

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 579 706**

51 Int. Cl.:

C12N 15/82 (2006.01)

C12N 9/88 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.08.2012 E 12753881 (7)**

97 Fecha y número de publicación de la concesión europea: **13.04.2016 EP 2744904**

54 Título: **Microorganismos recombinantes para la producción de ácidos C4-dicarboxílicos**

30 Prioridad:

19.08.2011 US 201161525345 P

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

16.08.2016

73 Titular/es:

**NOVOZYMES, INC. (100.0%)
1445 Drew Avenue
Davis, CA 95618, US**

72 Inventor/es:

**MCFARLAND, SARAH;
BROWN, STEPHEN y
LUTTRINGER, SHERYL**

74 Agente/Representante:

TOMAS GIL, Tesifonte Enrique

ES 2 579 706 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Microorganismos recombinantes para la producción de ácidos C4-dicarboxílicos

5 Referencia a un listado de secuencias

[0001] Esta solicitud contiene un listado de secuencias en forma legible por ordenador, que se incorpora aquí por referencia.

10 Antecedentes

[0002] Los ácidos orgánicos tienen una larga historia de uso comercial en una variedad de industrias. Por ejemplo, los ácidos orgánicos se usan en las industrias de la alimentación y los piensos (ácido cítrico, ácido ascórbico, ácido láctico, ácido acético y ácido glucónico) como monómeros para la producción de varios polímeros (ácido adípico, ácido láctico, ácido acrílico y ácido itacónico), como queladores metálicos (ácido glucónico) y como solventes "verdes" (ácido acético) (Sauer *et al.*, 2008, Trends in Biotechnology 26: 100-108). Los ácidos orgánicos pueden ser productos comerciales o bloques de construcción química usados en la producción de otros productos químicos. Además de aplicaciones de especialidad, hace mucho que se ha reconocido que los ácidos C4-dicarboxílicos también pueden servir como compuestos de bloque de construcción para la producción de productos químicos industriales de gran volumen, tal como 1,4-butanodiol, tetrahidrofurano y gamma-butirolactona.

[0003] Los ácidos orgánicos se pueden producir comercialmente bien por síntesis química de materias primas derivadas de petróleo (por ejemplo, ácido fumárico, ácido málico, ácido acrílico y ácido adípico) o por fermentación microbiana (por ejemplo, ácido cítrico, ácido láctico, ácido glucónico y ácido itacónico). Algunos ácidos orgánicos tales como ácido fumárico y ácido málico también pueden ser producidos por fermentación microbiana, pero se producen actualmente comercialmente por síntesis química de materias primas petroquímicas debido a costes de producción más bajos. Sin embargo, el coste en aumento de los químicos de bloque de construcción derivados del petróleo, la inestabilidad geopolítica que afecta a los precios del petróleo y el deseo para implementar los procesos de fabricación que utilizan materias primas derivadas de recursos renovables han estimulado un interés renovado en la producción de ácidos orgánicos de producción y otros químicos por fermentación microbiana.

[0004] Mientras que los ácidos C4-dicarboxílicos tal como el ácido málico se producen comercialmente hoy por síntesis química a partir de fuentes petroquímicas, también se pueden producir por fermentación microbiana. El ácido málico se ha producido a altos niveles en levadura genéticamente modificada (*Saccharomyces cerevisiae*) (Zelle *et al.*, 2008, Appl. Environ. Microbiol. 74: 2766-2777) y los hongos filamentosos de origen natural tales como *Aspergillus* spp. (patente de EE.UU. nº 3.063.910; Bercovitz *et al.*, 1990, Appl. Environ. Microbiol. 56: 1594-1597). Abe *et al.* (patente de EE.UU. nº 3.063.910) y Bercovitz *et al.* (1990, Appl. Environ. Microbiol. 56: 1594-1597) reportaron altos niveles de producción de ácido málico en diferentes especies de *Aspergillus*. Además, Battat *et al.* (1991, Biotechnol. Bioengineering, 37: 1108-1116) reportaron producción de ácido málico tan alta como 113 g/L por *Aspergillus flavus* en un fermentador agitado bajo condiciones optimizadas. La producción de ácido dicarboxílico por fermentación microbiana en la levadura se describe en la WO 2010/003728. La producción de ácido málico por fermentación microbiana se describe también en la WO 2009/011974, WO 2009/155382 y WO2010/111344. La mejora de la producción de ácidos C4-dicarboxílicos tales como ácido málico por ingeniería genética puede permitir la producción de ácido málico comercial económico por fermentación.

[0005] La sobreproducción de ácido málico en un huésped tal como *Aspergillus* spp. ocurre bajo condiciones de cultivo específicas (condiciones aeróbicas y alta proporción C:N; el carbonato de calcio también se puede añadir como un agente neutralizante y como fuente de CO₂ para biosíntesis de ácido málico). Bajo estas condiciones, rebosar el metabolismo vía el ciclo de ácido tricarboxílico reductivo citosólico (TCA) produce una biosíntesis de ácido málico aumentada y secreción en el medio de cultivo. La producción de ácido málico aumentada ha sido reportada en *Saccharomyces cerevisiae* mediante el aumento del nivel de piruvato carboxilasa (Bauer *et al.*, 1999, FEMS Microbiol Lett. 179: 107-113) o malato deshidrogenasa (Pines *et al.*, 1997, Appl. Microbiol. Biotechnol. 48: 248-255) utilizando ingeniería genética y aumentando la expresión de un transportador de ácido málico (Zelle *et al.*, 2008, *supra*). Se ha sugerido, basado en evidencia bioquímica, que la actividad de malato deshidrogenasa está limitando la producción de ácido málico en la cepa de *Aspergillus flavus* ATCC 13697 (Peleg *et al.*, 1988, Appl. Microbiol. Biotechnol. 28: 69-75). WO 2011/028643, PCT/US11/38881, PCT/US11/38881, solicitud de EE.UU. nº.: 13/165.696 y 13/165.719 y solicitud provisional EE.UU. nº 61/447.286 describen la producción de ácido C4-dicarboxílico.

[0006] Sería ventajoso en la técnica mejorar la producción de ácido C4-dicarboxílico, tal como la producción de ácido málico, como resultado de la ingeniería genética usando técnicas de ADN recombinante. La presente invención proporciona, entre otras cosas, métodos para mejorar la producción de ácido C4-dicarboxílico (por ejemplo, la producción de ácido málico).

65 Resumen

[0007] Se describen aquí células huésped recombinantes que comprenden actividad de anhidrasa carbónica, donde la célula huésped produce (o es capaz de producir) una cantidad aumentada de un ácido C4-dicarboxílico (por ejemplo, ácido málico). En un aspecto, la célula huésped recombinante es una célula huésped fúngica y comprende un polinucleótido heterólogo que codifica un polipéptido que tiene actividad de anhidrasa carbónica, donde la célula huésped produce (o es capaz de producir) y/o secreta (o es capaz de secretar) una cantidad superior de ácido C4-dicarboxílico en comparación con la célula huésped sin el polinucleótido heterólogo cuando se cultiva bajo las mismas condiciones. En algunos aspectos, la anhidrasa carbónica es una anhidrasa carbónica citosólica. En algunos aspectos, la célula huésped comprende además un polinucleótido heterólogo que codifica un transportador de bicarbonato, un polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico, un polinucleótido heterólogo que codifica una malato deshidrogenasa, y/o un polinucleótido heterólogo que codifica una piruvato carboxilasa. En algunos aspectos, la célula huésped es una célula huésped de *Aspergillus*, tal como una célula huésped de *Aspergillus oryzae*.

[0008] También se describen los métodos del uso de células huésped recombinantes para la producción de ácidos C4-dicarboxílicos. En un aspecto, un método para producir un ácido C4-dicarboxílico (por ejemplo, ácido málico), comprende: (a) cultivo de una célula huésped recombinante (por ejemplo, una célula huésped de *Aspergillus*) que tiene actividad de anhidrasa carbónica en un medio bajo condiciones adecuadas para producir el ácido C4-dicarboxílico; y (b) recuperación del ácido C4-dicarboxílico. En algunos aspectos, la célula huésped recombinante comprende un polinucleótido heterólogo que codifica una anhidrasa carbónica. En otro aspecto, un método para producir un ácido C4-dicarboxílico (por ejemplo, ácido málico) comprende (a) transformación en una célula huésped (por ejemplo, una célula huésped de *Aspergillus*) de unos polinucleótidos heterólogos que codifican un polipéptido que tiene actividad de anhidrasa carbónica descrita aquí; (b) cultivo del organismo transformado en un medio bajo condiciones adecuadas para producir el ácido C4-dicarboxílico; y (c) recuperación del ácido C4-dicarboxílico. En algunos aspectos de los métodos, la anhidrasa carbónica es una anhidrasa carbónica citosólica. En algunos aspectos de los métodos, la célula huésped recombinante comprende además un polinucleótido heterólogo que codifica un transportador de bicarbonato, un polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico, un polinucleótido heterólogo que codifica una malato deshidrogenasa y/o un polinucleótido heterólogo que codifica una piruvato carboxilasa. En algunos aspectos de los métodos, la célula huésped es una célula huésped de *Aspergillus*, tal como una célula huésped de *Aspergillus oryzae*.

30 Breve descripción de las figuras

[0009]

La figura 1 muestra un mapa de restricción de pShTh60.

La figura 2 muestra un mapa de restricción de pAmFs69.

Las figuras 3A y 3B muestran la secuencia de constructo de nucleótidos genómicos y la secuencia de aminoácidos deducida de un gen transportador de bicarbonato (*bt1*) de *Aspergillus oryzae* NRRL 3488 (SEC ID n.º: 1 y 2, respectivamente).

Las figuras 4A y 4B muestran la secuencia de constructo de nucleótidos genómicos y la secuencia de aminoácidos deducida de un gen transportador de bicarbonato de *Aspergillus oryzae* NRRL 3488 (SEC ID n.º: 3 y 4, respectivamente).

La figura 5 muestra un mapa de restricción de pSaMF36.

La figura 6 muestra la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un gen transportador de ácido (*c4t521*) C4-dicarboxílico de *Aspergillus aculeatus* (SEC ID n.º: 5 y 6, respectivamente).

La figura 7 muestra la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un gen de malato-deshidrogenasa (*mdh3*) de *Aspergillus oryzae* NRRL 3488 (SEC ID n.º: 7 y 8, respectivamente).

La figura 8 muestra un mapa de restricción de pSaMF21.

Las figuras 9A y 9B juntas muestran la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un gen de piruvato-carboxilasa (*pyc*) de *Aspergillus oryzae* NRRL 3488 (SEC ID n.º: 9 y 10, respectivamente).

La Figura 10 muestra un mapa de restricción de pRYAN1.

La Figura 11 muestra un mapa de restricción de pShTh76.

La Figura 12 muestra un mapa de restricción de pSaMF48.

La Figura 13 muestra un mapa de restricción de pSaMF58.

La figura 14 muestra la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un gen de anhidrasa carbónica de *Aspergillus aculeatus* (ACLA 007930) (SEC ID n.º: 54 y 55, respectivamente).

La Figura 15 muestra un mapa de restricción de pShTh66.

La Figura 16 muestra un mapa de restricción de pShTh77.

La Figura 17 muestra un mapa de restricción de pShTh147.

60 Definiciones

[0010] Anhidrasa carbónica: el término "anhidrasa carbónica" se define aquí como una metaloenzima de zinc que cataliza la reacción de dióxido de carbono (CO₂) y agua (H₂O) en bicarbonato (HCO₃⁻) (EC 4.2.1.1). Clases de proteínas de anhidrasa carbónica no limitadoras incluyen las familias α, β, γ, δ y ε. La actividad de anhidrasa carbónica se puede determinar a partir de extractos libres de células como se describe en la técnica, por ejemplo, como se describe en la técnica, por ejemplo, en Khalifah, 1970, J. Biol. Chem., 246: 2561-2573.

5 [0011] Anhidrasa carbónica citosólica: el término "anhidrasa carbónica citosólica" se refiere a una anhidrasa carbónica que hace no contiene una secuencia objetivo mitocondrial (MTS) N-terminal funcional. Ejemplos no limitativos incluyen una anhidrasa carbónica que carece de una MTS o una anhidrasa carbónica que contiene una MTS modificada pero no funcional (tal como por truncamiento y/o alteración de secuencia). Secuencias MTS N-terminales sus sitios de escisión correspondientes pueden ser predichos utilizando los programas MitoProtil v1,101 y TargetP 1,1 como se describe en Elleuche, 2009, *Curr Genet*, 55, 211-222.

10 [0012] Transportador de bicarbonato: el término "transportador de bicarbonato" se definen aquí como una proteína tal como una proteína de membrana integrada capaz de facilitar la transferencia de HCO_3^- a través de la membrana biológica, tal como una membrana celular y/o la membrana de un orgánulo celular. Clases no limitativas de proteínas transportadoras de bicarbonato incluyen la familia intercambiadora aniónica (AE) de intercambiadores de $\text{Cl}^-/\text{HCO}_3^-$, la familia de NBC de cotransportadores de $\text{Na}^+/\text{HCO}_3^-$, y los intercambiadores de $\text{Cl}^-/\text{HCO}_3^-$ dependientes de Na^+ . En algunos aspectos descritos aquí, el transportador de bicarbonato es un transportador de bicarbonato de sulfato, donde el transportador es capaz de facilitar la transferencia tanto de aniones de HCO_3^- como de SO_4^{2-} a través de la membrana biológica. La actividad de intercambio de bicarbonato se puede determinar a partir de extractos libres de células como se describe en la técnica, por ejemplo, como se describe en Sterling *et al.*, 2002, *Am J Physiol Cell Physiol* 283: C1522-1529.

20 [0013] Transportador de ácido C4-dicarboxílico: el término "transportador de ácido C4-dicarboxílico" se define aquí como una permeasa de ácido dicarboxílico que puede transportar ácido málico, ácido succínico, ácido oxaloacético, ácido malónico y/o ácido fumárico fuera de una célula (Grobler *et al.*, 1995, *Yeast* 11: 1485-1491; Camarasa *et al.*, 2001, *Applied and Environmental Microbiology* 67: 4144-4151). Un método computacional para predecir proteínas importadas mitocondriales y sus secuencias objetivo es descrito por Claros and Vincens, 1996, *Eur. J. Biochem.* 241: 779-786.

25 [0014] Malato-deshidrogenasa: el término "malato-deshidrogenasa" se define aquí como una malato: NAD^+ oxidorreductasa (EC 1.1.1.37) que cataliza la reducción de oxaloacetato en presencia de $\text{NADH} + \text{H}^+$ en malato y NAD^+ . Para fines de la presente invención, la actividad de malato-deshidrogenasa se puede determinar a partir de extractos libres de células según el procedimiento siguiente. La solución de ensayo consiste en 1 mM de ácido oxaloacético, 100 mM de Tris pH 8,0, 10 mM de NaHCO_3 , 5 mM de MgCl_2 y 0,1 mM de NADH (Sigma Chemical Co., St. Louis, MO, EE.UU.). La solución de ensayo sin ácido oxaloacético como sustrato se realiza como un control para medir índices de degradación NADH de fondo. Diluciones de 1/100, 1/500, 1/2500 y 1/12500 de cada sobrenadante se preparan con agua doblemente destilada. Partes alícuotas de 270 μl de la solución de ensayo se dispensan en placas de 96 pocillos de fondo plano de poliestireno. Una muestra de 30 μl de cada sobrenadante diluido se añade para iniciar el ensayo. Las reacciones son controladas utilizando un lector de placa SPECTRAMAX® 340PC (Molecular Devices, Sunnyvale, CA, EE.UU.) con la configuración siguiente: 340 nm, lectura cinética. Una serie de concentración de NADH se utiliza para construir una curva estándar y una serie de dilución de deshidrogenasa málica purificada (Sigma Chemical Co., St. Louis, MO, EE.UU.) se usa como un control positivo. Una unidad de actividad de malato-deshidrogenasa es igual a la cantidad de enzima capaz de convertir 1 μmol de oxaloacetato y $\text{NADH} + \text{H}^+$ en malato y NAD^+ por minuto a pH 8,0, 25°C.

40 [0015] Piruvato-carboxilasa: el término "piruvato-carboxilasa" se define aquí como una piruvato:carbono-dióxido ligasa (formadora de ADP) (EC 6.4.1.1) que cataliza la carboxilación de piruvato en presencia de ATP y HCO_3^- en oxaloacetato, ADP y fosfato. Para fines de la presente invención, la actividad de piruvato-carboxilasa se puede determinar a partir de extractos libres de célula según el procedimiento del ensayo de control de la calidad de SIGMA® piruvato-carboxilasa (Sigma Chemical Co., St. Louis, MO, EE.UU.) sustituyendo el tampón Tris a pH 8,0. Una unidad de actividad de piruvato-carboxilasa es igual a la cantidad de enzima capaz de convertir 1 μmol de piruvato y CO_2 en oxaloacetato por minuto a pH 7,8, 30°C.

50 [0016] Polinucleótido heterólogo: el término "polinucleótido heterólogo" se define aquí como un polinucleótido que no es nativo de la célula huésped; un polinucleótido nativo donde una o más (por ejemplo, dos, diferentes) modificaciones estructurales se han hecho en la región codificante; un polinucleótido nativo cuya expresión es alterada de forma cuantitativa como resultado de la manipulación del ADN por técnicas de ADN recombinante, por ejemplo, un promotor (foráneo) diferente enlazado al polinucleótido; o un polinucleótido nativo cuya expresión es alterada de forma cuantitativa por la introducción de una o más copias extra del polinucleótido en la célula huésped.

60 [0017] Secuencia codificante: el término "secuencia codificante" se refiere a una secuencia de polinucleótidos, que especifica la secuencia de aminoácidos de un polipéptido. Los límites de la secuencia codificante son generalmente determinados por un marco de lectura abierto, que empieza normalmente con el codón de inicio ATG o codones de inicio alternativos tales como GTG y TTG y termina con un codón de terminación tal como, TAA, TAG y TGA. La secuencia codificante puede ser una secuencia de ADN genómico, ADNc, un polinucleótido sintético y/o un polinucleótido recombinante.

65 [0018] Secuencia de ADNc: el término "secuencia de ADNc" se refiere a una secuencia de ADN después de la transcripción inversa a partir de una molécula de ARNm madura, empalmada, obtenida a partir de una célula eucariota. La transcripción de ARN primario inicial a partir de ADN genómico es un precursor de ARNm que se procesa a través de

una serie de pasos, incluyendo empalme, antes de aparecer como ARNm empalmado maduro. Una secuencia de ADNc carece de secuencias de intrones de intervención que pueden estar presente en la secuencia de ADN genómico correspondiente. Por consiguiente, el frase "la secuencia de ADNc de SEC ID n.º: X" se refiere a la secuencia resultante después de que las secuencias de intrones de intervención de SEC ID n.º: X, si están presentes, se hayan quitado. En algunos casos, cuando una secuencia de ADN genómico referenciada carece de secuencias de intrones de intervención, una secuencia de ADNc pueden ser idéntica a su secuencia de ADN genómico correspondiente.

