10440

aaggtttaaa aaatttcaac aaaaagcgta ctttacatat atatttatta gacaagaaaa 10500

gcagattaaa tagatataca ttcgattaac gataagtaaa atgtaaaatc acaggatttt 10560

cgtgtgtggt cttctacaca gacaagatga aacaattcgg cattaatacc tgagagcagg 10620

aagagcaaga taaaaggtag tatttgttgg cgatcccccc agagtctttt acatcttcgg 10680

aaaacaaaaa ctatttttc tttaatttct ttttttactt tctattttta atttatatat 10740

ttatattaaa aaatttaaat tataattatt tttatagcac gtgatgaaaa ggacccaggt 10800

ggcacttttc ggggaaatgt gcgcggaacc cctatttgtt tatttttcta aatacattca 10860

aatatgtatc cgctcatgag acaataaccc tgataaatgc ttcaataata ttgaaaaagg 10920

aagagtatga gtattcaaca tttccgtgtc gcccttattc ccttttttgc ggcattttgc 10980

cttcctgttt ttgctcaccc agaaacgctg gtgaaagtaa aagatgctga agatcagttg 11040

ggtgcacgag tgggttacat cgaactggat ctcaacagcg gtaagatcct tgagagtttt 11100

cgccccgaag aacgttttcc aatgatgagc acttttaaag ttctgctatg tggcgcggta 11160

ttatcccgta ttgacgccgg gcaagagcaa ctcggtcgcc gcatacacta ttctcagaat 11220

gacttggttg agtactcacc agtcacagaa aagcatctta cggatggcat gacagtaaga 11280

gaattatgca gtgctgccat aaccatgagt gataacactg cggccaactt acttctgaca 11340

acgatcggag gaccgaagga gctaaccgct ttttttcaca acatggggga tcatgtaact 11400

cgccttgatc gttgggaacc ggagctgaat gaagccatac caaacgacga gcgtgacacc 11460

acgatgcctg tagcaatggc aacaacgttg cgcaaactat taactggcga actacttact 11520

ctagcttccc ggcaacaatt aatagactgg atggaggcgg ataaagttgc aggaccactt 11580

ctgcgctcgg cccttccggc tggctggttt attgctgata aatctggagc cggtgagcgt 11640

gggtctcgcg gtatcattgc agcactgggg ccagatggta agccctcccg tatcgtagtt 11700

atctacacga cgggcagtca ggcaactatg gatgaacgaa atagacagat cgctgagata 11760

ggtgcctcac tgattaagca ttggtaactg tcagaccaag tttactcata tatactttag 11820

attgatttaa aacttcattt ttaatttaaa aggatctagg tgaagatcct ttttgataat 11880

ctcatgacca aaatccctta acgtgagttt tcgttccact gagcgtcaga ccccgtagaa 11940

aagatcaaag gatcttcttg agatcctttt tttctgcgcg taatctgctg cttgcaaaca 12000

aaaaaaccac cgctaccagc ggtggtttgt ttgccggatc aagagctacc aactcttttt 12060

ccgaaggtaa ctggcttcag cagagcgcag ataccaaata ctgtccttct agtgtagccg 12120

tagttaggcc accacttcaa gaactctgta gcaccgccta catacctcgc tctgctaatc 12180

ctgttaccag tggctgctgc cagtggcgat aagtcgtgtc ttaccgggtt ggactcaaga 12240

cgatagttac cggataaggc gcagcggtcg ggctgaacgg ggggttcgtg cacacagccc 12300

agcttggagc gaacgaccta caccgaactg agatacctac agcgtgagca ttgagaaagc 12360

gccacgcttc ccgaagggag aaaggcggac aggtatccgg taagcggcag ggtcggaaca 12420

ggagagcgca cgagggagct tccagggggg aacgcctggt atctttatag tcctgtcggg 12480

tttcgccacc tctgacttga gcgtcgattt ttgtgatgct cgtcaggggg gccgagccta 12540

tggaaaaacg ccagcaacgc ggccttttta cggttcctgg ccttttgctg gccttttgct 12600

cacatgttct ttcctgcgtt atcccctgat tctgtggata accgtattac cgcctttgag 12660

tgagctgata ccgctcgccg cagccgaacg accgagcgca gcgagtcagt gagcgaggaa 12720

gcggaagagc gcccaatacg caaaccgcct ctccccgcgc gttggccgat tcattaatgc 12780

agctggcacg acaggtttcc cgactggaaa gcgggcagtg agcgcaacgc aattaatgtg 12840

agttacctca ctcattaggc accccaggct ttacacttta tgcttccggc tcctatgttg 12900

tgtggaattg tgagcggata acaatttcac acaggaaaca gctatgacca tgat 12954

<210> 30
<211> 10329
<212> ADN
<213> Artificial

<220>
<223> plásmido pRN608 para la sobre-expresión de DAK1 y GCY1

<400> 30
agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg 60

gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga 120

gatacctaca gcgtgagcat tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca 180

ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt ccagggggga 240

acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt 300

tgtgatgctc gtcaggggg ccgagcctat ggaaaaacgc cagcaacgcg gccttttac 360

ggttcctggc cttttgctgg ccttttgctc acatgttctt tcctgcgtta tcccctgatt 420

ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc agccgaacga 480

ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg cccaatacgc aaaccgcctc 540

tccccgcgcg ttggccgatt cattaatgca gctggcacga caggtttccc gactggaaag 600

cgggcagtga gcgcaacgca attaatgtga gttacctcac tcattaggca ccccaggctt 660

tacactttat gcttccggct cctatgttgt gtggaattgt gagcggataa caatttcaca 720

caggaaacag ctatgaccat gattacgcca agctcggaat taaccctcac taaagggaac 780

aaaagctggg taccgggccc ccctcgagt accggtttct tcttcagatt ccctcatgga 840

gaaagtgcgg cagatgtata tgacagagtc gccagtttcc aagagacttt attcaggcac 900

ttccatgata ggcaagagag aagacccaga gatgttgttg tcctagttac acatggtatt 960

tattccagag tattcctgat gaaatggttt agatggacat acgaagagtt tgaatcgttt 1020

accaatgttc ctaacgggag cgtaatggtg atggaactgg acgaatccat caatagatac 1080

gtcctgagga ccgtgctacc caaatggact gattgtgagg gagacctaac tacatagtgt 1140

ttaaagatta cggatattta acttacttag aataatgcca ttttttgag ttataataat 1200

cctacgttag tgtgagcggg atttaaactg tgaggacctt aatacattca gacacttctg 1260

cggtatcacc ctacttattc ccttcgagat tatatctagg aacccatcag gttggtggaa 1320

gattacccgt tctaagactt ttcagcttcc tctattgatg ttacacctgg acaccccttt 1380

tctggcatcc agtttttaat cttcagtggc atgtgagatt ctccgaaatt aattaaagca 1440

atcacacaat tctctcggat accacctcgg ttgaaactga caggtggttt gttacgcatg 1500

ctaatgcaaa ggagcctata tacctttggc tcggctgctg taacagggaa tataaagggc 1560

agcataattt aggagtttag tgaacttgca acatttacta ttttccottc ttacgtaaat 1620

atttttcttt ttaattctaa atcaatcttt ttcaattttt tgtttgtatt cttttcttgc 1680

ttaaatctat aactacaaaa aacacataca taaatctaga aaatgtccgc taaatcgttt 1740

gaagtcacag atccagtcaa ttcaagtctc aaagggtttg cccttgctaa cccctccatt 1800

acgctggtcc ctgaagaaaa aattctcttc agaaagaccg attccgacaa gatcgcatta 1860

atttctggtg gtggtagtgg acatgaacct acacacgccg gtttcattgg taagggtatg 1920

ttgagtggcg ccgtggttgg cgaaattttt gcatcccctt caacaaaaca gattttaaat 1980

gcaatccgtt tagtcaatga aaatgcgtct ggcgttttat tgattgtgaa gaactacaca 2040

ggtgatgttt tgcattttgg tctgtccgct gagagagcaa gagccttggg tattaactgc 2100

cgcgttgctg tcataggtga tgatgttgca gttggcagag aaaagggtgg tatggttggt 2160

agaagagcat tggcaggtac cgttttggtt cataagattg taggtgcctt cgcagaagaa 2220

tattctagta agtatggctt agacggtaca gctaaagtgg ctaaaattat caacgacaat 2280

ttggtgacca ttggatcttc tttagaccat tgtaaagttc ctggcaggaa attcgaaagt 2340

gaattaaacg aaaaacaaat ggaattgggt atgggtattc ataacgaacc tggtgtgaaa 2400

gttttagacc ctattccttc taccgaagac ttgatctcca agtatatgct accaaaacta 2460

ttggatccaa acgataagga tagagctttt gtaaagtttg atgaagatga tgaagttgtc 2520

ttgttagtta acaatctcgg cggtgtttct aattttgtta ttagttctat cacttccaaa 2580

actacggatt tcttaaagga aaattacaac ataacccogg ttcaaacaat tgctggcaca 2640

ttgatgacct ccttcaatgg taatgggttc agtatcacat tactaaacgc cactaaggct 2700

acaaaggctt tgcaatctga ttttgaggag atcaaatcag tactagactt gttgaacgca 2760

tttacgaacg caccgggctg gccaattgca gattttgaaa agacttctgc cccatctgtt 2820

aacgatgact tgttacataa tgaagtaaca gcaaaggccg tcggtaccta tgactttgac 2880

aagtttgctg agtggatgaa gagtggtgct gaacaagtta tcaagagcga accgcacatt 2940

acggaactag acaatcaagt tggtgatggt gattgtggtt acactttagt ggcaggagtt 3000

aaaggcatca ccgaaaacct tgacaagctg tcgaaggact cattatctca ggcggttgcc 3060

caaatttcag atttcattga aggctcaatg ggaggtactt ctggtggttt atattctatt 3120

cttttgtcgg gtttttcaca cggattaatt caggtttgta aatcaaagga tgaacccgtc 3180

actaaggaaa ttgtggctaa gtcactcgga attgcattgg atactttata caaatataca 3240

aaggcaagga agggatcatc caccatgatt gatgctttag aaccattcgt taaagaattt 3300

actgcatcta aggatttcaa taaggcggta aaagctgcag aggaaggtgc taaatccact 3360

gctacattcg aggccaaatt tggcagagct tcgtatgtcg gcgattcatc tcaagtagaa 3420

gatcctggtg cagtaggcct atgtgagttt ttgaaggggg ttcaaagcgc cttgtaagtc 3480

gagacaaatc gctcttaaat atatacctaa agaacattaa agctatatta taagcaaaga 3540

tacgtaaatt ttgcttatat tattatacac atatcatatt tctatatttt taagatttgg 3600

ttatataatg tacgtaatgc aaaggaaata aattttatac attattgaac agcgtccaag 3660

taactacatt atgtgcacta atagtttagc gtcgtgaaga ctttattgtg tcgcgaaaag 3720

taaaaatttt aaaaattaga gcaccttgaa cttgcgaaaa aggttctcat caactgttta 3780

aaacgtacgt gtggaagaac gattacaaca ggtgttgtcc tctgaggaca taaaatacac 3840

accgagattc atcaactcat tgctggagtt agcatatcta caattgggtg aaatggggag 3900

cgatttgcag gcatttgctc ggcatgccgg tagaggtgtg gtcaataaga gcgacctcat 3960

gctatacctg agaaagcaac ctgacctaca ggaaagagtt actcaagaat aagaattttc 4020

gttttaaaac ctaagagtca ctttaaaatt tgtatacact tatttttttt ataacttatt 4080

taataataaa aatcataaat cataagaaat tcgcgcgctt acttgaatac ttcgaaagga 4140

gaccaatttg gatgtacaac cctttttcg ccctttcct tcgatatgtt attgatagct 4200

tcaaagtcct cagtagacaa agtaaatatt ttcctgttcg ttttgattcg atcgggattc 4260

acagattttg gcaagacaac ataacctctt tggacgtgcc agctaataac aacgtgtccg 4320

ggttgaacgt tatttttctt cgcaatttca aggataaccg gttccttcaa tagtggagca 4380

tcggtactac ctaacggaga ataagcttca accacaatgc ctttactttt acaaaaatta 4440

atcaattcgt cttgaggtag taatggatgt atttcgactt ggttagcagc tggcgtaagc 4500

ttattacctt gagatgctaa tagatctttc aggttattta tagaaaagtt ggagactcca 4560

acggccttag ttttaccagt ctttggtagt tcctgcatta attcccaggt tttgatgaaa 4620

ttccaattgg tgatatccac tgcacgagaa ccatccttct ttgttggcac actcaagatg 4680

tcttcatttt tgatgtaggc tggatctaat ctggcaggcc aatgcatcaa atataagtct 4740

acgtagtcca atcctaacct ctttagtgat tgatccagcg ctacttcagg ttcgtggtgt 4800

tgtgtacacc ataactttgt agtaacaaag atttcttccc gaggaacacc tgaatccttg 4860

atggcttgac cgacttggtc ttcattacgg taaatagcag cagtatcaat gtgtcggtag 4920

ccatctttca aagcggttaa aacagcctta taagcatcgt tctctttcga ctgccacgta 4980

cctaaaccta tttgagggat ttgggctcca gtatttagag aaaggatttt cgtagaatca 5040

tgtaaagtag caggcatttt ctgcagttaa ttcagtaaat tttcgatctt gggaagaaaa 5100

aagcagtaag cgtgaaaaat ctaaaagctg atgtagtaga agatcctatt ctttaacaaa 5160

gattgacctt ttctttttct tcttggtttg agtagaaagg ggaaggaaga atacaagaga 5220

gaggaaaaaa aggaagataa aaagagagcg tgatataaat gaatatatat taaacaagag 5280

agattgggaa ggaaaggatc aaacaaaccc aaaaatattt caaaaaggag agagagaggc 5340

gagtttggtt tcaaaacggt ttatttattt atgcaagagg acgtggaaga aaaagaagaa 5400

ggaagaaaaa aatttgaaag aaaaaaacgc gtggcgggta aagaagaaaa tggaaaatag 5460

aggccgggtg acagagaaat attgagggtt aattggaaaa tatgttaggg tgaggcatat 5520

gtttttaagg gttttgagga tccgataagg aagaatgtag gttaaatgtt gtgcattaat 5580

tgctgtggca gcttacccgc ttccccacac atttactagt tctagagcgg ccgccaccgc 5640

ggtggagctc caattcgccc tatagtgagt cgtattacaa ttcactggcc gtcgttttac 5700

aacgtcgtga ctgggaaaac cctggcgtta cccaacttaa tcgccttgca gcacatcccc 5760

ccttcgccag ctggcgtaat agcgaagagg cccgcaccga tcgcccttcc caacagttgc 5820

gcagcctgaa tggcgaatgg cgcgacgcgc cctgtagcgg cgcattaagc gcggcgggtg 5880

tggtggttac gcgcagcgtg accgctacac ttgccagcgc cctagcgccc gctcctttcg 5940

ctttcttccc ttcctttctc gccacgttcg ccggtagtgt tagacctgaa caaggtttac 6000

taaaaatccg taaagaactt caattgtacg ccaacttaag gcctcgtccc cgccgggtca 6060

cccggccagc gacatggagg cccagaatac cctccttgac agtcttgacg tgcgcagctc 6120

aggggcatga tgtgactgtc gcccgtacat ttagcccata catccccatg tataatcatt 6180

tgcatccata cattttgatg gccgcacggc gcgaagcaaa aattacggct cctcgctgca 6240

gacctgcgag cagggaaacg ctcccctcac agacgcgtga attgtcccca cgccgcgccc 6300

ctgtagagaa atataaaagg ttaggatttg ccactgaggt tcttctttca tatacttcct 6360

tttaaaatct tgctaggata cagttctcac atcacatccg aacataaaca accatgggta 6420

aaaagcctga actcaccgcg acgtctgtcg agaagtttct gatcgaaaag ttcgacagcg 6480

tctccgacct gatgcagctc tcggagggcg aagaatctcg tgctttcagc ttcgatgtag 6540

gagggcgtgg atatgtcctg cgggtaaata gctgcgccga tggtttctac aaagatcgtt 6600

atgtttatcg gcactttgca tcggccgcgc tcccgattcc ggaagtgctt gacattgggg 6660

aattcagcga gagcctgacc tattgcatct cccgccgtgc acagggtgtc acgttgcaag 6720

acctgcctga aaccgaactg cccgctgttc tgcagccggt cgcggaggcc atggatgcga 6780

tcgctgcggc cgatcttagc cagacgagcg ggttcggccc attcggaccg caaggaatcg 6840

gtcaatacac tacatggcgt gatttcatat gcgcgattgc tgatccccat gtgtatcact 6900

ggcaaactgt gatggacgac accgtcagtg cgtccgtcgc gcaggctctc gatgagctga 6960

tgctttgggc cgaggactgc cccgaagtcc ggcacctcgt gcacgcggat ttcggctcca 7020

acaatgtcct gacggacaat ggccgcataa cagcggtcat tgactggagc gaggcgatgt 7080

tcggggattc ccaatacgag gtcgccaaca tcttcttctg gaggccgtgg ttggcttgta 7140

tggagcagca gacgcgctac ttcgagcgga ggcatccgga gcttgcagga tcgccgcggc 7200

tccgggcgta tatgctccgc attggtcttg accaactcta tcagagcttg gttgacggca 7260

atttcgatga tgcagcttgg gcgcagggtc gatgcgacgc aatcgtccga tccggagccg 7320

ggactgtcgg gcgtacacaa atcgcccgca gaagcgcggc cgtctggacc gatggctgtg 7380

tagaagtact cgccgatagt ggaaaccgac gccccagcac tcgtccgagg gcaaaggaat 7440

aatcagtact gacaataaaa agattcttgt tttcaagaac ttgtcatttg tatagttttt 7500

ttatattgta gttgttctat tttaatcaaa tgttagcgtg atttatattt tttttcgcct 7560

cgacatcatc tgcccagatg cgaagttaag tgcgcagaaa gtaatatcat gcgtcaatcg 7620

tatgtgaatg ctggtcgcta tactgctgtc gattcgatac taacgccgcc atccagtgtc 7680

gacggatcct aggtgtacat aaactttata aatgaaattc ataatagaaa cgacacgaaa 7740

ttacaaaatg gaatatgttc atagggtaga cgaaactata tacgcaatct acatacattt 7800

atcaagaagg agaaaaagga ggatagtaaa ggaatacagg taagcaaatt gatactaatg 7860

gctcaacgtg ataaggaaaa agaattgcac tttaacatta atattgacaa ggaggagggc 7920

accacacaaa aagttaggtg taacagaaaa tcatgaaact acgattccta atttgatatt 7980

ggaggatttt ctctaaaaaa aaaaaaatac aacaaataaa aaacactcaa tgacctgacc 8040

atttgatgga gtttaagtca ataccttctt gaaccatttc ccataatggt gaaagttccc 8100

tcaagaattt tactctgtca gaaacggcct tacgacgtag tcgatatggt gcactctcag 8160

tacaatctgc tctgatgccg catagttaag ccagccccga cacccgccaa cacccgctga 8220

cgcgccctga cgggcttgtc tgctcccggc atccgcttac agacaagctg tgaccgtctc 8280

cgggagctgc atgtgtcaga ggttttcacc gtcatcaccg aaacgcgcga gacgaaaggg 8340

cctcgtgata cgcctatttt tataggttaa tgtcatgata ataatggttt cttaggacgg 8400

atcgcttgcc tgtaacttac acgcgcctcg tatcttttaa tgatggaata atttgggaat 8460

ttactctgtg tttatttatt tttatgtttt gtatttggat tttagaaagt aaataaagaa 8520

ggtagaagag ttacggaatg aagaaaaaaa aataaacaaa ggtttaaaaa atttcaacaa 8580

aaagcgtact ttacatatat atttattaga caagaaaagc agattaaata gatatacatt 8640

cgattaacga taagtaaaat gtaaaatcac aggattttcg tgtgtggtct tctacacaga 8700

caagatgaaa caattcggca ttaatacctg agagcaggaa gagcaagata aaaggtagta 8760

tttgttggcg atccccctag agtcttttac atcttcggaa aacaaaaact atttttttctt 8820

taatttcttt ttttactttc tatttttaat ttatatattt atattaaaaa atttaaatta 8880

taattatttt tatagcacgt gatgaaaagg acccaggtgg cacttttcgg ggaaatgtgc 8940

gcggaacccc tatttgttta tttttctaaa tacattcaaa tatgtatccg ctcatgagac 9000

aataaccctg ataaatgctt caataatatt gaaaaaggaa gagtatgagt attcaacatt 9060

tccgtgtcgc ccttattccc ttttttgcgg cattttgcct tcctgttttt gctcacccag 9120

aaacgctggt gaaagtaaaa gatgctgaag atcagttggg tgcacgagtg ggttacatcg 9180

aactggatct caacagcggt aagatccttg agagttttcg ccccgaagaa cgttttccaa 9240

tgatgagcac ttttaaagtt ctgctatgtg gcgcggtatt atcccgtatt gacgccgggc 9300

aagagcaact cggtcgccgc atacactatt ctcagaatga cttggttgag tactcaccag 9360

tcacagaaaa gcatcttacg gatggcatga cagtaagaga attatgcagt gctgccataa 9420

ccatgagtga taacactgcg gccaacttac ttctgacaac gatcggagga ccgaaggagc 9480

taaccgcttt ttttcacaac atgggggatc atgtaactcg ccttgatcgt tgggaaccgg 9540

agctgaatga agccatacca aacgacgagc gtgacaccac gatgcctgta gcaatggcaa 9600

caacgttgcg caaactatta actggcgaac tacttactct agcttcccgg caacaattaa 9660

tagactggat ggaggcggat aaagttgcag gaccacttct gcgctcggcc cttccggctg 9720

gctggtttat tgctgataaa tctggagccg gtgagcgtgg gtctcgcggt atcattgcag 9780

cactggggcc agatggtaag ccctcccgta tcgtagttat ctacacgacg ggcagtcagg 9840

caactatgga tgaacgaaat agacagatcg ctgagatagg tgcctcactg attaagcatt 9900

ggtaactgtc agaccaagtt tactcatata tactttagat tgatttaaaa cttcattttt 9960

aatttaaaag gatctaggtg aagatccttt ttgataatct catgaccaaa atcccttaac 10020

gtgagttttc gttccactga gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag 10080

atcctttttt tctgcgcgta atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg 10140

tggtttgttt gccggatcaa gagctaccaa ctcttttcc gaaggtaact ggcttcagca 10200

gagcgcagat accaaatact gtccttctag tgtagccgta gttaggccac cacttcaaga 10260

actctgtagc accgcctaca tacctcgctc tgctaatcct gttaccagtg gctgctgcca 10320

gtggcgata 10329

<210> 31
<211> 11163
<212> ADN
<213> Artificial

<220>
<223> pRN616

<400> 31

```
tttgattaaa atagaacaac tacaatataa aaaaactata caaatgacaa gttcttgaaa     60
acaagaatct ttttattgtc agtgtgtcag tcctgctcct cggccacgaa gtgcacgcag    120
ttgccggccg ggtcgcgcag ggcgaactcc cgcccccacg gctgctcgcc gatctcggtc    180
atggccggcc cggaggcgtc ccggaagttc gtggacacga cctccgacca ctcggcgtac    240
agctcgtcca ggccgcgcac ccacacccag gccagggtgt tgtccggcac cacctggtcc    300
tggaccgcgc tgatgaacag ggtcacgtcg tcccggacca caccggcgaa gtcgtcctcc    360
acgaagtccc gggagaaccc gagccggtcg gtccagaact cgaccgctcc ggcgacgtcg    420
cgcgcggtga gcaccggaac ggcactggtc aacttggcca tggttgttta tgttcggatg    480
tgatgtgaga actgtatcct agcaagattt taaaaggaag tatatgaaag aagaacctca    540
gtggcaaatc ctaacctttt atatttctct acaggggcgc ggcgtgggga caattcaacg    600
cgtctgtgag gggagcgttt ccctgctcgc aggtctgcag cgaggagccg taatttttgc    660
ttcgcgccgt gcggccatca aaatgtatgg atgcaaatga ttatacatgg ggatgtatgg    720
gctaaatgta cgggcgacag tcacatcatg cccctgagct gcgcacgtca agactgtcaa    780
ggagggtatt ctgggcctcc atgtcgctgg ccgggtgacc cggcggggac gaggccttaa    840
gcggccgcat gctagctccg gattatcgat gataagctgt caaacatgag aattaattcc    900
acggactata gactatacct agtatactcc gtctactgta cgatacactt ccgctcaggt    960
ccttgtcctt taacgaggcc ttaccactct tttgttactc tattgatcca gctcagcaaa   1020
ggcagtgtga tctaagattc tatcttcgcg atgtagtaaa actagctaga ccgagaaaga   1080
gactagaaat gcaaaaggca cttctacaat ggctgccatc attattatcc gatgtgacgc   1140
tgcagcttct caatgatatt cgaatacgct ttgaggagat acagcctaat atccgacaaa   1200
ctgttttaca gatttacgat cgtacttgtt acccatcatt gaattttgaa catccgaacc   1260
tgggagtttt ccctgaaaca gatagtatat ttgaacctgt ataataatat atagtctagc   1320
gctttacgga agacaatgta tgtatttcgg ttcctggaga aactattgca tctattgcat   1380
aggtaatctt gcacgtcgca tccccggttc attttctgcg tttccatctt gcacttcaat   1440
agcatatctt tgttaacgaa gcatctgtgc ttcattttgt agaacaaaaa tgcaacgcga   1500
gagcgctaat ttttcaaaca aagaatctga gctgcatttt tacagaacag aaatgcaacg   1560
```

cgaaagcgct attttaccaa cgaagaatct gtgcttcatt tttgtaaaac aaaaatgcaa 1620

cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc atttttacag aacagaaatg 1680