[0019] Secuencia de ADN genómico: el término "secuencia de ADN genómico" se refiere a una secuencia de ADN que se encuentra en el genoma de un organismo fuente (por ejemplo, un genoma eucariótico o procariótico). En algunos casos, una secuencia de ADN genómico a partir de un genoma eucariota contiene una o más secuencias de intrones de intervención que son quitadas del transcrito de ARN primario como resultado del empalme de ARN. Por consiguiente, el frase "la secuencia de ADN genómico de SEC ID n.º: Y" se refiere a la secuencia de ADN correspondiente a partir del organismo fuente que incluye secuencias de intrones de intervención, si las hay, que están presentes antes del empalme de ARN.

[0020] Secuencia polipeptídica madura: el término "secuencia polipeptídica madura" se refiere a la porción de la secuencia de polipéptidos referenciada después de cualquier modificación de secuencia postraduccional (tal como tratamiento N-terminal y/o truncamiento C-terminal). La secuencia polipeptídica madura puede ser predicha, por ejemplo, basada en el programa SignalP (Nielsen *et al.*, 1997, Protein Engineering 10: 1-6) o el programa InterProScan (The European Bioinformatics Institute). En algunos casos, la secuencia polipeptídica madura puede ser idéntica a la secuencia polipeptídica referenciada entera. Se conoce en la técnica que una célula huésped puede producir una mezcla de dos o más secuencias polipéptidas maduras diferentes (es decir, con un aminoácido C-terminal y/o N-terminal diferente) expresadas por el mismo polinucleótido.

[0021] Fragmento: el término "fragmento" se refiere a un polipéptido que tiene uno o varios (por ejemplo, dos, diferentes) aminoácidos eliminados del amino y/o carboxilo terminal de una secuencia polipeptídica referenciada. En un aspecto, el fragmento tiene actividad de anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato-deshidrogenasa o piruvato-carboxilasa. En otro aspecto, el número de residuos de aminoácidos en el fragmento es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% de cualquier anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato-deshidrogenasa o piruvato-carboxilasa descrito aquí, por ejemplo, al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de aminoácidos en la SEC ID n.º: 2, 4, 6, 8, 10, 27, 29, 32, 34, 36, 39, 41, 43, 45 o 55.

[0022] Subsecuencia: el término "subsecuencia" se refiere a un polinucleótido que tiene uno o varios (por ejemplo, dos, diferentes) nucleótidos eliminados del extremo 5' y/o 3' de la secuencia de nucleótidos referenciada. En un aspecto, la subsecuencia codifica un fragmento que tiene actividad de anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato-deshidrogenasa o piruvato-carboxilasa. En otro aspecto, el número de residuos de nucleótidos en la subsecuencia es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de nucleótidos en cualquier secuencia que codifica de una anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato-deshidrogenasa o piruvato-carboxilasa descrito aquí, por ejemplo, al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de nucleótidos en las SEC ID n.º: 1, 3, 5, 7, 9, 26, 28, 30, 31, 33, 35, 37, 38, 40, 42, 44 o 54.

[0023] Variante alélica: el término "variante alélica" se refiere a cualquiera de dos o más formas alternativas de un gen que ocupa el mismo locus cromosómico. La variación alélica surge naturalmente a través de la mutación, y puede dar como resultado el polimorfismo dentro de poblaciones. La mutaciones de gen pueden ser silenciosas (sin cambios en el polipéptido codificado) o puede codificar polipéptidos que tienen secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

[0024] Identidad de secuencia: la relación entre dos secuencias de aminoácidos o entre dos secuencias de nucleótidos se describe por el parámetro "identidad de secuencia".

[0025] Para fines descritos aquí, el grado de identidad de secuencia entre dos secuencias de aminoácidos se determina utilizando el algoritmo de Needleman-Wunsch (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453) como implementado en el programa v del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, Trends Genet. 16: 276-277), preferiblemente versión 3.0.0 o superior. Los parámetros opcionales usados son penalización por apertura de espacio de 10, penalización por extensión de espacio de 0,5 y la matriz de sustitución EBLOSUM62 (versión EMBOSS de BLOSUM62). El resultado de Needle etiquetado "identidad más larga" (obtenido utilizando la opción -nobrief) se usa como la identidad en porcentaje y se calcula de la siguiente manera:

$$\text{(Residuos idénticos x 100)} / (\text{Longitud del alineamiento} - \text{Número total de espacios en el alineamiento})$$

[0026] Para fines descritos aquí, el grado de identidad de secuencia entre dos secuencias desoxirribonucleótidas se determina utilizando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, *supra*) como implementado en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*,

2000, *supra*), preferiblemente versión 3.0.0 o superior. Los parámetros opcionales usados son penalización por apertura de espacio de 10, penalización por extensión de espacio de 05 y la matriz de sustitución EDNAFULL (versión EMBOSS de NCBI NUC4.4). El resultado de Needle etiquetado "identidad más larga" (se obtiene utilizando la opción -nobrief) se usa como la identidad en porcentaje y se calcula de la siguiente manera:

5
$$\frac{(\text{Desoxirribonucleótidos idénticos} \times 100)}{(\text{Longitud del alineamiento} - \text{Número total de espacios en el alineamiento})}$$

[0027] Expresión: el término "expresión" incluye cualquier paso implicado en la producción del polipéptido incluyendo, pero no limitado a, transcripción, modificación postranscripcional, traducción, modificación postraduccional, y secreción. La expresión se puede medir, por ejemplo, para detectar expresión aumentada por técnicas que se conocen en la técnica, tales como niveles de medición de ARNm y/o polipéptido traducido.

[0028] Constructo de ácidos nucleicos: el término "constructo de ácidos nucleicos" se refiere a un polinucleótido que comprende una o varias (por ejemplo, dos, diferentes) secuencias de control. El polinucleótido puede ser monocatenario o bicatenario, y se puede aislar de un gen de origen natural, modificado para contener segmentos de ácidos nucleicos de manera que no existirían de otro modo en la naturaleza, o sintético.

[0029] Secuencia de control: el término "secuencia de control" se refiere a una secuencia de ácidos nucleicos necesaria para la expresión del polipéptido. Las secuencias de control pueden ser nativas o foráneas del polinucleótido que codifica el polipéptido, y nativa o foráneas entre sí. Tales secuencias de control incluyen, pero de forma no limitativa, una secuencia líder, secuencia de poliadenilación, secuencia de propéptido, secuencia promotora, secuencia de péptido señal y secuencia terminadora de transcripción. Las secuencias de control pueden estar provistas de enlaces con motivo de introducir sitios de restricción específicos que faciliten el ligamiento de las secuencias de control con la región de codificación del polinucleótido que codifica un polipéptido.

[0030] Operativamente enlazado: el término "operativamente enlazado" se refiere a una configuración en la que una secuencia de control se coloca en una posición apropiada con respecto a la secuencia de codificación de un polinucleótido tal que la secuencia de control dirija la expresión de la secuencia codificante.

[0031] Vector de expresión: el término "vector de expresión" se refiere a una molécula de ADN lineal o circular que comprende un polinucleótido que codifica un polipéptido y está operativamente enlazado a secuencias de control, donde las secuencias de control proporcionan expresión del polinucleótido que codifica el polipéptido. Como mínimo, el vector de expresión comprende una secuencia promotora y secuencias de señal de parada transcripcional y traduccional.

[0032] Célula huésped: el término "célula huésped" se refiere a cualquier tipo de célula que sea susceptible de transformación, transfección, transducción y similar con un constructo de ácidos nucleicos o vector de expresión que comprenda uno o más (por ejemplo, dos, diferentes) polinucleótidos descritos aquí (por ejemplo, un polinucleótido que codifica una anhidrasa carbónica). El término "célula huésped" abarca cualquier descendiente de una célula madre que no sea idéntico a la célula madre debido a mutaciones que ocurren durante la replicación.

[0033] Interrupción: el término "interrupción" significa que un promotor, región de codificación y/o terminador de un polinucleótido que codifica un polipéptido que tiene actividad enzimática dentro de una célula huésped es parcialmente o totalmente modificado (tal como por modificación, eliminación, inserción y/o sustitución de uno o varios nucleótidos) dando como resultado la ausencia o reducción de dicha actividad enzimática de la célula huésped. La ausencia o reducción de la actividad enzimática se puede medir directamente por técnicas conocidas en la técnica (tales como mediciones de extracto libre de células referenciadas aquí); o por la ausencia o reducción de ARNm correspondiente (por ejemplo, al menos 50% de reducción, al menos 60% de reducción, al menos 70% de reducción, al menos 80% de reducción o al menos 90% de reducción); la ausencia o reducción en la cantidad de polipéptido correspondiente que tiene actividad enzimática (por ejemplo, al menos 50% de reducción, al menos 60% de reducción, al menos 70% de reducción, al menos 80% de reducción o al menos 90% de reducción); o la ausencia o reducción de la actividad específica del polipéptido correspondiente que tiene actividad enzimática (por ejemplo, al menos 50% de reducción, al menos 60% de reducción, al menos 70% de reducción, al menos 80% de reducción o al menos 90% de reducción) si está presente. Interrupciones de un gen particular de interés se pueden generar por métodos conocidos en la técnica, por ejemplo, por recombinación homóloga dirigida (véase *Methods in Yeast Genetics* (1997 edition), Adams, Gottschling, Kaiser, and Stems, Cold Spring Harbor Press (1998)) o por ARNi o tecnología antisentido.

[0034] Productividad volumétrica: el término "productividad volumétrica" se refiere a la cantidad de producto referenciado producida (por ejemplo, la cantidad de ácido C4-dicarboxílico producido) por volumen del sistema usado (por ejemplo, el volumen total de medios y contenidos en ello) por unidad de tiempo.

[0035] Medio Fermentable: el término "medio fermentable" se refiere a un medio que comprende uno o varios (por ejemplo, dos, diferentes) azúcares, tales como glucosa, fructosa, sacarosa, celobiosa, xilosa, xilulosa, arabinosa, manosa, galactosa y/o oligosacáridos solubles, donde el medio es capaz, en parte, de ser (fermentado) convertido por una célula huésped en un producto deseado, tal como ácido C4-dicarboxílico. En algunos casos, el medio de fermentación se deriva de una fuente natural, tal como azúcar de caña, almidón o celulosa, y puede ser el resultado del pretratamiento de la fuente por hidrólisis enzimática (sacarificación).

[0036] Referencia a "aproximadamente" un valor o parámetro aquí incluye aspectos que se dirigen a ese valor o parámetro *per se*. Por ejemplo, la descripción en referencia a "aproximadamente X" incluye el aspecto "X". Cuando se usa en combinación con valores medidos, "aproximadamente" incluye una gama que abarca al menos la incertidud asociada al método de la medición del valor particular, y puede incluir una gama de más o menos dos desviaciones estándar alrededor del valor declarado.

[0037] Como se utiliza en este caso y en las reivindicaciones anexas, las formas singulares "un/una" o "el/la" incluyen referentes plurales a menos que el contexto dicte claramente lo contrario. Se entiende que los aspectos descritos aquí incluyen aspectos "que consisten" y/o "que consisten esencialmente en".

[0038] A menos que se defina de otro modo o claramente se indique por contexto, todos los términos técnicos y científicos usados aquí tienen el mismo sentido que entiende comúnmente uno técnico en la materia.

15 Descripción detallada

[0039] Se describe aquí, entre otras cosas, la expresión aumentada de genes específicos en una célula huésped recombinante, tal como un hongo filamentoso (por ejemplo, *Aspergillus*) para mejorar la producción de ácidos C4-dicarboxílicos (por ejemplo, ácido málico). La célula huésped comprende un polinucleótido heterólogo para la expresión de una anhidrasa carbónica. En un aspecto, la anhidrasa carbónica es sobreexpresada bajo condiciones de cultivo para producir ácido C4-dicarboxílico en altos títulos. La célula huésped recombinante puede comprender además un polinucleótido heterólogo que codifica un transportador de bicarbonato, un polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico, un polinucleótido heterólogo que codifica una malato-deshidrogenasa y/o un polinucleótido heterólogo que codifica una piruvato-carboxilasa.

25 Anhidrasas carbónicas y polinucleótidos que codifican anhidrasas carbónicas

[0040] En algunos aspectos de las células huésped recombinantes y métodos descritos aquí, las células huésped tienen actividad de anhidrasa carbónica. La anhidrasa carbónica puede ser cualquier anhidrasa carbónica que sea adecuada para las células huésped y sus métodos de uso descritos aquí, tal como una anhidrasa carbónica de origen natural o una variante de la misma que retenga actividad de anhidrasa carbónica. En un aspecto, la anhidrasa carbónica está presente en el citosol de las células huésped. En algunos aspectos, la anhidrasa carbónica es una anhidrasa carbónica citosólica, tal y como se define aquí. En algunos aspectos, las células huésped comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una anhidrasa carbónica.

[0041] En algunos aspectos, las células huésped que comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una anhidrasa carbónica tienen un nivel aumentado de actividad de anhidrasa carbónica en comparación con las células huésped sin uno o varios polinucleótidos que codifican una anhidrasa carbónica, cuando se cultivan bajo las mismas condiciones. En algunos aspectos, las células huésped que comprenden uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica tienen un nivel aumentado de actividad de anhidrasa carbónica de al menos 5%, por ejemplo, al menos 10%, al menos 15%, al menos 20%, al menos 25%, al menos 50%, al menos 100%, al menos 150%, al menos 200%, al menos 300% o a 500% en comparación con las células huésped sin uno o varios polinucleótidos que codifican una anhidrasa carbónica, cuando se cultivan bajo las mismas condiciones.

[0042] La célula huésped comprende un polinucleótido heterólogo que codifica una anhidrasa carbónica. En un aspecto, el polinucleótido heterólogo que codifica una anhidrasa carbónica se selecciona de: (a) un polinucleótido que codifica una anhidrasa carbónica que tiene al menos 80% de identidad de secuencia con SEC ID n.º: 55; (b) un polinucleótido que hibridiza bajo condiciones de astringencia media con (i) SEC ID n.º: 54, (ii) la secuencia de ADNc de SEC ID n.º: 54, o (iii) la cadena complementaria en toda su longitud de (i) o (ii); y (c) un polinucleótido que tiene al menos 80% de identidad de secuencia con SEC ID n.º: 54, o la secuencia de ADNc de SEC ID n.º: 54. Como puede apreciar un experto en la técnica, en algunos casos el polinucleótido heterólogo que codifica una anhidrasa carbónica puede calificar bajo más de una de las selecciones (a), (b) y (c) mencionadas por encima.

[0043] En un aspecto, el polinucleótido heterólogo codifica una anhidrasa carbónica que tiene al menos 80%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 55. En un aspecto, la secuencia de anhidrasa carbónica difiere en no más de diez aminoácidos, por ejemplo, en no más de cinco aminoácidos, en no más de cuatro aminoácidos, en no más de tres aminoácidos, en no más de dos aminoácidos o en un aminoácido de SEC ID n.º: 55.

[0044] En un aspecto, la anhidrasa carbónica comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 55, una variante alélica de la misma o un fragmento de la anterior que tiene actividad de anhidrasa carbónica. En otro aspecto, la anhidrasa carbónica comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 55. En otro aspecto, la anhidrasa carbónica comprende o consiste en los aminoácidos 1 a 225 de SEC ID n.º: 55.

5 [0045] En un aspecto, el polinucleótido heterólogo codifica una anhidrasa carbónica que tiene una sustitución de aminoácido, eliminación y/o inserción de uno o varios (por ejemplo, dos, diferentes) aminoácidos de SEC ID n.º: 55. Una sustitución de aminoácido significa que un aminoácido que corresponde a una posición de la secuencia referenciada es diferente; una eliminación de aminoácido significa que un aminoácido que corresponde a una posición de la secuencia referenciada no está presente; y una inserción de aminoácido significa que está presente un aminoácido que no está presente en una posición correspondiente de la secuencia referenciada. En algunos de estos aspectos, el número total de sustituciones, deleciones y/o inserciones de aminoácidos de SEC ID n.º: 55 no es más de 10, por ejemplo, no más de 9, 8, 7, 6, 5, 4, 3, 2 o 1.

10 [0046] Los cambios en los aminoácidos son generalmente de naturaleza menor, esto es sustituciones de aminoácidos conservadoras o inserciones que no afectan significativamente al plegado y/o la actividad de la proteína; deleciones pequeñas, típicamente de uno a aproximadamente 30 aminoácidos; extensiones aminotermiales pequeñas o carboxilotermiales, tales como un residuo de metionina aminoterminal; un péptido enlazador pequeño de hasta aproximadamente 20-25 residuos; o una extensión pequeña que facilite la purificación cambiando la carga neta u otra función, tal como un tracto de polihistidina, un epítipo antigénico o un dominio de unión.

20 [0047] Ejemplos de sustituciones conservadoras están dentro del grupo de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrofóbicos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano y tirosina) y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina). Sustituciones de aminoácidos que generalmente no alteran la actividad específica se conocen en la técnica y son descritas, por ejemplo, por H. Neurath and R.L. Hill, 1979, In, *The Proteins*, Academic Press, Nueva York. Los intercambios que se producen más frecuentemente son Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu y Asp/Gly.

25 [0048] Alternativamente, los cambios de aminoácidos son de tal naturaleza que las propiedades fisicoquímicas de los polipéptidos son alteradas. Por ejemplo, los cambios en los aminoácidos pueden mejorar la termoestabilidad de la anhidrasa carbónica, alterar la especificidad de sustrato, cambiar el pH óptimo y similares.

30 [0049] Los aminoácidos esenciales en una anhidrasa carbónica se pueden identificar según procedimientos conocidos en la técnica, tales como mutagénesis dirigida al sitio o mutagénesis de escaneado de alanina (Cunningham and Wells, 1989, *Science* 244: 1081-1085). En la técnica anterior, mutaciones de alanina única se introducen en cada residuo de la molécula, y las moléculas mutantes resultantes se evalúan para actividad de anhidrasa carbónica para identificar residuos de aminoácidos que sean críticos para la actividad de la molécula. Véase también, Hilton *et al.*, 1996, *J. Biol. Chem.* 271: 4699-4708. El sitio activo de la anhidrasa carbónica u otra interacción biológica puede también ser determinado por análisis físico de estructura, como se determina por técnicas tales como resonancia magnética nuclear, cristalografía, difracción electrónica o marcaje por fotoafinidad, conjuntamente con mutación de aminoácidos de sitio de contacto putativo. Véase, por ejemplo, de Vos *et al.*, 1992, *Science* 255: 306-312; Smith *et al.*, 1992, *J. Mol. Biol.* 224: 899-904; Wlodaver *et al.*, 1992, *FEBS Lett.* 309: 59-64. Las identidades de los aminoácidos esenciales también pueden ser inferidas a partir de análisis de identidades con otra anhidrasa carbónica que esté relacionada con la anhidrasa carbónica referenciada.

45 [0050] Sustituciones, deleciones, y/o inserciones de aminoácidos únicos o múltiples se pueden hacer y evaluar usando métodos conocidos de mutagénesis, recombinación y/o redistribución, seguidos de un procedimiento de selección pertinente, tal como los descritos por Reidhaar-Olson and Sauer, 1988, *Science* 241: 53-57; Bowie and Sauer, 1989, *Proc. Natl. Acad. Sci. USA* 86: 2152-2156; WO 95/17413; o WO 95/22625. Otros métodos que se pueden usar incluyen PCR con tendencia al error, presentación en fagos (por ejemplo, Lowman *et al.*, 1991, *Biochemistry* 30: 10832-10837; patente de EE.UU. n.º 5.223.409; WO 92/06204), y mutagénesis dirigida a la región (Derbyshire *et al.*, 1986, *Gene* 46: 145; Ner *et al.*, 1988, *DNA* 7: 127).

50 [0051] Métodos de mutagénesis/redistribución se pueden combinar con métodos de selección automatizados de alto rendimiento para detectar actividad de polipéptidos mutagenizados clonados expresados por células huésped (Ness *et al.*, 1999, *Nature Biotechnology* 17: 893-896). Moléculas de ADN mutagenizadas que codifican anhidrasas carbónicas activas se pueden recuperar de las células huésped y rápidamente secuenciar usando métodos estándar en la técnica. Estos métodos permiten la determinación rápida de la importancia de residuos de aminoácidos individuales en un polipéptido.

60 [0052] En un aspecto, el polinucleótido heterólogo que codifica una anhidrasa carbónica hibridiza bajo al menos condiciones de astringencia medias, por ejemplo, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia altísimas con (i) SEC ID n.º: 54, (ii) la secuencia de ADNc de SEC ID n.º: 54, o (iii) la cadena complementaria en toda su longitud de (i) o (ii) (véase, por ejemplo, J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, *Molecular Cloning, A Laboratory Manual*, 2d edition, Cold Spring Harbor, New York). En un aspecto, el polinucleótido heterólogo que codifica una anhidrasa carbónica hibridiza bajo al menos condiciones de astringencia mitad, por ejemplo, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia altísimas con SEC ID n.º: 54, o la cadena complementaria en toda su longitud de la misma. En otro aspecto, el polinucleótido heterólogo que codifica una anhidrasa carbónica hibridiza bajo al menos condiciones de

astriencia medias, por ejemplo, condiciones de astriencia medio altas, condiciones de astriencia altas o condiciones de astriencia altísimas con la secuencia de ADNc de SEC ID n.º: 54, o la cadena complementaria en toda su longitud de la misma.

5 [0053] En un aspecto, el polinucleótido heterólogo que codifica una anhidrasa carbónica tiene al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 54 o la secuencia de ADNc de SEC ID n.º: 54. En un aspecto, el polinucleótido heterólogo que codifica una anhidrasa carbónica
10 tiene al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 54. En otro aspecto, el polinucleótido heterólogo que codifica una anhidrasa carbónica tiene al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con la secuencia de ADNc de SEC ID n.º: 54.

15 [0054] En un aspecto, el polinucleótido heterólogo que codifica una anhidrasa carbónica comprende SEC ID n.º: 54. En un aspecto, el polinucleótido heterólogo comprende los nucleótidos 1 a 750 de SEC ID n.º: 54. En un aspecto, el polinucleótido heterólogo comprende una subsecuencia de SEC ID n.º: 54, donde la subsecuencia codifica un polipéptido que tiene actividad de anhidrasa carbónica. En un aspecto, el número de residuos de nucleótidos en la subsecuencia es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de nucleótidos en SEC ID n.º: 54. En un aspecto, el polinucleótido heterólogo codifica un fragmento de SEC ID n.º: 55, donde el fragmento tiene actividad de anhidrasa carbónica. En un aspecto, el número de residuos de aminoácidos en el fragmento es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de aminoácidos en SEC ID n.º: 55.

25 [0055] La anhidrasa carbónica puede ser un polipéptido fusionado o polipéptido de fusión divisible donde otro polipéptido se fusiona en el N-término o el C-término de la anhidrasa carbónica. Un polipéptido fusionado se puede producir por la fusión de un polinucleótido que codifica otro polipéptido a un polinucleótido que codifica la anhidrasa carbónica. Técnicas para producir polipéptidos de fusión se conocen en la técnica, e incluyen ligamiento de las secuencias codificantes que codifican los polipéptidos de modo que éstas estén en marco y que la expresión del polipéptido fusionado esté bajo control del mismo promotor(es) y terminador. Las proteínas de fusión también se pueden construir utilizando tecnología de inteína donde las fusiones se crean postraduccionalmente (Cooper *et al.*, 1993, EMBO J. 12: 2575-2583; Dawson *et al.*, 1994, Science 266: 776-779).

30 [0056] Un polipéptido de fusión puede además comprender un sitio de escisión entre los dos polipéptidos. En la secreción de la proteína de fusión, el sitio es dividido liberando los dos polipéptidos. Ejemplos de sitios de escisión incluyen, pero de forma no limitativa, los sitios descritos en Martin *et al.*, 2003, J. Ind. Microbiol. Biotechnol. 3: 568-576; Svetina *et al.*, 2000, J. Biotechnol. 76: 245-251; Rasmussen-Wilson *et al.*, 1997, Appl. Environ. Microbiol. 63: 3488-3493; Ward *et al.*, 1995, Biotechnology 13: 498-503; y Contreras *et al.*, 1991, Biotechnology 9: 378-381; Eaton *et al.*, 1986, Biochemistry 25: 505-512; Collins-Racie *et al.*, 1995, Biotechnology 13: 982-987; Carter *et al.*, 1989, Proteins: Structure, Function, and Genetics 6: 240-248; y Stevens, 2003, Drug Discovery World 4: 35-48.

35 [0057] Técnicas usadas para aislar o clonar un polinucleótido tal como un polinucleótido que codifica una anhidrasa carbónica al igual que cualquier otro polipéptido usado en cualquiera de los aspectos mencionados aquí, se conocen en la técnica e incluyen aislamiento a partir de ADN genómico, preparación a partir de ADNc o una combinación de las mismas. La clonación de los polinucleótidos de tal ADN genómico se puede efectuar, por ejemplo, usando la bien conocida reacción en cadena de polimerasa (PCR) o selección de anticuerpo a partir de bibliotecas de expresión para detectar fragmentos de ADN clonados que comparten características estructurales. Véase, por ejemplo, Innis *et al.*, 1990, PCR: A Guide to Methods and Application, Academic Press, Nueva York. Otros procedimientos de amplificación de ácidos nucleicos tales como reacción en cadena de ligasa (LCR), transcripción activada ligada (LAT) y amplificación basada en secuencias de nucleótidos (NASBA) se pueden utilizar. Los polinucleótidos se pueden clonar a partir de una cepa de *Aspergillus*, u otro organismo u organismo relacionado, y así, por ejemplo, puede ser una variante alélica o de especies del polipéptido que codifica la región de la secuencia de nucleótidos.

45 [0058] El polinucleótido de SEC ID n.º: 54, o una subsecuencia de la misma, al igual que la secuencia de aminoácidos de SEC ID n.º: 55, o un fragmento de la misma, se pueden usar para diseñar sondas de ácidos nucleicos para identificar y clonar una anhidrasa carbónica a partir de cepas de diferentes géneros o especies según métodos bien conocidos en la técnica. En particular, tales sondas se pueden usar para hibridación con el ADN genómico o ADNc del género o especie de interés, siguiendo procedimientos de transferencia de Southern estándar, para identificar y aislar el gen correspondiente en el mismo. Tales sondas pueden ser considerablemente más cortas que la secuencia entera, por ejemplo, al menos 14 nucleótidos, al menos 25 nucleótidos, al menos 35 nucleótidos, al menos 70 nucleótidos de longitud. Las sondas pueden ser más largas, por ejemplo, al menos 100 nucleótidos, al menos 200 nucleótidos, al menos 300 nucleótidos, al menos 400 nucleótidos, al menos 500 nucleótidos de longitud. Incluso sondas más largas se pueden utilizar, por ejemplo, al menos 600 nucleótidos, al menos 700 nucleótidos, al menos 800 nucleótidos o al menos
60 900 nucleótidos de longitud. Sondas tanto de ADN como de ARN se pueden usar. Las sondas se marcan típicamente para la detección del gen correspondiente (por ejemplo, con ³²P, ³H, ³⁵S, biotina o avidina).

- [0059] Una genoteca de ADN genómico o ADNc obtenido a partir de tales otras cepas se puede seleccionar para ADN que hibridiza con las sondas anteriormente descritas y que codifica un polipéptido que tiene actividad de anhidrasa carbónica. ADN genómico u otro de tales otras cepas se puede separar por electroforesis en gel de poliacrilamida o agarosa, u otras técnicas de separación. ADN de las bibliotecas o el ADN separado se puede transferir e inmovilizar en la nitrocelulosa u otro material portador adecuado. Para identificar un clon o ADN que sea homólogo de la SEC ID n.º: 54, o una subsecuencia de la misma, el material portador se puede utilizar en una transferencia de Southern.
- [0060] Para fines de las sondas anteriormente descritas, hibridación indica que el polinucleótido hibridiza a una sonda de ácidos nucleicos marcada que corresponde a SEC ID n.º: 54, la cadena complementaria en toda su longitud de la misma, o una subsecuencia de la anterior; bajo condiciones de astringencia muy bajas a muy altas. Moléculas a las que la sonda de ácidos nucleicos hibridiza bajo estas condiciones se pueden detectar usando, por ejemplo, película radiográfica.
- [0061] En un aspecto, la sonda de ácidos nucleicos es SEC ID n.º: 54. En otro aspecto, la sonda de ácidos nucleicos es un polinucleótido que codifica el polipéptido de SEC ID n.º: 55, o un fragmento del mismo.
- [0062] Para sondas largas de al menos 100 nucleótidos de longitud, condiciones de astringencia muy baja a muy altas son definidas como prehibridación e hibridación a 42°C en 5X SSPE, 0,3% SDS, 200 microgramos/ml ADN de esperma de salmón cortado y desnaturalizado, y bien 25% de formamida para astringencias muy bajas y bajas, 35% de formamida para astringencias media y medio altas o 50% de formamida para astringencias alto y altísimas, siguiendo procedimientos de transferencia de Southern estándar durante 12 a 24 horas óptimamente. El material portador es finalmente lavado tres veces cada uno durante 15 minutos utilizando 2X SSC, 0,2% SDS a 45°C (astringencia muy baja), a 50°C (astringencia baja), a 55°C (astringencia media), a 60°C (astringencia medio alta), a 65°C (astringencia alta) y a 70°C (astringencia muy alta).
- [0063] Para sondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, las condiciones de astringencia se definen como prehibridación e hibridación en aproximadamente 5°C a aproximadamente 10°C por debajo de la T_m calculada utilizando el cálculo según Bolton y McCarthy (1962, Proc. Natl. Acad. Sci. USA 48:1390) en 0,9 M NaCl, 0,09 M Tris-HCl pH 7,6, 6 mM EDTA, 0,5% NP-40, 1X solución de Denhardt, 1 mM pirofosfato de sodio, 1 mM fosfato monobásico de sodio, 0,1 mM ATP y 0,2 mg de ARN de levadura por mL siguiendo procedimientos de transferencia de Southern estándar durante 12 a 24 horas óptimamente. El material portador es finalmente lavado una vez en 6X SSC más 0,1% SDS durante 15 minutos y dos veces cada uno durante 15 minutos utilizando 6X SSC a 5°C a 10°C por debajo de la T_m calculada.
- [0064] Polinucleótidos que codifican la anhidrasa carbónica descrita aquí se pueden obtener a partir de un microorganismo de cualquier género. Como se utiliza en este caso, el término "obtenido a partir de" en relación con una fuente dada significará que el polipéptido codificado por un polinucleótido es producido por la fuente o por una célula donde el polinucleótido de la fuente ha sido insertado.
- [0065] La anhidrasa carbónica puede ser una anhidrasa carbónica fúngica. En un aspecto, la anhidrasa carbónica fúngica es una anhidrasa carbónica de levadura tal como una anhidrasa carbónica de *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces* o *Yarrowia*.
- [0066] En otro aspecto, la anhidrasa carbónica fúngica es una anhidrasa carbónica fúngica filamentosa tal como una anhidrasa carbónica de *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Aureobasidium*, *Botryosphaeria*, *Ceriporiopsis*, *Chaetomidium*, *Chrysosporium*, *Claviceps*, *Cochliobolus*, *Coprinopsis*, *Coptotermes*, *Corynascus*, *Cryphonectria*, *Cryptococcus*, *Diplodia*, *Exidia*, *Filibasidium*, *Fusarium*, *Gibberella*, *Holomastigotoides*, *Humicola*, *Irpex*, *Lentinula*, *Leptosphaeria*, *Magnaporthe*, *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*, *Pseudoplectania*, *Pseudotriconympha*, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trichoderma*, *Trichophaea*, *Verticillium*, *Volvariella* o *Xylaria*.
- [0067] En otro aspecto, la anhidrasa carbónica es una anhidrasa carbónica de *Acremonium cellulolyticus*, *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus flavus*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Aspergillus sojae*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium tropicum*, *Chrysosporium merdarium*, *Chrysosporium inops*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium zonatum*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochromum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola grisea*, *Humicola insolens*, *Humicola lanuginosa*, *Irpex lacteus*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium funiculosum*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia spededonium*, *Thielavia setosa*,

Thielavia subthermophila, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* o *Trichoderma viride*.

5 [0068] En otro aspecto, la anhidrasa carbónica es una anhidrasa carbónica de *Aspergillus*, como la anhidrasa carbónica de *Aspergillus clavatus* de SEC ID n.º: 55.

10 [0069] Se entiende que para las especies anteriormente mencionadas, los estados perfectos e imperfectos, y otros equivalentes taxonómicos, por ejemplo, anamorfos, están abarcados independientemente del nombre de la especie por el que son conocidos. Expertos en la técnica fácilmente reconocerán la identidad de equivalentes apropiados.

15 [0070] Cepas de estas especies son fácilmente accesibles para el público en un número de colecciones de cultivos, tales como American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM), Centraalbureau Voor Schimmelcultures (CBS) y Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

20 [0071] En algunos aspectos, la anhidrasa carbónica tiene al menos 20%, por ejemplo, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de la actividad de anhidrasa carbónica del polipéptido maduro de SEC ID n.º: 55 bajo las mismas condiciones.

25 [0072] La anhidrasa carbónica también se puede identificar y obtener a partir de otras fuentes incluyendo microorganismos aislados de la naturaleza (por ejemplo, suelo, abonos, agua, ensilaje, etc.) o muestras de ADN obtenidas directamente de materiales naturales (por ejemplo, suelo, abonos, agua, ensilaje, etc.) utilizando las sondas mencionadas arriba. Técnicas para el aislamiento de microorganismos y ADN directamente de hábitats naturales son bien conocidos en la técnica. El polinucleótido que codifica una anhidrasa carbónica se puede derivar luego de forma similar por selección de una genoteca de ADNc o genómica de otros microorganismos o muestra de ADN mezclado. Una vez que un polinucleótido que codifica una anhidrasa carbónica ha sido detectado con sonda(s) adecuadas como se describe en este caso, la secuencia se puede aislar o clonar utilizando técnicas que son bien conocidas por los técnicos en la materia (véase, por ejemplo, J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, New York).

Transportadores de bicarbonato y polinucleótidos que codifican transportadores de bicarbonato

35 [0073] En algunos aspectos de las células huésped recombinantes y métodos descritos aquí, las células huésped tienen actividad de transportador de bicarbonato. El transportador de bicarbonato puede ser cualquier transportador de bicarbonato que sea adecuado para las células huésped y sus métodos de uso descritos aquí, tal como un transportador de bicarbonato de origen natural o una variante del mismo que retenga actividad de transportador de bicarbonato. En un aspecto, el transportador de bicarbonato está presente en el citosol de las células huésped. En algunos aspectos, las células huésped comprenden uno o más (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican un transportador de bicarbonato.

40 [0074] En algunos aspectos, las células huésped que comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican un transportador de bicarbonato tienen un nivel aumentado de actividad de transportador de bicarbonato en comparación con las células huésped sin uno o varios polinucleótidos que codifican un transportador de bicarbonato, cuando se cultivan bajo las mismas condiciones. En algunos aspectos, las células huésped que comprenden uno o varios polinucleótidos heterólogos que codifican un transportador de bicarbonato tienen un nivel aumentado de actividad de transportador de bicarbonato de al menos 5%, por ejemplo, al menos 10%, al menos 15%, al menos 20%, al menos 25%, al menos 50%, al menos 100%, al menos 150%, al menos 200%, al menos 300% o a 500% en comparación con las células huésped sin uno o varios polinucleótidos que codifican un transportador de bicarbonato, cuando se cultivan bajo las mismas condiciones.

45 [0075] En algunos aspectos, la célula huésped comprende un polinucleótido heterólogo que codifica un transportador de bicarbonato. En un aspecto, el polinucleótido heterólogo que codifica un transportador de bicarbonato se selecciona de:
55 (a) un polinucleótido que codifica un transportador de bicarbonato que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 2 o 4; (b) un polinucleótido que hibridiza bajo condiciones de astringencia bajas con (i) SEC ID n.º: 1 o 3, (ii) la secuencia de ADNc de SEC ID n.º: 1 o 3, o (iii) la cadena complementaria en toda su longitud de (i) o (ii); y (c) un polinucleótido que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 1 o 3, o la secuencia de ADNc de SEC ID n.º: 1 o 3. Como puede apreciarse uno de habilidad en la técnica, en algunos casos el polinucleótido heterólogo que codifica un transportador de bicarbonato pueden calificar bajo más de una de las selecciones (a), (b) y (c)
60 mencionadas por encima.

[0076] En un aspecto, el polinucleótido heterólogo codifica un transportador de bicarbonato que tiene al menos 60%, por ejemplo, al menos 65%, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%
65 o 100% de identidad de secuencia con SEC ID n.º: 2 o 4. En un aspecto, la secuencia del transportador de bicarbonato

difiere en no más de diez aminoácidos, por ejemplo, no más de cinco aminoácidos, no más de cuatro aminoácidos, no más de tres aminoácidos, no más de dos aminoácidos o un aminoácido de la SEC ID n.º: 2 o 4.

5 [0077] En un aspecto, el transportador de bicarbonato comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 2 o 4, una variante alélica de la misma o un fragmento de la anterior que tiene actividad de transportador de bicarbonato. En otro aspecto, el transportador de bicarbonato comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 2 o 4. En otro aspecto, el transportador de bicarbonato comprende o consiste en los aminoácidos 1 a 770 de SEC ID n.º: 2 o aminoácidos 1 a 843 de SEC ID n.º: 4.