caacgcgaga gcgctatttt accaacaaag aatctatact tctttttttgt tctacaaaaa 1740

tgcatcccga gagcgctatt tttctaacaa agcatcttag attacttttt ttctcctttg 1800

tgcgctctat aatgcagtct cttgataact ttttgcactg taggtccgtt aaggttagaa 1860

gaaggctact ttggtgtcta ttttctcttc cataaaaaaa gcctgactcc acttcccgcg 1920

tttactgatt actagcgaag ctgcgggtgc atttttttcaa gataaaggca tccccgatta 1980

tattctatac cgatgtggat tgcgcatact ttgtgaacag aaagtgatag cgttgatgat 2040

tcttcattgg tcagaaaatt atgaacggtt tcttctattt tgtctctata tactacgtat 2100

aggaaatgtt tacattttcg tattgttttc gattcactct atgaatagtt cttactacaa 2160

ttttttttgtc taaagagtaa tactagagat aaacataaaa aatgtagagg tcgagtttag 2220

atgcaagttc aaggagcgaa aggtggatgg gtaggttata tagggatata gcacagagat 2280

atatagcaaa gagatacttt tgagcaatgt ttgtggaaag cggtattcgc aatgccggca 2340

aaagggcgaa ttgattttga agagaatgtg gattttgatg taattgttgg gattccattt 2400

ttaataaggc aataatatta ggtatgtgga tatactagaa gttctcctcg accgtcgata 2460

tgcggtgtga aataccgcac agatgcgtaa ggagaaaata ccgcatcagg aaattgtaaa 2520

cgttaatatt ttgttaaaat tcgcgttaaa tttttgttaa atcagctcat tttttaacca 2580

ataggccgaa atcggcaaaa tcccttataa atcaaaagaa tagaccgaga tagggttgag 2640

tgttgttcca gtttggaaca agagtccact attaaagaac gtggactcca acgtcaaagg 2700

gcgaaaaacc gtctatcagg gcgatggccc actacgtgaa ccatcaccct aatcaagttt 2760

tttggggtcg aggtgccgta aagcactaaa tcggaaccct aaagggagcc cccgatttag 2820

agcttgacgg ggaaagccgg cgaacgtggc gagaaaggaa gggaagaaag cgaaaggagc 2880

gggcgctagg gcgctggcaa gtgtagcggt cacgctgcgc gtaaccacca cacccgccgc 2940

gcttaatgcg ccgctacagg gcgcgtcgcg ccattcgcca ttcaggctgc gcaactgttg 3000

ggaagggcga tcggtgcggg cctcttcgct attacgccag ctggcgaagg ggggatgtgc 3060

tgcaaggcga ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac 3120

ggccagtgaa ttgtaatacg actcactata gggcgaattg gagctccacc gcggtggcgg 3180

ccgctctaga actagtctcg agctcttcaa ctcaagacgc acagatatta taacatctgc 3240

ataataggca tttgcaagaa ttactcgtga gtaaggaaag agtgaggaac tatcgcatac 3300

ctgcatttaa agatgccgat ttgggcgcga atcctttatt ttggcttcac cctcatacta 3360

ttatcagggc cagaaaaagg aagtgtttcc ctccttcttg aattgatgtt accctcataa 3420

agcacgtggc ctcttatcga gaaagaaatt accgtcgctc gtgatttgtt tgcaaaaaga 3480

acaaaactga aaaaacccag acacgctcga cttcctgtct tcctattgat tgcagcttcc 3540

aatttcgtca cacaacaagg tcctagcgac ggctcacagg ttttgtaaca agcaatcgaa 3600

ggttctggaa tggcgggaaa gggtttagta ccacatgcta tgatgcccac tgtgatctcc 3660

agagcaaagt tcgttcgatc gtactgttac tctctctctt tcaaacagaa ttgtccgaat 3720

cgtgtgacaa caacagcctg ttctcacaca ctcttttctt ctaaccaagg gggtggttta 3780

gtttagtaga acctcgtgaa acttacattt acatatatat aaacttgcat aaattggtca 3840

atgcaagaaa tacatatttg gtcttttcta attcgtagtt tttcaagttc ttagatgctt 3900

tctttttctc ttttttacag atcatcaagg aagtaattat ctacttttta caacaaatat 3960

atctagactg cagaaaatgt ctgaattgaa cgagaagttg gctaccgctt gggaaggttt 4020

caccaagggt gactggcaaa acgaagttaa cgttagagac ttcatccaaa agaactacac 4080

cccatacgaa ggtgacgaat ctttcttggc tggtgctacc gaagctacca ccaccttgtg 4140

ggacaaggtt atggaaggtg ttaagttgga aaacagaacc cacgctccag ttgacttcga 4200

caccgctgtt gcttctacca tcacctctca cgacgctggt tacatcaaca agcaattgga 4260

aaagatcgtt ggtttacaaa ccgaagctcc attgaagaga gctttgatcc cattcggtgg 4320

tatcaagatg atcgaaggtt cttgtaaggc ttacaacaga gaattggacc caatgatcaa 4380

gaagattttc accgaataca gaaagaccca caaccaaggt gttttcgacg tttacactcc 4440

agacatcttg agatgtagaa agtctggtgt tttgactggt ttgccagacg cttacggtag 4500

aggtagaatc atcggtgact acagaagagt tgctttgtac ggtatcgact acttgatgaa 4560

ggacaagttg gctcaattca cctctttgca agctgacttg gaaaacggtg ttaacttgga 4620

acaaaccatc agattgagag aagaaatcgc tgaacaacac agagctttgg gtcaaatgaa 4680

ggaaatggct gctaagtacg gttacgacat ctctggtcca gctaccaacg ctcaagaagc 4740

tatccaatgg acctacttcg gttacttggc tgctgttaag tctcaaaacg gtgctgctat 4800

gtctttcggt aggacctcta ccttcttgga cgtttacatc gaaagagact tgaaggctgg 4860

taagatcacc gaacaagaag ctcaagaaat ggttgaccac ttggttatga agttgagaat 4920

ggttagattc ttgagaaccc cagaatacga cgaattgttc tctggtgacc caatctgggc 4980

taccgaatct atcggtggta tgggtttgga cggtagaacc ttggttacca agaactcttt 5040

cagattcttg aacaccttat acaccatggg tccatctcca gaaccaaaca tgaccatctt 5100

gtggtctgaa aagttaccat tgaacttcaa gaagttcgct gctaaggttt ctatcgacac 5160

ctcttctttg caatacgaaa acgacgactt gatgagacca gacttcaaca acgacgacta 5220

cgctatcgct tgttgtgttt ctccaatgat cgttggtaag caaatgcaat tcttcggtgc 5280

tagagctaac ttggctaaga ccatgttgta cgctatcaac ggtggtgttg acgaaaagtt 5340

gaagatgcaa gttggtccaa agtctgaacc aatcaagggt gacgttttga actacgacga 5400

agttatggaa agaatggacc acttcatgga ctggttggct aagcaataca tcaccgcttt 5460

gaacatcatc cactacatgc acgacaagta ctcttacgaa gcatcattga tggctttgca 5520

cgacagagac gtaatcagaa ccatggcttg tggtatcgct ggtttgtctg ttgctgctga 5580

ctctttgtct gctatcaagt acgctaaggt taagccaatc agagacgaag acggtttggc 5640

tatcgacttc gaaatcgaag gtgaataccc tcaattcggt aacaacgacc caagagttga 5700

cgacttggct gttgacttgg ttgaaagatt tatgaagaag atccaaaagt tgcacaccta 5760

cagagacgct atcccaaccc aatctgtttt gactatcaca tctaacgttg tttacggtaa 5820

gaagactggt aacaccccag acggtagaag agctggtgct ccattcggtc caggtgctaa 5880

cccaatgcac ggtagagacc aaaagggtgc tgtagcatct ttgacctctg ttgctaagtt 5940

gccattcgct tacgctaagg acggtatctc ttacaccttc tctatcgttc caaacgcttt 6000

gggtaaggac gatgaagtta gaaagaccaa cttggctggt ttgatggacg gttacttcca 6060

ccacgaagca tctatcgaag gtggtcaaca cttgaacgta aatgttatga acagagaaat 6120

gttgttggac gctatggaaa acccagaaaa gtacccacaa ttgaccatca gagtttctgg 6180

ttacgctgtt agattcaact ctttgaccaa ggaacaacaa caagacgtta tcaccagaac 6240

cttcacccaa tctatgtaag tcgagacaaa tcgctcttaa atatatacct aaagaacatt 6300

aaagctatat tataagcaaa gatacgtaaa ttttgcttat attattatac acatatcata 6360

tttctatatt tttaagattt ggttatataa tgtacgtaat gcaaaggaaa taaattttat 6420

acattattga acagcgtcca agtaactaca ttatgtgcac taatagttta gcgtcgtgaa 6480

gactttattg tgtcgcgaaa agtaaaaatt ttaaaaatta gagcaccttg aacttgcgaa 6540

aaaggttctc atcaactgtt taaaacgtac gtgtggaaga acgattacaa caggtgttgt 6600

cctctgagga cataaaatac acaccgagat tcatcaactc attgctggag ttagcatatc 6660

tacaattggg tgaaatgggg agcgatttgc aggcatttgc tcggcatgcc ggtagaggtg 6720

tggtcaataa gagcgacctc atgctatacc tgagaaagca acctgaccta caggaaagag 6780

ttactcaaga ataagaattt tcgttttaaa acctaagagt cactttaaaa tttgtataca 6840

cttatttttt ttataactta tttaataata aaaatcataa atcataagaa attcgcgcgc 6900

ttagaacata accttgtgac cgtattgttc caagataccc ttaactcttt ccatggtttc 6960

cttctttggt ggcttaacac cgtccaactt gtattcttca cccatagcaa cccacttgtg 7020

cttacccaat tcgtggtatg gcaacaattc gatcttttca acgttaccca tgtctctggt 7080

gaactcaccc aatctgtgag cagagtcgtc gtcgtcagac caacctggaa caacaacgta 7140

tctgatccaa accttaacgt tcttgttagc caagtactta gcaaattcca aggttctgtg 7200

gttagaaaca ccaaccaagt tttggtggat ttcgtcgttc atttgcttca agtccaacat 7260

aaccaagtcg gtaacttcca acaattcgtc gataactggg tcgtatcttc taacgaaacc 7320

gttggtgtcc aaacaggtgt ggataccttc cttcttacaa gctctgaacc agtctctaac 7380

aaattcagct tgcaagatag cttcaccacc agaagcggta acaccacctc ctgaagcgtt 7440

cataaagtgt ctgtaggtaa caacttcctt catcaagtct tcaacagtaa cttccttacc 7500

accgtgggtg tcccaggtgt ctctgttgtg acagtacaaa catctcatca aacaaccttg 7560

gaagaaggtg atgaatctga tacctggacc gtcaacagta ccacaagatt cgaaagagtg 7620

gattctaccg ataacagaca ttttaagctt ctgcagctta gattagattg ctatgctttc 7680

tttctaatga gcaagaagta aaaaaagttg taatagaaca agaaaaatga aactgaaact 7740

tgagaaattg aagaccgttt attaacttaa atatcaatgg gaggtcatcg aaagagaaaa 7800

aaatcaaaaa aaaaaatttt caagaaaaag aaacgtgata aaaattttta ttgccttttt 7860

cgacgaagaa aaagaaacga ggcggtctct tttttctttt ccaaaccttt agtacgggta 7920

attaacgaca ccctagagga agaaagaggg gaaatttagt atgctgtgct tgggtgtttt 7980

gaagtggtac ggcgatgcgc ggagtccgag aaaatctgga agagtaaaaa aggagtagaa 8040

acattttgaa gctatggtgt gtgggaccgg tcgaggggggg gcccggtacc cagcttttgt 8100

tccctttagt gagggttaat tccgagcttg gcgtaatcat ggtcatagct gtttcctgtg 8160

tgaaattgtt atccgctcac aattccacac aacataggag ccggaagcat aaagtgtaaa 8220

gcctggggtg cctaatgagt gaggtaactc acattaattg cgttgcgctc actgcccgct 8280

ttccagtcgg gaaacctgtc gtgccagctg cattaatgaa tcggccaacg cgcggggaga 8340

ggcggtttgc gtattgggcg ctcttccgct tcctcgctca ctgactcgct gcgctcggtc 8400

gttcggctgc ggcgagcggt atcagctcac tcaaaggcgg taatacggtt atccacagaa 8460

tcaggggata acgcaggaaa gaacatgtga gcaaaaggcc agcaaaaggc caggaaccgt 8520

aaaaaggccg cgttgctggc gtttttccat aggctcggcc cccctgacga gcatcacaaa 8580

aatcgacgct caagtcagag gtggcgaaac ccgacaggac tataaagata ccaggcgttc 8640

cccctggaa gctccctcgt gcgctctcct gttccgaccc tgccgcttac cggatacctg 8700

tccgcctttc tcccttcggg aagcgtggcg ctttctcaat gctcacgctg taggtatctc 8760

agttcggtgt aggtcgttcg ctccaagctg ggctgtgtgc acgaaccccc cgttcagccc 8820

gaccgctgcg ccttatccgg taactatcgt cttgagtcca acccggtaag acacgactta 8880

tcgccactgg cagcagccac tggtaacagg attagcagag cgaggtatgt aggcggtgct 8940

acagagttct tgaagtggtg gcctaactac ggctacacta gaaggacagt atttggtatc 9000

tgcgctctgc tgaagccagt taccttcgga aaaagagttg gtagctcttg atccggcaaa 9060

caaaccaccg ctggtagcgg tggttttttt gtttgcaagc agcagattac gcgcagaaaa 9120

aaaggatctc aagaagatcc tttgatcttt tctacggggt ctgacgctca gtggaacgaa 9180

aactcacgtt aagggatttt ggtcatgaga ttatcaaaaa ggatcttcac ctagatcctt 9240

ttaaattaaa aatgaagttt taaatcaatc taaagtatat atgagtaaac ttggtctgac 9300

agttaccaat gcttaatcag tgaggcacct atctcagcga tctgtctatt tcgttcatcc 9360

atagttgcct gactgcccgt cgtgtagata actacgatac gggagggctt accatctggc 9420

cccagtgctg caatgatacc gcgagaccca cgctcaccgg ctccagattt atcagcaata 9480

aaccagccag ccggaagggc cgagcgcaga agtggtcctg caactttatc cgcctccatc 9540

cagtctatta attgttgccg ggaagctaga gtaagtagtt cgccagttaa tagtttgcgc 9600

aacgttgttg ccattgctac aggcatcgtg gtgtcacgct cgtcgtttgg tatggcttca 9660

ttcagctccg gttcccaacg atcaaggcga gttacatgat cccccatgtt gtgaaaaaaa 9720

gcggttagct ccttcggtcc tccgatcgtt gtcagaagta agttggccgc agtgttatca 9780

ctcatggtta tggcagcact gcataattct cttactgtca tgccatccgt aagatgcttt 9840

tctgtgactg gtgagtactc aaccaagtca ttctgagaat agtgtatgcg gcgaccgagt 9900

tgctcttgcc cggcgtcaat acgggataat accgcgccac atagcagaac tttaaaagtg 9960

ctcatcattg gaaaacgttc ttcggggcga aaactctcaa ggatcttacc gctgttgaga 10020

tccagttcga tgtaacccac tcgtgcaccc aactgatctt cagcatcttt tactttcacc 10080

agcgtttctg ggtgagcaaa aacaggaagg caaaatgccg caaaaaaggg aataagggcg 10140

acacggaaat gttgaatact catactcttc ctttttcaat attattgaag catttatcag 10200

ggttattgtc tcatgagcgg atacatattt gaatgtattt agaaaaataa acaaataggg 10260

gttccgcgca catttccccg aaaagtgcca cctgacgtct aagaaaccat tattatcatg 10320

acattaacct ataaaaatag gcgtatcacg aggccctttc gtctcgcgcg tttcggtgat 10380

gacggtgaaa acctctgaca catgcagctc ccggagacgg tcacagcttg tctgtaagcg 10440

gatgccggga gcagacaagc ccgtcagggc gcgtcagcgg gtgttggcgg gtgtcggggc 10500

tggcttaact atgcggcatc agagcagatt gtactgagag tgcaccatat cgactacgtc 10560

gtaaggccgt ttctgacaga gtaaaattct tgagggaact ttcaccatta tgggaaatgg 10620

ttcaagaagg tattgactta aactccatca aatggtcagg tcattgagtg ttttttattt 10680

gttgtatttt ttttttttta gagaaaatcc tccaatatca aattaggaat cgtagtttca 10740

tgattttctg ttacacctaa ctttttgtgt ggtgccctcc tccttgtcaa tattaatgtt 10800

aaagtgcaat tctttttcct tatcacgttg agccattagt atcaatttgc ttacctgtat 10860

tcctttacta tcctcctttt tctccttctt gataaatgta tgtagattgc gtatatagtt 10920

tcgtctaccc tatgaacata ttccattttg taatttcgtg tcgtttctat tatgaatttc 10980

atttataaag tttatgtaca cctaggatcc gtcgacactg gatggcggcg ttagtatcga 11040

atcgacagca gtatagcgac cagcattcac atacgattga cgcatgatat tactttctgc 11100

gcacttaact tcgcatctgg gcagatgatg tcgaggcgaa aaaaaatata aatcacgcta 11160

aca 11163

<210> 32

<211> 7830

<212> ADN

<213> Artificial

<220>
<223> pRN619

<400> 32

aataccgcgc cacatagcag aactttaaaa gtgctcatca ttggaaaacg ttcttcgggg 60

cgaaaactct caaggatctt accgctgttg agatccagtt cgatgtaacc cactcgtgca 120

cccaactgat cttcagcatc ttttactttc accagcgttt ctgggtgagc aaaaacagga 180

aggcaaaatg ccgcaaaaaa gggaataagg gcgacacgga aatgttgaat actcatactc 240

ttccttttc aatattattg aagcatttat cagggttatt gtctcatgag cggatacata 300

tttgaatgta tttagaaaaa taaacaaata ggggttccgc gcacatttcc ccgaaaagtg 360

ccacctgacg tctaagaaac cattattatc atgacattaa cctataaaaa taggcgtatc 420

acgaggccct ttcgtctcgc gcgtttcggt gatgacggtg aaaacctctg acacatgcag 480

ctcccggaga cggtcacagc ttgtctgtaa gcggatgccg ggagcagaca agcccgtcag 540

ggcgcgtcag cgggtgttgg cgggtgtcgg ggctggctta actatgcggc atcagagcag 600

attgtactga gagtgcacca tatcgactac gtcgtaaggc cgtttctgac agagtaaaat 660

tcttgaggga actttcacca ttatgggaaa tggttcaaga aggtattgac ttaaactcca 720

tcaaatggtc aggtcattga gtgtttttta tttgttgtat ttttttttt ttagagaaaa 780

tcctccaata tcaaattagg aatcgtagtt tcatgatttt ctgttacacc taacttttg 840

tgtggtgccc tcctccttgt caatattaat gttaaagtgc aattcttttt ccttatcacg 900

ttgagccatt agtatcaatt tgcttacctg tattccttta ctatcctcct ttttctcctt 960

cttgataaat gtatgtagat tgcgtatata gtttcgtcta ccctatgaac atattccatt 1020

ttgtaatttc gtgtcgtttc tattatgaat ttcatttata aagtttatgt acacctagga 1080

tccgtcgaca ctggatggcg gcgttagtat cgaatcgaca gcagtatagc gaccagcatt 1140

cacatacgat tgacgcatga tattactttc tgcgcactta acttcgcatc tgggcagatg 1200

atgtcgaggc gaaaaaaaat ataaatcacg ctaacatttg attaaaatag aacaactaca 1260

atataaaaaa actatacaaa tgacaagttc ttgaaaacaa gaatcttttt attgtcagtg 1320

tgtcagtcct gctcctcggc cacgaagtgc acgcagttgc cggccgggtc gcgcagggcg 1380

aactcccgcc cccacggctg ctcgccgatc tcggtcatgg ccggcccgga ggcgtcccgg 1440

aagttcgtgg acacgacctc cgaccactcg gcgtacagct cgtccaggcc gcgcacccac 1500

acccaggcca gggtgttgtc cggcaccacc tggtcctgga ccgcgctgat gaacagggtc 1560

acgtcgtccc ggaccacacc ggcgaagtcg tcctccacga agtcccggga gaacccgagc 1620

cggtcggtcc agaactcgac cgctccggcg acgtcgcgcg cggtgagcac cggaacggca 1680

ctggtcaact tggccatggt tgtttatgtt cggatgtgat gtgagaactg tatcctagca 1740

agattttaaa aggaagtata tgaaagaaga acctcagtgg caaatcctaa ccttttatat 1800

ttctctacag gggcgcggcg tggggacaat tcaacgcgtc tgtgagggga gcgtttccct 1860

gctcgcaggt ctgcagcgag gagccgtaat ttttgcttcg cgccgtgcgg ccatcaaaat 1920

gtatggatgc aaatgattat acatggggat gtatgggcta aatgtacggg cgacagtcac 1980

atcatgcccc tgagctgcgc acgtcaagac tgtcaaggag ggtattctgg gcctccatgt 2040

cgctggccgg gtgacccggc ggggacgagg ccttaagcgg ccgcatgcta gctccggatt 2100

atcgatgata agctgtcaaa catgagaatt aattccacgg actatagact atacctagta 2160

tactccgtct actgtacgat acacttccgc tcaggtcctt gtcctttaac gaggccttac 2220

cactcttttg ttactctatt gatccagctc agcaaaggca gtgtgatcta agattctatc 2280

ttcgcgatgt agtaaaacta gctagaccga gaaagagact agaaatgcaa aaggcacttc 2340

tacaatggct gccatcatta ttatccgatg tgacgctgca gcttctcaat gatattcgaa 2400

tacgctttga ggagatacag cctaatatcc gacaaactgt tttacagatt tacgatcgta 2460

cttgttaccc atcattgaat tttgaacatc cgaacctggg agttttccct gaaacagata 2520

gtatatttga acctgtataa taatatatag tctagcgctt tacggaagac aatgtatgta 2580

tttcggttcc tggagaaact attgcatcta ttgcataggt aatcttgcac gtcgcatccc 2640

cggttcattt tctgcgtttc catcttgcac ttcaatagca tatctttgtt aacgaagcat 2700

ctgtgcttca ttttgtagaa caaaaatgca acgcgagagc gctaattttt caaacaaaga 2760

atctgagctg catttttaca gaacagaaat gcaacgcgaa agcgctattt taccaacgaa 2820

gaatctgtgc ttcatttttg taaaacaaaa atgcaacgcg agagcgctaa tttttcaaac 2880

aaagaatctg agctgcattt ttacagaaca gaaatgcaac gcgagagcgc tattttacca 2940

acaaagaatc tatacttctt ttttgttcta caaaaatgca tcccgagagc gctatttttc 3000

taacaaagca tcttagatta ctttttttct cctttgtgcg ctctataatg cagtctcttg 3060

ataacttttt gcactgtagg tccgttaagg ttagaagaag gctactttgg tgtctatttt 3120

ctcttccata aaaaaagcct gactccactt cccgcgttta ctgattacta gcgaagctgc 3180

gggtgcattt tttcaagata aaggcatccc cgattatatt ctataccgat gtggattgcg 3240

catactttgt gaacagaaag tgatagcgtt gatgattctt cattggtcag aaaattatga 3300

acggtttctt ctattttgtc tctatatact acgtatagga aatgtttaca ttttcgtatt 3360

gttttcgatt cactctatga atagttctta ctacaatttt tttgtctaaa gagtaatact 3420

agagataaac ataaaaaatg tagaggtcga gtttagatgc aagttcaagg agcgaaaggt 3480

ggatgggtag gttatatagg gatatagcac agagatatat agcaaagaga tacttttgag 3540

caatgtttgt ggaaagcggt attcgcaatg ccggcaaaag ggcgaattga ttttgaagag 3600

aatgtggatt ttgatgtaat tgttgggatt ccatttttaa taaggcaata atattaggta 3660

tgtggatata ctagaagttc tcctcgaccg tcgatatgcg gtgtgaaata ccgcacagat 3720

gcgtaaggag aaaataccgc atcaggaaat tgtaaacgtt aatattttgt taaaattcgc 3780

gttaaatttt tgttaaatca gctcattttt taaccaatag gccgaaatcg gcaaaatccc 3840

ttataaatca aaagaataga ccgagatagg gttgagtgtt gttccagttt ggaacaagag 3900

tccactatta aagaacgtgg actccaacgt caaagggcga aaaaccgtct atcagggcga 3960

tggcccacta cgtgaaccat caccctaatc aagtttttg gggtcgaggt gccgtaaagc 4020

actaaatcgg aaccctaaag ggagcccccg atttagagct tgacggggaa agccggcgaa 4080

cgtggcgaga aaggaaggga agaaagcgaa aggagcgggc gctagggcgc tggcaagtgt 4140

agcggtcacg ctgcgcgtaa ccaccacacc cgccgcgctt aatgcgccgc tacagggcgc 4200

gtcgcgccat tcgccattca ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc 4260

ttcgctatta cgccagctgg cgaaggggg atgtgctgca aggcgattaa gttgggtaac 4320

gccagggttt tcccagtcac gacgttgtaa aacgacggcc agtgaattgt aatacgactc 4380

actatagggc gaattggagc tccaccgcgg tggcggccgc tctagaacta gtggatcccc 4440

cgggctgcag gaattcgata tcaagcttat cgataccgtc gacctcgacc ggtcccacac 4500

accatagctt caaaatgttt ctactccttt tttactcttc cagattttct cggactccgc 4560

gcatcgccgt accacttcaa aacacccaag cacagcatac taaatttccc ctctttcttc 4620

ctctagggtg tcgttaatta cccgtactaa aggtttggaa aagaaaaaag agaccgcctc 4680

gtttcttttt cttcgtcgaa aaaggcaata aaaatttta tcacgtttct ttttcttgaa 4740

aatttttttt tttgattttt ttctctttcg atgacctccc attgatattt aagttaataa 4800

acggtcttca atttctcaag tttcagtttc atttttcttg ttctattaca actttttta 4860

cttcttgctc attagaaaga aagcatagca atctaatcta agctgcagaa gcttaaaatg 4920

tctgttatcg gtagaatcca ctctttcgaa tcttgtggta ctgttgacgg tccaggtatc 4980

agattcatca ccttcttcca aggttgtttg atgagatgtt tgtactgtca caacagagac 5040

acctgggaca cccacggtgg taaggaagtt actgttgaag acttgatgaa ggaagttgtt 5100

acctacagac actttatgaa cgcttcagga ggtggtgtta ccgcttctgg tggtgaagct 5160

atcttgcaag ctgaatttgt tagagactgg ttcagagctt gtaagaagga aggtatccac 5220

acctgtttgg acaccaacgg tttcgttaga agatacgacc cagttatcga cgaattgttg 5280

gaagttaccg acttggttat gttggacttg aagcaaatga acgacgaaat ccaccaaaac 5340

ttggttggtg tttctaacca cagaaccttg gaatttgcta agtacttggc taacaagaac 5400

gttaaggttt ggatcagata cgttgttgtt ccaggttggt ctgacgacga cgactctgct 5460

cacagattgg gtgagttcac cagagacatg ggtaacgttg aaaagatcga attgttgcca 5520

taccacgaat tgggtaagca caagtgggtt gctatgggtg aagaatacaa gttggacggt 5580

gttaagccac caaagaagga aaccatggaa agagttaagg gtatcttgga acaatacggt 5640

cacaaggtta tgttctaagc gcgcgaattt cttatgattt atgattttta ttattaaata 5700

agttataaaa aaaataagtg tatacaaatt ttaaagtgac tcttaggttt taaaacgaaa 5760

attcttattc ttgagtaact ctttcctgta ggtcaggttg ctttctcagg tatagcatga 5820

ggtcgctctt attgaccaca cctctaccgg catgccgagc aaatgcctgc aaatcgctcc 5880

ccatttcacc caattgtaga tatgctaact ccagcaatga gttgatgaat ctcggtgtgt 5940

attttatgtc ctcagaggac aacacctgtt gtaatcgttc ttccacacgt acccagcttt 6000

tgttcccttt agtgagggtt aattccgagc ttggcgtaat catggtcata gctgtttcct 6060

gtgtgaaatt gttatccgct cacaattcca cacaacatag gagccggaag cataaagtgt 6120

aaagcctggg gtgcctaatg agtgaggtaa ctcacattaa ttgcgttgcg ctcactgccc 6180

gctttccagt cgggaaacct gtcgtgccag ctgcattaat gaatcggcca acgcgcgggg 6240

agaggcggtt tgcgtattgg gcgctcttcc gcttcctcgc tcactgactc gctgcgctcg 6300

gtcgttcggc tgcggcgagc ggtatcagct cactcaaagg cggtaatacg gttatccaca 6360

gaatcagggg ataacgcagg aaagaacatg tgagcaaaag gccagcaaaa ggccaggaac 6420

cgtaaaaagg ccgcgttgct ggcgtttttc cataggctcg gccccctga cgagcatcac 6480

aaaaatcgac gctcaagtca gaggtggcga aacccgacag gactataaag ataccaggcg 6540

ttcccccctg gaagctccct cgtgcgctct cctgttccga ccctgccgct taccggatac 6600

ctgtccgcct ttctcccttc gggaagcgtg gcgctttctc aatgctcacg ctgtaggtat 6660

ctcagttcgg tgtaggtcgt tcgctccaag ctgggctgtg tgcacgaacc ccccgttcag 6720

cccgaccgct gcgccttatc cggtaactat cgtcttgagt ccaacccggt aagacacgac 6780

ttatcgccac tggcagcagc cactggtaac aggattagca gagcgaggta tgtaggcggt 6840

gctacagagt tcttgaagtg gtggcctaac tacggctaca ctagaaggac agtatttggt 6900

atctgcgctc tgctgaagcc agttaccttc ggaaaaagag ttggtagctc ttgatccggc 6960

aaacaaacca ccgctggtag cggtggtttt tttgtttgca agcagcagat tacgcgcaga 7020

aaaaaaggat ctcaagaaga tcctttgatc ttttctacgg ggtctgacgc tcagtggaac 7080

gaaaactcac gttaagggat tttggtcatg agattatcaa aaaggatctt cacctagatc 7140

cttttaaatt aaaaatgaag ttttaaatca atctaaagta tatatgagta aacttggtct 7200

gacagttacc aatgcttaat cagtgaggca cctatctcag cgatctgtct atttcgttca 7260

tccatagttg cctgactgcc cgtcgtgtag ataactacga tacgggaggg cttaccatct 7320

ggccccagtg ctgcaatgat accgcgagac ccacgctcac cggctccaga tttatcagca 7380

ataaaccagc cagccggaag ggccgagcgc agaagtggtc ctgcaacttt atccgcctcc 7440

atccagtcta ttaattgttg ccgggaagct agagtaagta gttcgccagt taatagtttg 7500

cgcaacgttg ttgccattgc tacaggcatc gtggtgtcac gctcgtcgtt tggtatggct 7560

tcattcagct ccggttccca acgatcaagg cgagttacat gatcccccat gttgtgaaaa 7620

aaagcggtta gctccttcgg tcctccgatc gttgtcagaa gtaagttggc cgcagtgtta 7680

tcactcatgg ttatggcagc actgcataat tctcttactg tcatgccatc cgtaagatgc 7740

ttttctgtga ctggtgagta ctcaaccaag tcattctgag aatagtgtat gcggcgaccg 7800

agttgctctt gcccggcgtc aatacgggat 7830

<210> 33
<211> 9695
<212> ADN
<213> Artificial

<220>
<223> pRN620

<400> 33
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accatatcga ctacgtcgta aggccgtttc tgacagagta aaattcttga gggaactttc 240

accattatgg gaaatggttc aagaaggtat tgacttaaac tccatcaaat ggtcaggtca 300

ttgagtgttt tttatttgtt gtatttttt ttttttagag aaaatcctcc aatatcaaat 360

taggaatcgt agtttcatga ttttctgtta cacctaactt tttgtgtggt gccctcctcc 420

ttgtcaatat taatgttaaa gtgcaattct ttttccttat cacgttgagc cattagtatc 480

aatttgctta cctgtattcc tttactatcc tccttttttct ccttcttgat aaatgtatgt 540

agattgcgta tatagtttcg tctaccctat gaacatattc cattttgtaa tttcgtgtcg 600

tttctattat gaatttcatt tataaagttt atgtacacct aggatccgtc gacactggat 660

ggcggcgtta gtatcgaatc gacagcagta tagcgaccag cattcacata cgattgacgc 720

atgatattac tttctgcgca cttaacttcg catctgggca gatgatgtcg aggcgaaaaa 780

aaatataaat cacgctaaca tttgattaaa atagaacaac tacaatataa aaaaactata 840

caaatgacaa gttcttgaaa acaagaatct ttttattgtc agtgtgtcag tcctgctcct 900

cggccacgaa gtgcacgcag ttgccggccg ggtcgcgcag ggcgaactcc cgcccccacg 960

gctgctcgcc gatctcggtc atggccggcc cggaggcgtc ccggaagttc gtggacacga 1020

cctccgacca ctcggcgtac agctcgtcca ggccgcgcac ccacacccag gccagggtgt 1080

tgtccggcac cacctggtcc tggaccgcgc tgatgaacag ggtcacgtcg tcccggacca 1140

caccggcgaa gtcgtcctcc acgaagtccc gggagaaccc gagccggtcg gtccagaact 1200

cgaccgctcc ggcgacgtcg cgcgcggtga gcaccggaac ggcactggtc aacttggcca 1260

tggttgttta tgttcggatg tgatgtgaga actgtatcct agcaagattt taaaaggaag 1320

tatatgaaag aagaacctca gtggcaaatc ctaacctttt atatttctct acaggggcgc 1380

ggcgtgggga caattcaacg cgtctgtgag gggagcgttt ccctgctcgc aggtctgcag 1440

cgaggagccg taatttttgc ttcgcgccgt gcggccatca aaatgtatgg atgcaaatga 1500

ttatacatgg ggatgtatgg gctaaatgta cgggcgacag tcacatcatg cccctgagct 1560

gcgcacgtca agactgtcaa ggagggtatt ctgggcctcc atgtcgctgg ccgggtgacc 1620

cggcggggac gaggccttaa gcggccgcat gctagctccg gattatcgat gataagctgt 1680

caaacatgag aattaattcc acggactata gactatacct agtatactcc gtctactgta 1740

cgatacactt ccgctcaggt ccttgtcctt taacgaggcc ttaccactct tttgttactc 1800

tattgatcca gctcagcaaa ggcagtgtga tctaagattc tatcttcgcg atgtagtaaa 1860

actagctaga ccgagaaaga gactagaaat gcaaaaggca cttctacaat ggctgccatc 1920

attattatcc gatgtgacgc tgcagcttct caatgatatt cgaatacgct ttgaggagat 1980

acagcctaat atccgacaaa ctgttttaca gatttacgat cgtacttgtt acccatcatt 2040

gaattttgaa catccgaacc tgggagtttt ccctgaaaca gatagtatat ttgaacctgt 2100

ataataatat atagtctagc gctttacgga agacaatgta tgtatttcgg ttcctggaga 2160

aactattgca tctattgcat aggtaatctt gcacgtcgca tccccggttc attttctgcg 2220

tttccatctt gcacttcaat agcatatctt tgttaacgaa gcatctgtgc ttcattttgt 2280

agaacaaaaa tgcaacgcga gagcgctaat ttttcaaaca aagaatctga gctgcatttt 2340

tacagaacag aaatgcaacg cgaaagcgct attttaccaa cgaagaatct gtgcttcatt 2400

tttgtaaaac aaaaatgcaa cgcgagagcg ctaatttttc aaacaaagaa tctgagctgc 2460

atttttacag aacagaaatg caacgcgaga gcgctatttt accaacaaag aatctatact 2520

tctttttgt tctacaaaaa tgcatcccga gagcgctatt tttctaacaa agcatcttag 2580

attacttttt ttctcctttg tgcgctctat aatgcagtct cttgataact ttttgcactg 2640

taggtccgtt aaggttagaa gaaggctact ttggtgtcta ttttctcttc cataaaaaaa 2700

gcctgactcc acttcccgcg tttactgatt actagcgaag ctgcgggtgc atttttttcaa 2760

gataaaggca tccccgatta tattctatac cgatgtggat tgcgcatact ttgtgaacag 2820

aaagtgatag cgttgatgat tcttcattgg tcagaaaatt atgaacggtt tcttctattt 2880

tgtctctata tactacgtat aggaaatgtt tacattttcg tattgttttc gattcactct 2940

atgaatagtt cttactacaa ttttttttgtc taaagagtaa tactagagat aaacataaaa 3000

aatgtagagg tcgagtttag atgcaagttc aaggagcgaa aggtggatgg gtaggttata 3060

tagggatata gcacagagat atatagcaaa gagatacttt tgagcaatgt ttgtggaaag 3120

cggtattcgc aatgccggca aaagggcgaa ttgattttga agagaatgtg gattttgatg 3180

taattgttgg gattccattt ttaataaggc aataatatta ggtatgtgga tatactagaa 3240

gttctcctcg accgtcgata tgcggtgtga aataccgcac agatgcgtaa ggagaaaata 3300

ccgcatcagg aaattgtaaa cgttaatatt ttgttaaaat tcgcgttaaa tttttgttaa 3360

atcagctcat tttttaacca ataggccgaa atcggcaaaa tcccttataa atcaaaagaa 3420

tagaccgaga tagggttgag tgttgttcca gtttggaaca agagtccact attaaagaac 3480

gtggactcca acgtcaaagg gcgaaaaacc gtctatcagg gcgatggccc actacgtgaa 3540

ccatcaccct aatcaagttt tttggggtcg aggtgccgta aagcactaaa tcggaaccct 3600

aaagggagcc cccgatttag agcttgacgg ggaaagccgg cgaacgtggc gagaaaggaa 3660

gggaagaaag cgaaaggagc gggcgctagg gcgctggcaa gtgtagcggt cacgctgcgc 3720

gtaaccacca cacccgccgc gcttaatgcg ccgctacagg gcgcgtcgcg ccattcgcca 3780

ttcaggctgc gcaactgttg ggaagggcga tcggtgcggg cctcttcgct attacgccag 3840

ctggcgaagg ggggatgtgc tgcaaggcga ttaagttggg taacgccagg gttttcccag 3900

tcacgacgtt gtaaaacgac ggccagtgaa ttgtaatacg actcactata gggcgaattg 3960

gagctccacc gcggtggcgg ccgctctaga actagtggat cccccgggct gcaggaattc 4020

gatatcaagc ttatcgatac cgtcgacctc gagctcttca actcaagacg cacagatatt 4080

ataacatctg cataataggc atttgcaaga attactcgtg agtaaggaaa gagtgaggaa 4140

ctatcgcata cctgcattta aagatgccga tttgggcgcg aatcctttat tttggcttca 4200

ccctcatact attatcaggg ccagaaaaag gaagtgtttc cctccttctt gaattgatgt 4260

taccctcata aagcacgtgg cctcttatcg agaaagaaat taccgtcgct cgtgatttgt 4320

ttgcaaaaag aacaaaactg aaaaaaccca gacacgctcg acttcctgtc ttcctattga 4380

ttgcagcttc caatttcgtc acacaacaag gtcctagcga cggctcacag gttttgtaac 4440

aagcaatcga aggttctgga atggcgggaa agggtttagt accacatgct atgatgccca 4500

ctgtgatctc cagagcaaag ttcgttcgat cgtactgtta ctctctctct ttcaaacaga 4560

attgtccgaa tcgtgtgaca acaacagcct gttctcacac actcttttct tctaaccaag 4620

ggggtggttt agtttagtag aacctcgtga aacttacatt tacatatata taaacttgca 4680

taaattggtc aatgcaagaa atacatattt ggtcttttct aattcgtagt ttttcaagtt 4740

cttagatgct ttctttttct ctttttttaca gatcatcaag gaagtaatta tctacttttt 4800

acaacaaata tatctagact gcagaaaatg tctgaattga acgagaagtt ggctaccgct 4860

tgggaaggtt tcaccaaggg tgactggcaa aacgaagtta acgttagaga cttcatccaa 4920

aagaactaca ccccatacga aggtgacgaa tctttcttgg ctggtgctac cgaagctacc 4980

accaccttgt gggacaaggt tatggaaggt gttaagttgg aaaacagaac ccacgctcca 5040

gttgacttcg acaccgctgt tgcttctacc atcacctctc acgacgctgg ttacatcaac 5100

aagcaattgg aaaagatcgt tggtttacaa accgaagctc cattgaagag agctttgatc 5160

ccattcggtg gtatcaagat gatcgaaggt tcttgtaagg cttacaacag agaattggac 5220

ccaatgatca agaagatttt caccgaatac agaaagaccc acaaccaagg tgttttcgac 5280

gtttacactc cagacatctt gagatgtaga aagtctggtg ttttgactgg tttgccagac 5340

gcttacggta gaggtagaat catcggtgac tacagaagag ttgctttgta cggtatcgac 5400

tacttgatga aggacaagtt ggctcaattc acctctttgc aagctgactt ggaaaacggt 5460

gttaacttgg aacaaaccat cagattgaga gaagaaatcg ctgaacaaca cagagctttg 5520

ggtcaaatga aggaaatggc tgctaagtac ggttacgaca tctctggtcc agctaccaac 5580

gctcaagaag ctatccaatg gacctacttc ggttacttgg ctgctgttaa gtctcaaaac 5640

ggtgctgcta tgtctttcgg taggacctct accttcttgg acgtttacat cgaaagagac 5700

ttgaaggctg gtaagatcac cgaacaagaa gctcaagaaa tggttgacca cttggttatg 5760

aagttgagaa tggttagatt cttgagaacc ccagaatacg acgaattgtt ctctggtgac 5820

ccaatctggg ctaccgaatc tatcggtggt atgggtttgg acggtagaac cttggttacc 5880

aagaactctt tcagattctt gaacacctta tacaccatgg gtccatctcc agaaccaaac 5940

atgaccatct tgtggtctga aaagttacca ttgaacttca agaagttcgc tgctaaggtt 6000

tctatcgaca cctcttcttt gcaatacgaa aacgacgact tgatgagacc agacttcaac 6060

aacgacgact acgctatcgc ttgttgtgtt tctccaatga tcgttggtaa gcaaatgcaa 6120

ttcttcggtg ctagagctaa cttggctaag accatgttgt acgctatcaa cggtggtgtt 6180

gacgaaaagt tgaagatgca agttggtcca aagtctgaac caatcaaggg tgacgttttg 6240

aactacgacg aagttatgga aagaatggac cacttcatgg actggttggc taagcaatac 6300

atcaccgctt tgaacatcat ccactacatg cacgacaagt actcttacga agcatcattg 6360

atggctttgc acgacagaga cgtaatcaga accatggctt gtggtatcgc tggtttgtct 6420

gttgctgctg actctttgtc tgctatcaag tacgctaagg ttaagccaat cagagacgaa 6480

gacggtttgg ctatcgactt cgaaatcgaa ggtgaatacc ctcaattcgg taacaacgac 6540

ccaagagttg acgacttggc tgttgacttg gttgaaagat ttatgaagaa gatccaaaag 6600

ttgcacacct acagagacgc tatcccaacc caatctgttt tgactatcac atctaacgtt 6660

gtttacggta agaagactgg taacacccca gacggtagaa gagctggtgc tccattcggt 6720

ccaggtgcta acccaatgca cggtagagac caaaagggtg ctgtagcatc tttgacctct 6780

gttgctaagt tgccattcgc ttacgctaag gacggtatct cttacacctt ctctatcgtt 6840

ccaaacgctt tgggtaagga cgatgaagtt agaaagacca acttggctgg tttgatggac 6900

ggttacttcc accacgaagc atctatcgaa ggtggtcaac acttgaacgt aaatgttatg 6960

aacagagaaa tgttgttgga cgctatggaa aacccagaaa agtacccaca attgaccatc 7020

agagtttctg gttacgctgt tagattcaac tctttgacca aggaacaaca acaagacgtt 7080

atcaccagaa ccttcaccca atctatgtaa gtcgagacaa atcgctctta aatatatacc 7140

taaagaacat taaagctata ttataagcaa agatacgtaa attttgctta tattattata 7200

cacatatcat atttctatat ttttaagatt tggttatata atgtacgtaa tgcaaaggaa 7260

ataaatttta tacattattg aacagcgtcc aagtaactac attatgtgca ctaatagttt 7320

agcgtcgtga agactttatt gtgtcgcgaa aagtaaaaat tttaaaaatt agagcacctt 7380

gaacttgcga aaaaggttct catcaactgt ttaaaacgta cccagctttt gttcccttta 7440

gtgagggtta attccgagct tggcgtaatc atggtcatag ctgtttcctg tgtgaaattg 7500

ttatccgctc acaattccac acaacatagg agccggaagc ataaagtgta aagcctgggg 7560

tgcctaatga gtgaggtaac tcacattaat tgcgttgcgc tcactgcccg ctttccagtc 7620

gggaaacctg tcgtgccagc tgcattaatg aatcggccaa cgcgcgggga gaggcggttt 7680

gcgtattggg cgctcttccg cttcctcgct cactgactcg ctgcgctcgg tcgttcggct 7740

gcggcgagcg gtatcagctc actcaaaggc ggtaatacgg ttatccacag aatcagggga 7800

taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc 7860

cgcgttgctg gcgtttttcc ataggctcgg ccccccctgac gagcatcaca aaaatcgacg 7920

ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt tcccccctgg 7980

aagctccctc gtgcgctctc ctgttccgac cctgccgctt accggatacc tgtccgcctt 8040

tctcccttcg ggaagcgtgg cgctttctca atgctcacgc tgtaggtatc tcagttcggt 8100

gtaggtcgtt cgctccaagc tgggctgtgt gcacgaaccc cccgttcagc ccgaccgctg 8160

cgccttatcc ggtaactatc gtcttgagtc caacccggta agacacgact tatcgccact 8220

ggcagcagcc actggtaaca ggattagcag agcgaggtat gtaggcggtg ctacagagtt 8280

cttgaagtgg tggcctaact acggctacac tagaaggaca gtatttggta tctgcgctct 8340

gctgaagcca gttaccttcg gaaaaagagt tggtagctct tgatccggca aacaaaccac 8400

cgctggtagc ggtggttttt ttgtttgcaa gcagcagatt acgcgcagaa aaaaaggatc 8460

tcaagaagat cctttgatct tttctacggg gtctgacgct cagtggaacg aaaactcacg 8520

ttaagggatt ttggtcatga gattatcaaa aaggatcttc acctagatcc ttttaaatta 8580

aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa acttggtctg acagttacca 8640