10 [0078] En un aspecto, el polinucleótido heterólogo codifica un transportador de bicarbonato que tiene una sustitución, eliminación y/o inserción de aminoácido de uno o varios (por ejemplo, dos, diferentes) aminoácidos de SEC ID n.º: 2 o 4, como se describe *supra*. En algunos aspectos, el número total de sustituciones, delecciones y/o inserciones de aminoácidos de SEC ID n.º: 2 o 4, no es más de 10, por ejemplo, no más de 9, 8, 7, 6, 5, 4, 3, 2 o 1.

15 [0079] En un aspecto, el polinucleótido heterólogo que codifica un transportador de bicarbonato hibridiza bajo al menos condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia muy altas con (i) SEC ID n.º: 1 o 3, (ii) la secuencia de ADNc de SEC ID n.º: 1 o 3 o (iii) la cadena complementaria en toda su longitud de (i) o (ii) (véase, por ejemplo, J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, *supra*). En un aspecto, el polinucleótido heterólogo que
20 codifica un transportador de bicarbonato hibridiza bajo al menos condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia muy altas con SEC ID n.º: 1 o 3, o la cadena complementaria en toda su longitud de la misma. En un aspecto, el polinucleótido heterólogo que codifica un transportador de bicarbonato hibridiza bajo al menos
25 condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia muy altas con la secuencia de ADNc de SEC ID n.º: 1 o 3, o la cadena complementaria en toda su longitud de la misma.

[0080] En un aspecto, el polinucleótido heterólogo que codifica un transportador de bicarbonato tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con la SEC ID n.º: 1 o 3, o la secuencia de ADNc de SEC ID n.º: 1 o 3. En un aspecto, el polinucleótido heterólogo que codifica un transportador de bicarbonato tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 1 o 3. En un aspecto, el polinucleótido heterólogo que codifica un transportador de bicarbonato tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con la secuencia de ADNc de SEC ID n.º: 1 o 3.
40

[0081] En un aspecto, el polinucleótido heterólogo que codifica un transportador de bicarbonato comprende SEC ID n.º: 1 o 3. En un aspecto, el polinucleótido heterólogo comprende los nucleótidos 1 a 2503 de SEC ID n.º: 1, o los nucleótidos 1 a 2657 de SEC ID n.º: 3. En un aspecto, el polinucleótido heterólogo comprende una subsecuencia de SEC ID n.º: 1 o 3, donde la subsecuencia codifica un polipéptido que tiene actividad de transportador de bicarbonato. En un aspecto, el número de residuos de nucleótidos en la subsecuencia es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de nucleótidos en SEC ID n.º: 1 o 3.
45

[0082] En un aspecto, el polinucleótido heterólogo codifica un fragmento de SEC ID n.º: 2 o 4, donde el fragmento tiene actividad de transportador de bicarbonato. En un aspecto, el número de residuos de aminoácidos en el fragmento es al menos 75%, por ejemplo, al menos 80%, 85%, 90%, o 95% del número de residuos de aminoácidos en SEC ID n.º: 2 o 4. En un aspecto, el fragmento contiene un dominio de transportador de bicarbonato, por ejemplo, el dominio de transportador de bicarbonato putativo de los aminoácidos 280 a 556 de SEC ID n.º: 2 o 192 a 480 de SEC ID n.º: 4.
50

[0083] El transportador de bicarbonato también puede ser una variante alélica o variante artificial de un transportador de bicarbonato.
55

[0084] El transportador de bicarbonato puede también incluir polipéptidos fusionados o polipéptidos de fusión divisibles, como se describe *supra*.

[0085] Técnicas usadas para aislar o clonar un polinucleótido que codifica un transportador de bicarbonato son descritas *supra*.
60

[0086] La secuencia de polinucleótidos de SEC ID n.º: 1, 3 o una subsecuencia de la misma; al igual que la secuencia de aminoácidos de SEC ID n.º: 2, 4 o un fragmento de la misma; se puede usar para diseñar sondas de ácidos nucleicos para identificar y clonar ADN que codifica un transportador de bicarbonato a partir de cepas de diferentes géneros o especies, como se describe *supra*. Una genoteca de ADN genómico o ADNc obtenida a partir de tales otros organismos
65

se puede seleccionar para ADN que hibridiza con las sondas anteriormente descritas y codifica un transportador de bicarbonato, como se describe *supra*.

[0087] En un aspecto, la sonda de ácidos nucleicos es SEC ID n.º: 1 o 3. En otro aspecto, la sonda de ácidos nucleicos es una secuencia de polinucleótidos que codifica SEC ID n.º: 2 o 4 o un fragmento de la misma.

[0088] Para sondas largas de al menos 100 nucleótidos de longitud, astringencia muy baja a altísima y condiciones de lavado son definidas como se describe *supra*. Para sondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, astringencia y condiciones de lavado son definidas como se describe *supra*.

[0089] Polinucleótidos que codifican el transportador de bicarbonato se pueden obtener a partir de microorganismos de cualquier género. En un aspecto, el transportador de bicarbonato puede ser un transportador de bicarbonato bacteriano, de levadura o fúngico filamentosos obtenidos a partir de los microorganismos descritos aquí. En otro aspecto, el transportador de bicarbonato es un transportador de bicarbonato de *Aspergillus*, tal como el transportador de bicarbonato de *Aspergillus oryzae* de SEC ID n.º: 2 o 4.

[0090] Otros transportadores de bicarbonato que se pueden usar con las células huésped y métodos de uso descritos aquí incluyen, pero de forma no limitativa, un transportador de bicarbonato de *H. sapiens* SLC4A1 (SEC ID n.º: 53 de WO2010/111344), un transportador de bicarbonato de *O. cuniculus* SLC4A8 (SEC ID n.º: 54 de WO2010/111344), y un transportador de bicarbonato de *S. cerevisiae* YNL275w (SEC ID n.º: 55 de WO2010/111344). Cualquier aspecto descrito aquí relacionado con la identidad de secuencia, hibridación, modificaciones de aminoácidos (por ejemplo, sustituciones, deleciones y/o inserciones), fragmentos o subsecuencias de las mismas está abarcado para los transportadores de bicarbonato anteriores.

[0091] En algunos aspectos, el transportador de bicarbonato tiene al menos 20%, por ejemplo, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de la actividad de transportador de bicarbonato del polipéptido maduro de SEC ID n.º: 2 o 4 bajo las mismas condiciones.

[0092] El transportador de bicarbonato también se puede identificar y obtener de otras fuentes incluyendo microorganismos aislados de la naturaleza (por ejemplo, suelo, abonos, agua, ensilaje, etc.) o muestras de ADN obtenidas directamente de materiales naturales (por ejemplo, suelo, abonos, agua, ensilaje, etc.) como se describe *supra*.

Transportadores de ácido C4-dicarboxílico y polinucleótidos que codifican transportadores de ácido C4-dicarboxílico

[0093] En algunos aspectos de las células huésped recombinantes y métodos descritos aquí, las células huésped tienen actividad de transportador de ácido C4-dicarboxílico. El transportador de ácido C4-dicarboxílico puede ser cualquier transportador de ácido C4-dicarboxílico que sea adecuado para las células huésped y sus métodos de uso descritos aquí, tal como un transportador de ácido C4-dicarboxílico de origen natural o una variante del mismo que retenga actividad de transportador de ácido C4-dicarboxílico. En un aspecto, el transportador de ácido C4-dicarboxílico está presente en el citosol de las células huésped. En algunos aspectos, las células huésped comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican un transportador de ácido C4-dicarboxílico.

[0094] En algunos aspectos, las células huésped que comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican un transportador de ácido C4-dicarboxílico tienen un nivel aumentado de actividad de transportador de ácido C4-dicarboxílico en comparación con las células huésped sin uno o varios polinucleótidos que codifican un transportador de ácido C4-dicarboxílico, cuando se cultivan bajo las mismas condiciones. En algunos aspectos, las células huésped que comprenden uno o varios polinucleótidos heterólogos que codifican un transportador de ácido C4-dicarboxílico tienen un nivel aumentado de actividad de transportador de ácido C4-dicarboxílico de al menos 5%, por ejemplo, al menos 10%, al menos 15%, al menos 20%, al menos 25%, al menos 50%, al menos 100%, al menos 150%, al menos 200%, al menos 300% o a 500% en comparación con las células huésped sin uno o varios polinucleótidos que codifican un transportador de ácido C4-dicarboxílico, cuando se cultivan bajo las mismas condiciones.

[0095] En algunos aspectos, la célula huésped comprende un polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico. En un aspecto, el polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico se selecciona de: (a) un polinucleótido que codifica un transportador de ácido C4-dicarboxílico que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 6; (b) un polinucleótido que hibridiza bajo condiciones de astringencia bajas con SEC ID n.º: 5, o la cadena complementaria en toda su longitud de la misma; y (c) un polinucleótido que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 5. Como puede ser apreciado por uno de habilidad en la técnica, en algunos casos el polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico puede calificar bajo más de una de las selecciones (a), (b) y (c) mencionadas por encima.

- 5 [0096] En un aspecto, el polinucleótido heterólogo codifica un transportador de ácido C4-dicarboxílico que tiene al menos 60%, por ejemplo, al menos 65%, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 6. En un aspecto, la secuencia del transportador de ácido C4-dicarboxílico difiere en no más de diez aminoácidos, por ejemplo, no más de cinco aminoácidos, no más de cuatro aminoácidos, no más de tres aminoácidos, no más de dos aminoácidos o un aminoácido de SEC ID n.º: 6.
- 10 [0097] En un aspecto, el transportador de ácido C4-dicarboxílico comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 6, una variante alélica de la misma o un fragmento de la anterior, que tiene actividad de transportador de ácido C4-dicarboxílico. En otro aspecto, el transportador de ácido C4-dicarboxílico comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 6. En otro aspecto, el transportador de ácido C4-dicarboxílico comprende o consiste en los aminoácidos 1 a 418 de SEC ID n.º: 6. En otro aspecto, el transportador de ácido C4-dicarboxílico comprende o consiste en los aminoácidos 18 a 418 de SEC ID n.º: 6.
- 15 [0098] En un aspecto, el polinucleótido heterólogo codifica un transportador de ácido C4-dicarboxílico que tiene una sustitución, eliminación, y/o inserción de aminoácido de uno o varios (por ejemplo, dos, diferentes) aminoácidos de SEC ID n.º: 6, como se describe *supra*. En algunos aspectos, el número total de sustituciones, deleciones y/o inserciones de aminoácidos de SEC ID n.º: 6, no es más de 10, por ejemplo, no más de 9, 8, 7, 6, 5, 4, 3, 2 o 1.
- 20 [0099] En un aspecto, el polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico hibridiza bajo al menos condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia muy altas con SEC ID n.º: 5, o la cadena complementaria en toda su longitud de la anterior (véase, por ejemplo, J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, *supra*).
- 25 [0100] En un aspecto, el polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 5.
- 30 [0101] En un aspecto, el polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico comprende SEC ID n.º: 5. En un aspecto, el polinucleótido heterólogo comprende los nucleótidos 1 a 1257 de SEC ID n.º: 5. En un aspecto, el polinucleótido heterólogo comprende los nucleótidos 52 a 1257 de SEC ID n.º: 5. En un aspecto, el polinucleótido heterólogo comprende una subsecuencia de SEC ID n.º: 5, donde la subsecuencia codifica un polipéptido que tiene actividad de transportador de ácido C4-dicarboxílico. En un aspecto, el número de residuos de nucleótidos en la subsecuencia es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de nucleótido en SEC ID n.º: 5.
- 35 [0102] En un aspecto, el polinucleótido heterólogo codifica un fragmento de SEC ID n.º: 6, donde el fragmento tiene actividad de transportador de ácido C4-dicarboxílico. En un aspecto, el número de residuos de aminoácidos en el fragmento es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de aminoácidos en SEC ID n.º: 6.
- 40 [0103] El transportador de ácido C4-dicarboxílico también puede ser una variante alélica o variante artificial de un transportador de ácido C4-dicarboxílico.
- 45 [0104] El transportador de ácido C4-dicarboxílico puede también incluir polipéptidos fusionados o polipéptidos divisibles de fusión, como se describe *supra*.
- 50 [0105] Técnicas usadas para aislar o clonar un polinucleótido que codifica un transportador de ácido C4-dicarboxílico son descritas *supra*.
- 55 [0106] La secuencia de polinucleótidos de SEC ID n.º: 5, o una subsecuencia de la misma, al igual que la secuencia de aminoácidos de SEC ID n.º: 6, o un fragmento de la misma; se pueden usar para diseñar sondas de ácidos nucleicos para identificar y clonar ADN que codifica un transportador de ácido C4-dicarboxílico a partir de cepas de diferentes géneros o especies, como se describe *supra*. Una genoteca de ADN genómico o ADNc obtenido a partir de tales otros organismos se puede seleccionar para ADN que hibridiza con las sondas anteriormente descritas y codifica un transportador de ácido C4-dicarboxílico, como se describe *supra*.
- 60 [0107] En un aspecto, la sonda de ácidos nucleicos es SEC ID n.º: 5. En otro aspecto, la sonda de ácidos nucleicos es una secuencia de polinucleótidos que codifica SEC ID n.º: 6, o un fragmento de la misma.
- 65 [0108] Para sondas largas de al menos 100 nucleótidos de longitud, astringencia muy baja a altísima y condiciones de lavado son definidas como se describe *supra*. Para sondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, astringencia y condiciones de lavado son definidas como se describe *supra*.

[0109] Polinucleótidos que codifican el transportador de ácido C4-dicarboxílico se pueden obtener a partir de microorganismos de cualquier género. En un aspecto, el transportador de ácido C4-dicarboxílico puede ser un transportador de ácido C4-dicarboxílico bacteriano, de levadura o fúngico filamentoso obtenido a partir de los microorganismos descritos aquí. En otro aspecto, el transportador de ácido C4-dicarboxílico es un transportador de ácido C4-dicarboxílico de *Aspergillus*, tal como el transportador de ácido C4-dicarboxílico de *Aspergillus aculeatus* de SEC ID n.º: 6.

[0110] Otro transportador de ácido C4-dicarboxílico que se puede usar con las células huésped y métodos de uso descritos aquí incluye, pero de forma no limitativa, el transportador de ácido dicarboxílico de *Aspergillus flavus* C4 (AFLA 107340), el transportador de ácido C4-dicarboxílico de *Aspergillus oryzae* de SEC ID n.º: 27 (codificado por la secuencia de polinucleótidos de SEC ID n.º: 26; véase US 2011/0053233), el transportador de ácido C4-dicarboxílico de *Aspergillus terreus* de SEC ID n.º: 29 (codificado por la secuencia de polinucleótidos de SEC ID n.º: 28; véase US 2011/0053233), el transportador de ácido C4-dicarboxílico de *Schizosaccharomyces pombe* de SEC ID n.º: 32 (codificado por la secuencia de polinucleótidos de SEC ID n.º: 30 o 31; véase US 2011/0053233), el transportador de ácido C4-dicarboxílico de *Aspergillus aculeatus* de SEC ID n.º: 34 (codificado por la secuencia de polinucleótidos de SEC ID n.º: 33; véase solicitud de EE.UU. n.º 13/165.696, titulada "Polypeptides Having C4-dicarboxylic acid Transporter Activity and Polynucleotides Encoding Same" solicitada el 21 de junio de 2011), el transportador de ácido C4-dicarboxílico de *Aspergillus aculeatus* de SEC ID n.º: 36 (codificado por la secuencia de polinucleótidos de SEC ID n.º: 35; véase la solicitud EE.UU. n.º 13/165.696, *supra*), el transportador de ácido C4-dicarboxílico de *Schizosaccharomyces japonicus* de SEC ID n.º: 39 (codificado por la secuencia de polinucleótidos de SEC ID n.º: 37 o 38; véase PCT/US11/38881, titulada "C4-dicarboxylic acid Production in Filamentous Fungi" solicitada el 2 de 2011), el transportador de ácido C4-dicarboxílico de *Aspergillus clavatus* de SEC ID n.º: 41 (codificado por la secuencia de polinucleótidos de SEC ID n.º: 40; véase la solicitud de EE.UU. n.º 13/165.719, titulada "Methods for Improving C4-dicarboxylic acid Production in Filamentous Fungi" solicitada el 21 de junio de 2011), el transportador de ácido C4-dicarboxílico de *Aspergillus fumigatus* de SEC ID n.º: 43 (codificado por la secuencia de polinucleótidos de SEC ID n.º: 42; véase la solicitud de EE.UU. n.º 13/165.719, *supra*), o cualquier aspecto del transportador de ácido C4-dicarboxílico descrito en la referencia respectiva en esta. Cualquier aspecto descrito aquí relacionado con la identidad de secuencia, hibridación, modificaciones de aminoácidos (por ejemplo, sustituciones, deleciones y/o inserciones), fragmentos o subsecuencias de las mismas está abarcado para los transportadores de ácido C4-dicarboxílico por encima.

[0111] En algunos aspectos, el transportador de ácido C4-dicarboxílico tiene al menos 20%, por ejemplo, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de la actividad de transportador de ácido C4-dicarboxílico del polipéptido maduro de SEC ID n.º: 6 bajo las mismas condiciones.

[0112] El transportador de ácido C4-dicarboxílico también se puede identificar y obtener de otras fuentes incluyendo microorganismos aislados de la naturaleza (por ejemplo, suelo, abonos, agua, ensilaje, etc.) o muestras de ADN obtenidas directamente de materiales naturales (por ejemplo, suelo, abonos, agua, ensilaje, etc.) como se describe *supra*.

Malato deshidrogenasas y polinucleótidos que codifican malato deshidrogenasas

[0113] En algunos aspectos de las células huésped recombinantes y métodos descritos aquí, las células huésped tienen actividad de malato deshidrogenasa. La malato deshidrogenasa puede ser cualquier malato deshidrogenasa que sea adecuada para las células huésped y sus métodos de uso descrita aquí, tal como una malato deshidrogenasa de origen natural o un variante de la misma que retenga actividad de malato deshidrogenasa. En un aspecto, la malato deshidrogenasa está presente en el citosol de las células huésped. En algunos aspectos, las células huésped comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una malato deshidrogenasa.

[0114] En algunos aspectos, las células huésped que comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una malato deshidrogenasa tienen un nivel aumentado de actividad de malato deshidrogenasa en comparación con las células huésped sin uno o varios polinucleótidos que codifican una malato deshidrogenasa, cuando se cultivan bajo las mismas condiciones. En algunos aspectos, las células huésped que comprenden uno o varios polinucleótidos heterólogos que codifican una malato deshidrogenasa tienen un nivel aumentado de actividad de malato deshidrogenasa de al menos 5%, por ejemplo, al menos 10%, al menos 15%, al menos 20%, al menos 25%, al menos 50%, al menos 100%, al menos 150%, al menos 200%, al menos 300% o a 500% en comparación con las células huésped sin uno o varios polinucleótidos que codifican una malato deshidrogenasa, cuando se cultivan bajo las mismas condiciones.