atgcttaatc agtgaggcac ctatctcagc gatctgtcta tttcgttcat ccatagttgc 8700

ctgactgccc gtcgtgtaga taactacgat acgggagggc ttaccatctg gccccagtgc 8760

tgcaatgata ccgcgagacc cacgctcacc ggctccagat ttatcagcaa taaaccagcc 8820

agccggaagg gccgagcgca gaagtggtcc tgcaacttta tccgcctcca tccagtctat 8880

taattgttgc cgggaagcta gagtaagtag ttcgccagtt aatagtttgc gcaacgttgt 8940

tgccattgct acaggcatcg tggtgtcacg ctcgtcgttt ggtatggctt cattcagctc 9000

cggttcccaa cgatcaaggc gagttacatg atcccccatg ttgtgaaaaa aagcggttag 9060

ctccttcggt cctccgatcg ttgtcagaag taagttggcc gcagtgttat cactcatggt 9120

tatggcagca ctgcataatt ctcttactgt catgccatcc gtaagatgct tttctgtgac 9180

tggtgagtac tcaaccaagt cattctgaga atagtgtatg cggcgaccga gttgctcttg 9240

cccggcgtca atacgggata ataccgcgcc acatagcaga actttaaaag tgctcatcat 9300

tggaaaacgt tcttcggggc gaaaactctc aaggatctta ccgctgttga gatccagttc 9360

gatgtaaccc actcgtgcac ccaactgatc ttcagcatct tttactttca ccagcgtttc 9420

tgggtgagca aaaacaggaa ggcaaaatgc cgcaaaaaag ggaataaggg cgacacggaa 9480

atgttgaata ctcatactct tccttttca atattattga agcatttatc agggttattg 9540

tctcatgagc ggatacatat ttgaatgtat ttagaaaaat aaacaaatag gggttccgcg 9600

cacatttccc cgaaaagtgc cacctgacgt ctaagaaacc attattatca tgacattaac 9660

ctataaaaat aggcgtatca cgaggccctt tcgtc 9695

<210> 34
<211> 14159
<212> ADN
<213> Artificial

<220>
<223> pRN618

<400> 34
tggtatacca acaattcttg aagaagaaga ttccattgaa actgcttcgc taggtgcgac 60

gacgactgat tctattgggt tatccgacac atcatcagaa gattcgcatt atggtaatgc 120

taagaaggta acatgaggat ccccttttcc tttgtcgata tcatgtaatt agttatgtca 180

cgcttacatt cacgccctcc tcccacatcc gctctaaccg aaaaggaagg agttagacaa 240

cctgaagtct aggtccctat ttatttttt taatagttat gttagtatta agaacgttat 300

ttatatttca aatttttctt ttttttctgt acaaacgcgt gtacgcatgt aacattatac 360

tgaaaacctt gcttgagaag gttttgggac gctcgaaggc ttcctaggct cgacactgga 420

tggcggcgtt agtatcgaat cgacagcagt atagcgacca gcattcacat acgattgacg 480

catgatatta ctttctgcgc acttaacttc gcatctgggc agatgatgtc gaggcgaaaa 540

aaaatataaa tcacgctaac atttgattaa aatagaacaa ctacaatata aaaaaactat 600

acaaatgaca agttcttgaa aacaagaatc tttttattgt cagtgtgtca gtcctgctcc 660

tcggccacga agtgcacgca gttgccggcc gggtcgcgca gggcgaactc ccgcccccac 720

ggctgctcgc cgatctcggt catggccggc ccggaggcgt cccggaagtt cgtggacacg 780

acctccgacc actcggcgta cagctcgtcc aggccgcgca cccacaccca ggccagggtg 840

ttgtccggca ccacctggtc ctggaccgcg ctgatgaaca gggtcacgtc gtcccggacc 900

acaccggcga agtcgtcctc cacgaagtcc cgggagaacc cgagccggtc ggtccagaac 960

tcgaccgctc cggcgacgtc gcgcgcggtg agcaccggaa cggcactggt caacttggcc 1020

atggttgttt atgttcggat gtgatgtgag aactgtatcc tagcaagatt ttaaaaggaa 1080

gtatatgaaa gaagaacctc agtggcaaat cctaaccttt tatatttctc tacaggggcg 1140

cggcgtgggg acaattcaac gcgtctgtga ggggagcgtt tccctgctcg caggtctgca 1200

gcgaggagcc gtaatttttg cttcgcgccg tgcggccatc aaaatgtatg gatgcaaatg 1260

attatacatg gggatgtatg ggctaaatgt acgggcgaca gtcacatcat gcccctgagc 1320

tgcgcacgtc aagactgtca aggagggtat tctgggcctc catgtcgctg gccgggtgac 1380

ccggcgggga cgaggcctta agcggccgca tgctagctcc ggattatcga tgataagctg 1440

tcaaacatga gaattaattc cacggactat agactatacc tagtatactc cgtctactgt 1500

acgatacact tccgctcagg tccttgtcct ttaacgaggc cttaccactc ttttgttact 1560

ctattgatcc agctcagcaa aggcagtgtg atctaagatt ctatcttcgc gatgtagtaa 1620

aactagctag accgagaaag agactagaaa tgcaaaaggc acttctacaa tggctgccat 1680

cattattatc cgatgtgacg ctgcagcttc tcaatgatat tcgaatacgc tttgaggaga 1740

tacagcctaa tatccgacaa actgttttac agatttacga tcgtacttgt tacccatcat 1800

tgaattttga acatccgaac ctgggagttt tccctgaaac agatagtata tttgaacctg 1860

tataataata tatagtctag cgctttacgg aagacaatgt atgtatttcg gttcctggag 1920

aaactattgc atctattgca taggtaatct tgcacgtcgc atccccggtt cattttctgc 1980

gtttccatct tgcacttcaa tagcatatct ttgttaacga agcatctgtg cttcattttg 2040

tagaacaaaa atgcaacgcg agagcgctaa tttttcaaac aaagaatctg agctgcattt 2100

ttacagaaca gaaatgcaac gcgaaagcgc tattttacca acgaagaatc tgtgcttcat 2160

ttttgtaaaa caaaaatgca acgcgagagc gctaattttt caaacaaaga atctgagctg 2220

catttttaca gaacagaaat gcaacgcgag agcgctattt taccaacaaa gaatctatac 2280

ttctttttg ttctacaaaa atgcatcccg agagcgctat ttttctaaca aagcatctta 2340

gattactttt tttctccttt gtgcgctcta taatgcagtc tcttgataac tttttgcact 2400

gtaggtccgt taaggttaga agaaggctac tttggtgtct attttctctt ccataaaaaa 2460

agcctgactc cacttcccgc gtttactgat tactagcgaa gctgcgggtg catttttca 2520

agataaaggc atccccgatt atattctata ccgatgtgga ttgcgcatac tttgtgaaca 2580

gaaagtgata gcgttgatga ttcttcattg gtcagaaaat tatgaacggt ttcttctatt 2640

ttgtctctat atactacgta taggaaatgt ttacattttc gtattgtttt cgattcactc 2700

tatgaatagt tcttactaca atttttttgt ctaaagagta atactagaga taaacataaa 2760

aaatgtagag gtcgagttta gatgcaagtt caaggagcga aaggtggatg ggtaggttat 2820

atagggatat agcacagaga tatatagcaa agagatactt ttgagcaatg tttgtggaaa 2880

gcggtattcg caatgccggc aaaagggcga attgattttg aagagaatgt ggattttgat 2940

gtaattgttg ggattccatt tttaataagg caataatatt aggtatgtgg atatactaga 3000

agttctcctc gaccgtcgat atgcggtgtg aaataccgca cagatgcgta aggagaaaat 3060

accgcatcag gaaattgtaa acgttaatat tttgttaaaa ttcgcgttaa atttttgtta 3120

aatcagctca ttttttaacc aataggccga aatcggcaaa atcccttata aatcaaaaga 3180

atagaccgag atagggttga gtgttgttcc agtttggaac aagagtccac tattaaagaa 3240

cgtggactcc aacgtcaaag ggcgaaaaac cgtctatcag ggcgatggcc cactacgtga 3300

accatcaccc taatcaagtt ttttggggtc gaggtgccgt aaagcactaa atcggaaccc 3360

taaagggagc ccccgattta gagcttgacg gggaaagccg gcgaacgtgg cgagaaagga 3420

agggaagaaa gcgaaaggag cgggcgctag ggcgctggca agtgtagcgg tcacgctgcg 3480

cgtaaccacc acacccgccg cgcttaatgc gccgctacag ggcgcgtcgc gccattcgcc 3540

attcaggctg cgcaactgtt gggaagggcg atcggtgcgg gcctcttcgc tattacgcca 3600

gctggcgaag gggggatgtg ctgcaaggcg attaagttgg gtaacgccag ggttttccca 3660

gtcacgacgt tgtaaaacga cggccagtga attgtaatac gactcactat agggcgaatt 3720

ggagctccac cgcggtggcg gccgctctag aactagtctc gagctcttca actcaagacg 3780

cacagatatt ataacatctg cataataggc atttgcaaga attactcgtg agtaaggaaa 3840

gagtgaggaa ctatcgcata cctgcattta aagatgccga tttgggcgcg aatcctttat 3900

tttggcttca ccctcatact attatcaggg ccagaaaaag gaagtgtttc cctccttctt 3960

gaattgatgt taccctcata aagcacgtgg cctcttatcg agaaagaaat taccgtcgct 4020

cgtgatttgt ttgcaaaaag aacaaaactg aaaaaaccca gacacgctcg acttcctgtc 4080

ttcctattga ttgcagcttc caatttcgtc acacaacaag gtcctagcga cggctcacag 4140

gttttgtaac aagcaatcga aggttctgga atggcgggaa agggtttagt accacatgct 4200

atgatgccca ctgtgatctc cagagcaaag ttcgttcgat cgtactgtta ctctctctct 4260

ttcaaacaga attgtccgaa tcgtgtgaca acaacagcct gttctcacac actcttttct 4320

tctaaccaag ggggtggttt agtttagtag aacctcgtga aacttacatt tacatatata 4380

taaacttgca taaattggtc aatgcaagaa atacatattt ggtcttttct aattcgtagt 4440

ttttcaagtt cttagatgct ttctttttct ctttttaca gatcatcaag gaagtaatta 4500

tctacttttt acaacaaata tatctagact gcagaaaatg tctgaattga acgagaagtt 4560

ggctaccgct tgggaaggtt tcaccaaggg tgactggcaa aacgaagtta acgttagaga 4620

cttcatccaa aagaactaca ccccatacga aggtgacgaa tctttcttgg ctggtgctac 4680

cgaagctacc accaccttgt gggacaaggt tatggaaggt gttaagttgg aaaacagaac 4740

ccacgctcca gttgacttcg acaccgctgt tgcttctacc atcacctctc acgacgctgg 4800

ttacatcaac aagcaattgg aaaagatcgt tggtttacaa accgaagctc cattgaagag 4860

agctttgatc ccattcggtg gtatcaagat gatcgaaggt tcttgtaagg cttacaacag 4920

agaattggac ccaatgatca agaagatttt caccgaatac agaaagaccc acaaccaagg 4980

tgttttcgac gtttacactc cagacatctt gagatgtaga aagtctggtg ttttgactgg 5040

tttgccagac gcttacggta gaggtagaat catcggtgac tacagaagag ttgctttgta 5100

cggtatcgac tacttgatga aggacaagtt ggctcaattc acctctttgc aagctgactt 5160

ggaaaacggt gttaacttgg aacaaaccat cagattgaga gaagaaatcg ctgaacaaca 5220

cagagctttg ggtcaaatga aggaaatggc tgctaagtac ggttacgaca tctctggtcc 5280

agctaccaac gctcaagaag ctatccaatg gacctacttc ggttacttgg ctgctgttaa 5340

gtctcaaaac ggtgctgcta tgtctttcgg taggacctct accttcttgg acgtttacat 5400

cgaaagagac ttgaaggctg gtaagatcac cgaacaagaa gctcaagaaa tggttgacca 5460

cttggttatg aagttgagaa tggttagatt cttgagaacc ccagaatacg acgaattgtt 5520

ctctggtgac ccaatctggg ctaccgaatc tatcggtggt atgggtttgg acggtagaac 5580

cttggttacc aagaactctt tcagattctt gaacacctta tacaccatgg gtccatctcc 5640

agaaccaaac atgaccatct tgtggtctga aaagttacca ttgaacttca agaagttcgc 5700

tgctaaggtt tctatcgaca cctcttcttt gcaatacgaa aacgacgact tgatgagacc 5760

agacttcaac aacgacgact acgctatcgc ttgttgtgtt tctccaatga tcgttggtaa 5820

gcaaatgcaa ttcttcggtg ctagagctaa cttggctaag accatgttgt acgctatcaa 5880

cggtggtgtt gacgaaaagt tgaagatgca agttggtcca aagtctgaac caatcaaggg 5940

tgacgttttg aactacgacg aagttatgga aagaatggac cacttcatgg actggttggc 6000

taagcaatac atcaccgctt tgaacatcat ccactacatg cacgacaagt actcttacga 6060

agcatcattg atggctttgc acgacagaga cgtaatcaga accatggctt gtggtatcgc 6120

tggtttgtct gttgctgctg actctttgtc tgctatcaag tacgctaagg ttaagccaat 6180

cagagacgaa gacggtttgg ctatcgactt cgaaatcgaa ggtgaatacc ctcaattcgg 6240

taacaacgac ccaagagttg acgacttggc tgttgacttg gttgaaagat ttatgaagaa 6300

gatccaaaag ttgcacacct acagagacgc tatcccaacc caatctgttt tgactatcac 6360

atctaacgtt gtttacggta agaagactgg taacacccca gacggtagaa gagctggtgc 6420

tccattcggt ccaggtgcta acccaatgca cggtagagac caaaagggtg ctgtagcatc 6480

tttgacctct gttgctaagt tgccattcgc ttacgctaag gacggtatct cttacacctt 6540

ctctatcgtt ccaaacgctt tgggtaagga cgatgaagtt agaaagacca acttggctgg 6600

tttgatggac ggttacttcc accacgaagc atctatcgaa ggtggtcaac acttgaacgt 6660

aaatgttatg aacagagaaa tgttgttgga cgctatggaa aacccagaaa agtacccaca 6720

attgaccatc agagtttctg gttacgctgt tagattcaac tctttgacca aggaacaaca 6780

acaagacgtt atcaccagaa ccttcaccca atctatgtaa gtcgagacaa atcgctctta 6840

aatatatacc taaagaacat taaagctata ttataagcaa agatacgtaa attttgctta 6900

tattattata cacatatcat atttctatat ttttaagatt tggttatata atgtacgtaa 6960

tgcaaaggaa ataaatttta tacattattg aacagcgtcc aagtaactac attatgtgca 7020

ctaatagttt agcgtcgtga agactttatt gtgtcgcgaa aagtaaaaat tttaaaaatt 7080

agagcacctt gaacttgcga aaaaggttct catcaactgt ttaaaacgta cgtgtggaag 7140

aacgattaca acaggtgttg tcctctgagg acataaaata cacaccgaga ttcatcaact 7200

cattgctgga gttagcatat ctacaattgg gtgaaatggg gagcgatttg caggcatttg 7260

ctcggcatgc cggtagaggt gtggtcaata agagcgacct catgctatac ctgagaaagc 7320

aacctgacct acaggaaaga gttactcaag aataagaatt ttcgttttaa aacctaagag 7380

tcactttaaa atttgtatac acttattttt tttataactt atttaataat aaaaatcata 7440

aatcataaga aattcgcgcg cttagaacat aaccttgtga ccgtattgtt ccaagatacc 7500

cttaactctt tccatggttt ccttctttgg tggcttaaca ccgtccaact tgtattcttc 7560

acccatagca acccacttgt gcttacccaa ttcgtggtat ggcaacaatt cgatcttttc 7620

aacgttaccc atgtctctgg tgaactcacc caatctgtga gcagagtcgt cgtcgtcaga 7680

ccaacctgga acaacaacgt atctgatcca aaccttaacg ttcttgttag ccaagtactt 7740

agcaaattcc aaggttctgt ggttagaaac accaaccaag ttttggtgga tttcgtcgtt 7800

catttgcttc aagtccaaca taaccaagtc ggtaacttcc aacaattcgt cgataactgg 7860

gtcgtatctt ctaacgaaac cgttggtgtc caaacaggtg tggatacctt ccttcttaca 7920

agctctgaac cagtctctaa caaattcagc ttgcaagata gcttcaccac cagaagcggt 7980

aacaccacct cctgaagcgt tcataaagtg tctgtaggta acaacttcct tcatcaagtc 8040

ttcaacagta acttccttac caccgtgggt gtcccaggtg tctctgttgt gacagtacaa 8100

acatctcatc aaacaacctt ggaagaaggt gatgaatctg atacctggac cgtcaacagt 8160

accacaagat tcgaaagagt ggattctacc gataacagac attttaagct tctgcagctt 8220

agattagatt gctatgcttt ctttctaatg agcaagaagt aaaaaaagtt gtaatagaac 8280

aagaaaaatg aaactgaaac ttgagaaatt gaagaccgtt tattaactta aatatcaatg 8340

ggaggtcatc gaaagagaaa aaaatcaaaa aaaaaaattt tcaagaaaaa gaaacgtgat 8400

aaaaattttt attgcctttt tcgacgaaga aaaagaaacg aggcggtctc ttttttcttt 8460

tccaaacctt tagtacgggt aattaacgac accctagagg aagaaagagg ggaaatttag 8520

tatgctgtgc ttgggtgttt tgaagtggta cggcgatgcg cggagtccga gaaaatctgg 8580

aagagtaaaa aaggagtaga aacattttga agctatggtg tgtgggaccg gtcgaggggg 8640

ggcccggtac ccagcttttg ttccctttag tgagggttaa ttccgagctt ggcgtaatca 8700

tggtcatagc tgtttcctgt gtgaaattgt tatccgctca caattccaca caacatagga 8760

gccggaagca taaagtgtaa agcctggggt gcctaatgag tgaggtaact cacattaatt 8820

gcgttgcgct cactgcccgc tttccagtcg ggaaacctgt cgtgccagct gcattaatga 8880

atcggccaac gcgcggggag aggcggtttg cgtattgggc gctcttccgc ttcctcgctc 8940

actgactcgc tgcgctcggt cgttcggctg cggcgagcgg tatcagctca ctcaaaggcg 9000

gtaatacggt tatccacaga atcaggggat aacgcaggaa agaacatgtg agcaaaaggc 9060

cagcaaaagg ccaggaaccg taaaaaggcc gcgttgctgg cgtttttcca taggctcggc 9120

ccccctgacg agcatcacaa aaatcgacgc tcaagtcaga ggtggcgaaa cccgacagga 9180

ctataaagat accaggcgtt ccccccctgga agctccctcg tgcgctctcc tgttccgacc 9240

ctgccgctta ccggatacct gtccgccttt ctcccttcgg gaagcgtggc gctttctcaa 9300

tgctcacgct gtaggtatct cagttcggtg taggtcgttc gctccaagct gggctgtgtg 9360

cacgaacccc ccgttcagcc cgaccgctgc gccttatccg gtaactatcg tcttgagtcc 9420

aacccggtaa gacacgactt atcgccactg gcagcagcca ctggtaacag gattagcaga 9480

gcgaggtatg taggcggtgc tacagagttc ttgaagtggt ggcctaacta cggctacact 9540

agaaggacag tatttggtat ctgcgctctg ctgaagccag ttaccttcgg aaaaagagtt 9600

ggtagctctt gatccggcaa acaaaccacc gctggtagcg gtggtttttt tgtttgcaag 9660

cagcagatta cgcgcagaaa aaaaggatct caagaagatc ctttgatctt ttctacgggg 9720

tctgacgctc agtggaacga aaactcacgt taagggattt tggtcatgag attatcaaaa 9780

aggatcttca cctagatcct tttaaattaa aaatgaagtt ttaaatcaat ctaaagtata 9840

tatgagtaaa cttggtctga cagttaccaa tgcttaatca gtgaggcacc tatctcagcg 9900

atctgtctat ttcgttcatc catagttgcc tgactgcccg tcgtgtagat aactacgata 9960

cgggagggct taccatctgg ccccagtgct gcaatgatac cgcgagaccc acgctcaccg 10020

gctccagatt tatcagcaat aaaccagcca gccggaaggg ccgagcgcag aagtggtcct 10080

gcaactttat ccgcctccat ccagtctatt aattgttgcc gggaagctag agtaagtagt 10140

tcgccagtta atagtttgcg caacgttgtt gccattgcta caggcatcgt ggtgtcacgc 10200

tcgtcgtttg gtatggcttc attcagctcc ggttcccaac gatcaaggcg agttacatga 10260

tcccccatgt tgtgaaaaaa agcggttagc tccttcggtc ctccgatcgt tgtcagaagt 10320

aagttggccg cagtgttatc actcatggtt atggcagcac tgcataattc tcttactgtc 10380

atgccatccg taagatgctt ttctgtgact ggtgagtact caaccaagtc attctgagaa 10440

tagtgtatgc ggcgaccgag ttgctcttgc ccggcgtcaa tacgggataa taccgcgcca 10500

catagcagaa ctttaaaagt gctcatcatt ggaaaacgtt cttcggggcg aaaactctca 10560

aggatcttac cgctgttgag atccagttcg atgtaaccca ctcgtgcacc caactgatct 10620

tcagcatctt ttactttcac cagcgtttct gggtgagcaa aaacaggaag gcaaaatgcc 10680

gcaaaaaagg gaataagggc gacacggaaa tgttgaatac tcatactctt cctttttcaa 10740

tattattgaa gcatttatca gggttattgt ctcatgagcg gatacatatt tgaatgtatt 10800

tagaaaaata aacaaatagg ggttccgcgc acatttcccc gaaaagtgcc acctgacgtc 10860

taagaaacca ttattatcat gacattaacc tataaaaata ggcgtatcac gaggcccttt 10920

cgtctcgcgc gtttcggtga tgacggtgaa aacctctgac acatgcagct cccggagacg 10980

gtcacagctt gtctgtaagc ggatgccggg agcagacaag cccgtcaggg cgcgtcagcg 11040

ggtgttggcg ggtgtcgggg ctggcttaac tatgcggcat cagagcagat tgtactgaga 11100

gtgcaccata tcgactacgt cgtaaggccg tttctgacag agtaaaattc ttgagggaac 11160

tttcaccatt atgggaaatg gttcaagaag gtattgactt aaactccatc aaatggtcag 11220

gtcattgagt gtttttttatt tgttgtattt ttttttttt agagaaaatc ctccaatatc 11280

aaattaggaa tcgtagtttc atgattttct gttacaccta acttttttgtg tggtgccctc 11340

ctccttgtca atattaatgt taaagtgcaa ttctttttcc ttatcacgtt gagccattag 11400

tatcaatttg cttacctgta ttcctttact atcctccttt ttctccttct tgataaatgt 11460

atgtagattg cgtatatagt ttcgtctacc ctatgaacat attccatttt gtaatttcgt 11520

gtcgtttcta ttatgaattt catttataaa gtttatgtac gaattctatc cttttgttgt 11580

ttccgggtgt acaatatgga cttcctcttt tctggcaacc aaacccatac atcgggattc 11640

ctataatacc ttcgttggtc tccctaacat gtaggtggcg gaggggagat atacaataga 11700

acagatacca gacaagacat aatgggctaa acaagactac accaattaca ctgcctcatt 11760

gatggtggta cataacgaac taatactgta gccctagact tgatagccat catcatatcg 11820

aagtttcact acccttttc catttgccat ctattgaagt aataataggc gcatgcaact 11880

tcttttcttt ttttttcttt tctctctccc ccgttgttgt ctcaccatat ccgcaatgac 11940

aaaaaaatga tggaagacac taaaggaaaa aattaacgac aaagacagca ccaacagatg 12000

tcgttgttcc agagctgatg aggggtatct cgaagcacac gaaacttttt ccttccttca 12060

ttcacgcaca ctactctcta atgagcaacg gtatacggcc ttccttccag ttacttgaat 12120

ttgaaataaa aaaagtttgc tgtcttgcta tcaagtataa atagacctgc aattattaat 12180

cttttgtttc ctcgtcattg ttctcgttcc ctttcttcct tgtttctttt tctgcacaat 12240

atttcaagct ataccaagca tacaatcaac tccagctgca ttaaaatgag taatcctcaa 12300

aaagctctaa acgactttct gtccagtgaa tctgttcata cacatgatag ttctaggaaa 12360

caatctaata agcagtcatc cgacgaagga cgctcttcat cacaaccttc acatcatcac 12420

tctggtggta ctaacaacaa taataacaat aataataata ataataacag taacaacaac 12480

aacaacggca acgatggggg aaatgatgac gactatgatt atgaaatgca agattataga 12540

ccttctccgc aaagtgcgcg gcctactccc acgtatgttc cacaatattc tgtagaaagt 12600

gggactgctt tcccgattca agaggttatt cctagcgcat acattaacac acaagatata 12660

aaccataaag ataacggtcc gccgagtgca agcagtaata gagcattcag gcctagaggg 12720

cagaccacag tgtcggccaa cgtgcttaac attgaagatt tttacaaaaa tgcagacgat 12780

gcgcatacca tcccggagtc acatttatcg agaaggagaa gtaggtcgag ggctacgagt 12840

aatgctgggc acagtgccaa tacaggcgcc acgaatggca ggactactgg tgcccaaact 12900

aatatggaaa gcaatgaatc accacgtaac gtccccatta tggtgaagcc aaagacatta 12960

taccagaacc ctcaaacacc tacagtcttg ccctccacat accatccaat taataaatgg 13020

tcttccgtca aaaacactta tttgaaggaa tttttagccg agtttatggg aacaatggtt 13080

atgattattt tcggtagtgc tgttgtttgt caggtcaatg ttgctgggaa aatacagcag 13140

gacaatttca acgtggcttt ggataacctt aacgttaccg ggtcttctgc agaaacgata 13200

gacgctatga agagtttaac atccttggtt tcatccgttg cgggcggtac ctttgatgat 13260

gtggcattgg gctgggctgc tgccgtggtg atgggctatt tctgcgctgg tggtagtgcc 13320

atctcaggtg ctcatttgaa tccgtctatt acattagcca atttggtgta tagaggtttt 13380

cccctgaaga aagttcctta ttactttgct ggacaattga tcggtgcctt cacaggcgct 13440

ttgatcttgt ttatttggta caaaagggtg ttacaagagg catatagcga ttggtggatg 13500

aatgaaagtg ttgcgggaat gttttgcgtt tttccaaagc cttatctaag ttcaggacgg 13560

caatttttt ccgaattttt atgtggagct atgttacaag caggaacatt tgcgctgacc 13620

gatccttata cgtgtttgtc ctctgatgtt ttcccattga tgatgtttat tttgattttc 13680

attatcaatg cttccatggc ttatcagaca ggtacagcaa tgaatttggc tcgtgatctg 13740

ggcccacgtc ttgcactata tgcagttgga tttgatcata aaatgctttg ggtgcatcat 13800

catcatttct tttgggttcc catggtaggc ccatttattg gtgcgttaat gggggggttg 13860

gtttacgatg tctgtattta tcagggtcat gaatctccag tcaactggtc tttaccagtt 13920

tataaggaaa tgattatgag agcctggttt agaaggcctg gttggaagaa gagaaataga 13980

gcaagaagaa catcggacct gagtgacttc tcatacaata acgatgatga tgaggaattt 14040

ggagaaagaa tggctcttca aaagacaaag accaagtcat ctatttcaga caacgaaaat 14100

gaagcaggag aaaagaaagt gcaatttaaa tctgttcagc gcggcaaaag aacgtttgg 14159

<210> 35
<211> 31
<212> ADN
<213> Artificial

<220>
<223> cebador FDHuf

<400> 35
tcgaagactc cgaatgaaaa agacatgcca g 31

<210> 36
<211> 33
<212> ADN
<213> Artificial

<220>
<223> cebador FDHur

<400> 36
tccggatacc aagttcattt tcaatacacc cca 33

<210> 37
<211> 27
<212> ADN
<213> Artificial

<220>
<223> cebador FDHdf

<400> 37
atgcatgcag aatggttctt atgccac 27

<210> 38
<211> 30

<212> ADN
<213> Artificial

<220>

<223> cebador FDHdr

<400> 38

gaagacagtt ctgttattaa cgacgagcca 30

<210> 39
<211> 5577
<212> ADN
<213> Artificial

<220>
<223> pRN621

<400> 39

tctgtgcggt atttcacacc gcatacaggt ggcacttttc ggggaaatgt gcgcggaacc 60

cctatttgtt tatttttcta aatacattca aatatgtatc cgctcatgag acaataaccc 120

tgataaatgc ttcaataata gcacgtgagg agggccacca tggccaagtt gaccagtgcc 180

gttccggtgc tcaccgcgcg cgacgtcgcc ggagcggtcg agttctggac cgaccggctc 240

gggttctccc gggacttcgt ggaggacgac ttcgccggtg tggtccggga cgacgtgacc 300

ctgttcatca gcgcggtcca ggaccaggtg gtgccggaca acaccctggc ctgggtgtgg 360

gtgcgcggcc tggacgagct gtacgccgag tggtcggagg tcgtgtccac gaacttccgg 420

gacgcctccg ggccggccat gaccgagatc ggcgagcagc cgtgggggcg ggagttcgcc 480

ctgcgcgacc cggccggcaa ctgcgtgcac ttcgtggccg aggagcagga ctgacacgtg 540

ctaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttttga taatctcatg 600

accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt agaaaagatc 660

aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca aacaaaaaaa 720

ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct ttttccgaag 780

gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta gccgtagtta 840

ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct aatcctgtta 900

ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc aagacgatag 960

ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca gcccagcttg 1020

gagcgaacga cctacaccga actgagatac ctacagcgtg agctatgaga aagcgccacg 1080

cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg aacaggagag 1140

cgcacgaggg agcttccagg gggaaacgcc tggtatcttt atagtcctgt cgggtttcgc 1200

cacctctgac ttgagcgtcg atttttgtga tgctcgtcag gggggcggag cctatggaaa 1260

aacgccagca acgcggcctt tttacggttc ctgggctttt gctggccttt tgctcacatg 1320

ttctttcctg cgttatcccc tgattctgtg gataaccgta ttaccgcctt tgagtgagct 1380

gataccgctc gccgcagccg aacgaccgag cgcagcgagt cagtgagcga ggaagcggaa 1440

gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta atgcagctgg 1500

cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa tgtgagttag 1560

ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcgtat gttgtgtgga 1620

attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta cgccaagcta 1680

tttaggtgac actatagaat actcaagcta tgcatcaagc ttggtaccga gctcggatcc 1740

actagtaacg gccgccagtg tgctggaatt cgcccttcga agactccgaa tgaaaaagac 1800

atgccagtaa taaaactatt ttgatgttat gcggaatata ctattcttgg attattcact 1860

gttaactaaa agttggagaa atcactctgc actgtcaatc attgaaaaaa agaacatata 1920

aaagggcaca aaattgagtc ttttttaatg agttcttgct gaggaaagtt tagttaatat 1980

atcatttacg taaaatatgc atattcttgt attgtgcttt ttttattcat tttaagcagg 2040

aacaatttac aagtattgca acgctaatca aatcaaaata acagctgaaa attaatatgt 2100

cgaagggaaa ggttttgctg gttctttacg aaggtggtaa gcatgctgaa gagcaggaaa 2160

agttattggg gtgtattgaa aatgaacttg gtatccggag ctcgtacgtt cgaacttaag 2220

gcctcgtccc cgccgggtca cccggccagc gacatggagg cccagaatac cctccttgac 2280

agtcttgacg tgcgcagctc aggggcatga tgtgactgtc gcccgtacat ttagcccata 2340

catccccatg tataatcatt tgcatccata cattttgatg gccgcacggc gcgaagcaaa 2400

aattacggct cctcgctgca gacctgcgag cagggaaacg ctcccctcac agacgcgttg 2460

aattgtcccc acgccgcgcc cctgtagaga aatataaaag gttaggattt gccactgagg 2520

ttcttctttc atatacttcc ttttaaaatc ttgctaggat acagttctca catcacatcc 2580

gaacataaac aaccgggtag cccagaacga cgcccggtcg agatccgtcc cgccaccgcc 2640

gccgacatgg cggcggtctg cgacatcgtc aatcactaca tcgagacgag cacggtcaac 2700

ttccgtacgg agccgcagac tccgcaggag tggatcgacg acctggagcg cctccaggac 2760

cgctacccct ggctcgtcgc cgaggtggag ggcgtcgtcg ccggcatcgc ctacgccggc 2820

ccctggaagg cccgcaacgc ctacgactgg accgtcgagt cgacggtgta cgtctcccac 2880

cggcaccagc ggctcggact gggctccacc ctctacaccc acctgctgaa gtccatggag 2940

acccagggct tcaagagcgt ggtcgccgtc atcggactgc ccaacgaccc gagcgtgcgc 3000

ctgcacgagg cgctcggata caccgcgcgc gggacgctgc gggcagccgg ctacaagcac 3060

gggggctggc acgacgtggg gttctggcag cgcgacttct agctgccggc cccgccccgc 3120

cccgtccggc ccgtcacaca gatctaatca atactgacaa taaaaagatt cttgttttca 3180

agaacttgtc atttgtatag ttttttttata ttgtagttgt tctattttaa tcaaatgtta 3240

gcgtgattta tatttttttt cgcctcgaca tcatctgccc agatgcgaag ttaagtgcgc 3300

agaaagtaat atcatgcgtc aatcgtatgt gaatgctggt cgctatactg ctgtcgattc 3360

gatactaacg ccgccatcca gtgtcgacgg atcctaggtg tacagggccc aagggcgaat 3420

tctgcagata tccatcacac tggcggccgc tcgagcatgc atgcagaatg gttcttatgc 3480

caccagagct tatggacaga agaaataaga gtgattatga gtatttgtga gcagaagttt 3540

tccggtctcc ttttgttctt gttttggcgt attctccact attcgtccat agcacattta 3600

taccttagct aaatattttg taaagcaaaa ttttcgttat ctcttaaaaa atagaagagc 3660

ggtttattaa tatcaaataa ttgaaactgc tgatatggta gctatataca aaatctgctg 3720

tcaaaatttg gcagtaaacg atcttcacgg tagcggttca aataaagagg aaaagtcttt 3780

ctcccttact gtttttctgg aatttggctc gtcgttaata acagaactgt cttcgaaggg 3840

cgaattctgc agatatccat cacactggcg gccgctcgag catgcatcta gagggcccaa 3900

ttcgccctat agtgagtcgt attacaattc actggccgtc gttttacaac gtcgtgactg 3960

ggaaaaccct ggcgttaccc aacttaatcg ccttgcagca catccccctt tcgccagctg 4020

gcgtaatagc gaagaggccc gcaccgatcg cccttcccaa cagttgcgca gcctatacgt 4080

acggcagttt aaggtttaca cctataaaag agagagccgt tatcgtctgt ttgtggatgt 4140

acagagtgat attattgaca cgccggggcg acggatggtg atcccccctgg ccagtgcacg 4200

tctgctgtca gataaagtct cccgtgaact ttacccggtg gtgcatatcg gggatgaaag 4260

ctggcgcatg atgaccaccg atatggccag tgtgccggtc tccgttatcg gggaagaagt 4320

ggctgatctc agccaccgcg aaaatgacat caaaaacgcc attaacctga tgttctgggg 4380

aatataaatg tcaggcatga gattatcaaa aaggatcttc acctagatcc ttttcacgta 4440

gaaagccagt ccgcagaaac ggtgctgacc ccggatgaat gtcagctact gggctatctg 4500

gacaagggaa aacgcaagcg caaagagaaa gcaggtagct tgcagtgggc ttacatggcg 4560

atagctagac tgggcggttt tatggacagc aagcgaaccg gaattgccag ctggggcgcc 4620

ctctggtaag gttgggaagc cctgcaaagt aaactggatg gctttctcgc cgccaaggat 4680

ctgatggcgc aggggatcaa gctctgatca agagacagga tgaggatcgt ttcgcatgat 4740

tgaacaagat ggattgcacg caggttctcc ggccgcttgg gtggagaggc tattcggcta 4800

tgactgggca caacagacaa tcggctgctc tgatgccgcc gtgttccggc tgtcagcgca 4860

ggggcgcccg gttctttttg tcaagaccga cctgtccggt gccctgaatg aactgcaaga 4920

cgaggcagcg cggctatcgt ggctggccac gacgggcgtt ccttgcgcag ctgtgctcga 4980

cgttgtcact gaagcgggaa gggactggct gctattgggc gaagtgccgg ggcaggatct 5040

cctgtcatct caccttgctc ctgccgagaa agtatccatc atggctgatg caatgcggcg 5100

gctgcatacg cttgatccgg ctacctgccc attcgaccac caagcgaaac atcgcatcga 5160

gcgagcacgt actcggatgg aagccggtct tgtcgatcag gatgatctgg acgaagagca 5220

tcaggggctc gcgccagccg aactgttcgc caggctcaag gcgagcatgc ccgacggcga 5280

ggatctcgtc gtgacccatg gcgatgcctg cttgccgaat atcatggtgg aaaatggccg 5340

cttttctgga ttcatcgact gtggccggct gggtgtggcg gaccgctatc aggacatagc 5400

gttggctacc cgtgatattg ctgaagagct tggcggcgaa tgggctgacc gcttcctcgt 5460

gctttacggt atcgccgctc ccgattcgca gcgcatcgcc ttctatcgcc ttcttgacga 5520

gttcttctga attattaacg cttacaattt cctgatgcgg tattttctcc ttacgca 5577

<210> 40
<211> 5830
<212> ADN
<213> Artificial

<220>
<223> pRN622

<400> 40
tatttcacac cgcatacagg tggcactttt cggggaaatg tgcgcggaac ccctatttgt 60

ttatttttct aaatacattc aaatatgtat ccgctcatga gacaataacc ctgataaatg 120

cttcaataat agcacgtgag gagggccacc atggccaagt tgaccagtgc cgttccggtg 180

ctcaccgcgc gcgacgtcgc cggagcggtc gagttctgga ccgaccggct cgggttctcc 240

cgggacttcg tggaggacga cttcgccggt gtggtccggg acgacgtgac cctgttcatc 300

agcgcggtcc aggaccaggt ggtgccggac aacaccctgg cctgggtgtg ggtgcgcggc 360

ctggacgagc tgtacgccga gtggtcggag gtcgtgtcca cgaacttccg ggacgcctcc 420

gggccggcca tgaccgagat cggcgagcag ccgtgggggc gggagttcgc cctgcgcgac 480

ccggccggca actgcgtgca cttcgtggcc gaggagcagg actgacacgt gctaaaactt 540

catttttaat ttaaaaggat ctaggtgaag atccttttttg ataatctcat gaccaaaatc 600

ccttaacgtg agttttcgtt ccactgagcg tcagaccccg tagaaaagat caaaggatct 660

tcttgagatc ctttttttct gcgcgtaatc tgctgcttgc aaacaaaaaa accaccgcta 720

ccagcggtgg tttgtttgcc ggatcaagag ctaccaactc tttttccgaa ggtaactggc 780

ttcagcagag cgcagatacc aaatactgtc cttctagtgt agccgtagtt aggccaccac 840

ttcaagaact ctgtagcacc gcctacatac ctcgctctgc taatcctgtt accagtggct 900

gctgccagtg gcgataagtc gtgtcttacc gggttggact caagacgata gttaccggat 960

aaggcgcagc ggtcgggctg aacgggggggt tcgtgcacac agcccagctt ggagcgaacg 1020

acctacaccg aactgagata cctacagcgt gagctatgag aaagcgccac gcttcccgaa 1080

gggagaaagg cggacaggta tccggtaagc ggcagggtcg gaacaggaga gcgcacgagg 1140

gagcttccag ggggaaacgc ctggtatctt tatagtcctg tcgggtttcg ccacctctga 1200

cttgagcgtc gatttttgtg atgctcgtca gggggggcgga gcctatggaa aaacgccagc 1260

aacgcggcct ttttacggtt cctgggcttt tgctggcctt ttgctcacat gttctttcct 1320

gcgttatccc ctgattctgt ggataaccgt attaccgcct ttgagtgagc tgataccgct 1380

cgccgcagcc gaacgaccga gcgcagcgag tcagtgagcg aggaagcgga agagcgccca 1440

atacgcaaac cgcctctccc cgcgcgttgg ccgattcatt aatgcagctg gcacgacagg 1500

tttcccgact ggaaagcggg cagtgagcgc aacgcaatta atgtgagtta gctcactcat 1560

taggcacccc aggctttaca ctttatgctt ccggctcgta tgttgtgtgg aattgtgagc 1620

ggataacaat ttcacacagg aaacagctat gaccatgatt acgccaagct atttaggtga 1680

cactatagaa tactcaagct atgcatcaag cttggtaccg agctcggatc cactagtaac 1740

ggccgccagt gtgctggaat tcgcccttc gaagactccg aatgaaaaag acatgccagt 1800

aataaaaata attgatgtta tgcggaatat actattcttg gattattcac tgttaactaa 1860

aagttggaga aatcactctg cactgtcaat cattgaaaaa aagaacatat aaaagggcac 1920

aaaatcgagt ctttttaat gagttcttgc tgaggaaaat ttagttaata tatcatttac 1980

ataaaacatg catattattg tgttgtactt tctttattca ttttaagcag gaataattac 2040

aagtattgca acgctaatca aatcgaaata acagctgaaa attaatatgt cgaagggaaa 2100

ggttttgctg gttctttatg aaggtggtaa gcatgctgaa gagcaggaaa agttattggg 2160

gtgtattgaa aatgaacttg gtatccggag ctcgtacgtt cgaacttaag gcctcgtccc 2220

cgccgggtca cccggccagc gacatggagg cccagaatac cctccttgac agtcttgacg 2280

tgcgcagctc aggggcatga tgtgactgtc gcccgtacat ttagcccata catccccatg 2340

tataatcatt tgcatccata cattttgatg gccgcacggc gcgaagcaaa aattacggct 2400

cctcgctgca gacctgcgag cagggaaacg ctcccctcac agacgcgttg aattgtcccc 2460

acgccgcgcc cctgtagaga aatataaaag gttaggattt gccactgagg ttcttctttc 2520

atatacttcc ttttaaaatc ttgctaggat acagttctca catcacatcc gaacataaac 2580

aaccatgggt aaggaaaaga ctcacgtttc gaggccgcga ttaaattcca acatggatgc 2640

tgatttatat gggtataaat gggctcgcga taatgtcggg caatcaggtg cgacaatcta 2700

tcgattgtat gggaagcccg atgcgccaga gttgtttctg aaacatggca aaggtagcgt 2760

tgccaatgat gttacagatg agatggtcag actaaactgg ctgacggaat ttatgcctct 2820

tccgaccatc aagcatttta tccgtactcc tgatgatgca tggttactca ccactgcgat 2880

ccccggcaaa acagcattcc aggtattaga agaatatcct gattcaggtg aaaatattgt 2940

tgatgcgctg gcagtgttcc tgcgccggtt gcattcgatt cctgtttgta attgtccttt 3000

taacagcgat cgcgtatttc gtctcgctca ggcgcaatca cgaatgaata acggtttggt 3060

tgatgcgagt gattttgatg acgagcgtaa tggctggcct gttgaacaag tctggaaaga 3120

aatgcataag cttttgccat tctcaccgga ttcagtcgtc actcatggtg atttctcact 3180

tgataacctt atttttgacg aggggaaatt aataggttgt attgatgttg gacgagtcgg 3240

aatcgcagac cgataccagg atcttgccat cctatggaac tgcctcggtg agttttctcc 3300

ttcattacag aaacggcttt ttcaaaaata tggtattgat aatcctgata tgaataaatt 3360

gcagtttcat ttgatgctcg atgagttttt ctaatcagta ctgacaataa aaagattctt 3420

gttttcaaga acttgtcatt tgtatagttt ttttatattg tagttgttct attttaatca 3480

aatgttagcg tgatttatat ttttttttcgc ctcgacatca tctgcccaga tgcgaagtta 3540

agtgcgcaga aagtaatatc atgcgtcaat cgtatgtgaa tgctggtcgc tatactgctg 3600

tcgattcgat actaacgccg ccatccagtg tcgacggatc ctaggtgtac agggcccaaa 3660

agggcgaatt ctgcagatat ccatcacact ggcggccgct cgagcatgca gaatggttct 3720

tatgccacca gagcttatgg acagaagaaa taagagtgat tatgagtatt tgtgagcaga 3780

agttttccgg tctccttttg ttcttgtttt ggcgtattct ccactattcg tccatagcac 3840

atttatacct tagctaaata ttttgtaaag caaaattttc gttatctctt aaaaaatага 3900

agagcggttt attaatatca aataattgaa actgctgata tggtagctat atacaaaatc 3960

tgctgtcaaa atttggcagt aaacgatctt cacggtagcg gttcaaataa agaggaaaag 4020

tccttctccc ttactgtttt tctggaattt ggctcgtcgt taataacaga actgtcttca 4080

agggcgaatt ctgcagatat ccatcacact ggcggccgct cgagcatgca tctagagggc 4140

ccaattcgcc ctatagtgag tcgtattaca attcactggc cgtcgtttta caacgtcgtg 4200

actgggaaaa ccctggcgtt acccaactta atcgccttgc agcacatccc cctttcgcca 4260

gctggcgtaa tagcgaagag gcccgcaccg atcgcccttc ccaacagttg cgcagcctat 4320

acgtacggca gtttaaggtt tacacctata aaagagagag ccgttatcgt ctgtttgtgg 4380

atgtacagag tgatattatt gacacgccgg ggcgacggat ggtgatcccc ctggccagtg 4440

cacgtctgct gtcagataaa gtctcccgtg aactttaccc ggtggtgcat atcggggatg 4500

aaagctggcg catgatgacc accgatatgg ccagtgtgcc ggtctccgtt atcggggaag 4560

aagtggctga tctcagccac cgcgaaaatg acatcaaaaa cgccattaac ctgatgttct 4620

ggggaatata aatgtcaggc atgagattat caaaaaggat cttcacctag atccttttca 4680

cgtagaaagc cagtccgcag aaacggtgct gaccccggat gaatgtcagc tactgggcta 4740

tctggacaag ggaaaacgca agcgcaaaga gaaagcaggt agcttgcagt gggcttacat 4800

ggcgatagct agactgggcg gttttatgga cagcaagcga accggaattg ccagctgggg 4860

cgccctctgg taaggttggg aagccctgca aagtaaactg gatggctttc tcgccgccaa 4920

ggatctgatg gcgcagggga tcaagctctg atcaagagac aggatgagga tcgtttcgca 4980

tgattgaaca agatggattg cacgcaggtt ctccggccgc ttgggtggag aggctattcg 5040

gctatgactg ggcacaacag acaatcggct gctctgatgc cgccgtgttc cggctgtcag 5100

cgcaggggcg cccggttctt tttgtcaaga ccgacctgtc cggtgccctg aatgaactgc 5160

aagacgaggc agcgcggcta tcgtggctgg ccacgacggg cgttccttgc gcagctgtgc 5220

tcgacgttgt cactgaagcg ggaagggact ggctgctatt gggcgaagtg ccggggcagg 5280

atctcctgtc atctcacctt gctcctgccg agaaagtatc catcatggct gatgcaatgc 5340

ggcggctgca tacgcttgat ccggctacct gcccattcga ccaccaagcg aaacatcgca 5400

tcgagcgagc acgtactcgg atggaagccg gtcttgtcga tcaggatgat ctggacgaag 5460

agcatcaggg gctcgcgcca gccgaactgt tcgccaggct caaggcgagc atgcccgacg 5520

gcgaggatct cgtcgtgacc catggcgatg cctgcttgcc gaatatcatg gtggaaaatg 5580

gccgcttttc tggattcatc gactgtggcc ggctgggtgt ggcggaccgc tatcaggaca 5640

tagcgttggc tacccgtgat attgctgaag agcttggcgg cgaatgggct gaccgcttcc 5700

tcgtgcttta cggtatcgcc gctcccgatt cgcagcgcat cgccttctat cgccttcttg 5760

acgagttctt ctgaattatt aacgcttaca atttcctgat gcggtatttt ctccttacgc 5820

atctgtgcgg 5830

<210> 41
<211> 7690
<212> ADN
<213> Artificial

<220>
<223> pRN558
<400> 41

tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 60

cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 120

ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 180

accataattc cgttttaaga gcttggtgag cgctaggagt cactgccagg tatcgtttga 240

acacggcatt agtcagggaa gtcataacac agtcctttcc cgcaattttc tttttctatt 300

actcttggcc tcctctagta cactctatat ttttttatgc ctcggtaatg attttcattt 360

tttttttcc acctagcgga tgactctttt tttttcttag cgattggcat tatcacataa 420

tgaattatac attatataaa gtaatgtgat ttcttcgaag aatatactaa aaaatgagca 480

ggcaagataa acgaaggcaa agatgacaga gcagaaagcc ctagtaaagc gtattacaaa 540

tgaaaccaag attcagattg cgatctcttt aaagggtggt cccctagcga tagagcactc 600

gatcttccca gaaaaagagg cagaagcagt agcagaacag gccacacaat cgcaagtgat 660

taacgtccac acaggtatag ggtttctgga ccatatgata catgctctgg ccaagcattc 720

cggctggtcg ctaatcgttg agtgcattgg tgacttacac atagacgacc atcacaccac 780

tgaagactgc gggattgctc tcggtcaagc ttttaaagag gccctactgg cgcgtggagt 840

aaaaaggttt ggatcaggat ttgcgccttt ggatgaggca ctttccagag cggtggtaga 900

tctttcgaac aggccgtacg cagttgtcga acttggtttg caaagggaga aagtaggaga 960

tctctcttgc gagatgatcc cgcattttct tgaaagcttt gcagaggcta gcagaattac 1020

cctccacgtt gattgtctgc gaggcaagaa tgatcatcac cgtagtgaga gtgcgttcaa 1080

ggctcttgcg gttgccataa gagaagccac ctcgcccaat ggtaccaacg atgttccctc 1140

caccaaaggt gttcttatgt agtgacaccg attatttaaa gctgcagcat acgatatata 1200

tacatgtgta tatatgtata cctatgaatg tcagtaagta tgtatacgaa cagtatgata 1260

ctgaagatga caaggtaatg catcattcta tacgtgtcat tctgaacgag gcgcgctttc 1320

ctttttttctt tttgcttttt ctttttttttt ctcttgaact cgacggatca tatgcggtgt 1380

gaaataccgc acagatgcgt aaggagaaaa taccgcatca ggaaattgta aacgttaata 1440

ttttgttaaa attcgcgtta aatttttgtt aaatcagctc atttttttaac caataggccg 1500

aaatcggcaa aatcccttat aaatcaaaag aatagaccga gatagggttg agtgttgttc 1560

cagtttggaa caagagtcca ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa 1620

ccgtctatca gggcgatggc ccactacgtg aaccatcacc ctaatcaagt tttttggggt 1680

cgaggtgccg taaagcacta aatcggaacc ctaaagggag cccccgattt agagcttgac 1740

ggggaaagcc ggcattgcga ataccgcttc cacaaacatt gctcaaaagt atctctttgc 1800

tatatatctc tgtgctatat ccctatataa cctacccatc caccttttcgc tccttgaact 1860

tgcatctaaa ctcgacctct acatttttta tgtttatctc tagtattact ctttagacaa 1920

aaaaattgta gtaagaacta ttcatagagt gaatcgaaaa caatacgaaa atgtaaacat 1980

ttcctatacg tagtatatag agacaaaata gaagaaaccg ttcataattt tctgaccaat 2040

gaagaatcat caacgctatc actttctgtt cacaaagtat gcgcaatcca catcggtata 2100

gaatataatc ggggatgcct ttatcttgaa aaaatgcacc cgcagcttcg ctagtaatca 2160

gtaaacgcgg gaagtggagt caggcttttt ttatggaaga gaaaatagac accaaagtag 2220

ccttcttcta accttaacgg acctacagtg caaaaagtta tcaagagact gcattataga 2280

gcgcacaaag gagaaaaaaa gtaatctaag atgctttgtt agaaaaatag cgctctcggg 2340

atgcattttt gtagaacaaa aaagaagtat agattctttg ttggtaaaat agcgctctcg 2400

cgttgcattt ctgttctgta aaaatgcagc tcagattctt tgtttgaaaa attagcgctc 2460

tcgtcgcgtt gcatttttgt tttacaaaaa tgaagcacag attcttcgtt ggtaaaatag 2520

cgctttcgcg ttgcatttct gttctgtaaa aatgcagctc agattctttg tttgaaaaat 2580

tagcgctctc gcgttgcatt tttgttctac aaaatgaagc acagatgctt cgttaacaaa 2640

gatatgctat tgaagtgcaa gatggaaacg cagaaaatga accggggatg cgacgtgcaa 2700

gattacctat gcaatagatg caatagtttc tccaggaacc gaaatacata cattgtcttc 2760

cgtaaagcgc tagactatat attattatac aggttcaaat atactatctg tttcagggaa 2820

aactcccagg ttcggatgtt caaaattcaa tgatgggtaa caagtacgat cgtaaatctg 2880

taaaacagtt tgtcggatat taggctgtat ctcctcaaag cgtattcgaa tatcattgag 2940

aagctgcagc gtcacatcgg ataataatga tggcagccat tgtagaagtg ccttttgcat 3000

ttctagtctc tttctcggtc tagctagttt tactacatcg cgaagataga atcttagatc 3060

acactgcctt tgctgagctg gatcaataga gtaacaaaag agtggtaagg cctcgttaaa 3120

ggacaaggac ctgagcggaa gtgtatcgta cagtagacgg agtatctagt atagtctata 3180

gtccgtggaa ttaattctca tctttgacag cttatcatcg ataatccgga gctagcatgc 3240

ggccgccagt gtgatggata tctgcagaat tcgcccttt aagcttcgta cgtgtggaag 3300

aacgattaca acaggtgttg tcctctgagg acataaaata cacaccgaga ttcatcaact 3360

cattgctgga gttagcatat ctacaattgg gtgaaatggg gagcgatttg caggcatttg 3420

ctcggcatgc cggtagaggt gtggtcaata agagcgacct catgctatac ctgagaaagc 3480

aacctgacct acaggaaaga gttactcaag aataagaatt ttcgttttaa aacctaagag 3540

tcactttaaa atttgtatac acttattttt tttataactt atttaataat aaaaatcata 3600

aatcataaga aattcgcgcg cttaagcagc ttcaccagcc tttctagcca aagattgagc 3660

catcttttca gcggtagcca aagcagaaga ggtcataatg tccaagttac cagcgtaagc 3720

tggcaagtag tgagcagcac cttcaacttc caaccaaaca gcggtcttca aaccagagaa 3780

ttgaccaaca cctggcaagt taactggctt gtcttgtggg ataacttcga attgaactct 3840

ttgcttcaat ctgtaacctg gaacgtatgc ttgaacagct tcagccattt cgttgattga 3900

agcttcgatg tcgtcttgag atgcttcgtc agacaaaacg taaacggtgt ctctcatcat 3960

caatggtggt tcagctgggt tcaaaacgat gatagcctta cccttagcag caccaccaac 4020

aacttcgata gccctagagg tggtttcggt gaattcgtcg atgttagctc tggtacctgg 4080

accagcagac ttagaagcga tagaagcgat gatttcagcg tagtgaactc tagcaactct 4140

tgaaacagca gcaaccattg ggatggtagc ttgaccacca caggtaacca tgttaacgtt 4200

taattggtca acgttagctt ccaagttaac aactggaaca cagtatggac cgatagcagc 4260

tggggtcaag tcgatcaatc tgatgtctgg cttagcttct ctcaaagcag cgtcgttctt 4320

aacgtgagca ccagcagagg tagcgtcgaa aacgatgtcg atgtcagcga attctggcat 4380

gttcatcaaa ccgataacac 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

aagaccaaat atgtatttct tgcattgacc aatttatgca agtttatata tatgtaaatg 4740

taagtttcac gaggttctac taaactaaac caccccttg gttagaagaa aagagtgtgt 4800

gagaacaggc tgttgttgtc acacgattcg gacaattctg tttgaaagag agagagtaac 4860

agtacgatcg aacgaacttt gctctggaga tcacagtggg catcatagca tgtggtacta 4920

aacccttcc cgccattcca gaaccttcga ttgcttgtta caaaacctgt gagccgtcgc 4980

taggaccttg ttgtgtgacg aaattggaag ctgcaatcaa taggaagaca ggaagtcgag 5040

cgtgtctggg ttttttcagt tttgttcttt ttgcaaacaa atcacgagcg acggtaattt 5100

ctttctcgat aagaggccac gtgctttatg agggtaacat caattcaaga aggagggaaa 5160

cacttccttt ttctggccct gataatagta tgagggtgaa gccaaaataa aggattcgcg 5220

cccaaatcgg catctttaaa tgcaggtatg cgatagttcc tcactctttc cttactcacg 5280

agtaattctt gcaaatgcct attatgcaga tgttataata tctgtgcgtc ttgagttgaa 5340

gagctcgaga ctagtggatc ccccgggctg caggaattcg atatcaagct tatcgatacc 5400

gtcgacctcg agggggggcc cggtacccag cttttgttcc ctttagtgag ggttaattcc 5460

gagcttggcg taatcatggt catagctgtt tcctgtgtga aattgttatc cgctcacaat 5520

tccacacaac ataggagccg gaagcataaa gtgtaaagcc tggggtgcct aatgagtgag 5580

gtaactcaca ttaattgcgt tgcgctcact gcccgctttc agtcgggaa acctgtcgtg 5640

ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta ttgggcgctc 5700

ttccgcttcc tcgctcactg actcgctgcg ctcggtcgtt cggctgcggc gagcggtatc 5760

agctcactca aaggcggtaa tacggttatc cacagaatca ggggataacg caggaaagaa 5820

catgtgagca aaaggccagc aaaaggccag gaaccgtaaa aaggccgcgt tgctggcgtt 5880

tttccatagg ctcggccccc ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg 5940

gcgaaacccg acaggactat aaagatacca ggcgttcccc cctggaagct ccctcgtgcg 6000

ctctcctgtt ccgaccctgc cgcttaccgg atacctgtcc gcctttctcc cttcgggaag 6060

cgtggcgctt tctcaatgct cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc 6120

caagctgggc tgtgtgcacg aacccccccgt tcagcccgac cgctgcgcct tatccggtaa 6180

ctatcgtctt gagtccaacc cggtaagaca cgacttatcg ccactggcag cagccactgg 6240

taacaggatt agcagagcga ggtatgtagg cggtgctaca gagttcttga agtggtggcc 6300

taactacggc tacactagaa ggacagtatt tggtatctgc gctctgctga agccagttac 6360

cttcggaaaa agagttggta gctcttgatc cggcaaacaa accaccgctg gtagcggtgg 6420

ttttttttgtt tgcaagcagc agattacgcg cagaaaaaaa ggatctcaag aagatccttt 6480

gatcttttct acggggtctg acgctcagtg gaacgaaaac tcacgttaag ggattttggt 6540

catgagatta tcaaaaagga tcttcaccta gatcctttta aattaaaaat gaagttttaa 6600

atcaatctaa agtatatatg agtaaacttg gtctgacagt taccaatgct taatcagtga 6660

ggcacctatc tcagcgatct gtctatttcg ttcatccata gttgcctgac tgcccgtcgt 6720

gtagataact acgatacggg agggcttacc atctggcccc agtgctgcaa tgataccgcg 6780

agacccacgc tcaccggctc cagatttatc 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 ccatgttgtg aaaaaaagcg gttagctcct tcggtcctcc 7080

gatcgttgtc agaagtaagt tggccgcagt gttatcactc atggttatgg cagcactgca 7140

taattctctt actgtcatgc catccgtaag atgcttttct gtgactggtg agtactcaac 7200

caagtcattc tgagaatagt gtatgcggcg accgagttgc tcttgcccgg cgtcaatacg 7260

ggataatacc gcgccacata gcagaacttt aaaagtgctc atcattggaa aacgttcttc 7320

ggggcgaaaa ctctcaagga tcttaccgct gttgagatcc agttcgatgt aacccactcg 7380

tgcacccaac tgatcttcag catcttttac tttcaccagc gtttctgggt gagcaaaaac 7440

aggaaggcaa aatgccgcaa aaaagggaat aagggcgaca cggaaatgtt gaatactcat 7500

actcttcctt tttcaatatt attgaagcat ttatcagggt tattgtctca tgagcggata 7560

catatttgaa tgtatttaga aaaataaaca aatagggggtt ccgcgcacat ttccccgaaa 7620

agtgccacct gacgtcttat tatcatgaca ttaacctata aaaataggcg tatcacgagg 7680

ccctttcgtc 7690

<210> 42
<211> 9416
<212> ADN
<213> Artificial

<220>
<223> pRN595
<400> 42

gacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata ataagacgtc 60

aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 120

ctcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 180
aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 240
ctgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 300
gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 360
ctttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 420
ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca 480
gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 540
aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct 600
gacaacgatc ggaggaccga aggagctaac cgcttttttt 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 cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 960
agttatctac acgacgggca gtcaggcaac tatggatgaa cgaaatagac agatcgctga 1020
gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 1080
ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttga 1140
taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 1200
agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca 1260
aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 1320
tttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta 1380
gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 1440
aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 1500
aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca 1560
gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agcattgaga 1620
aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 1680
aacaggagag cgcacgaggg agcttccagg ggggaacgcc tggtatcttt atagtcctgt 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 cgcagcgagt cagtgagcga 1980
ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 2040

atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 2100

tgtgagttac ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcctat 2160

gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 2220

cgccaagctc ggaattaacc ctcactaaag ggaacaaaag ctgggtaccg ggcccccccct 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

atcgagaaag aaattaccgt cgctcgtgat ttgtttgcaa aaagaacaaa actgaaaaaa 2640

cccagacacg ctcgacttcc tgtcttccta ttgattgcag cttccaattt cgtcacacaa 2700

caaggtccta gcgacggctc acaggttttg taacaagcaa tcgaaggttc tggaatggcg 2760

ggaaagggtt tagtaccaca tgctatgatg cccactgtga tctccagagc aaagttcgtt 2820

cgatcgtact gttactctct ctctttcaaa cagaattgtc cgaatcgtgt gacaacaaca 2880

gcctgttctc acacactctt ttcttctaac caaggggtg gtttagttta gtagaacctc 2940

gtgaaactta catttacata tatataaact tgcataaatt ggtcaatgca agaaatacat 3000

atttggtctt ttctaattcg tagtttttca agttcttaga tgctttcttt ttctcttttt 3060

tacagatcat caaggaagta attatctact ttttacaaca aatatatcta gaaaatggct 3120

gttaccaacg ttgctgaatt gaacgctttg gttgaaaggg ttaagaaggc tcaaagagaa 3180

tacgcttctt tcacccaaga acaagttgac aagatcttca gagctgctgc tttggctgct 3240

gctgacgcta gaatcccatt ggctaagatg gctgttgctg aatctggtat gggtatcgtt 3300

gaagacaagg ttatcaagaa ccacttcgct tctgaataca tctacaacgc ttacaaggac 3360

gaaaagacct gtggtgtttt gtcagaagac gacaccttcg gtaccatcac catcgctgaa 3420

ccaatcggta tcatctgtgg tatcgttcca accaccaacc caacctctac cgctatcttc 3480

aagtctttga tctctttgaa gaccagaaac gctatcatct tctctccaca cccaagagct 3540

aaagacgcta ccaacaaggc tgctgacatc gttttgcaag ctgctatcgc tgctggtgct 3600

ccaaaggact tgatcggttg gatcgaccaa ccatctgttg aattgtctaa cgctttgatg 3660

caccacccag acatcaactt gatcttggct accggtggtc caggtatggt taaggctgct 3720

tactcttctg gtaagccagc tatcggtgtt ggtgctggta acaccccagt tgttatcgac 3780

gaaaccgctg acatcaagag agctgttgct tctgttttga tgtctaagac cttcgacaac 3840

ggtgttatct gtgcttctga acaatctgtt gttgttgttg actctgttta cgacgctgtt 3900

agagaaagat tcgctaccca cggtggttac ttgttgcaag gtaaggaatt gaaggctgtt 3960

caagacgtta tcttgaagaa cggtgctttg aacgctgcta tcgttggtca accagcttac 4020

aagatcgctg aattagctgg tttctctgtt ccagaaaaca ccaagatctt gatcggtgaa 4080

gttaccgttg ttgacgaatc tgaaccattc gctcacgaaa agttgtctcc aaccttggct 4140

atgtacagag ctaaggactt cgaagacgct gttgaaaaag ctgaaaagtt ggttgctatg 4200

ggtggtattg gtcacacctc ttgtttgtac accgaccaag acaaccaacc agctagagtt 4260

tcttacttcg gtcaaaagat gaagaccgct agaatcttga tcaacacccc agcttctcaa 4320

ggtggtatcg gtgacttgta caacttcaag ttggctccat ctttgacctt gggttgtggt 4380

tcttggggtg gtaactctat ctctgaaaac gttggtccaa agcacttgat caacaagaag 4440

accgttgcta agagagctga aaacatgttg tggcacaagt tgccaaaatc tatctacttc 4500

agaagaggtt ctttgccaat cgctttggac gaagttatca ccgacggtca caagagagct 4560

ttgatcgtta ccgacagatt cttgttcaac aacggttacg ctgaccaaat cacctctgtt 4620

ttgaaggctg ctggtgttga aaccgaagtt ttcttcgaag ttgaagctga cccaaccttg 4680

tctatcgtta gaaagggtgc tgaattggct aactctttca agccagacgt tatcatcgct 4740

ttgggtggtg gttctccaat ggacgctgct aagatcatgt gggttatgta cgaacaccca 4800

gaaacccact tcgaagaatt ggctttgaga ttcatggaca tcagaaagag aatctacaag 4860

ttcccaaaga tgggtgttaa ggctaagatg atcgctgtta ccaccacctc tggtaccggt 4920

tctgaagtta ccccattcgc tgttgttacc gacgacgcta ccggtcaaaa gtacccattg 4980

gctgactacg ctttgacccc agacatggct atcgttgacg ctaacttggt tatggacatg 5040

ccaaagtctt tgtgtgcttt cggtggtttg gacgctgtta cccacgctat ggaagcttac 5100

gtttctgttt tggcttctga attctctgac ggtcaagctt tgcaagcttt gaagttgttg 5160

aaggaatact tgccagcttc ttaccacgaa ggttctaaga acccagttgc tagagaaaga 5220

gttcactctg ctgctaccat cgctggtatc gctttcgcta acgctttctt gggtgtttgt 5280

cactctatgg ctcacaagtt gggttctcaa ttccacatcc cacacggttt ggctaacgct 5340

ttgttgatct gtaacgttat cagatacaac gctaacgaca acccaaccaa gcaaaccgct 5400

ttctctcaat acgacagacc acaagctaga agaagatacg ctgaaatcgc tgaccacttg 5460

ggtttgtctg ctccaggtga cagaaccgct gcaaagatcg aaaagttgtt ggcttggttg 5520

gaaaccttga aggctgaatt gggtatccca aagtctatca gagaagctgg tgttcaagaa 5580

gctgacttct tggctaacgt tgacaagttg tctgaagacg ctttcgacga ccaatgtacc 5640

ggtgctaacc caagataccc attgatctct gaattgaagc aaatcttgtt ggacacctac 5700

tacggtagag actacgttga aggtgaaacc gctgctaaga aggaagctgc tccagctaag 5760

gctgaaaaga aggctaagaa gtctgcttag cttaagcgcg cgaatttctt atgatttatg 5820

attttttatta ttaaataagt tataaaaaaa ataagtgtat acaaatttta aagtgactct 5880

taggttttaa aacgaaaatt cttattcttg agtaactctt tcctgtaggt caggttgctt 5940

tctcaggtat agcatgaggt cgctcttatt gaccacacct ctaccggcat gccgagcaaa 6000

tgcctgcaaa tcgctcccca tttcacccaa ttgtagatat gctaactcca gcaatgagtt 6060

gatgaatctc ggtgtgtatt ttatgtcctc agaggacaac acctgttgta atcgttcttc 6120

cacacgtacg aagcttaaaa gggcgaattc tgcagatatc catcacactg gcggccgcat 6180

gctagctccg gattatcgat gataagctgt caaagatgag aattaattcc acggactata 6240

gactatacta gatactccgt ctactgtacg atacacttcc gctcaggtcc ttgtccttta 6300

acgaggcctt accactcttt tgttactcta ttgatccagc tcagcaaagg cagtgtgatc 6360

taagattcta tcttcgcgat gtagtaaaac tagctagacc gagaaagaga ctagaaatgc 6420

aaaaggcact tctacaatgg ctgccatcat tattatccga tgtgacgctg cagcttctca 6480

atgatattcg aatacgcttt gaggagatac agcctaatat ccgacaaact gttttacaga 6540

tttacgatcg tacttgttac ccatcattga attttgaaca tccgaacctg ggagttttcc 6600

ctgaaacaga tagtatattt gaacctgtat aataatatat agtctagcgc tttacggaag 6660

acaatgtatg tatttcggtt cctggagaaa ctattgcatc tattgcatag gtaatcttgc 6720

acgtcgcatc cccggttcat tttctgcgtt tccatcttgc acttcaatag catatctttg 6780

ttaacgaagc atctgtgctt cattttgtag aacaaaaatg caacgcgaga gcgctaattt 6840

ttcaaacaaa gaatctgagc tgcattttta cagaacagaa atgcaacgcg aaagcgctat 6900

tttaccaacg aagaatctgt gcttcatttt tgtaaaacaa aaatgcaacg cgacgagagc 6960

gctaattttt caaacaaaga atctgagctg catttttaca gaacagaaat gcaacgcgag 7020

agcgctattt taccaacaaa gaatctatac ttctttttg ttctacaaaa atgcatcccg 7080

agagcgctat ttttctaaca aagcatctta gattactttt tttctccttt gtgcgctcta 7140

taatgcagtc tcttgataac tttttgcact gtaggtccgt taaggttaga agaaggctac 7200

tttggtgtct attttctctt ccataaaaaa agcctgactc cacttcccgc gtttactgat 7260

tactagcgaa gctgcgggtg catttttttca agataaaggc atccccgatt atattctata 7320

ccgatgtgga ttgcgcatac tttgtgaaca gaaagtgata 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 aaggtggatg ggtaggttat atagggatat agcacagaga tatatagcaa 7620

agagatactt ttgagcaatg tttgtggaag cggtattcgc aatgccggct ttccccgtca 7680

agctctaaat cgggggctcc ctttagggtt ccgatttagt gctttacggc acctcgaccc 7740

caaaaaactt gattagggtg atggttcacg tagtgggcca tcgccctgat agacggtttt 7800

tcgccctttg acgttggagt ccacgttctt taatagtgga ctcttgttcc aaactggaac 7860

aacactcaac cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc 7920

ctattggtta aaaaatgagc tgatttaaca aaaatttaac gcgaatttta acaaaatatt 7980

aacgtttaca atttcctgat gcggtatttt ctccttacgc atctgtgcgg tatttcacac 8040

cgcatatgat ccgtcgagtt caagagaaaa aaaaagaaaa agcaaaaaga aaaaaggaaa 8100

gcgcgcctcg ttcagaatga cacgtataga atgatgcatt accttgtcat cttcagtatc 8160

atactgttcg tatacatact tactgacatt cataggtata catatataca catgtatata 8220

tatcgtatgc tgcagcttta aataatcggt gtcactacat aagaacacct ttggtggagg 8280

gaacatcgtt ggtaccattg ggcgaggtgg cttctcttat ggcaaccgca agagccttga 8340

acgcactctc actacggtga tgatcattct tgcctcgcag acaatcaacg tggagggtaa 8400

ttctgctagc ctctgcaaag ctttcaagaa aatgcgggat catctcgcaa gagagatctc 8460

ctactttctc cctttgcaaa ccaagttcga caactgcgta cggcctgttc gaaagatcta 8520

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

taatacgctt tactagggct ttctgctctg tcatctttgc cttcgtttat cttgcctgct 8940

catttttag tatattcttc gaagaaatca cattacttta tataatgtat aattcattat 9000

gtgataatgc caatcgctaa gaaaaaaaaa gagtcatccg ctaggtggaa aaaaaaaaat 9060

gaaaatcatt accgaggcat aaaaaaatat agagtgtact agaggaggcc aagagtaata 9120

gaaaaagaaa attgcgggaa aggactgtgt tatgacttcc ctgactaatg ccgtgttcaa 9180

acgatacctg gcagtgactc ctagcgctca ccaagctctt aaaacggaat tatggtgcac 9240

tctcagtaca atctgctctg atgccgcata gttaagccag ccccgacacc cgccaacacc 9300

cgctgacgcg ccctgacggg cttgtctgct cccggcatcc gcttacagac aagctgtgac 9360

cgtctccggg agctgcatgt gtcagaggtt ttcaccgtca tcaccgaaac gcgcga 9416

<210> 43
<211> 9352
<212> ADN
<213> Artificial

<220>
<223> pRN596

<400> 43

gacgaaaggg cctcgtgata cgcctatttt ataggttaa tgtcatgata ataagacgtc 60

aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 120

ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 180

aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 240

ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 300

gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 360

ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 420

ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca 480

gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 540

aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct 600

gacaacgatc ggaggaccga aggagctaac cgcttttttt 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 cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 960

agttatctac acgacgggca gtcaggcaac tatggatgaa cgaaatagac agatcgctga 1020

gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 1080

ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttttga 1140

taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 1200

agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca 1260

aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 1320

ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta 1380

gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 1440

aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 1500

aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca 1560

gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agcattgaga 1620

aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 1680

aacaggagag cgcacgaggg agcttccagg ggggaacgcc tggtatcttt atagtcctgt 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 cgcagcgagt cagtgagcga 1980

ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 2040

atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 2100

tgtgagttac ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcctat 2160

gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 2220

cgccaagctc ggaattaacc ctcactaaag ggaacaaaag ctgggtaccg ggcccccccct 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

atcgagaaag aaattaccgt cgctcgtgat ttgtttgcaa aaagaacaaa actgaaaaaa 2640

cccagacacg ctcgacttcc tgtcttccta ttgattgcag cttccaattt cgtcacacaa 2700

caaggtccta gcgacggctc acaggttttg taacaagcaa tcgaaggttc tggaatggcg 2760

ggaaagggtt tagtaccaca tgctatgatg cccactgtga tctccagagc aaagttcgtt 2820

cgatcgtact gttactctct ctctttcaaa cagaattgtc cgaatcgtgt gacaacaaca 2880

gcctgttctc acacactctt ttcttctaac caaggggtg gtttagttta gtagaacctc 2940

gtgaaactta catttacata tatataaact tgcataaatt ggtcaatgca agaaatacat 3000

atttggtctt ttctaattcg tagtttttca agttcttaga tgctttcttt ttctcttttt 3060

tacagatcat caaggaagta attatctact ttttacaaca aatatatcta gaaaatgtct 3120

acccaacaaa ccatgaccgt tgacgaacac atcaaccaat tagttagaaa ggctcaagtt 3180

gctttgaagg aatacttgaa gccagaatac acccaagaaa agatcgacta catcgttaag 3240

aaggcttctg ttgctgcttt ggaccaacac tgtgctttgg ctgctgctgc tgttgaagaa 3300

accggtagag gtatcttcga agacaaggct accaagaaca tcttcgcttg tgaacacgtt 3360

acccacgaaa tgagacacgc taagaccgtt ggtatcatca acgttgaccc attgtacggt 3420

atcaccgaaa tcgctgaacc agttggtgtt gtttgtggtg ttaccccagt taccaaccca 3480

acctctaccg ctatcttcaa gtctttgatc tctatcaaga ccagaaaccc aatcgttttc 3540

tctttccacc catctgcttt gaagtgttct attatggctg ctaaaatcgt tagagacgct 3600

gctatcgctg ctggtgctcc agaaaactgt atccaatgga tcgaattcgg tggtatcgaa 3660

gcttctaaca agttgatgaa ccacccaggt gttgctacta tcttggctac cggtggtaac 3720

gctatggtta aggctgcata ctcttctggt aagccagctt tgggtgttgg tgctggtaac 3780

gttccaacct acatcgaaaa gacctgtaac atcaagcaag ctgctaacga cgttgttatg 3840

tctaagtctt tcgacaacgg tatgatctgt gcttctgaac aagctgctat catcgacaag 3900

gaaatctacg accaagttgt tgaagaaatg aagaccttgg gtgcttactt catcaacgaa 3960

gaagaaaagg ctaagttgga aaagttcatg ttcggtgtta acgcttactc tgctgacgtt 4020

aacaacgcta gattgaaccc aaagtgtcca ggtatgtctc cacaatggtt cgctgaacaa 4080

gttggtatca aggtaccaga agactgtaac atcatctgtg ctgtttgtaa ggaagttggt 4140

ccaaacgaac cattgaccag agaaaagttg tctccagttt tggctatctt gaaagctgaa 4200

aacacccaag acggtatcga caaggctgaa gctatggttg aatttaacgg tagaggtcac 4260

tctgctgcta tccactctaa cgacaaggct gttgttgaaa agtacgcttt gaccatgaag 4320

gcttgtagaa tcttgcacaa caccccatct tctcaaggtg gtatcggttc tatctacaac 4380

tacatctggc catctttcac cttgggttgt ggttcttacg gtggtaactc tgtttctgct 4440

aacgttacct accacaactt gttgaacatc aagagattgg ctgacagaag aaacaacttg 4500

caatggttca gagttccacc aaagatcttc ttcgaaccac actctatcag atacttggct 4560

gaattgaagg aattgtctaa gatcttcatc gtttctgaca gaatgatgta caagttgggt 4620

tacgttgaca gagttatgga cgttttgaag agaagatcta acgaagttga aatcgaaatc 4680

ttcatcgacg ttgaaccaga cccatctatc caaaccgttc aaaagggttt ggctgttatg 4740

aacaccttcg gtccagacaa catcatcgct atcggtggtg gttctgctat ggacgctgct 4800

aagatcatgt ggttgttgta cgaacaccca gaagctgact tcttcgctat gaagcaaaag 4860

ttcatcgact tgagaaagag agctttcaag ttcccaacca tgggtaagaa ggctagattg 4920

atctgtatcc caaccacctc tggtaccggt tctgaagtta ccccattcgc tgttatctct 4980

gaccacgaaa ccggtaagaa gtacccattg gctgactact ctttgacccc atctgttgct 5040

atcgttgacc caatgttcac catgtctttg ccaaagagag ctatcgctga caccggtttg 5100

gacgttttgg ttcacgctac cgaagcttac gtttctgtta tggctaacga atacaccgac 5160

ggtttggcta gagaagctgt taagttggtt tttgaaaact tgttgaagtc ttacaacggt 5220

gacttggaag ctagagaaaa gatgcacaac gctgctacca tcgctggtat ggctttcgct 5280

tctgctttct tgggtatgga ccactctatg gctcacaagg ttggtgctgc tttccacttg 5340

ccacacggta gatgtgttgc tgttttgttg ccacacgtta tcagatacaa cggtcaaaag 5400

ccaagaaagt tggctatgtg gccaaagtac aacttctaca aggctgacca aagatacatg 5460

gaattggctc aaatggttgg tttgaagtgt aacaccccag ctgaaggtgt tgaagctttc 5520

gctaaggctt gtgaagaatt gatgaaggct accgaaacca tcaccggttt caagaaggct 5580

aacatcgacg aagctgcttg gatgtctaag gttccagaaa tggctttgtt ggctttcgaa 5640

gaccaatgtt ctccagctaa cccaagagtt ccaatggtta aggacatgga aaagatcttg 5700

aaggctgctt actacccaat cgcttagctt aagcgcgcga atttcttatg atttatgatt 5760

cttattatta aataagttat aaaaaaaata agtgtataca aattttaaag tgactcttag 5820

gttttaaaac gaaaattctt attcttgagt aactctttcc tgtaggtcag gttgctttct 5880

caggtatagc atgaggtcgc tcttattgac cacacctcta ccggcatgcc gagcaaatgc 5940

ctgcaaatcg ctccccattt cacccaattg tagatatgct aactccagca atgagttgat 6000

gaatctcggt gtgtatttta tgtcctcaga ggacaacacc tgttgtaatc gttcttccac 6060

acgtacgaag ctaaaagggc gaattctgca gatatccatc acactggcgg ccgcatgcta 6120

gctccggatt atcgatgata agctgtcaaa gatgagaatt aattccacgg actatagact 6180

atactagata ctccgtctac tgtacgatac acttccgctc aggtccttgt cctttaacga 6240

ggccttacca ctcttttgtt actctattga tccagctcag caaaggcagt gtgatctaag 6300

attctatctt cgcgatgtag taaaactagc tagaccgaga aagagactag aaatgcaaaa 6360

ggcacttcta caatggctgc catcattatt atccgatgtg acgctgcagc ttctcaatga 6420

tattcgaata cgctttgagg agatacagcc taatatccga caaactgttt tacagattta 6480

cgatcgtact tgttacccat cattgaattt tgaacatccg aacctgggag ttttccctga 6540

aacagatagt atatttgaac ctgtataata atatatagtc tagcgcttta cggaagacaa 6600

tgtatgtatt tcggttcctg gagaaactat tgcatctatt gcataggtaa tcttgcacgt 6660

cgcatccccg gttcattttc tgcgtttcca tcttgcactt caatagcata tctttgttaa 6720

cgaagcatct gtgcttcatt ttgtagaaca aaaatgcaac gcgagagcgc taatttttca 6780

aacaaagaat ctgagctgca tttttacaga acagaaatgc aacgcgaaag cgctatttta 6840

ccaacgaaga atctgtgctt catttttgta aaacaaaaat gcaacgcgac gagagcgcta 6900

atttttcaaa caaagaatct gagctgcatt tttacagaac agaaatgcaa cgcgagagcg 6960

ctatttttacc aacaaagaat ctatacttct tttttgttct acaaaaatgc atcccgagag 7020

cgctattttt ctaacaaagc atcttagatt actttttttc tcctttgtgc gctctataat 7080

gcagtctctt gataactttt tgcactgtag gtccgttaag gttagaagaa ggctactttg 7140

gtgtctattt tctcttccat aaaaaaagcc tgactccact tcccgcgttt actgattact 7200

agcgaagctg cgggtgcatt ttttcaagat aaaggcatcc ccgattatat tctataccga 7260

tgtggattgc gcatactttg tgaacagaaa gtgatagcgt tgatgattct tcattggtca 7320

gaaaattatg aacggtttct tctattttgt ctctatatac tacgtatagg aaatgtttac 7380

attttcgtat tgttttcgat tcactctatg aatagttctt actacaattt ttttgtctaa 7440

agagtaatac tagagataaa cataaaaaat gtagaggtcg agtttagatg caagttcaag 7500

gagcgaaagg tggatgggta ggttatatag ggatatagca cagagatata tagcaaagag 7560

atacttttga gcaatgtttg tggaagcggt attcgcaatg ccggctttcc ccgtcaagct 7620

ctaaatcggg ggctcccttt agggttccga tttagtgctt tacggcacct cgaccccaaa 7680

aaacttgatt agggtgatgg ttcacgtagt gggccatcgc cctgatagac ggtttttcgc 7740

cctttgacgt tggagtccac gttctttaat agtggactct tgttccaaac tggaacaaca 7800

ctcaacccta tctcggtcta ttcttttgat ttataaggga ttttgccgat ttcggcctat 7860

tggttaaaaa atgagctgat ttaacaaaaa tttaacgcga attttaacaa aatattaacg 7920

tttacaattt cctgatgcgg tattttctcc ttacgcatct gtgcggtatt tcacaccgca 7980

tatgatccgt cgagttcaag agaaaaaaaa agaaaaagca aaaagaaaaa aggaaagcgc 8040

gcctcgttca gaatgacacg tatagaatga tgcattacct tgtcatcttc agtatcatac 8100

tgttcgtata catacttact gacattcata ggtatacata tatacacatg tatatatatc 8160

gtatgctgca gctttaaata atcggtgtca ctacataaga acacctttgg tggagggaac 8220

atcgttggta ccattgggcg aggtggcttc tcttatggca accgcaagag ccttgaacgc 8280

actctcacta cggtgatgat cattcttgcc tcgcagacaa tcaacgtgga gggtaattct 8340

gctagcctct gcaaagcttt caagaaaatg cgggatcatc tcgcaagaga gatctcctac 8400

tttctccctt tgcaaaccaa gttcgacaac tgcgtacggc ctgttcgaaa gatctaccac 8460

cgctctggaa agtgcctcat ccaaaggcgc aaatcctgat ccaaaccttt ttactccacg 8520

cgccagtagg gcctctttaa aagcttgacc gagagcaatc ccgcagtctt cagtggtgtg 8580

atggtcgtct atgtgtaagt caccaatgca ctcaacgatt agcgaccagc cggaatgctt 8640

ggccagagca tgtatcatat ggtccagaaa ccctatacct gtgtggacgt taatcacttg 8700

cgattgtgtg gcctgttctg ctactgcttc tgcctctttt tctgggaaga tcgagtgctc 8760

tatcgctagg 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 ctaatgccgt gttcaaacga 9120