[0115] En algunos aspectos, la célula huésped comprende un polinucleótido heterólogo que codifica una malato deshidrogenasa. En un aspecto, el polinucleótido heterólogo que codifica una malato deshidrogenasa se selecciona de: (a) un polinucleótido que codifica una malato deshidrogenasa que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 8; (b) un polinucleótido que hibridiza bajo condiciones de astringencia bajas con (i) SEC ID n.º: 7, (ii) la secuencia de ADNc de SEC ID n.º: 7 o (iii) la cadena complementaria en toda su longitud de (i) o (ii); y (c) un

polinucleótido que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 7 o la secuencia de ADNc de SEC ID n.º: 7. Como puede ser apreciado por uno de habilidad en la técnica, en algunos casos el polinucleótido heterólogo que codifica una malato deshidrogenasa pueden calificar bajo más de una de las selecciones (a), (b) y (c) mencionadas por encima.

5

[0116] En un aspecto, el polinucleótido heterólogo codifica una malato deshidrogenasa que tiene al menos 60%, por ejemplo, al menos 65%, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 8. En un aspecto, la secuencia de malato deshidrogenasa difiere en no más de diez aminoácidos, por ejemplo, no más de cinco aminoácidos, no más de cuatro aminoácidos, no más de tres aminoácidos, no más de dos aminoácidos o un aminoácido de SEC ID n.º: 8.

10

[0117] En un aspecto, la malato deshidrogenasa comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 8, una variante alélica de la misma, o un fragmento de la anterior, que tiene actividad de malato deshidrogenasa. En otro aspecto, la malato deshidrogenasa comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 8. En otro aspecto, la malato deshidrogenasa comprende o consiste en los aminoácidos 1 a 330 de SEC ID n.º: 8.

15

[0118] En un aspecto, el polinucleótido heterólogo codifica una malato deshidrogenasa que tiene una sustitución, eliminación, y/o inserción de aminoácidos de uno o varios (por ejemplo, dos, diferentes) aminoácidos de SEC ID n.º: 8, como se describe *supra*. En algunos aspectos, el número total de sustituciones, deleciones y/o inserciones de aminoácidos de SEC ID n.º: 8, no es más de 10, por ejemplo, no más de 9, 8, 7, 6, 5, 4, 3, 2 o 1.

20

[0119] En un aspecto, el polinucleótido heterólogo que codifica una malato deshidrogenasa hibridiza bajo al menos condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia muy altas con (i) SEC ID n.º: 7, (ii) la secuencia de ADNc de SEC ID n.º: 7 o (iii) la cadena complementaria en toda su longitud de (i) o (ii) (véase, por ejemplo, J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, *supra*). En un aspecto, el polinucleótido heterólogo que codifica una malato deshidrogenasa hibridiza bajo al menos condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia muy altas con SEC ID n.º: 7, o la cadena complementaria en toda su longitud de la misma. En otro aspecto, el polinucleótido heterólogo que codifica una malato deshidrogenasa hibridiza bajo al menos condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas, o condiciones de astringencia muy altas con la secuencia de ADNc de SEC ID n.º: 7, o la cadena complementaria en toda su longitud de la misma.

25

30

35

[0120] En un aspecto, el polinucleótido heterólogo que codifica una malato deshidrogenasa tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 7 o la secuencia de ADNc de SEC ID n.º: 7. En un aspecto, el polinucleótido heterólogo que codifica una malato deshidrogenasa tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 7. En un aspecto, el polinucleótido heterólogo que codifica una malato deshidrogenasa tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con la secuencia de ADNc de SEC ID n.º: 7.

40

45

[0121] En un aspecto, el polinucleótido heterólogo que codifica una malato deshidrogenasa comprende SEC ID n.º: 7. En un aspecto, el polinucleótido heterólogo comprende los nucleótidos 1 a 1430 de SEC ID n.º: 7. En un aspecto, el polinucleótido heterólogo comprende una subsecuencia de SEC ID n.º: 7, donde la subsecuencia codifica un polipéptido que tiene actividad de malato deshidrogenasa. En un aspecto, el número de residuos de nucleótidos en la subsecuencia es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de nucleótidos en SEC ID n.º: 7.

50

[0122] En un aspecto, el polinucleótido heterólogo codifica un fragmento de SEC ID n.º: 8, donde el fragmento tiene actividad de malato deshidrogenasa. En un aspecto, el número de residuos de aminoácidos en el fragmento es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de aminoácidos en SEC ID n.º: 8.

55

[0123] La malato deshidrogenasa también puede ser una variante alélica o variante artificial de una malato deshidrogenasa.

60

[0124] La malato deshidrogenasa puede también incluir polipéptidos fusionados o polipéptidos divisibles de fusión, como se describe *supra*.

[0125] Técnicas usadas para aislar o clonar un polinucleótido que codifica una malato deshidrogenasa son descritas *supra*.

65

[0126] La secuencia de polinucleótidos de SEC ID n.º: 7, o una subsecuencia de la misma; al igual que la secuencia de aminoácidos de SEC ID n.º: 8, o un fragmento de la misma; se puede usar para diseñar sondas de ácidos nucleicos para identificar y clonar ADN que codifica una malato deshidrogenasa a partir de cepas de diferentes géneros o especies, como se describe *supra*. Una genoteca de ADN genómico o ADNc obtenido a partir de tales otros organismos se puede seleccionar para ADN que hibridiza con las sondas anteriormente descritas y codifica una malato deshidrogenasa, como se describe *supra*.

[0127] En un aspecto, la sonda de ácidos nucleicos es SEC ID n.º: 7. En otro aspecto, la sonda de ácidos nucleicos es una secuencia de polinucleótidos que codifica SEC ID n.º: 8, o un fragmento de la misma.

[0128] Para sondas largas de al menos 100 nucleótidos de longitud, astringencia muy baja a altísima y condiciones de lavado son definidas como se describe *supra*. Para sondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, astringencia y condiciones de lavado son definidas como se describe *supra*.

[0129] Polinucleótidos que codifican la malato deshidrogenasa se pueden obtener a partir de microorganismos de cualquier género. En un aspecto, la malato deshidrogenasa puede ser una malato deshidrogenasa bacteriana, de levadura o fúngica filamentosa obtenida a partir de los microorganismos descritos aquí. En otro aspecto, la malato deshidrogenasa es una malato deshidrogenasa de *Aspergillus*, como la malato deshidrogenasa de *Aspergillus oryzae* de SEC ID n.º: 8.

[0130] Otras malato deshidrogenasas que se pueden usar con las células huésped y métodos de uso descritos aquí incluyen, pero de forma no limitativa, una malato deshidrogenasa de *Aspergillus nidulans* (AN6717.1; SIMS *et al.*, 2004, Mycol. Res. 108: 853-857) malato-deshidrogenasa de *Aspergillus niger* (An16g00120; Pel *et al.*, 2007, Nature Biotechnology 25: 221-231) malato deshidrogenasa de *Phytophthora infestans* (PITG 13614.1; Calcagno *et al.*, 2009, Mycological Research 113: 771-781); malato deshidrogenasa de *Saccharomyces cerevisiae* (YKL085W; McAlister-Henn and Thompson, 1987, J Bacteriol. 169: 5157-5166); malato deshidrogenasa de *Talaromyces emersonii* (AF439996, AF487682; Maloney *et al.*, 2004, Eur. J. Biochem. 271: 3115-3126); y malato deshidrogenasa de *Ustilago maydis* (um00403, um11161; McCann and Snetselaar, 2008, Fungal Genetics and Biology 45: S77-S87), la malato deshidrogenasa de *Aspergillus oryzae* de SEC ID n.º: 45 (codificado por la secuencia de polinucleótidos de SEC ID n.º: 44; véase solicitud de EE.UU. n.º 12/870,523, titulada "Methods for Improving Malic Acid Production in Filamentous Fungi" solicitada el 27 de agosto de 2010) o cualquier aspecto de la malato deshidrogenasa descrita en la referencia respectiva en esta. Cualquier aspecto descrito aquí relacionado con la identidad de secuencia, hibridación, modificaciones de aminoácidos (por ejemplo, sustituciones, deleciones y/o inserciones), fragmentos o subsecuencias de las mismas está abarcado para las malato deshidrogenasas por encima.

[0131] En algunos aspectos, la malato deshidrogenasa tiene al menos 20%, por ejemplo, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de la actividad de malato deshidrogenasa del polipéptido maduro de SEC ID n.º: 8 bajo las mismas condiciones.

[0132] La malato deshidrogenasa también se puede identificar y obtener a partir de otras fuentes, incluyendo microorganismos aislados de la naturaleza (por ejemplo, suelo, abonos, agua, ensilaje, etc.) o muestras de ADN obtenidas directamente de materiales naturales (por ejemplo, suelo, abonos, agua, ensilaje, etc.) como se describe *supra*.

Piruvato carboxilasas y polinucleótidos que codifican piruvato carboxilasas

[0133] En algunos aspectos de las células huésped recombinantes y métodos descritos aquí, las células huésped tienen actividad de piruvato carboxilasa. La piruvato carboxilasa puede ser cualquier piruvato carboxilasa que sea adecuada para las células huésped y sus métodos de uso descritos aquí, tal como una piruvato carboxilasa de origen natural o una variante de la misma que retenga actividad de piruvato carboxilasa. En un aspecto, la piruvato carboxilasa está presente en el citosol de las células huésped. En algunos aspectos, las células huésped comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una piruvato carboxilasa.

[0134] En algunos aspectos, las células huésped que comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una piruvato carboxilasa tienen un nivel aumentado de actividad de piruvato carboxilasa en comparación con las células huésped sin uno o varios polinucleótidos que codifican una piruvato carboxilasa, cuando se cultivan bajo las mismas condiciones. En algunos aspectos, las células huésped que comprenden uno o varios polinucleótidos heterólogos que codifican una piruvato carboxilasa tienen un nivel aumentado de actividad de piruvato carboxilasa de al menos 5%, por ejemplo, al menos 10%, al menos 15%, al menos 20%, al menos 25%, al menos 50%, al menos 100%, al menos 150%, al menos 200%, al menos 300% o a 500% en comparación con las células huésped sin uno o varios polinucleótidos que codifican una piruvato carboxilasa, cuando se cultivan bajo las mismas condiciones.

- 5 [0135] En algunos aspectos, la célula huésped comprende un polinucleótido heterólogo que codifica una piruvato carboxilasa. En un aspecto, el polinucleótido heterólogo que codifica una piruvato carboxilasa se selecciona de: (a) un polinucleótido que codifica una piruvato carboxilasa que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 10; (b) un polinucleótido que hibridiza bajo condiciones de astringencia bajas con (i) SEC ID n.º: 9, (ii) la secuencia de ADNc de SEC ID n.º: 9 o (iii) la cadena complementaria en toda su longitud de (i) o (ii); y (c) un polinucleótido que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 9 o la secuencia de ADNc de SEC ID n.º: 9. Como puede ser apreciado por uno de habilidad en la técnica, en algunos casos el polinucleótido heterólogo que codifica una piruvato carboxilasa puede calificar bajo más de unas de las selecciones (a), (b) y (c) mencionadas por encima.
- 10 [0136] En un aspecto, el polinucleótido heterólogo codifica una piruvato carboxilasa que tiene al menos 60%, por ejemplo, al menos 65%, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 10. En un aspecto, la secuencia de piruvato carboxilasa difiere en no más de diez aminoácidos, por ejemplo, no más de cinco aminoácidos, no más de cuatro aminoácidos, no más de tres aminoácidos, no más de dos aminoácidos o un aminoácido de SEC ID n.º: 10.
- 20 [0137] En un aspecto, la piruvato carboxilasa comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 10, una variante alélica de la misma, o un fragmento de la anteriormente mencionada, que tiene actividad de piruvato carboxilasa. En otro aspecto, la piruvato carboxilasa comprende o consiste en la secuencia de aminoácidos de SEC ID n.º: 10. En otro aspecto, la piruvato carboxilasa comprende o consiste en los aminoácidos 1 a 1193 de SEC ID n.º: 10.
- 25 [0138] En un aspecto, el polinucleótido heterólogo codifica una piruvato carboxilasa que tiene una sustitución, eliminación, y/o inserción de aminoácidos de uno o varios (por ejemplo, dos, diferentes) aminoácidos de SEC ID n.º: 10, como se describe *supra*. En algunos aspectos, el número total de sustituciones, deleciones y/o inserciones de aminoácidos de SEC ID n.º: 10, no es más de 10, por ejemplo, no mayor de 9, 8, 7, 6, 5, 4, 3, 2 o 1.
- 30 [0139] En un aspecto, el polinucleótido heterólogo que codifica una piruvato carboxilasa hibridiza bajo al menos condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia muy altas con (i) SEC ID n.º: 9, (ii) la secuencia de ADNc de SEC ID n.º: 9 o (iii) la cadena complementaria en toda su longitud de (i) o (ii) (véase, por ejemplo, J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, *supra*). En un aspecto, el polinucleótido heterólogo que codifica una piruvato carboxilasa hibridiza bajo al menos condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia muy altas con SEC ID n.º: 9, o la cadena complementaria en toda su longitud de la misma. En un aspecto, el polinucleótido heterólogo que codifica una piruvato carboxilasa hibridiza bajo al menos condiciones de astringencia bajas, por ejemplo, condiciones de astringencia medias, condiciones de astringencia medio altas, condiciones de astringencia altas o condiciones de astringencia muy altas con la secuencia de ADNc de SEC ID n.º: 9, o la cadena complementaria en toda su longitud de la misma.
- 40 [0140] En un aspecto, el polinucleótido heterólogo que codifica una piruvato carboxilasa tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 9 o la secuencia de ADNc de SEC ID n.º: 9. En un aspecto, el polinucleótido heterólogo que codifica una piruvato carboxilasa tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de identidad de secuencia con SEC ID n.º: 9. En un aspecto, el polinucleótido heterólogo que codifica una piruvato carboxilasa tiene al menos 65%, por ejemplo, al menos 70%, al menos 75%, al menos 80%, al menos 85%, al menos 85%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% DE identidad de secuencia CON la secuencia de ADNc de SEC ID n.º: 9.
- 45 [0141] En un aspecto, el polinucleótido heterólogo que codifica una piruvato carboxilasa comprende SEC ID n.º: 9. En un aspecto, el polinucleótido heterólogo comprende los nucleótidos 1 a 3643 de SEC ID n.º: 9. En un aspecto, el polinucleótido heterólogo comprende una subsecuencia de SEC ID n.º: 9, donde la subsecuencia codifica un polipéptido que tiene actividad de piruvato carboxilasa. En un aspecto, el número de residuos de nucleótidos en la subsecuencia es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de nucleótidos en SEC ID n.º: 9.
- 50 [0142] En un aspecto, el polinucleótido heterólogo codifica un fragmento de SEC ID n.º: 10, donde el fragmento tiene actividad de piruvato carboxilasa. En un aspecto, el número de residuos de aminoácidos en el fragmento es al menos 75%, por ejemplo, al menos 80%, 85%, 90% o 95% del número de residuos de aminoácidos en SEC ID n.º: 10.
- 60 [0143] La piruvato carboxilasa también puede ser una variante alélica o variante artificial de una piruvato carboxilasa.
- 65 [0144] La piruvato carboxilasa puede también incluir polipéptidos fusionados o polipéptidos divisibles de fusión, como se describe *supra*.

[0145] Técnicas usadas para aislar o clonar un polinucleótido que codifica una piruvato carboxilasa son descritas *supra*.

5 [0146] La secuencia de polinucleótidos de SEC ID n.º: 9, o una subsecuencia de la misma, al igual que la secuencia de aminoácidos de SEC ID n.º: 10, o un fragmento de la mismo, se puede usar para diseñar sondas de ácido nucleico para identificar y clonar ADN que codifica una piruvato carboxilasa a partir de cepas de diferentes géneros o especies, como se describe *supra*. Una genoteca de ADN genómico o ADNc obtenido a partir de tales otros organismos se puede seleccionar para ADN que hibridiza con las sondas anteriormente descritas y que codifica una piruvato carboxilasa, como se describe *supra*.

10 [0147] En un aspecto, la sonda de ácidos nucleicos es SEC ID n.º: 9. En otro aspecto, la sonda de ácidos nucleicos es una secuencia de polinucleótidos que codifica SEC ID n.º: 10, o un fragmento de la anteriormente mencionada.

15 [0148] Para sondas largas de al menos 100 nucleótidos de longitud, astringencia muy baja a altísima y condiciones de lavado son definidas como se describe *supra*. Para sondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, astringencia y condiciones de lavado son definidas como se describe *supra*.

20 [0149] Polinucleótidos que codifican la piruvato carboxilasa se pueden obtener a partir de microorganismos de cualquier género. En un aspecto, la piruvato carboxilasa puede ser una piruvato carboxilasa bacteriana, de levadura o fúngica filamentosa obtenida a partir de los microorganismos descritos aquí. En otro aspecto, la piruvato carboxilasa es una piruvato carboxilasa de *Aspergillus*, como la piruvato carboxilasa de *Aspergillus oryzae* de SEC ID n.º: 10.

25 [0150] Otras piruvato carboxilasas que se pueden usar con las células huésped y métodos de uso descritos aquí incluyen, pero de forma no limitativa, una piruvato carboxilasa de *Aspergillus clavatus* NRRL 1 (XP_001271664; presentación directa, presentada (26-OCT-2006), The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, EE.UU); piruvato carboxilasa de *Aspergillus fumigatus* Af293 (XP_752054; Nierman *et al.*, 2005, Nature 438: 1151-1156); piruvato carboxilasa de *Aspergillus nidulans* FGSC A4 (XP_662066; Galagan *et al.*, 2005, Nature 438: 1105-1115) piruvato carboxilasa de *Aspergillus niger* (An15g02820; Pel *et al.*, 2007, Nature Biotechnology 25: 221 - 231; ASPNG 5061; Panneman *et al.*, presentada (JUL-1998) en las bases de datos EMBL/GenBank/DDBJ);
30 piruvato carboxilasa de *Aspergillus terreus* (093918; presentación directa, presentada (OCT-1998) The Institute for Genomic Research, 9712 Medical Center Drive, Rockville, MD 20850, EE.UU); piruvato carboxilasa de *Magnaporthe grisea* 70-15 (XP_367852 presentación directa, presentada (26-SEP-2005) Broad Institute of MIT and Harvard, 320 Charles Street, Cambridge, MA 02142, EE.UU); piruvato carboxilasa de *Neurospora crassa* OR74A (XP_965636; Galagan *et al.*, 2003, Nature 422: 859-868); carboxilasa de *Rhizopus orizaepiruvate* (R03G_06931.1); piruvato
35 carboxilasa de *Saccharomyces cerevisiae* (NP_009777; Gaffeau *et al.*, 1996, Science 274: 546-547); piruvato carboxilasa de *Schizosaccharomyces pombe* (NP_595900; presentación directa, presentada (29-JUN-2007) European Schizosaccharomyces genome sequencing project, Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA); y piruvato carboxilasa de *Ustilago maydis* (um01054; McCann and Snetselaar, 2008, Fungal Genetics and Biology 45: S77-S87). La invención abarca cualquier aspecto de identidad de secuencia, hibridación, variantes y fragmentos descrito aquí como aplicado a las secuencias de polipéptidos de malato deshidrogenasa y secuencias de polinucleótidos anteriormente descritas.