tacctggcag tgactcctag cgctcaccaa gctcttaaaa cggaattatg gtgcactctc 9180

agtacaatct gctctgatgc cgcatagtta agccagcccc gacacccgcc aacacccgct 9240

gacgcgccct gacgggcttg tctgctcccg gcatccgctt acagacaagc tgtgaccgtc 9300

tccgggagct gcatgtgtca gaggttttca ccgtcatcac cgaaacgcgc ga 9352

<210> 44
<211> 29
<212> ADN
<213> Artificial

<220>
<223> cebador GPD2uf

<400> 44
ggtaccagat cttttgcggc gaggtgccg 29

<210> 45
<211> 42
<212> ADN
<213> Artificial

<220>
<223> cebador GPD2ur

<400> 45
tctagactta aggaatgtgt atcttgttaa tcttctgaca gc 42

<210> 46
<211> 28
<212> ADN
<213> Artificial

<220>
<223> cebador GPD2df

<400> 46
ctcgagatag tctacaacaa cgtccgca 28

<210> 47
<211> 39
<212> ADN
<213> Artificial

<220>
<223> cebador GPD2dr

<400> 47
ccatggagat ctgcagtgaa aaagctcgaa gaaacagct 39

<210> 48
<211> 4397
<212> ADN
<213> Artificial

<220>
<223> pRN594

<400> 48
gacgacggcg acctggcggg cttcgtggtc gtctcgtact ccggctggaa ccgccggctg 60

accgtcgagg acatcgaggt cgccccggag caccgggggc acggggtcgg gcgcgcgttg 120

atggggctcg cgacggagtt cgcccgcgag cggggcgccg ggcacctctg gctggaggtc 180

accaacgtca acgcaccggc gatccacgcg taccggcgga tggggttcac cctctgcggc 240

ctggacaccg ccctgtacga cggcaccgcc tcggacggcg agcaggcgct ctacatgagc 300

atgccctgcc cctagtactg acaataaaaa gattcttgtt ttcaagaact tgtcatttgt 360

atagtttttt tatattgtag ttgttctatt ttaatcaaat gttagcgtga tttatatttt 420

ttttcgcctc gacatcatct gcccagatgc gaagttaagt gcgcagaaag taatatcatg 480

cgtcaatcgt atgtgaatgc tggtcgctat actgctgtcg attcgatact aacgccgcca 540

tccagtgtcg acggatccta ggtgtacagg gcccaaaagg gcgaattctg cagatatcca 600

tcacactggc ggccgctcga gatagtctac aacaacgtcc gcatggaaga cctaccggag 660

atgattgaag agctagacat cgatgacgaa tagacactct ccccccccct ccccctctga 720

tctttcctgt tgcctctttt tcccccaacc aatttatcat tatacacaag ttctacaact 780

actactagta acattactac agttattata attttctatt ctctttttct ttaagaatct 840

atcattaacg ttaatttcta tatatacata actaccatta tacacgctat tatcgtttac 900

atatcacatc accgttaatg aaagatacga caccctgtac actaacacaa ttaaataatc 960

gccataacct tttctgttat ctatagccct taaagctgtt tcttcgagct ttttcactgc 1020

agatctccat ggcgatgcct gcttgccgaa tatcatggtg gaaaatggcc gcttttctgg 1080

attcatcgac tgtggccggc tgggtgtggc ggaccgctat caggacatag cgttggctac 1140

ccgtgatatt gctgaagagc ttggcggcga atgggctgac cgcttcctcg tgctttacgg 1200

tatcgccgct cccgattcgc agcgcatcgc cttctatcgc cttcttgacg agttcttctg 1260

aattgaaaaa ggaagagtat gagtattcaa catttccgtg tcgcccttat tccctttttt 1320

gcggcatttt gccttcctgt ttttgctcac ccagaaacgc tggtgaaagt aaaagatgct 1380

gaagatcagt tgggtgcacg agtgggttac atcgaactgg atctcaacag cggtaagatc 1440

cttgagagtt ttcgccccga agaacgtttt ccaatgatga gcacttttaa agttctgcta 1500

tgtggcgcgg tattatcccg tattgacgcc gggcaagagc aactcggtcg ccgcatacac 1560

tattctcaga atgacttggt tgagtactca ccagtcacag aaaagcatct tacggatggc 1620

atgacagtaa gagaattatg cagtgctgcc ataaccatga gtgataacac tgcggccaac 1680

ttacttctga caacgatcgg aggaccgaag gagctaaccg ctttttttgca caacatgggg 1740

gatcatgtaa ctcgccttga tcgttgggaa ccggagctga atgaagccat accaaacgac 1800

gagcgtgaca ccacgatgcc tgtagcaatg gcaacaacgt tgcgcaaact attaactggc 1860

gaactactta ctctagcttc ccggcaacaa ttaatagact ggatggaggc ggataaagtt 1920

gcaggaccac ttctgcgctc ggcccttccg gctggctggt ttattgctga taaatctgga 1980

gccggtgagc gtgggtctcg cggtatcatt gcagcactgg ggccagatgg taagccctcc 2040

cgtatcgtag ttatctacac gacggggagt caggcaacta tggatgaacg aaatagacag 2100

atcgctgaga 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 tacatacctc 2520

gctctgctaa tcctgttacc agtggctgct gccagtggcg ataagtcgtg tcttaccggg 2580

ttggactcaa gacgatagtt accggataag gcgcagcggt cgggctgaac gggggttcg 2640

tgcacacagc ccagcttgga gcgaacgacc tacaccgaac tgagatacct acagcgtgag 2700

ctatgagaaa gcgccacgct tcccgaaggg agaaaggcgg acaggtatcc ggtaagcggc 2760

agggtcggaa caggagagcg cacgagggag cttccagggg gaaacgcctg gtatctttat 2820

agtcctgtcg ggtttcgcca cctctgactt gagcgtcgat ttttgtgatg ctcgtcaggg 2880

gggcggagcc tatggaaaaa cgccagcaac gcggcctttt tacggttcct ggccttttgc 2940

tggccttttg ctcacatgtt ctttcctgcg ttatcccctg attctgtgga taaccgtatt 3000

accgccttg agtgagctga taccgctcgc cgcagccgaa cgaccgagcg cagcgagtca 3060

gtgagcgagg aagcggaaga gcgcccaata cgcaaaccgc ctctccccgc gcgttggccg 3120

attcattaat gcagctggca cgacaggttt cccgactgga aagcgggcag tgagcgcaac 3180

gcaattaatg tgagttagct cactcattag gcaccccagg ctttacactt tatgcttccg 3240

gctcgtatgt tgtgtggaat tgtgagcgga taacaatttc acacaggaaa cagctatgac 3300

catgattacg ccaagctatt taggtgacac tatagaatac tcaagctatg catcaagctt 3360

ggtaccagat cttttgcggc gaggtgccga tgggttgctg aggggaagag tgtttagctt 3420

acggacctat tgccattgtt attccgatta atctattgtt cagcagctct tctctaccct 3480

gtcattctag tatttttttt ttttttttt ggttttactt ttttttcttc ttgccttttt 3540

ttcttgttac tttttttcta gttttttttc cttccactaa gctttttcct tgatttatcc 3600

ttgggttctt ctttctactc ctttagattt tttttttata tattaatttt taagtttatg 3660

tattttggta gattcaattc tctttccctt tccttttcct tcgctcccct tccttatcaa 3720

tgcttgctgt cagaagatta acaagataca cattccttaa ggcctcgtcc ccgccgggtc 3780

acccggccag cgacatggag gcccagaata ccctccttga cagtcttgac gtgcgcagct 3840

caggggcatg atgtgactgt cgcccgtaca tttagcccat acatccccat gtataatcat 3900

ttgcatccat acattttgat ggccgcacgg cgcgaagcaa aaattacggc tcctcgctgc 3960

agacctgcga gcagggaaac gctcccctca cagacgcgtt gaattgtccc cacgccgcgc 4020

ccctgtagag aaatataaaa ggttaggatt 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 cggtggaccc gcccctgacc aaggtgttcc 4320

ccgacgacga atcggacgac gaatcggacg ccggggagga cggcgacccg gactcccgga 4380

cgttcgtcgc gtacggg 4397

<210> 49
<211> 367
<212> PRT
<213> Escherichia coli

<400> 49

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

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

ttgcatccat acattttgat ggccgcacgg cgcgaagcaa aaattacggc tcctcgctgc 3960

agacctgcga gcagggaaac gctcccctca cagacgcgtt gaattgtccc cacgccgcgc 4020

ccctgtagag aaatataaaa ggttaggatt 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 cggtggaccc gcccctgacc aaggtgttcc 4320

ccgacgacga atcggacgac gaatcggacg ccggggagga cggcgacccg gactcccgga 4380

cgttcgtcgc gtacggg 4397

<210> 49
<211> 367
<212> PRT
<213> Escherichia coli

<400> 49

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

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
        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> 50
<211> 1120
<212> ADN
<213> Artificial

<220>
<223> Secuencia de nucleótidos que codifica gldA de E. coli optimizada en codones para levadura

<400> 50
ctgcagaaaa tggacagaat catccaatct ccaggtaagt acatccaagg tgctgacgtt 60

atcaacagat tgggtgaata cttgaagcca ttggctgaaa gatggttggt tgttggtgac 120

aagttcgttt tgggtttcgc tcaatctacc gttgaaaagt ctttcaagga cgctggtttg 180

gttgttgaaa tcgctccatt cggtggtgaa tgttctcaaa acgaaatcga cagattgaga 240

ggtatcgctg aaaccgctca atgtggtgct atcttgggta tcggtggtgg taagaccttg 300

gacaccgcta aggctttggc tcacttcatg ggtgttccag ttgctatcgc tccaaccatc 360

gcttctaccg acgctccatg ttctgctttg tctgttatct acaccgacga aggtgaattc 420

gacagatact tgttgttgcc aaacaaccca aacatggtta tcgttgacac caagatcgtt 480

gctggtgctc cagctagatt gttggcagct ggtatcggtg acgctttggc tacctggttc 540

gaagctagag cttgttctag atctggtgct accaccatgg ctggtggtaa gtgtacccaa 600

gctgctttgg ctttggctga attgtgttac aacaccttgt tggaagaagg 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 cgttgctgaa gctgcttgtg ctgaaggtga aaccatccac 1020

aacatgccag gtggtgctac cccagaccaa gtttacgctg ctttgttggt tgctgaccaa 1080

tacggtcaaa gattcctaca agaatgggaa taaggcgcgc 1120

<210> 51
<211> 1949
<212> ADN
<213> Artificial

<220>
<223> pRNgldA

<400> 51
actagtaaat gtgtggggaa gcgggtaagc tgccacagca attaatgcac aacatttaac 60

ctacattctt ccttatcgga tcctcaaaac ccttaaaaac atatgcctca ccctaacata 120

ttttccaatt aaccctcaat atttctctgt cacccggcct ctattttcca ttttcttctt 180

tacccgccac gcgtttttt ctttcaaatt tttttcttcc ttcttctttt tcttccacgt 240

cctcttgcat aaataaataa accgttttga aaccaaactc gcctctctct ctccttttg 300

aaatattttt gggtttgttt gatcctttcc ttcccaatct ctcttgttta atatatattc 360

atttatatca cgctctcttt ttatcttcct ttttttcctc tctcttgtat tcttccttcc 420

cctttctact caaaccaaga agaaaaagaa aaggtcaatc tttgttaaag aataggatct 480

tctactacat cagcttttag atttttcacg cttactgctt ttttcttccc aagatcgaaa 540

atttactgaa ttaactgcag aaaatggaca gaatcatcca atctccaggt aagtacatcc 600

aaggtgctga cgttatcaac agattgggtg aatacttgaa gccattggct gaaagatggt 660

tggttgttgg tgacaagttc gttttgggtt tcgctcaatc taccgttgaa aagtctttca 720

aggacgctgg tttggttgtt gaaatcgctc cattcggtgg tgaatgttct caaaacgaaa 780

tcgacagatt gagaggtatc gctgaaaccg ctcaatgtgg tgctatcttg ggtatcggtg 840

gtggtaagac cttggacacc gctaaggctt tggctcactt catgggtgtt ccagttgcta 900

tcgctccaac catcgcttct accgacgctc catgttctgc tttgtctgtt atctacaccg 960

acgaaggtga attcgacaga tacttgttgt tgccaaacaa cccaaacatg gttatcgttg 1020

acaccaagat cgttgctggt gctccagcta gattgttggc agctggtatc ggtgacgctt 1080

tggctacctg gttcgaagct agagcttgtt ctagatctgg tgctaccacc atggctggtg 1140

gtaagtgtac ccaagctgct ttggctttgg ctgaattgtg ttacaacacc ttgttggaag 1200

aaggtgaaaa ggctatgttg gctgctgaac aacacgttgt taccccagct ttggaaagag 1260

ttatcgaagc taacacctac ttgtctggtg ttggtttcga atctggtggt ttggctgctg 1320

ctcacgctgt tcacaacggt ttgaccgcta tcccagacgc tcaccactac taccacggtg 1380

aaaaggttgc tttcggtacc ttgacccaat tggttttgga aaacgctcca gttgaagaaa 1440

tcgaaaccgt tgctgctttg tctcacgctg ttggtttgcc aatcaccttg gctcaattgg 1500

acatcaagga agacgttcca gctaagatga gaatcgttgc tgaagctgct tgtgctgaag 1560

gtgaaaccat ccacaacatg ccaggtggtg ctaccccaga ccaagtttac gctgctttgt 1620

tggttgctga ccaatacggt caaagattcc tacaagaatg ggaataaggc gcgccccttt 1680

tcctttgtcg atatcatgta attagttatg tcacgcttac attcacgccc tcctcccaca 1740

tccgctctaa ccgaaaagga aggagttaga caacctgaag tctaggtccc tatttatttt 1800

ttttaatagt tatgttagta ttaagaacgt tatttatatt tcaaattttt ctttttttc 1860

tgtacaaacg cgtgtacgca tgtaacatta tactgaaaac cttgcttgag aaggttttgg 1920

gacgctcgaa ggcttcctag gcgtacgtt 1949

<210> 52
<211> 552
<212> PRT
<213> Citrobacter freundii

<400> 52
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

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

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> 53
<211> 1678
<212> ADN
<213> Artificial

<220>
<223> Secuencia de nucleótidos que codifica dhaK de C. freundii optimizada en codones para levadura

<400> 53
tctagaaaaa tgtctcaatt cttcttcaac cagagaaccc acttggtttc tgacgttatc 60

gacggtgcta tcatcgcttc accatggaac aatttggcta gattggaatc tgacccagct 120

atcagaatcg ttgttagaag agacttgaac aagaacaacg ttgctgttat ctctggtggt 180

ggttctggtc acgaaccagc tcacgttggt ttcatcggta agggtatgtt gaccgctgct 240

gtttgtggtg acgttttcgc ttctccatct gttgacgctg ttttgactgc tatccaagct 300

gttaccggtg aagctggttg tttgttgatc gttaagaact acaccggtga cagattgaac 360

ttcggtttgg ctgctgaaaa ggctagaaga ttgggttaca acgttgaaat gttgatcgtt 420

ggtgacgaca tctctttgcc agacaacaag cacccaagag gtatcgctgg taccatcttg 480

gttcacaaga tcgctggtta cttcgctgaa agaggttaca acttagctac cgttttgaga 540

gaagctcaat acgctgcttc taacaccttc tctttgggtg ttgctttgtc ttcttgtcac 600

ttgccacaag aaaccgacgc tgctccaaga caccacccag gtcacgctga attgggtatg 660

ggtatccacg gtgaaccagg tgcttctgtt atcgacaccc aaaactctgc tcaagttgtt 720

aacttgatgg ttgacaagtt gttggctgct ttgccagaaa ccggtagatt ggctgttatg 780

atcaacaact tgggtggtgt ttctgttgct gaaatggcta tcatcaccag agaattggct 840

tcttctccat tgcactcaag aatcgactgg ttgatcggtc cagcttcttt ggtaaccgct 900

ttggacatga agggtttctc tttgaccgct atcgttttgg aagaatctat cgaaaaggct 960

ttgttgaccg aagttgaaac ctctaactgg ccaaccccag ttccaccaag agaaatcacc 1020

tgtgttgttt cttctcacgc ttctgctaga gttgaattcc aaccatctgc taacgctttg 1080

gttgctggta tcgttgaatt ggttaccgct accttgtctg acttggaaac ccacttgaac 1140

gctttggacg ctaaggttgg tgacggtgac accggttcta ccttcgctgc tgctgctaga 1200

gaaatcgctt ctttgttgca cagacaacaa ttgccattga acaacttggc taccttgttc 1260

gctttgatcg gtgaaagatt gaccgttgtt atgggtggtt cttctggtgt tttgatgtct 1320

atcttcttca ccgctgctgg tcaaaagttg gaacaaggtg ctaacgttgt tgaagctttg 1380

aacaccggtt tggctcaaat gaagttctac ggtggtgctg acgaaggtga cagaaccatg 1440

atcgacgctt tgcaaccagc tttgacctct ttgttggctc aaccaaagaa cttgcaagct 1500

gctttcgacg ctgctcaagc tggtgctgaa agaacctgtt tgtcttctaa ggctaacgct 1560

ggtagagctt cttacttgtc ttctgaatct ttgttgggta acatggaccc aggtgctcaa 1620

agattggcta tggttttcaa ggctttggct gaatctgaat tgggttaata aggtcgac 1678

<210> 54
<211> 2951
<212> ADN
<213> Artificial

<220>
<223> pRNdhaK

<400> 54
ggatccacta gtaacggccg ccagtgtgct ggaattcgcc cttctcgagc ttaagacgcg 60

tttcttcttc agattccctc atggagaaag tgcggcagat gtatatgaca gagtcgccag 120

tttccaagag actttattca ggcacttcca tgataggcaa gagagaagac ccagagatgt 180

tgttgtccta gttacacatg gtatttattc cagagtattc ctgatgaaat ggtttagatg 240

gacatacgaa gagtttgaat cgtttaccaa tgttcctaac gggagcgtaa tggtgatgga 300

actggacgaa tccatcaata gatacgtcct gaggaccgtg ctacccaaat ggactgattg 360

tgagggagac ctaactacat agtgtttaaa gattacggat atttaactta cttagaataa 420

tgccattttt ttgagttata ataatcctac gttagtgtga gcgggattta aactgtgagg 480

accttaatac attcagacac ttctgcggta tcaccctact tattcccttc gagattatat 540

ctaggaaccc atcaggttgg tggaagatta cccgttctaa gacttttcag cttcctctat 600

tgatgttaca cctggacacc ccttttctgg catccagttt ttaatcttca gtggcatgtg 660

agattctccg aaattaatta aagcaatcac acaattctct cggataccac ctcggttgaa 720

actgacaggt ggtttgttac gcatgctaat gcaaaggagc ctatatacct ttggctcggc 780

tgctgtaaca gggaatataa agggcagcat aatttaggag tttagtgaac ttgcaacatt 840

tactattttc ccttcttacg taaatatttt tctttttaat tctaaatcaa tctttttcaa 900

ttttttgttt gtattctttt cttgcttaaa tctataacta caaaaaacac atacataaat 960

ctagaaaaat gtctcaattc ttcttcaacc agagaaccca cttggtttct gacgttatcg 1020

acggtgctat catcgcttca ccatggaaca atttggctag attggaatct gacccagcta 1080

tcagaatcgt tgttagaaga gacttgaaca agaacaacgt tgctgttatc tctggtggtg 1140

gttctggtca cgaaccagct cacgttggtt tcatcggtaa gggtatgttg accgctgctg 1200

tttgtggtga cgttttcgct tctccatctg ttgacgctgt tttgactgct atccaagctg 1260

ttaccggtga agctggttgt ttgttgatcg ttaagaacta caccggtgac agattgaact 1320

tcggtttggc tgctgaaaag gctagaagat tgggttacaa cgttgaaatg ttgatcgttg 1380

gtgacgacat ctctttgcca gacaacaagc acccaagagg tatcgctggt accatcttgg 1440

ttcacaagat cgctggttac ttcgctgaaa gaggttacaa cttagctacc gttttgagag 1500

aagctcaata cgctgcttct aacaccttct ctttgggtgt tgctttgtct tcttgtcact 1560

tgccacaaga aaccgacgct gctccaagac accacccagg tcacgctgaa ttgggtatgg 1620

gtatccacgg tgaaccaggt gcttctgtta tcgacaccca aaactctgct caagttgtta 1680

acttgatggt tgacaagttg ttggctgctt tgccagaaac cggtagattg gctgttatga 1740

tcaacaactt gggtggtgtt tctgttgctg aaatggctat catcaccaga gaattggctt 1800

cttctccatt gcactcaaga atcgactggt tgatcggtcc agcttctttg gtaaccgctt 1860

tggacatgaa gggtttctct ttgaccgcta tcgttttgga agaatctatc gaaaaggctt 1920

tgttgaccga agttgaaacc tctaactggc caaccccagt tccaccaaga gaaatcacct 1980

gtgttgtttc ttctcacgct tctgctagag ttgaattcca accatctgct aacgctttgg 2040

ttgctggtat cgttgaattg gttaccgcta ccttgtctga cttggaaacc cacttgaacg 2100

ctttggacgc taaggttggt gacggtgaca ccggttctac cttcgctgct gctgctagag 2160

aaatcgcttc tttgttgcac agacaacaat tgccattgaa caacttggct accttgttcg 2220

ctttgatcgg tgaaagattg accgttgtta tgggtggttc ttctggtgtt ttgatgtcta 2280

tcttcttcac cgctgctggt caaaagttgg aacaaggtgc taacgttgtt gaagctttga 2340

acaccggttt ggctcaaatg aagttctacg gtggtgctga cgaaggtgac agaaccatga 2400

tcgacgcttt gcaaccagct ttgacctctt tgttggctca accaaagaac ttgcaagctg 2460

ctttcgacgc tgctcaagct ggtgctgaaa gaacctgttt gtcttctaag gctaacgctg 2520

gtagagcttc ttacttgtct tctgaatctt tgttgggtaa catggaccca ggtgctcaaa 2580

gattggctat ggttttcaag gctttggctg aatctgaatt gggttaataa ggtcgagaca 2640

aatcgctctt aaatatatac ctaaagaaca ttaaagctat attataagca aagatacgta 2700

aattttgctt atattattat acacatatca tatttctata tttttaagat ttggttatat 2760

aatgtacgta atgcaaagga ataaaatttt atacattatt gaacagcgtc caagtaacta 2820

cattatgtgc actaatagtt tagcgtcgtg aagactttat tgtgtcgcga aaagtaaaaa 2880

ttttaaaaat tagagcacct tgaacttgcg aaaaaggttc tcatcaactg tttaaaacgt 2940

acgaagctta a 2951

<210> 55
<211> 14200
<212> ADN
<213> Artificial

<220>
<223> pRN957

<400> 55

gacgaaaggg cctcgtgata cgcctatttt tataggttaa tgtcatgata ataagacgtc 60

aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 120

ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 180

aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 240

ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 300

gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 360

ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 420

ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca 480

gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 540

aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct 600

gacaacgatc ggaggaccga aggagctaac cgcttttttt 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 cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 960

agttatctac acgacgggca gtcaggcaac tatggatgaa cgaaatagac agatcgctga 1020

gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 1080

ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttttga 1140

taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 1200

agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca 1260

aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 1320

ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta 1380

gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 1440

aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 1500

aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca 1560

gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agcattgaga 1620

aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 1680

aacaggagag cgcacgaggg agcttccagg ggggaacgcc tggtatcttt atagtcctgt 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 cgcagcgagt cagtgagcga 1980

ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 2040

atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 2100

tgtgagttac ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcctat 2160

gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 2220

cgccaagctc ggaattaacc ctcactaaag ggaacaaaag ctgggtaccg ggcccccccct 2280

cgaggtcgag cttaagacgc gtttcttctt cagattccct catggagaaa gtgcggcaga 2340

tgtatatgac agagtcgcca gtttccaaga gactttattc aggcacttcc atgataggca 2400

agagagaaga cccagagatg ttgttgtcct agttacacat ggtatttatt ccagagtatt 2460

cctgatgaaa tggtttagat ggacatacga agagtttgaa tcgtttacca atgttcctaa 2520

cgggagcgta atggtgatgg aactggacga atccatcaat agatacgtcc tgaggaccgt 2580

gctacccaaa tggactgatt gtgagggaga cctaactaca tagtgtttaa agattacgga 2640

tatttaactt acttagaata atgccatttt tttgagttat aataatccta cgttagtgtg 2700

agcgggattt aaactgtgag gaccttaata cattcagaca cttctgcggt atcaccctac 2760

ttattccctt cgagattata tctaggaacc catcaggttg gtggaagatt acccgttcta 2820

agacttttca gcttcctcta ttgatgttac acctggacac ccctttttctg gcatccagtt 2880

tttaatcttc agtggcatgt gagattctcc gaaattaatt aaagcaatca cacaattctc 2940

tcggatacca cctcggttga aactgacagg tggtttgtta cgcatgctaa tgcaaaggag 3000

cctatatacc tttggctcgg ctgctgtaac agggaatata aagggcagca taatttagga 3060

gtttagtgaa cttgcaacat ttactatttt cccttcttac gtaaatattt ttctttttaa 3120

ttctaaatca atctttttca atttttttgtt tgtattcttt tcttgcttaa atctataact 3180

acaaaaaaca catacataaa tctagaaaaa tgtctcaatt cttcttcaac cagagaaccc 3240

acttggtttc tgacgttatc gacggtgcta tcatcgcttc accatggaac aatttggcta 3300

gattggaatc tgacccagct atcagaatcg ttgttagaag agacttgaac aagaacaacg 3360

ttgctgttat ctctggtggt ggttctggtc acgaaccagc tcacgttggt ttcatcggta 3420

agggtatgtt gaccgctgct gtttgtggtg acgttttcgc ttctccatct gttgacgctg 3480

ttttgactgc tatccaagct gttaccggtg aagctggttg tttgttgatc gttaagaact 3540

acaccggtga cagattgaac ttcggtttgg ctgctgaaaa ggctagaaga ttgggttaca 3600

acgttgaaat gttgatcgtt ggtgacgaca tctctttgcc agacaacaag cacccaagag 3660

gtatcgctgg taccatcttg gttcacaaga tcgctggtta cttcgctgaa agaggttaca 3720

acttagctac cgttttgaga gaagctcaat acgctgcttc taacaccttc tctttgggtg 3780

ttgctttgtc ttcttgtcac ttgccacaag aaaccgacgc tgctccaaga caccacccag 3840

gtcacgctga attgggtatg ggtatccacg gtgaaccagg tgcttctgtt atcgacaccc 3900

aaaactctgc tcaagttgtt aacttgatgg ttgacaagtt gttggctgct ttgccagaaa 3960

ccggtagatt ggctgttatg atcaacaact tgggtggtgt ttctgttgct gaaatggcta 4020

tcatcaccag agaattggct tcttctccat tgcactcaag aatcgactgg ttgatcggtc 4080

cagcttcttt ggtaaccgct ttggacatga agggtttctc tttgaccgct atcgttttgg 4140

aagaatctat cgaaaaggct ttgttgaccg aagttgaaac ctctaactgg ccaaccccag 4200

ttccaccaag agaaatcacc tgtgttgttt cttctcacgc ttctgctaga gttgaattcc 4260

aaccatctgc taacgctttg gttgctggta tcgttgaatt ggttaccgct accttgtctg 4320

acttggaaac ccacttgaac gctttggacg ctaaggttgg tgacggtgac accggttcta 4380

ccttcgctgc tgctgctaga gaaatcgctt ctttgttgca cagacaacaa ttgccattga 4440

acaacttggc taccttgttc gctttgatcg gtgaaagatt gaccgttgtt atgggtggtt 4500

cttctggtgt tttgatgtct atcttcttca ccgctgctgg tcaaaagttg gaacaaggtg 4560

ctaacgttgt tgaagctttg aacaccggtt tggctcaaat gaagttctac ggtggtgctg 4620

acgaaggtga cagaaccatg atcgacgctt tgcaaccagc tttgacctct ttgttggctc 4680

aaccaaagaa cttgcaagct gctttcgacg ctgctcaagc tggtgctgaa agaacctgtt 4740

tgtcttctaa ggctaacgct ggtagagctt cttacttgtc ttctgaatct ttgttgggta 4800

acatggaccc aggtgctcaa agattggcta tggttttcaa ggctttggct gaatctgaat 4860

tgggttaata aggtcgagac aaatcgctct taaatatata cctaaagaac attaaagcta 4920

tattataagc aaagatacgt aaattttgct tatattatta tacacatatc 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 gcgtcccaaa accttctcaa 5220

gcaaggtttt cagtataatg ttacatgcgt acacgcgttt gtacagaaaa aaaagaaaaa 5280

tttgaaatat aaataacgtt cttaatacta acataactat taaaaaaaat aaatagggac 5340

ctagacttca ggttgtctaa ctccttcctt ttcggttaga gcggatgtgg gaggagggcg 5400

tgaatgtaag cgtgacataa ctaattacat gatatcgaca aaggaaaagg ggcgcgcctt 5460

attcccattc ttgtaggaat ctttgaccgt attggtcagc aaccaacaaa gcagcgtaaa 5520

cttggtctgg ggtagcacca cctggcatgt tgtggatggt ttcaccttca gcacaagcag 5580

cttcagcaac gattctcatc ttagctggaa cgtcttcctt gatgtccaat tgagccaagg 5640

tgattggcaa accaacagcg tgagacaaag cagcaacggt ttcgatttct tcaactggag 5700

cgttttccaa aaccaattgg gtcaaggtac cgaaagcaac cttttcaccg tggtagtagt 5760

ggtgagcgtc tgggatagcg gtcaaaccgt tgtgaacagc gtgagcagca gccaaaccac 5820

cagattcgaa accaacacca gacaagtagg tgttagcttc gataactctt tccaaagctg 5880

gggtaacaac gtgttgttca gcagccaaca tagccttttc accttcttcc aacaaggtgt 5940

tgtaacacaa ttcagccaaa gccaaagcag cttgggtaca cttaccacca gccatggtgg 6000

tagcaccaga tctagaacaa gctctagctt cgaaccaggt agccaaagcg tcaccgatac 6060

cagctgccaa caatctagct ggagcaccag caacgatctt ggtgtcaacg ataaccatgt 6120

ttgggttgtt tggcaacaac aagtatctgt cgaattcacc ttcgtcggtg tagataacag 6180

acaaagcaga acatggagcg tcggtagaag cgatggttgg agcgatagca actggaacac 6240

ccatgaagtg agccaaagcc ttagcggtgt ccaaggtctt accaccaccg atacccaaga 6300

tagcaccaca ttgagcggtt tcagcgatac ctctcaatct gtcgatttcg ttttgagaac 6360

attcaccacc gaatggagcg atttcaacaa ccaaaccagc gtccttgaaa gacttttcaa 6420

cggtagattg agcgaaaccc aaaacgaact tgtcaccaac aaccaaccat ctttcagcca 6480

atggcttcaa gtattcaccc aatctgttga taacgtcagc accttggatg tacttacctg 6540

gagattggat gattctgtcc attttctgca gttaattcag taaattttcg atcttgggaa 6600

gaaaaaagca gtaagcgtga aaaatctaaa agctgatgta gtagaagatc ctattcttta 6660

acaaagattg acctttttctt tttcttcttg gtttgagtag aaaggggaag gaagaataca 6720

agagagagga aaaaaaggaa gataaaaaga gagcgtgata taaatgaata tatattaaac 6780

aagagagatt gggaaggaaa ggatcaaaca aacccaaaaa tatttcaaaa aggagagaga 6840

gaggcgagtt tggtttcaaa acggtttatt tatttatgca agaggacgtg gaagaaaaag 6900

aagaaggaag aaaaaaattt gaaagaaaaa aacgcgtggc gggtaaagaa gaaaatggaa 6960

aatagaggcc gggtgacaga gaaatattga gggttaattg gaaaatatgt tagggtgagg 7020

catatgtttt taagggtttt gaggatccga taaggaagaa tgtaggttaa atgttgtgca 7080

ttaattgctg tggcagctta cccgcttccc cacacattta ctagtctcga gctcttcaac 7140

tcaagacgca cagatattat aacatctgca taataggcat ttgcaagaat tactcgtgag 7200

taaggaaaga gtgaggaact atcgcatacc tgcatttaaa gatgccgatt tgggcgcgaa 7260

tcctttattt tggcttcacc ctcatactat tatcagggcc agaaaaagga agtgtttccc 7320

tccttcttga attgatgtta ccctcataaa gcacgtggcc tcttatcgag aaagaaatta 7380

ccgtcgctcg tgatttgttt gcaaaaagaa caaaactgaa aaaacccaga cacgctcgac 7440

ttcctgtctt cctattgatt gcagcttcca atttcgtcac acaacaaggt cctagcgacg 7500

gctcacaggt tttgtaacaa gcaatcgaag gttctggaat ggcgggaaag ggtttagtac 7560

cacatgctat gatgcccact gtgatctcca gagcaaagtt cgttcgatcg tactgttact 7620

ctctctcttt caaacagaat tgtccgaatc gtgtgacaac aacagcctgt tctcacacac 7680

tcttttcttc taaccaaggg ggtggtttag tttagtagaa cctcgtgaaa cttacattta 7740

catatatata aacttgcata aattggtcaa tgcaagaaat acatatttgg tcttttctaa 7800

ttcgtagttt ttcaagttct tagatgcttt ctttttctct tttttacaga tcatcaagga 7860

agtaattatc tacttttttac aacaaatata tctagaaaat ggctgttacc aacgttgctg 7920

aattgaacgc tttggttgaa agggttaaga aggctcaaag agaatacgct tctttcaccc 7980

aagaacaagt tgacaagatc ttcagagctg ctgctttggc tgctgctgac gctagaatcc 8040

cattggctaa gatggctgtt gctgaatctg gtatgggtat cgttgaagac aaggttatca 8100

agaaccactt cgcttctgaa tacatctaca acgcttacaa ggacgaaaag acctgtggtg 8160

ttttgtcaga agacgacacc ttcggtacca tcaccatcgc tgaaccaatc ggtatcatct 8220

gtggtatcgt tccaaccacc aacccaacct ctaccgctat cttcaagtct ttgatctctt 8280

tgaagaccag aaacgctatc atcttctctc cacacccaag agctaaagac gctaccaaca 8340

aggctgctga catcgttttg caagctgcta tcgctgctgg tgctccaaag gacttgatcg 8400

gttggatcga ccaaccatct gttgaattgt ctaacgcttt gatgcaccac ccagacatca 8460

acttgatctt ggctaccggt ggtccaggta tggttaaggc tgcttactct tctggtaagc 8520

cagctatcgg tgttggtgct ggtaacaccc cagttgttat cgacgaaacc gctgacatca 8580

agagagctgt tgcttctgtt ttgatgtcta agaccttcga caacggtgtt atctgtgctt 8640

ctgaacaatc tgttgttgtt gttgactctg tttacgacgc tgttagagaa agattcgcta 8700

cccacggtgg ttacttgttg caaggtaagg aattgaaggc tgttcaagac gttatcttga 8760

agaacggtgc tttgaacgct gctatcgttg gtcaaccagc ttacaagatc gctgaattag 8820

ctggtttctc tgttccagaa aacaccaaga tcttgatcgg tgaagttacc gttgttgacg 8880

aatctgaacc attcgctcac gaaaagttgt ctccaacctt ggctatgtac agagctaagg 8940

acttcgaaga cgctgttgaa aaagctgaaa agttggttgc tatgggtggt attggtcaca 9000

cctcttgttt gtacaccgac caagacaacc aaccagctag agtttcttac ttcggtcaaa 9060

agatgaagac cgctagaatc ttgatcaaca ccccagcttc tcaaggtggt atcggtgact 9120

tgtacaactt caagttggct ccatctttga ccttgggttg tggttcttgg ggtggtaact 9180

ctatctctga aaacgttggt ccaaagcact tgatcaacaa gaagaccgtt gctaagagag 9240

ctgaaaacat gttgtggcac aagttgccaa aatctatcta cttcagaaga ggttctttgc 9300

caatcgcttt ggacgaagtt atcaccgacg gtcacaagag agctttgatc gttaccgaca 9360

gattcttgtt caacaacggt tacgctgacc aaatcacctc tgttttgaag gctgctggtg 9420

ttgaaaccga agttttcttc gaagttgaag ctgacccaac cttgtctatc gttagaaagg 9480

gtgctgaatt ggctaactct ttcaagccag acgttatcat cgctttgggt ggtggttctc 9540

caatggacgc tgctaagatc atgtgggtta tgtacgaaca cccagaaacc cacttcgaag 9600

aattggcttt gagattcatg gacatcagaa agagaatcta caagttccca aagatgggtg 9660

ttaaggctaa gatgatcgct gttaccacca cctctggtac cggttctgaa gttaccccat 9720

tcgctgttgt taccgacgac gctaccggtc aaaagtaccc attggctgac tacgctttga 9780

ccccagacat ggctatcgtt gacgctaact tggttatgga catgccaaag tctttgtgtg 9840

ctttcggtgg tttggacgct gttacccacg ctatggaagc ttacgtttct gttttggctt 9900

ctgaattctc tgacggtcaa gctttgcaag ctttgaagtt gttgaaggaa tacttgccag 9960

cttcttacca cgaaggttct aagaacccag ttgctagaga aagagttcac tctgctgcta 10020

ccatcgctgg tatcgctttc gctaacgctt tcttgggtgt ttgtcactct atggctcaca 10080

agttgggttc tcaattccac atcccacacg gtttggctaa cgctttgttg atctgtaacg 10140

ttatcagata caacgctaac gacaacccaa ccaagcaaac cgctttctct caatacgaca 10200

gaccacaagc tagaagaaga tacgctgaaa tcgctgacca cttgggtttg tctgctccag 10260

gtgacagaac cgctgcaaag atcgaaaagt tgttggcttg gttggaaacc ttgaaggctg 10320

aattgggtat cccaaagtct atcagagaag ctggtgttca agaagctgac ttcttggcta 10380

acgttgacaa gttgtctgaa gacgctttcg acgaccaatg taccggtgct aacccaagat 10440

acccattgat ctctgaattg aagcaaatct tgttggacac ctactacggt agagactacg 10500

ttgaaggtga aaccgctgct aagaaggaag ctgctccagc taaggctgaa aagaaggcta 10560

agaagtctgc ttagcttaag cgcgcgaatt tcttatgatt tatgattttt attattaaat 10620

aagttataaa aaaaataagt gtatacaaat tttaaagtga ctcttaggtt ttaaaacgaa 10680

aattcttatt cttgagtaac tctttcctgt aggtcaggtt gctttctcag gtatagcatg 10740

aggtcgctct tattgaccac acctctaccg gcatgccgag caaatgcctg caaatcgctc 10800

cccatttcac ccaattgtag atatgctaac tccagcaatg agttgatgaa tctcggtgtg 10860

tattttatgt cctcagagga caacacctgt tgtaatcgtt cttccacacg tacgaagctt 10920

aaaagggcga attctgcaga tatccatcac actggcggcc gcatgctagc tccggattat 10980

cgatgataag ctgtcaaaga tgagaattaa ttccacggac tatagactat actagatact 11040

ccgtctactg tacgatacac ttccgctcag gtccttgtcc tttaacgagg ccttaccact 11100

cttttgttac tctattgatc cagctcagca aaggcagtgt gatctaagat tctatcttcg 11160

cgatgtagta aaactagcta gaccgagaaa gagactagaa atgcaaaagg cacttctaca 11220

atggctgcca tcattattat ccgatgtgac gctgcagctt ctcaatgata ttcgaatacg 11280

ctttgaggag atacagccta atatccgaca aactgtttta cagatttacg atcgtacttg 11340

ttacccatca ttgaattttg aacatccgaa cctgggagtt ttccctgaaa cagatagtat 11400

atttgaacct gtataataat atatagtcta gcgctttacg gaagacaatg tatgtatttc 11460

ggttcctgga gaaactattg catctattgc ataggtaatc ttgcacgtcg catccccggt 11520

tcattttctg cgtttccatc ttgcacttca atagcatatc tttgttaacg aagcatctgt 11580

gcttcatttt gtagaacaaa aatgcaacgc gagagcgcta atttttcaaa caaagaatct 11640

gagctgcatt tttacagaac agaaatgcaa cgcgaaagcg ctattttacc aacgaagaat 11700

ctgtgcttca tttttgtaaa acaaaaatgc aacgcgacga gagcgctaat ttttcaaaca 11760

aagaatctga gctgcatttt tacagaacag aaatgcaacg cgagagcgct attttaccaa 11820

caaagaatct atacttcttt tttgttctac aaaaatgcat cccgagagcg ctatttttct 11880

aacaaagcat cttagattac tttttttctc ctttgtgcgc tctataatgc agtctcttga 11940

taacttttttg cactgtaggt ccgttaaggt tagaagaagg ctactttggt gtctattttc 12000

tcttccataa aaaaagcctg actccacttc ccgcgtttac tgattactag cgaagctgcg 12060

ggtgcatttt ttcaagataa aggcatcccc 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 gaagcggtat tcgcaatgcc ggctttcccc gtcaagctct aaatcggggg 12480

ctccctttag ggttccgatt tagtgcttta cggcacctcg accccaaaaa acttgattag 12540

ggtgatggtt cacgtagtgg gccatcgccc tgatagacgg tttttcgccc tttgacgttg 12600

gagtccacgt tctttaatag tggactcttg ttccaaactg gaacaacact caaccctatc 12660

tcggtctatt cttttgattt ataagggatt ttgccgattt cggcctattg gttaaaaaat 12720

gagctgattt aacaaaaatt taacgcgaat tttaacaaaa tattaacgtt tacaatttcc 12780

tgatgcggta ttttctcctt acgcatctgt gcggtatttc acaccgcata tgatccgtcg 12840

agttcaagag aaaaaaaaag aaaaagcaaa aagaaaaaag gaaagcgcgc ctcgttcaga 12900

atgacacgta tagaatgatg cattaccttg tcatcttcag tatcatactg ttcgtataca 12960

tacttactga cattcatagg tatacatata tacacatgta tatatatcgt atgctgcagc 13020

tttaaataat cggtgtcact acataagaac acctttggtg gagggaacat cgttggtacc 13080

attgggcgag gtggcttctc ttatggcaac cgcaagagcc ttgaacgcac tctcactacg 13140

gtgatgatca ttcttgcctc gcagacaatc aacgtggagg gtaattctgc tagcctctgc 13200

aaagctttca agaaaatgcg ggatcatctc gcaagagaga tctcctactt tctccctttg 13260

caaaccaagt tcgacaactg cgtacggcct gttcgaaaga tctaccaccg ctctggaaag 13320

tgcctcatcc aaaggcgcaa atcctgatcc aaacctttt actccacgcg ccagtagggc 13380

ctctttaaaa gcttgaccga gagcaatccc gcagtcttca gtggtgtgat ggtcgtctat 13440

gtgtaagtca ccaatgcact caacgattag cgaccagccg gaatgcttgg ccagagcatg 13500

tatcatatgg tccagaaacc ctatacctgt gtggacgtta atcacttgcg attgtgtggc 13560

ctgttctgct actgcttctg cctctttttc tgggaagatc gagtgctcta tcgctagggg 13620

accacccttt aaagagatcg caatctgaat cttggtttca tttgtaatac gctttactag 13680

ggctttctgc tctgtcatct ttgccttcgt ttatcttgcc tgctcatttt ttagtatatt 13740

cttcgaagaa atcacattac tttatataat gtataattca ttatgtgata atgccaatcg 13800

ctaagaaaaa aaaagagtca tccgctaggt ggaaaaaaaa aaatgaaaat cattaccgag 13860

gcataaaaaa atatagagtg tactagagga ggccaagagt aatagaaaaa gaaaattgcg 13920

ggaaaggact gtgttatgac ttccctgact aatgccgtgt tcaaacgata cctggcagtg 13980

actcctagcg ctcaccaagc tcttaaaacg gaattatggt gcactctcag tacaatctgc 14040

tctgatgccg catagttaag ccagccccga cacccgccaa cacccgctga cgcgccctga 14100

cgggcttgtc tgctcccggc atccgcttac agacaagctg tgaccgtctc cgggagctgc 14160

atgtgtcaga ggttttcacc gtcatcaccg aaacgcgcga 14200

<210> 56
<211> 11289
<212> ADN
<213> Artificial

<220>
<223> pRN958

<400> 56
gacgaaaggg cctcgtgata cgcctatttt ataggttaa tgtcatgata ataagacgtc 60

aggtggcact tttcggggaa atgtgcgcgg aacccctatt tgtttatttt tctaaataca 120

ttcaaatatg tatccgctca tgagacaata accctgataa atgcttcaat aatattgaaa 180

aaggaagagt atgagtattc aacatttccg tgtcgccctt attccctttt ttgcggcatt 240

ttgccttcct gtttttgctc acccagaaac gctggtgaaa gtaaaagatg ctgaagatca 300

gttgggtgca cgagtgggtt acatcgaact ggatctcaac agcggtaaga tccttgagag 360

ttttcgcccc gaagaacgtt ttccaatgat gagcactttt aaagttctgc tatgtggcgc 420

ggtattatcc cgtattgacg ccgggcaaga gcaactcggt cgccgcatac actattctca 480

gaatgacttg gttgagtact caccagtcac agaaaagcat cttacggatg gcatgacagt 540

aagagaatta tgcagtgctg ccataaccat gagtgataac actgcggcca acttacttct 600

gacaacgatc ggaggaccga aggagctaac cgcttttttt 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 cgcggtatca ttgcagcact ggggccagat ggtaagccct cccgtatcgt 960

agttatctac acgacgggca gtcaggcaac tatggatgaa cgaaatagac agatcgctga 1020

gataggtgcc tcactgatta agcattggta actgtcagac caagtttact catatatact 1080

ttagattgat ttaaaacttc atttttaatt taaaaggatc taggtgaaga tccttttttga 1140

taatctcatg accaaaatcc cttaacgtga gttttcgttc cactgagcgt cagaccccgt 1200

agaaaagatc aaaggatctt cttgagatcc ttttttttctg cgcgtaatct gctgcttgca 1260

aacaaaaaaa ccaccgctac cagcggtggt ttgtttgccg gatcaagagc taccaactct 1320

ttttccgaag gtaactggct tcagcagagc gcagatacca aatactgtcc ttctagtgta 1380

gccgtagtta ggccaccact tcaagaactc tgtagcaccg cctacatacc tcgctctgct 1440

aatcctgtta ccagtggctg ctgccagtgg cgataagtcg tgtcttaccg ggttggactc 1500

aagacgatag ttaccggata aggcgcagcg gtcgggctga acgggggggtt cgtgcacaca 1560

gcccagcttg gagcgaacga cctacaccga actgagatac ctacagcgtg agcattgaga 1620

aagcgccacg cttcccgaag ggagaaaggc ggacaggtat ccggtaagcg gcagggtcgg 1680

aacaggagag cgcacgaggg agcttccagg ggggaacgcc tggtatcttt atagtcctgt 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 cgcagcgagt cagtgagcga 1980

ggaagcggaa gagcgcccaa tacgcaaacc gcctctcccc gcgcgttggc cgattcatta 2040

atgcagctgg cacgacaggt ttcccgactg gaaagcgggc agtgagcgca acgcaattaa 2100

tgtgagttac ctcactcatt aggcacccca ggctttacac tttatgcttc cggctcctat 2160

gttgtgtgga attgtgagcg gataacaatt tcacacagga aacagctatg accatgatta 2220

cgccaagctc ggaattaacc ctcactaaag ggaacaaaag ctgggtaccg ggccctagga 2280

agccttcgag cgtcccaaaa ccttctcaag caaggttttc agtataatgt tacatgcgta 2340

cacgcgtttg tacagaaaaa aaagaaaaat ttgaaatata aataacgttc ttaatactaa 2400

cataactatt aaaaaaaata aatagggacc tagacttcag gttgtctaac tccttccttt 2460

tcggttagag cggatgtggg aggagggcgt gaatgtaagc gtgacataac taattacatg 2520

atatcgacaa aggaaaaggg gcgcgcctta ttcccattct tgtaggaatc tttgaccgta 2580

ttggtcagca accaacaaag cagcgtaaac ttggtctggg gtagcaccac ctggcatgtt 2640

gtggatggtt tcaccttcag cacaagcagc ttcagcaacg attctcatct tagctggaac 2700

gtcttccttg atgtccaatt gagccaaggt gattggcaaa ccaacagcgt gagacaaagc 2760

agcaacggtt tcgatttctt caactggagc gttttccaaa accaattggg tcaaggtacc 2820

gaaagcaacc ttttcaccgt ggtagtagtg gtgagcgtct gggatagcgg tcaaaccgtt 2880

gtgaacagcg tgagcagcag ccaaaccacc agattcgaaa ccaacaccag acaagtaggt 2940

gttagcttcg ataactcttt ccaaagctgg ggtaacaacg tgttgttcag cagccaacat 3000

agccttttca ccttcttcca acaaggtgtt gtaacacaat tcagccaaag ccaaagcagc 3060

ttgggtacac ttaccaccag ccatggtggt agcaccagat ctagaacaag ctctagcttc 3120

gaaccaggta gccaaagcgt caccgatacc agctgccaac aatctagctg gagcaccagc 3180

aacgatcttg gtgtcaacga taaccatgtt tgggttgttt ggcaacaaca agtatctgtc 3240

gaattcacct tcgtcggtgt agataacaga caaagcagaa catggagcgt cggtagaagc 3300

gatggttgga gcgatagcaa ctggaacacc catgaagtga gccaaagcct tagcggtgtc 3360

caaggtctta ccaccaccga tacccaagat agcaccacat tgagcggttt cagcgatacc 3420

tctcaatctg tcgatttcgt tttgagaaca ttcaccaccg aatggagcga tttcaacaac 3480

caaaccagcg tccttgaaag acttttcaac ggtagattga gcgaaaccca aaacgaactt 3540

gtcaccaaca accaaccatc tttcagccaa tggcttcaag tattcaccca atctgttgat 3600

aacgtcagca ccttggatgt acttacctgg agattggatg attctgtcca ttttctgcag 3660

ttaattcagt aaattttcga tcttgggaag aaaaaagcag taagcgtgaa aaatctaaaa 3720

gctgatgtag tagaagatcc tattctttaa caaagattga ccttttcttt ttcttcttgg 3780

tttgagtaga aaggggaagg aagaatacaa gagagaggaa aaaaaggaag ataaaaagag 3840

agcgtgatat aaatgaatat atattaaaca agagagattg ggaaggaaag gatcaaacaa 3900

acccaaaaat atttcaaaaa ggagagagag aggcgagttt ggtttcaaaa cggtttattt 3960

atttatgcaa gaggacgtgg aagaaaaaga agaaggaaga aaaaaatttg aaagaaaaaa 4020

acgcgtggcg ggtaaagaag aaaatggaaa atagaggccg ggtgacagag aaatattgag 4080

ggttaattgg aaaatatgtt agggtgaggc atatgttttt aagggttttg aggatccgat 4140

aaggaagaat gtaggttaaa tgttgtgcat taattgctgt ggcagcttac ccgcttcccc 4200

acacatttac tagtctcgag ctcttcaact caagacgcac agatattata acatctgcat 4260

aataggcatt tgcaagaatt actcgtgagt aaggaaagag tgaggaacta tcgcatacct 4320

gcatttaaag atgccgattt gggcgcgaat cctttatttt ggcttcaccc tcatactatt 4380

atcagggcca gaaaaaggaa gtgtttccct ccttcttgaa ttgatgttac cctcataaag 4440

cacgtggcct cttatcgaga aagaaattac cgtcgctcgt gatttgtttg caaaaagaac 4500

aaaactgaaa aaacccagac acgctcgact tcctgtcttc ctattgattg cagcttccaa 4560

tttcgtcaca caacaaggtc ctagcgacgg ctcacaggtt ttgtaacaag caatcgaagg 4620

ttctggaatg gcgggaaagg gtttagtacc acatgctatg atgcccactg tgatctccag 4680

agcaaagttc gttcgatcgt actgttactc tctctctttc aaacagaatt gtccgaatcg 4740

tgtgacaaca acagcctgtt ctcacacact cttttcttct aaccaagggg gtggtttagt 4800

ttagtagaac ctcgtgaaac ttacatttac atatatataa acttgcataa attggtcaat 4860

gcaagaaata catatttggt cttttctaat tcgtagtttt tcaagttctt agatgctttc 4920

tttttctctt ttttacagat catcaaggaa gtaattatct actttttaca acaaatatat 4980

ctagaaaatg gctgttacca acgttgctga attgaacgct ttggttgaaa gggttaagaa 5040

ggctcaaaga gaatacgctt ctttcaccca agaacaagtt gacaagatct tcagagctgc 5100

tgctttggct gctgctgacg ctagaatccc attggctaag atggctgttg ctgaatctgg 5160

tatgggtatc gttgaagaca aggttatcaa gaaccacttc gcttctgaat acatctacaa 5220

cgcttacaag gacgaaaaga cctgtggtgt tttgtcagaa gacgacacct tcggtaccat 5280

caccatcgct gaaccaatcg gtatcatctg tggtatcgtt ccaaccacca acccaacctc 5340

taccgctatc ttcaagtctt tgatctcttt gaagaccaga aacgctatca tcttctctcc 5400

acacccaaga gctaaagacg ctaccaacaa ggctgctgac atcgttttgc aagctgctat 5460

cgctgctggt gctccaaagg acttgatcgg ttggatcgac caaccatctg ttgaattgtc 5520

taacgctttg atgcaccacc cagacatcaa cttgatcttg gctaccggtg gtccaggtat 5580

ggttaaggct gcttactctt ctggtaagcc agctatcggt gttggtgctg gtaacacccc 5640

agttgttatc gacgaaaccg ctgacatcaa gagagctgtt gcttctgttt tgatgtctaa 5700

gaccttcgac aacggtgtta tctgtgcttc tgaacaatct gttgttgttg ttgactctgt 5760

ttacgacgct gttagagaaa gattcgctac ccacggtggt tacttgttgc aaggtaagga 5820

attgaaggct gttcaagacg ttatcttgaa gaacggtgct ttgaacgctg ctatcgttgg 5880

tcaaccagct tacaagatcg ctgaattagc tggtttctct gttccagaaa acaccaagat 5940

cttgatcggt gaagttaccg ttgttgacga atctgaacca ttcgctcacg aaaagttgtc 6000

tccaaccttg gctatgtaca gagctaagga cttcgaagac gctgttgaaa aagctgaaaa 6060

gttggttgct atgggtggta ttggtcacac ctcttgtttg tacaccgacc aagacaacca 6120

accagctaga gtttcttact tcggtcaaaa gatgaagacc gctagaatct tgatcaacac 6180

cccagcttct caaggtggta tcggtgactt gtacaacttc aagttggctc catctttgac 6240

cttgggttgt ggttcttggg gtggtaactc tatctctgaa aacgttggtc caaagcactt 6300

gatcaacaag aagaccgttg ctaagagagc tgaaaacatg ttgtggcaca agttgccaaa 6360

atctatctac ttcagaagag gttctttgcc aatcgctttg gacgaagtta tcaccgacgg 6420

tcacaagaga gctttgatcg ttaccgacag attcttgttc aacaacggtt acgctgacca 6480

aatcacctct gttttgaagg ctgctggtgt tgaaaccgaa gttttcttcg aagttgaagc 6540

tgacccaacc ttgtctatcg ttagaaaggg tgctgaattg gctaactctt tcaagccaga 6600

cgttatcatc gctttgggtg gtggttctcc aatggacgct gctaagatca tgtgggttat 6660

gtacgaacac ccagaaaccc acttcgaaga attggctttg agattcatgg acatcagaaa 6720

gagaatctac aagttcccaa agatgggtgt taaggctaag atgatcgctg ttaccaccac 6780

ctctggtacc ggttctgaag ttaccccatt cgctgttgtt accgacgacg ctaccggtca 6840

aaagtaccca ttggctgact acgctttgac cccagacatg gctatcgttg acgctaactt 6900

ggttatggac atgccaaagt ctttgtgtgc tttcggtggt ttggacgctg ttacccacgc 6960

tatggaagct tacgtttctg ttttggcttc tgaattctct gacggtcaag ctttgcaagc 7020

tttgaagttg ttgaaggaat acttgccagc ttcttaccac gaaggttcta agaacccagt 7080

tgctagagaa agagttcact ctgctgctac catcgctggt atcgctttcg ctaacgcttt 7140

cttgggtgtt tgtcactcta tggctcacaa gttgggttct caattccaca tcccacacgg 7200

tttggctaac gctttgttga tctgtaacgt tatcagatac aacgctaacg acaacccaac 7260

caagcaaacc gctttctctc aatacgacag accacaagct agaagaagat acgctgaaat 7320

cgctgaccac ttgggtttgt ctgctccagg tgacagaacc gctgcaaaga tcgaaaagtt 7380

gttggcttgg ttggaaacct tgaaggctga attgggtatc ccaaagtcta tcagagaagc 7440

tggtgttcaa gaagctgact tcttggctaa cgttgacaag ttgtctgaag acgctttcga 7500

cgaccaatgt accggtgcta acccaagata cccattgatc tctgaattga agcaaatctt 7560

gttggacacc tactacggta gagactacgt tgaaggtgaa accgctgcta agaaggaagc 7620

tgctccagct aaggctgaaa agaaggctaa gaagtctgct tagcttaagc gcgcgaattt 7680

cttatgattt atgattttta ttattaaata agttataaaa aaaataagtg tatacaaatt 7740

ttaaagtgac tcttaggttt taaaacgaaa attcttattc ttgagtaact ctttcctgta 7800

ggtcaggttg ctttctcagg tatagcatga ggtcgctctt attgaccaca cctctaccgg 7860

catgccgagc aaatgcctgc aaatcgctcc ccatttcacc caattgtaga tatgctaact 7920

ccagcaatga gttgatgaat ctcggtgtgt attttatgtc ctcagaggac aacacctgtt 7980

gtaatcgttc ttccacacgt acgaagctta aaagggcgaa ttctgcagat atccatcaca 8040

ctggcggccg catgctagct ccggattatc gatgataagc tgtcaaagat gagaattaat 8100

tccacggact atagactata ctagatactc cgtctactgt acgatacact tccgctcagg 8160

tccttgtcct ttaacgaggc cttaccactc ttttgttact ctattgatcc agctcagcaa 8220

aggcagtgtg atctaagatt ctatcttcgc gatgtagtaa aactagctag accgagaaag 8280

agactagaaa tgcaaaaggc acttctacaa tggctgccat cattattatc cgatgtgacg 8340

ctgcagcttc tcaatgatat tcgaatacgc tttgaggaga tacagcctaa tatccgacaa 8400

actgttttac agatttacga tcgtacttgt tacccatcat tgaattttga acatccgaac 8460

ctgggagttt tccctgaaac agatagtata tttgaacctg tataataata tatagtctag 8520

cgctttacgg aagacaatgt atgtatttcg gttcctggag aaactattgc atctattgca 8580

taggtaatct tgcacgtcgc atccccggtt cattttctgc gtttccatct tgcacttcaa 8640

tagcatatct ttgttaacga agcatctgtg cttcattttg tagaacaaaa atgcaacgcg 8700

agagcgctaa tttttcaaac aaagaatctg agctgcattt ttacagaaca gaaatgcaac 8760

gcgaaagcgc tattttacca acgaagaatc tgtgcttcat ttttgtaaaa caaaaatgca 8820

acgcgacgag agcgctaatt tttcaaacaa agaatctgag ctgcattttt acagaacaga 8880

aatgcaacgc gagagcgcta ttttaccaac aaagaatcta tacttctttt ttgttctaca 8940

aaaatgcatc ccgagagcgc tatttttcta acaaagcatc ttagattact ttttttctcc 9000

tttgtgcgct ctataatgca gtctcttgat aacttttgc actgtaggtc cgttaaggtt 9060

agaagaaggc tactttggtg tctattttct cttccataaa aaaagcctga ctccacttcc 9120

cgcgtttact gattactagc gaagctgcgg gtgcattttt tcaagataaa ggcatccccg 9180

attatattct ataccgatgt ggattgcgca tactttgtga acagaaagtg atagcgttga 9240

tgattcttca ttggtcagaa aattatgaac ggtttcttct attttgtctc tatatactac 9300

gtataggaaa tgtttacatt ttcgtattgt tttcgattca ctctatgaat agttcttact 9360

acaatttttt tgtctaaaga gtaatactag agataaacat aaaaaatgta gaggtcgagt 9420

ttagatgcaa gttcaaggag cgaaaggtgg atgggtaggt tatataggga tatagcacag 9480

agatatatag caaagagata cttttgagca atgtttgtgg aagcggtatt cgcaatgccg 9540

gctttccccg tcaagctcta aatcgggggc tccctttagg gttccgattt agtgctttac 9600

ggcacctcga ccccaaaaaa cttgattagg gtgatggttc acgtagtggg ccatcgccct 9660

gatagacggt ttttcgccct ttgacgttgg agtccacgtt ctttaatagt ggactcttgt 9720

tccaaactgg aacaacactc aaccctatct cggtctattc ttttgattta taagggattt 9780

```
tgccgatttc ggcctattgg ttaaaaaatg agctgattta acaaaaattt aacgcgaatt   9840
ttaacaaaat attaacgttt acaatttcct gatgcggtat tttctcctta cgcatctgtg   9900
cggtatttca caccgcatat gatccgtcga gttcaagaga aaaaaaaaga aaaagcaaaa   9960
agaaaaaagg aaagcgcgcc tcgttcagaa tgacacgtat agaatgatgc attaccttgt  10020
catcttcagt atcatactgt tcgtatacat acttactgac attcataggt atacatatat  10080
acacatgtat atatatcgta tgctgcagct ttaaataatc ggtgtcacta cataagaaca  10140
cctttggtgg agggaacatc gttggtacca ttgggcgagg tggcttctct tatggcaacc  10200
gcaagagcct tgaacgcact ctcactacgg tgatgatcat tcttgcctcg cagacaatca  10260
acgtggaggg taattctgct agcctctgca aagctttcaa gaaaatgcgg gatcatctcg  10320
caagagagat ctcctacttt ctccctttgc aaaccaagtt cgacaactgc gtacggcctg  10380
ttcgaaagat ctaccaccgc tctggaaagt gcctcatcca aaggcgcaaa tcctgatcca  10440
aacctttta ctccacgcgc cagtagggcc tctttaaaag cttgaccgag agcaatcccg  10500
cagtcttcag tggtgtgatg gtcgtctatg tgtaagtcac caatgcactc aacgattagc  10560
gaccagccgg aatgcttggc cagagcatgt atcatatggt ccagaaaccc tatacctgtg  10620
tggacgttaa tcacttgcga ttgtgtggcc tgttctgcta ctgcttctgc ctctttttct  10680
gggaagatcg agtgctctat cgctagggga ccacccttta aagagatcgc aatctgaatc  10740
ttggtttcat ttgtaatacg ctttactagg gctttctgct ctgtcatctt tgccttcgtt  10800
tatcttgcct gctcattttt tagtatattc ttcgaagaaa tcacattact ttatataatg  10860
tataattcat tatgtgataa tgccaatcgc taagaaaaaa aaagagtcat ccgctaggtg  10920
gaaaaaaaaa aatgaaaatc attaccgagg cataaaaaaa tatagagtgt actagaggag  10980
gccaagagta atagaaaaag aaaattgcgg gaaaggactg tgttatgact tccctgacta  11040
atgccgtgtt caaacgatac ctggcagtga ctcctagcgc tcaccaagct cttaaaacgg  11100
aattatggtg cactctcagt acaatctgct ctgatgccgc atagttaagc cagccccgac  11160
acccgccaac acccgctgac gcgccctgac gggcttgtct gctcccggca tccgcttaca  11220
gacaagctgt gaccgtctcc gggagctgca tgtgtcagag gttttcaccg tcatcaccga  11280
aacgcgcga                                                          11289
```