45 [0151] En algunos aspectos, la piruvato carboxilasa tiene al menos 20%, por ejemplo, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100% de la actividad de piruvato carboxilasa del polipéptido maduro de SEC ID n.º: 10 bajo las mismas condiciones.

50 [0152] La piruvato carboxilasa también se puede identificar y obtener a partir de otras fuentes incluyendo microorganismos aislados de la naturaleza (por ejemplo, suelo, abonos, agua, ensilaje, etc.) o muestras de ADN obtenidas directamente a partir de materiales naturales (por ejemplo, suelo, abonos, agua, ensilaje, etc.) como se describe *supra*.

Vectores de expresión y constructos de ácidos nucleicos

55 [0153] Las células huésped recombinantes y métodos utilizan vectores de expresión que comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato deshidrogenasa y/o piruvato carboxilasa enlazada para una o varias secuencias de control que dirigen la expresión en una célula huésped adecuada bajo condiciones compatibles con la(s) secuencia(s) de control. Tales vectores de expresión se pueden utilizar en cualquiera de las células huésped y métodos descritos aquí. Los polinucleótidos descritos aquí se pueden manipular en una variedad de formas para proporcionar expresión de un polipéptido deseado. La manipulación del polinucleótido antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar polinucleótidos utilizando métodos de ADN recombinante se conocen bien en la técnica.

65 [0154] Los varios nucleótidos y secuencias de control se pueden juntar para producir un vector de expresión recombinante que puede incluir uno o varios (por ejemplo, dos, diferentes) sitios de restricción convenientes para

5 permitir la inserción o sustitución del polinucleótido que codifica el polipéptido en tales sitios. Alternativamente, el polinucleótido(s) se puede expresar por inserción del polinucleótido(s) o un constructo de ácidos nucleicos que comprende la secuencia en un vector apropiado para la expresión. Al crear el vector de expresión, la secuencia codificante se localiza en el vector de modo que la secuencia codificante esté operativamente enlazada a las secuencias de control apropiadas para la expresión.

10 [0155] El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido o virus) que puede ser convenientemente sometido a procedimientos de ADN recombinante y puede provocar la expresión del polinucleótido. La elección del vector típicamente dependerá de la compatibilidad del vector con la célula huésped en la que el vector debe ser introducido. El vector puede ser un plásmido lineal o cerrado circular.

15 [0156] En un aspecto, cada polinucleótido que codifica una anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato deshidrogenasa y/o piruvato carboxilasa descrito aquí está contenido en un vector independiente. En un aspecto, al menos dos de los polinucleótidos están contenidos en un vector único. En un aspecto, al menos cuatro de los polinucleótidos están contenidos en un vector único. En un aspecto, todos los polinucleótidos que codifican la anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato deshidrogenasa y piruvato carboxilasa están contenidos en un vector único.

20 [0157] El vector puede ser un vector de replicación autónoma, es decir, un vector que exista como un entidad extracromosómica, cuya replicación sea independiente de la replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un minicromosoma o un cromosoma artificial. El vector puede contener cualquier medio para asegurar la autorreplicación. Alternativamente, el vector puede ser uno que, cuando se introduce en la célula huésped, se integre en el genoma y se replique con el cromosoma(s) en el que ha sido integrado. Además, un vector único o plásmido o dos o más vectores o plásmidos que juntos contengan el ADN total que se va a introducir en el genoma de la célula huésped, o un transposón, puede ser utilizado.

30 [0158] El vector de expresión puede contener cualquier secuencia promotora adecuada que sea reconocida por una célula huésped para la expresión de un polinucleótido que codifica cualquier polipéptido descrito aquí (por ejemplo, una anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato deshidrogenasa o piruvato carboxilasa). La secuencia promotora contiene secuencias de control transcripcionales que median la expresión del polipéptido. El promotor puede ser cualquier polinucleótido que muestre actividad transcripcional en la célula huésped de elección incluyendo promotores mutantes, truncados e híbridos, y se puede obtener a partir de polipéptidos extracelulares o intracelulares de genes que codifican bien homólogos o heterólogos de la célula huésped.

35 [0159] Cada polinucleótido descrito aquí puede estar operativamente enlazado a un promotor que sea foráneo del polinucleótido. Por ejemplo, en un aspecto, el polinucleótido heterólogo que codifica una anhidrasa carbónica o subunidad de la misma está operativamente enlazado a un promotor que es foráneo del polinucleótido. En otro aspecto, el polinucleótido heterólogo que codifica un transportador de bicarbonato o subunidad del mismo está operativamente enlazado al promotor foráneo al polinucleótido. En otro aspecto, el polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico o subunidad del mismo está operativamente enlazado al promotor foráneo al polinucleótido. En otro aspecto, el polinucleótido heterólogo que codifica una malato deshidrogenasa o subunidad de la misma está operativamente enlazado al promotor foráneo al polinucleótido. En otro aspecto, el polinucleótido heterólogo que codifica una piruvato carboxilasa o subunidad de la misma está operativamente enlazado al promotor foráneo del polinucleótido.

40 [0160] Ejemplos de promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos en una célula huésped fúngica filamentosa son los promotores obtenidos a partir de los genes para acetamidasa de *Aspergillus nidulans*, alfa-amilasa neutra de *Aspergillus niger*, alfa-amilasa estable en ácido de *Aspergillus niger*, glucoamilasa de *Aspergillus niger* o *Aspergillus awamori* (glaA), TAKA amilasa de *Aspergillus oryzae*, proteasa alcalina de *Aspergillus oryzae*, triosa fosfato isomerasa de *Aspergillus oryzae*, proteasa de tipo tripsina de *Fusarium oxysporum* (WO 96/00787), amiloglucosidasa de *Fusarium venenatum* (WO 00/56900), *Fusarium venenatum Daria* (WO 00/56900), *Fusarium venenatum Quinn* (WO 00/56900), lipasa de *Rhizomucor miehei*, proteinasa aspártica de *Rhizomucor miehei*, beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa IV de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasas I de *Trichoderma reesei*, xilanasas II de *Trichoderma reesei*, beta-xilosidasa de *Trichoderma reesei*, al igual que el promotor NA2-tpi (un promotor modificado a partir de un gen que codifica una alfa-amilasa neutra en *Aspergilli* donde el líder no traducido ha sido sustituido por un líder no traducido de un gen que codifica triosa fosfato isomerasa en *Aspergilli*; ejemplos no limitativos incluyen promotores modificados del gen que codifica la alfa-amilasa neutra en *Aspergillus niger* donde el líder no traducido ha sido sustituido por un líder no traducido del gen que codifica triosa fosfato isomerasa en *Aspergillus nidulans* o *Aspergillus oryzae*); y promotores mutantes, truncados e híbridos de los mismos.

65 [0161] En un huésped de levadura, promotores útiles se obtienen a partir de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), galactocinasa de *Saccharomyces cerevisiae* (GAL1), alcohol deshidrogenasa/gliceraldehído-3' fosfato dehidrogenasa de *Saccharomyces cerevisiae* (ADH1, aDH2/GAP) triosa fosfato isomerasa de *Saccharomyces*

cerevisiae, (TPI) metalotioneína de *Saccharomyces cerevisiae* (CUP1) y 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para células huésped de levadura son descritos por Romanos *et al.*, 1992, *Yeast* 8: 423-488.

5 [0162] La secuencia de control también puede ser una secuencia terminadora de transcripción adecuada, que es reconocida por una célula huésped para terminar la transcripción. La secuencia terminadora está operativamente enlazada al terminal 3' del polinucleótido que codifica el polipéptido. Cualquier terminador que sea funcional en la célula huésped de elección puede ser utilizado.

10 [0163] Terminadores preferidos para células huésped fúngicas filamentosas son obtenidos a partir de los genes para antranilato sintasa de *Aspergillus nidulans*, glucoamilasa de *Aspergillus niger*, alfa-glicosidasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae* y proteasa de tipo tripsina de *Fusarium oxysporum*.

15 [0164] Terminadores preferidos para células huésped de levadura son obtenidos a partir de los genes para enolasa de *Saccharomyces cerevisiae*, citocromo C de *Saccharomyces cerevisiae* (CYC1) y gliceraldehído-3-fosfato-deshidrogenasa de *Saccharomyces cerevisiae* (GPD). Otros terminadores útiles para células huésped de levadura son descritos por Romanos *et al.*, 1992, *supra*.

20 [0165] La secuencia de control también puede ser una secuencia líder adecuada, cuando es transcrita es una región no traducida de un ARNm que es importante para la traducción por la célula huésped. La secuencia líder está operativamente enlazada al término 5' del polinucleótido que codifica el polipéptido. Cualquier secuencia líder que sea funcional en la célula huésped de elección puede ser utilizada.

25 [0166] Líderes preferidos para células huésped fúngicas filamentosas son obtenidas de los genes para TAKA amilasa de *Aspergillus oryzae* y triosa fosfato isomerasa de *Aspergillus nidulans*.

30 [0167] Líderes adecuados para células huésped de levadura son obtenidos a partir de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*, alfa-factor de *Saccharomyces cerevisiae* y alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH2/GAP).

35 [0168] La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia operativamente enlazada al término 3' del polinucleótido y, cuando transcrita, es reconocida por la célula huésped como una señal para añadir residuos de poliadenosina al ARNm transcrito. Cualquier secuencia de poliadenilación que sea funcional en la célula huésped de elección puede ser utilizada.

40 [0169] Secuencias de poliadenilación preferidas para células huésped fúngicas filamentosas son obtenidas de los genes para TAKA amilasa de *Aspergillus oryzae*, glucoamilasa de *Aspergillus niger*, antranilato sintasa de *Aspergillus nidulans*, proteasa de tipo tripsina de *Fusarium oxysporum* y alfa-glicosidasa de *Aspergillus niger*.

[0170] Secuencias de poliadenilación útiles para células huésped de levadura son descritas por Guo and Sherman, 1995, *Mol. Cellular Biol.* 15: 5983-5990.

45 [0171] La secuencia de control también puede ser una región codificante del péptido señal que codifica un péptido señal enlazado al N-término de un polipéptido y dirige el polipéptido en la vía secretora de la célula. El extremo 5' de la secuencia codificante del polinucleótido puede intrínsecamente contener una secuencia codificante del péptido señal naturalmente enlazada en el marco de lectura de la traducción con el segmento de la secuencia codificante que codifica el polipéptido. Alternativamente, el extremo 5' de la secuencia codificante puede contener una secuencia codificante del péptido señal que sea foránea a la secuencia codificante. La secuencia codificante del péptido señal foráneo se puede requerir donde la secuencia codificante no contenga naturalmente una secuencia codificante del péptido señal. Alternativamente, la secuencia codificante del péptido señal foráneo puede sencillamente reemplazar la secuencia codificante del péptido señal natural para mejorar la secreción del polipéptido. Sin embargo, cualquier secuencia codificante del péptido señal que dirija el polipéptido expresado en la vía secretora de una célula huésped de elección puede ser utilizada.

55 [0172] Secuencias codificantes del péptido señal eficaces para células huésped fúngicas filamentosas son las secuencias codificantes del péptido señal obtenidas de los genes para amilasa neutra de *Aspergillus niger*, glucoamilasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae*, celulasa de *Humicola insolens*, endoglucanasa V de *Humicola insolens*, lipasa de *Humicola lanuginosa* y proteinasa aspártica de *Rhizomucor miehei*.

60 [0173] Péptidos señal útiles para células huésped de levadura son obtenidos de los genes para alfa-factor de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*. Otras secuencias codificantes del péptido señal útiles son descritas por Romanos *et al.*, 1992, *supra*.

65 [0174] La secuencia de control también puede ser una secuencia codificante del propéptido que codifica un propéptido posicionado en el N-término de un polipéptido. El polipéptido resultante se conoce como una proenzima o propolipéptido

(o un zimógeno en algunos casos). Un propeptido es generalmente inactivo y se puede convertir en un polipeptido activo por escisión catalítica o autocatalítica del propeptido del propeptido. La secuencia codificante del propeptido se puede obtener a partir de los genes para proteasa alcalina de *Bacillus subtilis* (*aprE*), proteasa neutra de *Bacillus subtilis* (*nprT*), lacasa de *Myceliophthora thermophila* (WO 95/33836), proteinasa aspártica de *Rhizomucor miehei* y alfa-factor de *Saccharomyces cerevisiae*.

[0175] Donde ambos el péptido señal y las secuencias de propeptido están presentes en el N-término de un polipeptido, la secuencia de propeptido está situada junto al N-término de un polipeptido y la secuencia de péptido señal está situada junto al N-término de la secuencia de propeptido.

[0176] También puede ser deseable añadir secuencias reguladoras que permitan la regulación de la expresión del polipeptido con respecto al crecimiento de la célula huésped. Ejemplos de sistemas reguladores son aquellos que provocan que la expresión del gen cambie a ON u OFF en respuesta a un estímulo químico o físico, incluyendo la presencia de un compuesto regulador. Sistemas reguladores en sistemas procarióticos incluyen los sistemas operadores *lac*, *tac* y *trp*. En la levadura, el sistema ADH2 o sistema GAL1 puede ser utilizado. En hongos filamentosos, el promotor de glucoamilasa de *Aspergillus niger*, promotor de TAKA alfa-amilasa de *Aspergillus oryzae* y promotor de glucoamilasa de *Aspergillus oryzae* pueden ser utilizados. Otros ejemplos de secuencias reguladoras son los que permiten amplificación génica. En sistemas eucarióticos, estas secuencias reguladoras incluyen el gen de dihidrofolato reductasa que se amplifica en presencia de metotrexato, y los genes de metalotioneína que se amplifican con metales pesados. En estos casos, el polinucleótido que codifica el polipeptido estaría operativamente enlazado con la secuencia reguladora.

[0177] Los vectores puede contener uno o varios (por ejemplo, dos, diferentes) marcadores seleccionables que permitan la selección fácil de células transformadas, modificadas, transducidas o similar. Una etiqueta seleccionable es un gen cuyo producto proporciona resistencia biocida o vírica, resistencia a metales pesados, prototrofia a auxótrofos y similares.

[0178] Marcadores adecuados para células huésped de levadura son ADE2, HIS3, LEU2, LYS2, MET3, TRP1 y URA3. Marcadores seleccionables para su uso en una célula huésped fúngica filamentosos incluyen, pero de forma no limitativa, *amdS* (acetamidasa), *argB* (ornitina carbamoiltransferasa), *bar* (fosfonitricina acetiltransferasa), *hph* (higromicina fosfotransferasa), *niaD* (nitrato reductasa), *pyrG* (orotidina-5'-fosfato-descarboxilasa), *sC* (sulfato adeniltransferasa) y *trpC* (antranilato sintasa), al igual que equivalentes de los mismos. Se prefiere para su uso en una célula de *Aspergillus* los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces hygroscopicus*.

[0179] Los vectores puede contener uno o varios (por ejemplo, dos, diferentes) elementos que permitan integración del vector en el genoma de la célula huésped o replicación autónoma del vector en la célula independiente del genoma.

[0180] Para integración en el genoma de la célula huésped, el vector puede contar con la secuencia del polinucleótido que codifica el polipeptido o cualquier otro elemento del vector para integración en el genoma por recombinación homólogos o no homóloga. Alternativamente, el vector puede contener polinucleótidos adicionales para dirigir la integración por recombinación homóloga en el genoma de la célula huésped en una ubicación(es) precisa(s) en el cromosoma(s). Para aumentar la probabilidad de integración en una ubicación precisa, los elementos integracionales deberían contener un número suficiente de ácidos nucleicos, tal como de 100 a 10.000 pares de bases, 400 a 10.000 pares de bases y 800 a 10.000 pares de bases, que tienen un grado alto de identidad con la secuencia objetivo correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos integracionales pueden ser cualquier secuencia que sea homóloga de la secuencia objetivo en el genoma de la célula huésped. Además, los elementos integracionales pueden ser polinucleótidos no codificantes o codificantes. Por otro lado, el vector se puede integrar en el genoma de la célula huésped por recombinación no homóloga.

[0181] Para replicación autónoma, el vector puede comprender además un origen de replicación que permita al vector replicar autónomamente en la célula huésped en cuestión. El origen de replicación puede ser cualquier replicador plásmido que medie la replicación autónoma que funciona en una célula. El término "origen de replicación" o "replicador plásmido" se refiere a un polinucleótido que permite a un plásmido o vector replicar in vivo.

[0182] Ejemplos de orígenes de replicación para su uso en una célula huésped de levadura son el origen de replicación de 2 micras, ARS1, ARS4, la combinación de ARS1 y CEN3 y la una combinación de ARS4 y CEN6.

[0183] Ejemplos de orígenes de replicación útiles en una célula fúngica filamentosos son AMA1 y ANS1 (Gems *et al.*, 1991, Gene 98: 61-67; Cullen *et al.*, 1987, Nucleic Acids Res. 15: 9163-9175; WO 00/24883). Aislamiento del gen AMA1 y construcción de plásmidos o vectores que comprenden el gen se pueden realizar según los métodos descritos en la WO 00/24883.

[0184] Más de una copia de un polinucleótido descrito aquí se puede insertar en una célula huésped para aumentar la producción de un polipeptido. Un aumento en el número de copias del polinucleótido puede ser obtenido por la integración de al menos una copia adicional de la secuencia en el genoma de la célula huésped o incluyendo un gen marcador seleccionable amplificable con el polinucleótido donde las células contienen copias amplificadas del gen

marcador seleccionable, y así copias adicionales del polinucleótido, se pueden seleccionar mediante el cultivo de las células en presencia del agente seleccionable apropiado.

5 [0185] Los procedimientos usados para enlazar los elementos anteriormente descritos para construir los vectores de expresión recombinantes descritos aquí son conocidos por un experto en la materia (véase, por ejemplo, Sambrook *et al.*, 1989, *supra*).