<210> 57
<211> 11102
<212> ADN
<213> Artificial

<220>
<223> pRN607

<400> 57

agtcgtgtct taccgggttg gactcaagac gatagttacc ggataaggcg cagcggtcgg 60

gctgaacggg gggttcgtgc acacagccca gcttggagcg aacgacctac accgaactga 120

gatacctaca gcgtgagcat tgagaaagcg ccacgcttcc cgaagggaga aaggcggaca 180

ggtatccggt aagcggcagg gtcggaacag gagagcgcac gagggagctt ccagggggga 240

acgcctggta tctttatagt cctgtcgggt ttcgccacct ctgacttgag cgtcgatttt 300

tgtgatgctc gtcaggggggg ccgagcctat ggaaaaacgc cagcaacgcg gccttttac 360

ggttcctggc cttttgctgg ccttttgctc acatgttctt tcctgcgtta tcccctgatt 420

ctgtggataa ccgtattacc gcctttgagt gagctgatac cgctcgccgc agccgaacga 480

ccgagcgcag cgagtcagtg agcgaggaag cggaagagcg cccaatacgc aaaccgcctc 540

tccccgcgcg ttggccgatt cattaatgca gctggcacga caggtttccc gactggaaag 600

cgggcagtga gcgcaacgca attaatgtga gttacctcac tcattaggca ccccaggctt 660

tacactttat gcttccggct cctatgttgt gtggaattgt gagcggataa caatttcaca 720

caggaaacag ctatgaccat gattacgcca agctcggaat taaccctcac taaagggaac 780

aaaagctggg tacgttttaa acagttgatg agaacctttt tcgcaagttc aaggtgctct 840

aatttttaaa atttttactt ttcgcgacac aataaagtct tcacgacgct aaactattag 900

tgcacataat gtagttactt ggacgctgtt caataatgta taaaatttat ttcctttgca 960

ttacgtacat tatataacca aatcttaaaa atatagaaat atgatatgtg tataataata 1020

taagcaaaat ttacgtatct ttgcttataa tatagcttta atgttcttta ggtatatatt 1080

taagagcgat ttgtctcgac ttacaaggcg ctttgaaccc ccttcaaaaa ctcacatagg 1140

cctactgcac caggatcttc tacttgagat gaatcgccga catacgaagc tctgccaaat 1200

ttggcctcga atgtagcagt ggatttagca ccttcctctg cagcttttac cgccttattg 1260

aaatccttag atgcagtaaa ttctttaacg aatggttcta aagcatcaat catggtggat 1320

gatcccttcc ttgcctttgt atatttgtat aaagtatcca atgcaattcc gagtgactta 1380

gccacaattt ccttagtgac gggttcatcc tttgatttac aaacctgaat taatccgtgt 1440

gaaaaacccg acaaaagaat agaatataaa ccaccagaag tacctcccat tgagccttca 1500

atgaaatctg aaatttgggc aaccgcctga gataatgagt ccttcgacag cttgtcaagg 1560

ttttcggtga tgcctttaac tcctgccact aaagtgtaac cacaatcacc atcaccaact 1620

tgattgtcta gttccgtaat gtgcggttcg ctcttgataa cttgttcagc accactcttc 1680

atccactcag caaacttgtc aaagtcatag gtaccgacgg cctttgctgt tacttcatta 1740

tgtaacaagt catcgttaac agatggggca gaagtctttt caaaatctgc aattggccag 1800

cccggtgcgt tcgtaaatgc gttcaacaag tctagtactg atttgatctc ctcaaaatca 1860

gattgcaaag cctttgtagc cttagtggcg tttagtaatg tgatactgaa cccattacca 1920

ttgaaggagg tcatcaatgt gccagcaatt gtttgaaccg gggttatgtt gtaattttcc 1980

tttaagaaat ccgtagtttt ggaagtgata gaactaataa caaaattaga aacaccgccg 2040

agattgttaa ctaacaagac aacttcatca tcttcatcaa actttacaaa agctctatcc 2100

ttatcgtttg gatccaatag ttttggtagc atatacttgg agatcaagtc ttcggtagaa 2160

ggaatagggt ctaaaacttt cacaccaggt tcgttatgaa tacccatacc caattccatt 2220

tgtttttcgt ttaattcact ttcgaatttc ctgccaggaa ctttacaatg gtctaaagaa 2280

gatccaatgg tcaccaaatt gtcgttgata attttagcca ctttagctgt accgtctaag 2340

ccatacttac tagaatattc ttctgcgaag gcacctacaa tcttatgaac caaaacggta 2400

cctgccaatg ctcttctacc aaccatacca ccctttttctc tgccaactgc aacatcatca 2460

cctatgacag caacgcggca gttaataccc aaggctcttg ctctctcagc ggacagacca 2520

aaatgcaaaa catcacctgt gtagttcttc acaatcaata aaacgccaga cgcattttca 2580

ttgactaaac ggattgcatt taaaatctgt tttgttgaag gggatgcaaa aatttcgcca 2640

accacggcgc cactcaacat acccttacca atgaaaccgg cgtgtgtagg ttcatgtcca 2700

ctaccaccac cagaaattaa tgcgatcttg tcggaatcgg tctttctgaa gagaattttt 2760

tcttcaggga ccagcgtaat ggaggggtta gcaagggcaa acccttgag acttgaattg 2820

actggatctg tgacttcaaa cgatttagcg gacattttct agatttatgt atgtgttttt 2880

tgtagttata gatttaagca agaaaagaat acaaacaaaa aattgaaaaa gattgattta 2940

gaattaaaaa gaaaaatatt tacgtaagaa gggaaaatag taaatgttgc aagttcacta 3000

aactcctaaa ttatgctgcc ctttatattc cctgttacag cagccgagcc aaaggtatat 3060

aggctccttt gcattagcat gcgtaacaaa ccacctgtca gtttcaaccg aggtggtatc 3120

cgagagaatt gtgtgattgc tttaattaat ttcggagaat ctcacatgcc actgaagatt 3180

aaaaactgga tgccagaaaa ggggtgtcca ggtgtaacat caatagagga agctgaaaag 3240

tcttagaacg ggtaatcttc caccaacctg atgggttcct agatataatc tcgaagggaa 3300

taagtagggt gataccgcag aagtgtctga atgtattaag gtcctcacag tttaaatccc 3360

gctcacacta acgtaggatt attataactc aaaaaaatgg cattattcta agtaagttaa 3420

atatccgtaa tctttaaaca ctatgtagtt aggtctccct cacaatcagt ccatttgggt 3480

agcacggtcc tcaggacgta tctattgatg gattcgtcca gttccatcac cattacgctc 3540

ccgttaggaa cattggtaaa cgattcaaac tcttcgtatg tccatctaaa ccatttcatc 3600

aggaatactc tggaataaat accatgtgta actaggacaa caacatctct gggtcttctc 3660

tcttgcctat catggaagtg cctgaataaa gtctcttgga aactggcgac tctgtcatat 3720

acatctgccg cactttctcc atgagggaat ctgaagaaga aaccggttta tcattatcaa 3780

tactgccatt tcaaagaata cgtaaataat taatagtagt gattttccta actttattta 3840

gtcaaaaaat tagcctttta attctgctgt aacccgtaca tgcccaaaat agggggcggg 3900

ttacacagaa tatataacat cgtaggtgtc tgggtgaaca gtttattcct ggcatccact 3960

aaatataatg gagcccgctt tttaagctgg catccagaaa aaaaaagaat cccagcacca 4020

aaatattgtt ttcttcacca accatcagtt cataggtcca ttctcttagc gcaactacag 4080

agaacagggg cacaaacagg caaaaaacgg gcacaacctc aatggagtga tgcaacctgc 4140

ctggagtaaa tgatgacaca aggcaattga cccacgcatg tatctatctc attttcttac 4200

accttctatt accttctgct ctctctgatt tggaaaaagc tgaaaaaaaa ggttgaaacc 4260

agttccctga aattattccc ctacttgact aataagtata taaagacggt aggtattgat 4320

tgtaattctg taaatctatt tcttaaactt cttaaattct acttttatag ttagtctttt 4380

ttttagtttt aaaacaccaa gaacttagtt tcgaataaac acacataaag aattcaaaat 4440

gtcgctgatc agcatcctgt ctcccctaat tacttccgag ggcttagatt caagaatcaa 4500

accttcacca aaaaaggatg cctctactac cactaagcca tcactatgga aaactactga 4560

gttcaaattc tactacattg catttctggt cgtggttccc ttgatgttct atgctgggtt 4620

acaagctagt tcgcccgaaa atccaaacta tgcaagatac gaacgtctcc tatctcaagg 4680

ttggttattt ggcagaaaag tagacaatag tgattctcaa tataggtttt tcagggacaa 4740

ttttgcgcta ttgtcagttt taatgctagt ccacacttct ataaaacgca ttgtacttta 4800

ttcaacaaat atcactaaat tgaggtttga tctgatattt ggtttgatct ttttagtggc 4860

cgctcatggt gtcaattcga taagaatttt agcccatatg ctaattttat atgccatcgc 4920

ccatgtacta aagaacttta gaagaatagc caccatcagc atttggattt atggtatttc 4980

tacgcttttt attaacgaca acttcagagc atatccattt ggtaatattt gctctttttt 5040

aagcccattg gaccattggt atagaggtat cattccaaga tgggatgtct ttttcaattt 5100

tactcttttg agagtcttaa gttacaactt ggacttctta gagaggtggg agaatttaca 5160

aaagaagaaa agtccatcct atgaatcaaa agaagctaaa tcagccattt tgctcaatga 5220

acgtgctaga ttaactgctg cacaccccat acaggactac agcttaatga attatattgc 5280

atatgttact tacacgccac ttttcattgc cggccccatt ataacattca atgattatgt 5340

ttaccaatcg aaacatacct tgccatcaat aaatttcaaa ttcatttttt actatgcggt 5400

gagattcgtt attgctctct tatctatgga gttcatttta cactttctcc acgttgtggc 5460

aatctcaaaa accaaagcgt gggaaaatga cacacctttc cagatttcca tgattggctt 5520

atttaatttg aatattattt ggctaaaact actgattccg tggaggctgt ttaggctgtg 5580

ggctttgcta gacggaatcg atacacctga aaatatgatc aggtgtgttg ataacaatta 5640

cagttcacta gcattctgga gagcttggca tagaagctac aataagtggg ttgtccgtta 5700

catatatatt cctctaggtg gttcaaaaaa tagagttttg acatcactag cagtcttttc 5760

cttcgtagct atatggcatg acatcgaact aaagttatta ttatggggtt ggctaatagt 5820

tttgttcctc ttaccagaaa tttttgctac ccaaattttc tctcattata ccgacgcagt 5880

ctggtacaga cacgtttgcg ctgtcggtgc tgttttcaac atatgggtta tgatgatcgc 5940

taatcttttt ggattctgct tgggctctga cggtactaaa aaattactaa gcgatatgtt 6000

ctgtaccgta tctggtttca aatttgtaat tttggcaagc gttagtttat tcatcgcagt 6060

acaaataatg tttgaaatca gagaagaaga aaagaggcac ggaatttacc taaaatgctg 6120

aggatcccct tttcctttgt cgatatcatg taattagtta tgtcacgctt acattcacgc 6180

cctcctccca catccgctct aaccgaaaag gaaggagtta gacaacctga agtctaggtc 6240

cctatttatt ttttttaata gttatgttag tattaagaac gttatttata tttcaaattt 6300

ttctttttt tctgtacaaa cgcgtgtacg catgtaacat tatactgaaa accttgcttg 6360

agaaggtttt gggacgctcg aaggcttcct agttctagag cggccgccac cgcggtggag 6420

ctccaattcg ccctatagtg agtcgtatta caattcactg gccgtcgttt tacaacgtcg 6480

tgactgggaa aaccctggcg ttacccaact taatcgcctt gcagcacatc ccccttcgc 6540

cagctggcgt aatagcgaag aggcccgcac cgatcgccct tcccaacagt tgcgcagcct 6600

gaatggcgaa tggcgcgacg cgccctgtag cggcgcatta agcgcggcgg gtgtggtggt 6660

tacgcgcagc gtgaccgcta cacttgccag cgccctagcg cccgctcctt tcgctttctt 6720

cccttccttt ctcgccacgt tcgccggtag tgttagacct gaacaaggtt tactaaaaat 6780

ccgtaaagaa cttcaattgt acgccaactt aaggcctcgt ccccgccggg tcacccggcc 6840

agcgacatgg aggcccagaa taccctcctt gacagtcttg acgtgcgcag ctcaggggca 6900

tgatgtgact gtcgcccgta catttagccc atacatcccc atgtataatc atttgcatcc 6960

atacattttg atggccgcac ggcgcgaagc aaaaattacg gctcctcgct gcagacctgc 7020

gagcagggaa acgctcccct cacagacgcg tgaattgtcc ccacgccgcg cccctgtaga 7080

gaaatataaa aggttaggat ttgccactga ggttcttctt tcatatactt ccttttaaaa 7140

tcttgctagg atacagttct cacatcacat ccgaacataa acaaccatgg gtaaaaagcc 7200

tgaactcacc gcgacgtctg tcgagaagtt tctgatcgaa aagttcgaca gcgtctccga 7260

cctgatgcag ctctcggagg gcgaagaatc tcgtgctttc agcttcgatg taggagggcg 7320

tggatatgtc ctgcgggtaa atagctgcgc cgatggtttc tacaaagatc gttatgttta 7380

tcggcacttt gcatcggccg cgctcccgat tccggaagtg cttgacattg gggaattcag 7440

cgagagcctg acctattgca tctcccgccg tgcacagggt gtcacgttgc aagacctgcc 7500

tgaaaccgaa ctgcccgctg ttctgcagcc ggtcgcggag gccatggatg cgatcgctgc 7560

ggccgatctt agccagacga gcgggttcgg cccattcgga ccgcaaggaa tcggtcaata 7620

cactacatgg cgtgatttca tatgcgcgat tgctgatccc catgtgtatc actggcaaac 7680

tgtgatggac gacaccgtca gtgcgtccgt cgcgcaggct ctcgatgagc tgatgctttg 7740

ggccgaggac tgccccgaag tccggcacct cgtgcacgcg gatttcggct ccaacaatgt 7800

cctgacggac aatggccgca taacagcggt cattgactgg agcgaggcga tgttcgggga 7860

ttcccaatac gaggtcgcca acatcttctt ctggaggccg tggttggctt gtatggagca 7920

gcagacgcgc tacttcgagc ggaggcatcc ggagcttgca ggatcgccgc ggctccgggc 7980

gtatatgctc cgcattggtc ttgaccaact ctatcagagc ttggttgacg gcaatttcga 8040

tgatgcagct tgggcgcagg gtcgatgcga cgcaatcgtc cgatccggag ccgggactgt 8100

cgggcgtaca caaatcgccc gcagaagcgc ggccgtctgg accgatggct gtgtagaagt 8160

actcgccgat agtggaaacc gacgccccag cactcgtccg agggcaaagg aataatcagt 8220

actgacaata aaaagattct tgttttcaag aacttgtcat ttgtatagtt tttttatatt 8280

gtagttgttc tattttaatc aaatgttagc gtgatttata tttttttttcg cctcgacatc 8340

atctgcccag atgcgaagtt aagtgcgcag aaagtaatat catgcgtcaa tcgtatgtga 8400

atgctggtcg ctatactgct gtcgattcga tactaacgcc gccatccagt gtcgacggat 8460

cctaggtgta cataaacttt ataaatgaaa ttcataatag aaacgacacg aaattacaaa 8520

atggaatatg ttcatagggt agacgaaact atatacgcaa tctacataca tttatcaaga 8580

aggagaaaaa ggaggatagt aaaggaatac aggtaagcaa attgatacta atggctcaac 8640

gtgataagga aaaagaattg cactttaaca ttaatattga caaggaggag ggcaccacac 8700

aaaaagttag gtgtaacaga aaatcatgaa actacgattc ctaatttgat attggaggat 8760

tttctctaaa aaaaaaaaaa tacaacaaat aaaaaacact caatgacctg accatttgat 8820

ggagtttaag tcaatacctt cttgaaccat ttcccataat ggtgaaagtt ccctcaagaa 8880

ttttactctg tcagaaacgg ccttacgacg tagtcgatat ggtgcactct cagtacaatc 8940

tgctctgatg ccgcatagtt aagccagccc cgacacccgc caacacccgc tgacgcgccc 9000

tgacgggctt gtctgctccc ggcatccgct tacagacaag ctgtgaccgt ctccgggagc 9060

tgcatgtgtc agaggttttc accgtcatca ccgaaacgcg cgagacgaaa gggcctcgtg 9120

atacgcctat ttttataggt taatgtcatg ataataatgg tttcttagga cggatcgctt 9180

gcctgtaact tacacgcgcc tcgtatcttt taatgatgga ataatttggg aatttactct 9240

gtgtttattt atttttatgt tttgtatttg gattttagaa agtaaataaa gaaggtagaa 9300

gagttacgga atgaagaaaa aaaaataaac aaaggtttaa aaaatttcaa caaaaagcgt 9360

actttacata tatatttatt agacaagaaa agcagattaa atagatatac attcgattaa 9420

cgataagtaa aatgtaaaat cacaggattt tcgtgtgtgg tcttctacac agacaagatg 9480

aaacaattcg gcattaatac ctgagagcag gaagagcaag ataaaaggta gtatttgttg 9540

gcgatccccc tagagtcttt tacatcttcg gaaaacaaaa actatttttt ctttaatttc 9600

ttttttttact ttctattttt aatttatata tttatattaa aaaatttaaa ttataattat 9660

ttttatagca cgtgatgaaa aggacccagg tggcactttt cggggaaatg tgcgcggaac 9720

ccctatttgt ttatttttct aaatacattc aaatatgtat ccgctcatga gacaataacc 9780

ctgataaatg cttcaataat attgaaaaag gaagagtatg agtattcaac atttccgtgt 9840

cgcccttatt ccctttttttg cggcattttg ccttcctgtt tttgctcacc cagaaacgct 9900

ggtgaaagta aaagatgctg aagatcagtt gggtgcacga gtgggttaca tcgaactgga 9960

tctcaacagc ggtaagatcc ttgagagttt tcgccccgaa gaacgttttc caatgatgag 10020

cacttttaaa gttctgctat gtggcgcggt attatcccgt attgacgccg ggcaagagca 10080

actcggtcgc cgcatacact attctcagaa tgacttggtt gagtactcac cagtcacaga 10140

aaagcatctt acggatggca tgacagtaag agaattatgc agtgctgcca taaccatgag 10200

tgataacact gcggccaact tacttctgac aacgatcgga ggaccgaagg agctaaccgc 10260

ttttttttcac aacatggggg atcatgtaac tcgccttgat cgttgggaac cggagctgaa 10320

tgaagccata ccaaacgacg agcgtgacac cacgatgcct gtagcaatgg caacaacgtt 10380

gcgcaaacta ttaactggcg aactacttac tctagcttcc cggcaacaat taatagactg 10440

gatggaggcg gataaagttg caggaccact tctgcgctcg gcccttccgg ctggctggtt 10500

tattgctgat aaatctggag ccggtgagcg tgggtctcgc ggtatcattg cagcactggg 10560

gccagatggt aagccctccc gtatcgtagt tatctacacg acgggcagtc aggcaactat 10620

ggatgaacga aatagacaga tcgctgagat aggtgcctca ctgattaagc attggtaact 10680

gtcagaccaa gtttactcat atatacttta gattgattta aaacttcatt tttaatttaa 10740

aaggatctag gtgaagatcc tttttgataa tctcatgacc aaaatccctt aacgtgagtt 10800

ttcgttccac tgagcgtcag accccgtaga aaagatcaaa ggatcttctt gagatccttt 10860

ttttctgcgc gtaatctgct gcttgcaaac aaaaaaacca ccgctaccag cggtggtttg 10920

tttgccggat caagagctac caactctttt tccgaaggta actggcttca gcagagcgca 10980

gataccaaat actgtccttc tagtgtagcc gtagttaggc caccacttca agaactctgt 11040

agcaccgcct acatacctcg ctctgctaat cctgttacca gtggctgctg ccagtggcga 11100

ta 11102

REIVINDICACIONES

1. Una célula de levadura, que comprende:

a) un gen exógeno que codifica una enzima con la capacidad de convertir piruvato y coenzima-A en formiato y acetil-CoA;

b) una modificación genética que reduce la actividad de formiato deshidrogenasa dependiente de NAD^+ específica en la célula;

c) un gen exógeno que codifica una enzima con actividad de acetaldehído deshidrogenasa, gen que confiere a la célula la capacidad de reducir acetil-CoA en acetaldehído; y

d) una modificación genética que aumenta la actividad específica de glicerol deshidrogenasa enlazada a NAD^+.

2. Una célula de levadura de acuerdo con la reivindicación 1, en donde la célula comprende, además, una modificación genética seleccionada del grupo que consiste en:

a) una modificación genética que aumenta la capacidad específica de dihidroxiacetona quinasa;

b) una modificación genética que aumenta el transporte de glicerol a la célula; codificando un gen exógeno una enzima para la activación de la piruvato formiato liasa;

c) una modificación genética que aumenta la actividad de acetil-CoA sintetasa específica en la célula; y

d) una modificación genética que reduce la actividad de glicerol 3-fosfato deshidrogenasa dependiente de NAD^+ específica en la célula, al tiempo que la célula de levadura comprende una ruta de respuesta de glicerol de alta osmolaridad funcional.

3. Una célula de levadura de acuerdo con la reivindicación 1 ó 2, en donde la célula comprende al menos uno de:

i) un gen xilosa isomerasa exógeno, gen que confiere a la célula la capacidad de isomerizar xilosa en xilulosa; y

ii) genes exógenos que codifican una L-arabinosa isomerasa, una L-ribuloquinasa 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.

4. Una célula de levadura de acuerdo con una cualquiera de las reivindicaciones 1 - 3, en donde al menos uno de:

a) la modificación genética que reduce la actividad de formiato deshidrogenasa dependiente de NAD^+ específica en la célula comprende la inactivación de al menos una copia endógena de un gen que codifica una formiato deshidrogenasa en el genoma de la célula;

b) la modificación genética que aumenta la actividad específica de la glicerol deshidrogenasa enlazada a NAD^+ es la expresión de un gen heterólogo que codifica una glicerol deshidrogenasa enlazada a NAD^+;

c) la modificación genética que aumenta el transporte del glicerol a la célula es la sobre-expresión de una secuencia de nucleótidos que codifica al menos uno de una proteína de absorción de glicerol y un canal de glicerol; y

d) la modificación genética que reduce la actividad de glicerol 3-fosfato deshidrogenasa dependiente de NAD^+ específica en la célula comprende la inactivación de al menos una copia endógena de un gen que codifica una glicerol 3-fosfato deshidrogenasa en el genoma de la célula; y

e) la modificación genética que aumenta la actividad de acetil-CoA sintetasa específica es la sobre-expresión de una secuencia de nucleótidos que codifica una acetil-CoA sintetasa; y

f) la modificación genética que aumenta la actividad específica de dihidroxiacetona quinasa es la sobre-expresión de una secuencia de nucleótidos que codifica una dihidroxiacetona quinasa.

5. Una célula de levadura de acuerdo con una cualquiera de las reivindicaciones 1-4, en donde:

a) el gen exógeno que codifica una enzima con actividad de piruvato formiato liasa comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 57% de identidad de secuencia de aminoácidos con SEQ ID NO: 1;

b) el gen exógeno que codifica una enzima con actividad de piruvato formiato liasa comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 70% de identidad de secuencia de aminoácidos con SEQ ID NO: 3;

c) el gen que codifica una formiato deshidrogenasa, cuya actividad se ha de reducir o inactivar en la célula de la invención es un gen que codifica una formiato deshidrogenasa que tiene una secuencia de aminoácidos con al menos 70% de identidad de secuencia con al menos una de SEQ ID NO: 5 y 6;

d) 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:

i) al menos 64% de identidad de secuencia de aminoácidos con SEQ ID NO: 7,

ii) al menos 76% de identidad de secuencia de aminoácidos con SEQ ID NO: 9 y

iii) al menos 61% de identidad de secuencia de aminoácidos con SEQ ID NO: 11;

e) la secuencia de nucleótidos que codifica la glicerol deshidrogenasa enlazada a NAD^+ comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 50% de identidad de secuencia de aminoácidos con SEQ ID NO: 49,

f) la secuencia de nucleótidos que codifica la dihidroxiacetona quinasa comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 50% de identidad de secuencia de aminoácidos con al menos una de SEQ ID NO: 14, 15 y 52:

g) la secuencia de nucleótidos que codifica la proteína de absorción de glicerol comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 50% de identidad de secuencia de aminoácidos con al menos una SEQ ID NO: 16 y 17, y en donde la secuencia de nucleótidos que codifica el canal de glicerol comprende una secuencia de nucleótidos que codifica una secuencia de aminoácidos con al menos 30% de identidad de secuencia de aminoácidos con la secuencia de aminoácidos entre los aminoácidos 250 y 530 de SEQ ID NO: 18;

h) la secuencia de nucleótidos que codifica la acetil-CoA sintetasa comprende una secuencia de aminoácidos con al menos 70% de identidad de secuencia de aminoácidos con al menos una de SEQ ID NOs: 19 y 20; e

i) el gen que codifica una glicerol fosfato deshidrogenasa, cuya actividad se ha de reducir o inactivar en la célula de la invención es un gen que codifica una glicerol fosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos 70% de identidad de secuencia con SEQ ID NO: 21, mientras que la célula tiene al menos una copia funcional de un gen endógeno que codifica una glicerol fosfato deshidrogenasa que tiene una secuencia de aminoácidos con al menos 70% de identidad de secuencia con SEQ ID NO: 22.

6. Una célula de levadura de acuerdo con una cualquiera de las reivindicaciones 1-5, en donde la célula de levadura comprende al menos una modificación genética adicional que resulta en una característica seleccionada del grupo que consiste en:

a) actividad específica para xilulosa quinasa incrementada;

b) flujo incrementado de la vía de pentosa fosfato

c) actividad específica para reductosa aldosa inespecífica reducida

d) transporte incrementado de al menos una de xilosa y arabinosa a la célula huésped;

e) sensibilidad disminuida a la represión de catabolitos;

f) tolerancia incrementada a etanol, osmolaridad o ácidos orgánicos; y

g) producción reducida de sub-productos.

7. Una célula de levadura de acuerdo con una cualquiera de las reivindicaciones 1-6, en donde la célula de levadura es de un género seleccionado del grupo que consiste en *Saccharomyces, Kluyveromyces, Candida, Pichia, Schizosaccharomyces, Hansenula, Kloeckera, Schwanniomyces* y *Yarrowia*.

8. Una célula de levadura de acuerdo con la reivindicación 7, en donde 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 pombes.*

9. Una célula de levadura de acuerdo con una cualquiera de las reivindicaciones precedentes, en donde al menos una de las secuencias de nucleótidos codificadoras de la enzima con actividad de piruvato formiato liasa, la enzima con actividad de piruvato formiato liasa, la enzima con actividad de acetaldehído deshidrogenasa, la xilosa isomerasa, la L-arabinosa isomerasa, la L-ribuloquinasa y la L-ribulosa-5-fosfato 4-epimerasa es una secuencia de nucleótidos, cuyo uso de codones ha sido optimizado para la expresión en la célula de levadura.

10. Un procedimiento para producir un producto de fermentación seleccionado del grupo que consiste en etanol, ácido láctico, ácido 3-hidroxi-propiónico, ácido acrílico, 1,3-propanodiol, un butanol y un producto derivado de isoprenoides, en donde el procedimiento comprende las etapas de:

a) fermentar un medio con una célula de levadura según se define en una cualquiera de las reivindicaciones 1-9, en donde el medio contiene o es alimentado con una fuente de glicerol y con una fuente de al menos una de una hexosa y una pentosa, en donde la célula de levadura fermenta el glicerol y la al menos una de la hexosa y pentosa para dar el producto de fermentación y formiato; y, opcionalmente,

b) recuperar al menos uno del producto de fermentación y formiato.

11. Un procedimiento de acuerdo con la reivindicación 11, en el que el producto de fermentación es etanol.

12. Un procedimiento de acuerdo con la reivindicación 10 u 11, en el que el medio contiene o es alimentado con un hidrolizado lignocelulósico.

13. Un procedimiento de acuerdo con una cualquiera de las reivindicaciones 10-12, en el que la célula de levadura fermenta en condiciones anaerobias.

14. Un procedimiento de acuerdo con una cualquiera de las reivindicaciones 10-12, en el que el pH del medio se regula durante el proceso de fermentación para mantener una concentración de ácido fórmico no disociado que no es mayor que 20,0 mM.