Células huésped

10 [0186] Células huésped recombinantes pueden comprender uno o varios (por ejemplo, dos, diferentes) polinucleótido(s) descrito(s) aquí que pueden estar operativamente enlazados a una o varias secuencias de control que dirijan la expresión de uno o varios de los polipéptidos descritos para la producción recombinante de ácido C4-dicarboxílico. La célula huésped puede comprender cualquier polinucleótido que codifique una anhidrasa carbónica descrita aquí, y opcionalmente comprende cualquiera o una combinación de una pluralidad de polinucleótidos adicionales descritos aquí. Por ejemplo, en un aspecto, la célula huésped recombinante comprende un polinucleótido heterólogo que codifica una anhidrasa carbónica descrita aquí, y opcionalmente comprende uno o varios polinucleótidos heterólogos que codifican un transportador de bicarbonato, un transportador de ácido C4-dicarboxílico, una malato deshidrogenasa o una piruvato carboxilasa descrita aquí; donde la célula huésped produce (o es capaz de producir) una cantidad superior de ácido C4-dicarboxílico comparada con la célula huésped sin los polinucleótidos heterólogos cuando se cultiva bajo las mismas condiciones.

25 [0187] En un aspecto, la célula huésped recombinante comprende uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una anhidrasa carbónica descritos aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, y uno o varios polinucleótidos heterólogos que codifican un bicarbonato transportador descrito aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, y uno o varios polinucleótidos heterólogos que codifican un ácido C4-dicarboxílico transportador descrito aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, y uno o varios polinucleótidos heterólogos que codifican una malato deshidrogenasa descrita aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, y uno o varios polinucleótidos heterólogos que codifican una piruvato carboxilasa descrita aquí.

35 [0188] En un aspecto, la célula huésped recombinante comprende uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un bicarbonato transportador descrito aquí, y uno o varios polinucleótidos heterólogos que codifican un ácido C4-dicarboxílico transportador descrito aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un bicarbonato transportador descrito aquí, y uno o varios polinucleótidos heterólogos que codifican una malato deshidrogenasa descrita aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un ácido C4-dicarboxílico transportador descrito aquí, y uno o varios polinucleótidos heterólogos que codifican una malato deshidrogenasa descrita aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican una piruvato carboxilasa descrita aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un ácido C4-dicarboxílico transportador descrito aquí, y uno o varios polinucleótidos heterólogos que codifican una piruvato carboxilasa descrita aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un bicarbonato transportador descrito aquí, y uno o varios polinucleótidos heterólogos que codifican una malato deshidrogenasa descrita aquí, y uno o varios polinucleótidos heterólogos que codifican una piruvato carboxilasa descrita aquí.

55 [0189] En un aspecto, la célula huésped recombinante comprende uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un bicarbonato transportador descrito aquí, uno o varios polinucleótidos heterólogos que codifican un ácido C4-dicarboxílico transportador descrito aquí, y uno o varios polinucleótidos heterólogos que codifican una malato deshidrogenasa descrita aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un bicarbonato transportador descrito aquí, uno o varios polinucleótidos heterólogos que codifican un ácido C4-dicarboxílico transportador descrito aquí, y uno o varios polinucleótidos heterólogos que codifican una piruvato carboxilasa descrita aquí. En un aspecto, la célula huésped recombinante comprende uno o varios polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un bicarbonato transportador descrito aquí, uno o varios polinucleótidos heterólogos que codifican un ácido C4-dicarboxílico transportador descrito aquí, uno o varios polinucleótidos heterólogos que codifican una malato deshidrogenasa descrita aquí, y uno o varios polinucleótidos heterólogos que codifican una piruvato carboxilasa descrita aquí. En un aspecto, la célula huésped recombinante comprende uno o varios

polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un ácido C4-dicarboxílico transportador descrito aquí, uno o varios polinucleótidos heterólogos que codifican una malato deshidrogenasa descrita aquí, y uno o varios polinucleótidos heterólogos que codifican una piruvato carboxilasa descrita aquí.

5

[0190] En un aspecto, la célula huésped recombinante comprende uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, uno o varios polinucleótidos heterólogos que codifican un bicarbonato transportador descrito aquí, uno o varios polinucleótidos heterólogos que codifican un ácido C4-dicarboxílico transportador descrito aquí, uno o varios polinucleótidos heterólogos que codifican una malato deshidrogenasa descrita aquí, y uno o varios polinucleótidos heterólogos que codifican una piruvato carboxilasa descrita aquí.

10

[0191] En algunos de estos aspectos, la célula huésped recombinante carece de una anhidrasa carbónica endógena, carece de un transportador de bicarbonato endógeno, carece de un transportador de ácido C4-dicarboxílico endógeno, carece de una malato deshidrogenasa endógena y/o carece de una piruvato carboxilasa endógena.

15

[0192] En un aspecto, la célula huésped recombinante comprende:

20

(1) un polinucleótido heterólogo que codifica una anhidrasa carbónica, donde el polinucleótido se selecciona de: (a) un polinucleótido que codifica una anhidrasa carbónica que tiene al menos 80% de identidad de secuencia con SEC ID n.º: 55; (b) un polinucleótido que hibridiza bajo condiciones de astringencia media con (i) SEC ID n.º: 54, (ii) la secuencia de ADNc de SEC ID n.º: 54, o (iii) la cadena complementaria en toda su longitud de (i) o (ii); y (c) un polinucleótido que tiene al menos 80% de identidad de secuencia con SEC ID n.º: 54 o la secuencia de ADNc de SEC ID n.º: 54;

25

y opcionalmente comprende uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos seleccionados de:

30

(2) un polinucleótido heterólogo que codifica un transportador de bicarbonato, donde el polinucleótido se selecciona de: (a) un polinucleótido que codifica un transportador de bicarbonato que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 2 o 4; (b) un polinucleótido que hibridiza bajo condiciones de astringencia baja con (i) SEC ID n.º: 1 o 3, (ii) la secuencia de ADNc de SEC ID n.º: 1 o 3, o (iii) la cadena complementaria en toda su longitud de (i) o (ii); y (c) un polinucleótido que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 1 o 3, o la secuencia de ADNc de SEC ID n.º: 1 o 3;

35

(3) un polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico, donde el polinucleótido se selecciona de: (a) un polinucleótido que codifica un transportador de ácido C4-dicarboxílico que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 6; (b) un polinucleótido que hibridiza bajo condiciones de astringencia bajas con SEC ID n.º: 5, o la cadena complementaria en toda su longitud de la misma; y (c) un polinucleótido que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 5;

40

(4) un polinucleótido heterólogo que codifica una malato deshidrogenasa, donde el polinucleótido se selecciona de: (a) un polinucleótido que codifica una malato deshidrogenasa que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 8; (b) un polinucleótido que hibridiza bajo condiciones de astringencia bajas con (i) SEC ID n.º: 7, (ii) la secuencia de ADNc de SEC ID n.º: 7, o (iii) la cadena complementaria en toda su longitud de (i) o (ii); y (c) un polinucleótido que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 7 o la secuencia de ADNc de SEC ID n.º: 7; y

45

(5) un polinucleótido heterólogo que codifica una piruvato carboxilasa, donde el polinucleótido se selecciona de: (a) un polinucleótido que codifica una piruvato carboxilasa que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 10; (b) un polinucleótido que hibridiza bajo condiciones de astringencia baja con (i) SEC ID n.º: 9, (ii) la secuencia de ADNc de SEC ID n.º: 9, o (iii) la cadena complementaria en toda su longitud de la misma; y (c) un polinucleótido que tiene al menos 65% de identidad de secuencia con SEC ID n.º: 9 o la secuencia de ADNc de SEC ID n.º: 9;

50

donde la célula huésped produce (o es capaz de producir) una cantidad superior de ácido C4-dicarboxílico en comparación con la célula huésped sin el polinucleótido heterólogo que codifica una anhidrasa carbónica, cuando se cultivan bajo las mismas condiciones.

55

[0193] Como puede apreciar uno de habilidad en la técnica, en algunos casos los polinucleótidos heterólogos que codifican los polipéptidos mencionados arriba pueden calificar bajo más de una de las selecciones respectivas (a), (b) y (c).

60

[0194] Un constructo o vector (o constructos o vectores múltiples) que comprenden uno o varios (por ejemplo, dos, diferentes) polinucleótidos se introduce en una célula huésped de modo que el constructo o vector se mantenga como un integrante cromosómico o como un vector extracromosomal autorreplicante como se ha descrito anteriormente. El término "célula huésped" abarca cualquier descendiente de una célula madre que no se idéntico a la célula madre debido a mutaciones que se produzcan durante la replicación. En algunos casos, la elección de una célula huésped depende del gen que codifica del polipéptido y su fuente. Los aspectos descritos abajo recurren a las células huésped, de por sí, al igual que métodos que utilizan las células huésped.

65

[0195] La célula huésped es una célula fúngica. "Hongos", como se utilizan en este caso incluye los filos *Ascomycota*, *Basidiomycota*, *Chytridiomycota* y *Zygomycota* (como se define por Hawkswort *et al.*, en Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK) al igual que Oomycota (como se cita en Hawkswort *et al.*, 1995, *supra*, página 171) y todos los hongos mitospóricos (Hawkswort *et al.*, 1995, *supra*).

[0196] La célula huésped fúngica puede ser una célula de levadura. "Levadura", como se utiliza en este caso, incluye levadura ascoesporógena (Endomicetales), levadura basidioesporogénea y levadura de los *Fungi Imperfecti* (Blastomicetos). Ya que la clasificación de levadura puede cambiar en el futuro, para los fines descritos aquí, la levadura debe ser definida como se describe en *Biology and Activities of Yeast* (Skinner, F.A., Passmore, S.M., and Davenport, R.R., eds, Soc. App. Bacteriol. Symposium Series No. 9, 1980).

[0197] La célula huésped de levadura puede ser una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, *Yarrowia* o *Issatchenkia*, tal como un célula de *Candida sonorensis*, *Candida methanosorbosa*, *Candida ethanolica*, *Kluyveromyces lactis*, *Kluyveromyces marxianus*, *Pichia fermentans*, *Pichia galeiformis*, *Pichia membranifaciens*, *Pichia deserticola*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces bulderi*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, *Yarrowia lipolytica* o *Issatchenkia orientalis*.

[0198] La célula huésped fúngica puede ser una célula fúngica filamentosa. Los "hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión Eumycota y Oomycota (como se define por Hawkswort *et al.*, 1995, *supra*). Los hongos filamentosos se caracterizan generalmente por una pared micelial compuesta por quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos. El crecimiento vegetativo es por alargamiento hifal y el catabolismo de carbono es estrictamente aeróbico. En cambio, el crecimiento vegetativo por levaduras tales como *Saccharomyces cerevisiae* es por injerto de un talo unicelular y el catabolismo de carbono puede ser fermentativo.

[0199] La célula huésped fúngica filamentosa puede ser una célula de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolyposcladium*, *Trametes* o *Trichoderma*. Por ejemplo, la célula huésped fúngica filamentosa puede ser una célula de *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermisporea*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *Coprinus cinereus*, *Coriolus hirsutus*, *Fusarium bacridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochromum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* o *Trichoderma viride*.

[0200] En un aspecto, la célula huésped es una célula huésped de *Aspergillus*. En otro aspecto, la célula huésped es *Aspergillus oryzae*.

[0201] Las células fúngicas se pueden transformar por un proceso que implica la formación de protoplasto, transformación de los protoplastos y regeneración de la pared celular conocida de por sí. Procedimientos adecuados para la transformación de células huésped de *Aspergillus* y *Trichoderma* son descritas en la EP 238023 y Yelton *et al.*, 1984, Proc. Natl. Acad. Sci. USA 81: 1470-1474. Métodos adecuados para transformar las especies de *Fusarium* son descritos por Malardier *et al.*, 1989, Gene 78: 147-156 y la WO 96/00787. La levadura se puede transformar utilizando los procedimientos descritos por Becker y Guarente, en Abelson, J.N. and Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito *et al.*, 1983, J. Bacteriol. 153: 163; y Hinnen *et al.*, 1978, Proc. Natl. Acad. Sci. USA 75: 1920.

[0202] En algunos aspectos, la célula huésped comprende uno o varios (por ejemplo, dos, diferentes) polinucleótidos descritos aquí, donde la célula huésped secreta (y/o es capaz de secretar) un nivel aumentado de ácido C4-dicarboxílico en comparación con la célula huésped sin el o más polinucleótidos cuando se cultiva bajo las mismas condiciones. En algunos aspectos, la célula huésped secreta y/o es capaz de secretar un nivel aumentado de ácido C4-dicarboxílico de al menos 5%, por ejemplo, al menos 10%, al menos 15%, al menos 20%, al menos 25%, al menos 50%, al menos 100%, al menos 150%, al menos 200%, al menos 300% o 500% en comparación con la célula huésped sin dicho o más polinucleótidos, cuando se cultiva bajo las mismas condiciones. Ejemplos de condiciones de cultivo adecuadas son descritas más adelante y serán fácilmente aparentes para uno de habilidad en la técnica según las instrucciones de la presente.

[0203] En cualquiera de los aspectos de las células huésped recombinantes y métodos descritos aquí, el ácido C4-dicarboxílico puede ser ácido málico, ácido succínico, ácido oxaloacético, ácido malónico o ácido fumárico, o combinaciones de los mismos. En algunos aspectos, el ácido C4-dicarboxílico es ácido málico, ácido succínico, o ácido fumárico o combinaciones de los mismos. En algunos aspectos, el ácido C4-dicarboxílico es ácido málico o ácido fumárico, o una combinación de ácido málico y ácido fumárico. En algunos aspectos, el ácido C4-dicarboxílico es ácido málico.

[0204] En cualquiera de estos aspectos, la célula huésped produce (y/o es capaz de producir) ácido C4-dicarboxílico en un rendimiento de al menos 10%, por ejemplo, al menos 20%, al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80% o al menos 90%, del teórico.

[0205] En cualquiera de estos aspectos, el huésped recombinante tiene una productividad volumétrica de ácido C4-dicarboxílico mayor de aproximadamente 0,1 g/L por hora, por ejemplo, mayor de aproximadamente 0,2 g/L por hora, 0,5 g/L por hora, 0,6 g/L por hora, 0,7 g/L por hora, 0,8 g/L por hora, 0,9 g/L por hora, 1,0 g/L por hora, 1,1 g/L por hora, 1,2 g/L por hora, 1,3 g/L por hora, 1,5 g/L por hora, 1,75 g/L por hora, 2,0 g/L por hora, 2,25 g/L por hora, 2,5 g/L por hora, o 3,0 g/L por hora; o entre aproximadamente 0,1 g/L por hora y aproximadamente 2,0 g/L por hora, por ejemplo, entre aproximadamente 0,3 g/L por hora y aproximadamente 1,7 g/L por hora, aproximadamente 0,5 g/L por hora y aproximadamente 1,5 g/L por hora, aproximadamente 0,7 g/L por hora y aproximadamente 1,3 g/L por hora, aproximadamente 0,8 g/L por hora y aproximadamente 1,2 g/L por hora, o aproximadamente 0,9 g/L por hora y aproximadamente 1,1 g/L por hora.

[0206] Las células huésped recombinantes se pueden cultivar en un medio nutritivo adecuado para la producción de la anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato deshidrogenasa o piruvato carboxilasa usando métodos bien conocidos en la técnica. Por ejemplo, la célula se puede cultivar por cultivo en matraz de agitación, y fermentación a pequeña escala o a gran escala (incluyendo, fermentaciones de lote continuo, lote alimentado o estado sólido) en laboratorio o fermentadores industriales realizadas en un medio adecuado y bajo condiciones que permitan al polipéptido deseado ser expresado y/o aislado. El cultivo tiene lugar en un medio nutritivo adecuado que comprende fuentes de nitrógeno y carbono y sales inorgánicas, como se describe en este caso, usando procedimientos conocidos en la técnica. Medios adecuados están disponibles de proveedores comerciales, se pueden preparar según composiciones publicadas (por ejemplo, en catálogos de la American Type Culture Collection), o se pueden preparar a partir de ingredientes disponibles comercialmente.

[0207] La anhidrasa carbónica, transportador de bicarbonato, transportador de ácido C4-dicarboxílico, malato deshidrogenasa y piruvato carboxilasa, y actividades de las mismas, se pueden detectar utilizando métodos conocidos en la técnica. Estos métodos de detección pueden incluir el uso de anticuerpos específicos, formación de un producto enzimático o desaparición de un sustrato enzimático. Véase, por ejemplo, *Molecular Cloning: A Laboratory Manual*, Third Ed., Cold Spring Harbor Laboratory, New York (2001); Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, MD (1999); and Hanai *et al.*, *Appl. Environ. Microbiol.* 73:7814-7818 (2007)).

Métodos

[0208] Las células huésped recombinantes descritas aquí se pueden utilizar para la producción de ácido C4-dicarboxílico. En un aspecto es un método de producción de ácido C4-dicarboxílico para producir (por ejemplo, ácido málico), que comprende: (a) cultivo de cualquier célula huésped recombinante descrita aquí (por ejemplo, cualquier célula huésped con actividad de anhidrasa carbónica, y opcionalmente, actividad de transportador de bicarbonato, actividad de transportador de ácido C4-dicarboxílico, actividad de malato deshidrogenasa y/o actividad de piruvato carboxilasa) en un medio bajo condiciones adecuadas para producir el ácido C4-dicarboxílico; y (b) recuperación del ácido C4-dicarboxílico. En un aspecto, es un método de producción de ácido C4-dicarboxílico, que comprende: (a) cultivo en un medio de cualquiera de las células huésped recombinantes descritas aquí, donde la célula huésped comprende uno o varios (por ejemplo, dos, diferentes) polinucleótidos heterólogos que codifican una anhidrasa carbónica descrita aquí, y opcionalmente, uno o varios polinucleótidos heterólogos que codifican un bicarbonato transportador descrito aquí, un ácido C4-dicarboxílico transportador descrito aquí, una malato deshidrogenasa descrita aquí, y/o una piruvato carboxilasa descrita aquí, bajo condiciones adecuadas para producir el ácido C4-dicarboxílico; y (b) recuperación del ácido C4-dicarboxílico.

[0209] Métodos para la producción de ácido C4-dicarboxílico se pueden realizar en un medio fermentable que comprende cualquiera o varios (por ejemplo, dos, diferentes) azúcares, tal como glucosa, fructosa, sacarosa, celobiosa, xilosa, xilulosa, arabinosa, manosa, galactosa y/o oligosacáridos solubles. En algunos casos, el medio de fermentación se deriva de una fuente natural, tal como azúcar de caña, almidón o celulosa, y puede ser el resultado del pretratamiento de la fuente por hidrólisis enzimática (sacarificación).

[0210] Además de las fuentes de carbono apropiadas de uno o varios (por ejemplo, dos, diferentes) azúcar(es), el medio fermentable puede contener otros nutrientes o estimuladores conocidos por los expertos en la técnica, tales como macronutrientes (por ejemplo, fuentes de nitrógeno) y micronutrientes (por ejemplo, vitaminas, sales minerales y cofactores metálicos). En algunos aspectos, la fuente de carbono puede estar preferentemente suplementada con al

5 menos una fuente de nitrógeno, tal como extracto de levadura, N₂, peptona (por ejemplo, Bacto™ peptona), o soytone (por ejemplo, Bacto™ Soytone). Ejemplos no limitativos de vitaminas incluyen multivitaminas, biotina, pantotenato, ácido nicotínico, meso-inositol, tiamina, piridoxina, ácido para-aminobenzóico, ácido fólico, riboflavina y vitaminas A, B, C, D y E. Ejemplos de sales minerales y cofactores metálicos incluyen, pero de forma no limitativa Na, P, K, Mg, S, Ca, Fe, Zn, Mn, y Cu.

10 [0211] Condiciones adecuadas usadas para los métodos de producción de ácido C4-dicarboxílico pueden ser determinadas por un experto en la técnica siguiendo las instrucciones aquí. En algunos aspectos de los métodos, las células huésped se cultivan durante aproximadamente 12 horas a aproximadamente 216 horas, tal como
 15 aproximadamente 24 horas a aproximadamente 144 horas, o aproximadamente 36 horas a aproximadamente 96 horas. La temperatura está típicamente entre aproximadamente 26°C a aproximadamente 60°C, por ejemplo, aproximadamente 34°C a aproximadamente 50°C, y a un pH de aproximadamente 3,0 a aproximadamente 8,0, tal como aproximadamente 3,0 a aproximadamente 7,0, aproximadamente 3,0 a aproximadamente 6,0, aproximadamente 3,0 a aproximadamente 5,0, aproximadamente 3,5 a aproximadamente 4,5, aproximadamente 4,0 a aproximadamente 8,0, aproximadamente 4,0 a aproximadamente 7,0, aproximadamente 4,0 a aproximadamente 6,0, aproximadamente 4,0 a aproximadamente 5,0, aproximadamente 5,0 a aproximadamente 8,0, aproximadamente 5,0 a aproximadamente 7,0 o aproximadamente 5,0 a aproximadamente 6,0. En algunos aspectos de los métodos, el pH intracelular resultante de la célula huésped es aproximadamente 3,0 a aproximadamente 8,0, tal como aproximadamente 3,0 a aproximadamente 7,0, aproximadamente 3,0 a aproximadamente 6,0, aproximadamente 3,0 a aproximadamente 5,0, aproximadamente 3,5 a aproximadamente 4,5, aproximadamente 4,0 a aproximadamente 8,0, aproximadamente 4,0 a aproximadamente 7,0, aproximadamente 4,0 a aproximadamente 6,0, aproximadamente 4,0 a aproximadamente 5,0, aproximadamente 5,0 a aproximadamente 8,0, aproximadamente 5,0 a aproximadamente 7,0, o aproximadamente 5,0 a aproximadamente 6,0. El cultivo se puede realizar bajo condiciones anaeróbicas, microaeróbicas o aeróbicas, como sea apropiado. En algunos aspectos, el cultivo se realiza bajo condiciones anaeróbicas. Agentes tamponadores adecuados se conocen en la técnica.

30 [0212] El cultivo se puede realizar bajo condiciones anaeróbicas, sustancialmente anaeróbicas (microaeróbicas) o aeróbicas, como sea apropiado. En resumen, anaeróbico se refiere a un ambiente desprovisto de oxígeno, sustancialmente anaeróbico (microaeróbico) se refiere a un ambiente donde la concentración de oxígeno es inferior al aire, y aeróbico se refiere a un ambiente donde la concentración de oxígeno es aproximadamente igual o mayor que el aire. Condiciones sustancialmente anaeróbicas incluyen, por ejemplo, un cultivo, fermentación por lotes o fermentación continua de manera que la concentración de oxígeno disuelto en el medio permanezca menor del 10% de saturación. Condiciones sustancialmente anaeróbicas también incluyen crecimiento o permanencia de células en el medio líquido o en el agar sólido dentro de una cámara hermética mantenida con una atmósfera inferior al 1 % de oxígeno. El porcentaje de oxígeno se puede mantener mediante, por ejemplo, burbujeo del cultivo con una mezcla de N₂/CO₂ u otro gas o gases sin oxígeno adecuados. En algunas formas de realización, el cultivo se realiza bajo condiciones anaeróbicas o condiciones sustancialmente anaeróbicas.

40 [0213] Los métodos descritos aquí pueden emplear cualquier modo de operación de fermentación adecuado. Por ejemplo, una fermentación de modo de lote se puede utilizar con un sistema cerrado donde medios de cultivo y microorganismo huésped, establecidos al principio de la fermentación, no tienen entrada adicional salvo por los reactivos ciertos reactivos, por ejemplo, para control de pH, control de espuma u otros requeridos para mantenimiento del proceso. El proceso descrito aquí puede también ser empleado en el modo lote alimentado o continuo.

45 [0214] Los métodos descritos aquí se pueden practicar en diferentes configuraciones de biorreactor, tales como tanque agitado, columna de burbuja, reactor de puente aéreo y otros conocidos por los expertos en la técnica.

50 [0215] Los métodos se pueden realizar en cultivo libre de células o en cultivo de células inmovilizadas como apropiado. Cualquier soporte material para cultivo de células inmovilizadas puede ser utilizado, tales como alginatos, lecho fibroso o materiales argile tales como crisotilo, montmorilonita KSF y montmorilonita K-10.

55 [0216] En un aspecto de los métodos, el ácido C4-dicarboxílico (por ejemplo, ácido málico) se produce en una graduación mayor de aproximadamente 10 g/L, por ejemplo, mayor de aproximadamente 25 g/L, 50 g/L, 75 g/L, 100 g/L, 125 g/L, 150 g/L, 160 g/L, 170 g/L, 180 g/L, 190 g/L, 200 g/L, 210 g/L, 225 g/L, 250 g/L, 275 g/L, 300 g/L, 325 g/L, 350 g/L, 400 g/L, o 500g/L; o entre aproximadamente 10 g/L y aproximadamente 500 g/L, por ejemplo, entre aproximadamente 50 g/L y aproximadamente 350 g/L, aproximadamente 100 g/L y aproximadamente 300 g/L, aproximadamente 150 g/L y aproximadamente 250 g/L, aproximadamente 175 g/L y aproximadamente 225 g/L, o aproximadamente 190 g/L y aproximadamente 210 g/L. En un aspecto de los métodos, el ácido C4-dicarboxílico se produce a una graduación mayor de aproximadamente 0,01 gramo por gramo de carbohidrato, por ejemplo, mayor de aproximadamente 0,02, 0,05, 0,75, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, o 1,0 gramo por gramo de carbohidrato.

60 [0217] En un aspecto de los métodos, la cantidad de ácido C4-dicarboxílico producido (por ejemplo, ácido málico) es al menos 5%, por ejemplo, al menos 10%, al menos 15%, al menos 20%, al menos 25%, al menos 30%, al menos 50% o al menos 100% superior en comparación con el cultivo de la célula huésped sin uno o varios (por ejemplo, dos, diferentes) polinucleótidos que codifican la anhidrasa carbónica bajo las mismas condiciones.

[0218] El ácido C4-dicarboxílico recombinante (por ejemplo, ácido málico) puede ser opcionalmente recuperado y purificado del medio de fermentación utilizando cualquier procedimiento conocido en la técnica (véase, por ejemplo, WO 1998/022611 y US 7.601.865) incluyendo, pero no limitado a, cromatografía (por ejemplo, cromatografía de exclusión por tamaño, cromatografía de adsorción, cromatografía de intercambio de iones), procedimientos electroforéticos, solubilidad diferencial, destilación, extracción (por ejemplo, extracción líquido-líquido), pervaporación, filtración extractiva, filtración de membrana, separación de membrana, ósmosis inversa, ultrafiltración o cristalización. En un ejemplo, el ácido C4-dicarboxílico es recuperado de otro material en el medio de fermentación por filtración.

[0219] En algunos aspectos de los métodos, el ácido C4-dicarboxílico recombinante antes y/o después de ser opcionalmente purificado es sustancialmente puro. Con respecto a los métodos de producción del ácido C4-dicarboxílico (o un ácido C4-dicarboxílico específico del mismo, tal como ácido málico), "sustancialmente puro" se refiere a una preparación recuperada de ácido C4-dicarboxílico que contiene no más de 15% de impurezas, donde la impureza se refiere a compuestos diferentes de ácidos C4-dicarboxílicos. En una variación, una preparación de ácido C4-dicarboxílico sustancialmente puro se proporciona donde la preparación contiene no más de 25% de impureza, o no más de 20% de impureza, o no más de 10% de impureza, o no más de 5% de impureza, o no más de 3% de impureza, o no más de 1% de impureza, o no más de 0,5% de impureza.

[0220] Ensayos adecuados para probar la producción de ácido C4-dicarboxílico para los métodos de producción y células huésped descritos aquí se pueden realizar utilizando métodos conocidos en la técnica. Por ejemplo, el ácido C4-dicarboxílico final (y otros compuestos orgánicos) se pueden analizar por métodos tales como HPLC (cromatografía líquida de alto rendimiento, por sus siglas en inglés), GC-MS (espectroscopia de gases - cromatografía de masas, por sus siglas en inglés) y LC-MS (cromatografía líquida - espectroscopia de masas, por sus siglas en inglés) u otros métodos analíticos adecuados usando procedimientos rutinarios bien conocidos en la técnica. La liberación de ácido C4-dicarboxílico en el caldo de fermentación también puede ser evaluada con el sobrenadante del cultivo. Subproductos y azúcar residual en el medio de fermentación (por ejemplo, glucosa) se pueden cuantificar por HPLC usando, por ejemplo, un detector de índice de refracción para glucosa y alcoholes, y un detector de UV para ácidos orgánicos Lin *et al.*, *Biotechnol. Bioeng.* 90:775 -779 (2005)) o usando otro ensayo adecuado y métodos de detección bien conocidos en la técnica.

[0221] Los ejemplos siguientes se proporcionan a modo de ilustración y no están destinados a ser limitación de la invención.

Ejemplos

[0222] Productos químicos usados como tampones y sustratos fueron productos comerciales de al menos calidad reactiva.

Cepas fúngicas

[0223] *Aspergillus clavatus* fue usado como fuente de un gen de anhidrasa carbónica (ACLA 007930). *Aspergillus oryzae* NRRL 3488 (o ATCC 56747) fue usado como fuente de un gen de transportador de bicarbonato (*bt1*), un gen de piruvato carboxilasa (*pyc*), un gen de malato deshidrogenasa (*mdh3*) y para la producción de los ácidos C4-dicarboxílicos. *Aspergillus aculeatus* fue usado como fuente de un gen de proteína de transporte de ácido C4-dicarboxílico (*c4t521*).

Medios y soluciones

[0224]

Medio YEG estaba compuesto por 20 g de glucosa, 5 g de extracto de levadura y agua desionizada hasta 1 litro. Placas de COVE estaban compuestas por 1 M de sacarosa, 2% de solución salina de COVE, 10 mM de acetamida, 15 mM de CsCl y 25 g/l de Agar Noble.

Solución salina de COVE estaba compuesta por 26 g de KCl, 26 g de MgSO₄·7H₂O, 76 g de KH₂PO₄, 50 ml de solución de oligoelementos de COVE y agua desionizada hasta 1 litro.

Solución de oligoelementos de COVE estaba compuesta por 0,04 g de Na₂B₄O₇·10H₂O, 0,04 g de CuSO₄·5H₂O, 1,2 g de FeSO₄·7H₂O, 0,7 g de MnSO₄·H₂O, 0,8 g de Na₂MoO₂·2H₂O, 10 g de ZnSO₄·7H₂O y agua desionizada hasta 1 litro.

Medio de semilla estaba compuesto por 40 g de glucosa, 6 g Bacto-peptona, 750 mg de KH₂PO₄, 750 mg de K₂HPO₄, 100 mg de MgSO₄·7H₂O, 100 mg de CaCl₂·H₂O, 5 mg de FeSO₄·7H₂O, 5 mg de NaCl y agua desionizada hasta 1 litro.

Medio de semilla B estaba compuesto por 30 g de glucosa, 3 g de Bacto-peptona, 560 mg de KH₂PO₄, 560 mg de K₂HPO₄, 925 mg de NaH₂PO₄·H₂O, 820 mg de Na₂HPO₄, 75 mg de MgSO₄·7H₂O, 75 mg de CaCl₂·H₂O, 0,75 ml de 1000X solución Micronutriente y agua desionizada hasta 1 litro.

Medio de producción ácido C estaba compuesto por 100 g de glucosa, 80 g de CaCO₃, 6 g de bacto-peptona, 150 mg de KH₂PO₄, 150 mg de K₂HPO₄, 100 mg de MgSO₄·7H₂O, 100 mg de CaCl₂·H₂O, 1 ml 1000X solución Micronutriente y agua desionizada hasta 1 litro.

Medio de lote fermentador estaba compuesto por 140 g de glucosa, 120 g de CaCO₃, 9 g de Bacto-peptona, 150 mg de KH₂PO₄, 150 mg de K₂HPO₄, 100 mg de MgSO₄·7H₂O, 100 mg de CaCl₂·2H₂O, 5 mg de FeSO₄·7H₂O, 5 mg de NaCl, 5 ml de L61 plurónico y agua desionizada hasta 1 litro.

1000X Solución Micronutriente estaba compuesta por 5 g de NaCl, 5 g de FeSO₄·7H₂O, 1 g de ácido cítrico y agua desionizada hasta 1 litro.

Placas de PDA estaban compuestas por 39 g/l de agar de dextrosa de patata.

Placas 2XiT+amp estaban compuestas por 16 g de triptona, 10 g de extracto de levadura, 5 g de NaCl, 100 mg de ampicilina, 15 g de Bacto agar y agua desionizada hasta 1 litro.

Placas MM estaban compuestas por 50 ml 20X MM solución salina, 1 mL de solución de oligoelementos de COVE, 10 g de glucosa, 20 ml de solución madre de biotina, 500mg de MgSO₄·7H₂O, 20g de agar noble, pH 6,5 con NaOH y agua desionizada hasta 1 litro.

Placas de sacarosa 2 MM + 1M estaban compuestas por 50ml 20X MM solución salina, 1 ml de solución de oligoelementos de COVE, 342,3g de sacarosa, 10g de glucosa, 20ml de solución madre de biotina, 500mg de MgSO₄·7H₂O, 20g de agar noble, pH 6,5 con NaOH y agua desionizada hasta 1 litro.

Solución salina 20X MM estaba compuesta por 120g de NaNO₃, 10,4g de KCl, 30,4g de KH₂PO₄ y agua desionizada hasta 1 litro.

Solución madre de biotina estaba compuesta por 5 mM de biotina en 100 mM de tampón Tris (pH 8,0).

Ejemplo 1: clonación de un gen de transportador de bicarbonato de *Aspergillus oryzae* (*bt1*) y construcción de vector de expresión pAmFs69

[0225] El gen de transportador de bicarbonato *bt1* (AO090012000782) fue clonado a partir de ADN genómico de *Aspergillus oryzae* NRRL3488 por amplificación de PCR usando cebadores homólogos del gen de transportador de bicarbonato predicho de *Aspergillus oryzae* número de modelo AO090012000782 encontrado en la secuencia genómica publicada de *A. oryzae* ATCC 42149 (Galagan *et al.*, 2005, Nature 438: 1105-1115).

[0226] ADN genómico de *A. oryzae* NRRL3488 fue aislado por inoculación de 100 ml de medio YEG en un matraz de agitación con 2X10⁶ de esporas e incubación del matraz a 37°C durante toda la noche con agitación a 200 r.p.m. Los micelios fueron cosechados en embudo forrado con MIRACLOTH® (Calbiochem, San Diego, CA, EE.UU.) y aproximadamente 2 gramos de tejido fue congelado en nitrógeno líquido. Los micelios fueron interrumpidos por trituración en un mortero frío y mano de mortero. El ADN genómico fue aislado de los micelios en polvo utilizando un kit DNeasy® Plant Maxi (QIAGEN Inc., Valencia, CA, EE.UU.) según las instrucciones del fabricante. El gen *bt1* de *Aspergillus oryzae* fue amplificado utilizando el cebador sentido 069824 y cebador antisentido 069825 mostrado por debajo:

Cebador 069824:

5'-GTGATAGAACATCGTCCATAATGGAATCCAGCGCTGTACA-3' (SEC ID n.º: 11)

Cebador 069825:

5'-GTGTCAGTCACCTCTAGTTATCAGATTTCAATCTCGTCTT-3' (SEC ID n.º: 12)

[0227] Las reacciones de amplificación fueron realizadas utilizando ADN polimerasa de alta fidelidad Phusion® Hot Start (Finnzymes Inc., Massachusetts, EE.UU.) según las instrucciones del fabricante. Cada reacción por PCR contenía 47 ng de ADN genómico de *Aspergillus oryzae* NRRL3488, 200 µM dNTPs, 50 pM de cebador sentido, 50 pM de cebador antisentido, 1X tampón de reacción Phusion® GC Buffer (Finnzymes Inc.), y 50 unidades de ADN polimerasa de alta fidelidad Phusion® Hot Start en un volumen final de 50 µl. Las reacciones de amplificación fueron incubadas en un MASTERCYCLER® de EPPENDORF® (Eppendorf Scientific Inc., Westbury, NY, EE.UU.) programado para 1 ciclo a 98°C durante 30 segundos; 35 ciclos a 98°C durante 10 segundos, 66°C durante 30 segundos y 72°C durante 2,5 minutos; y 1 ciclo a 72°C durante 10 minutos. El producto de PCR fue purificado por 1% de electroforesis en gel de agarosa en 50 mM de base Tris-50 mM de acetate-0,1 mM tampón de disodio EDTA (TAE). Un fragmento de aproximadamente 2,5 kb fue extirpado del gel y extraído de la agarosa utilizando un kit de extracción en gel QIAQUICK® (QIAGEN Inc.).

[0228] El plásmido pShTh60 (figura 1, véase también el número de solicitud PCT: PCT/US10/47002, solicitada el 27 de agosto de 2010) fue digerido con *Sex AI* y *Pac I*, separado por 0,8% de electroforesis en gel de agarosa en tampón TBE (10,8 g/L Tris Base, 5,5 g/L ácido bórico, 2 mM EDTA, pH 8,0) y purificado utilizando un kit de extracción en gel QIAQUICK® (QIAGEN Inc.). El producto de PCR purificado anteriormente fue luego insertado en el fragmento pShTh60 digerido utilizando un equipo de reacción Advantage de In-Fusion™ (Clontech, Mountain View, CA, EE.UU.) según las instrucciones del fabricante dando como resultado el plásmido pAmFs69 (figura 2).

[0229] Una parte alícuota de 2,5 µl de la reacción de ligamiento anterior fue transformada en células de *E. coli* químicamente competentes ONE SHOT® TOP10 (Invitrogen, San Diego, CA, EE.UU.) según las instrucciones del fabricante. Los transformantes fueron colocados en placas sobre placas 2XiT+amp e incubados a 37°C durante toda la noche.

[0230] El análisis de secuencia de ADN fue usado en los transformantes resultantes para confirmar la integridad de la secuencia codificante de *bt1*. Los cebadores 610849, 610851, 610853, 610855, 610857, 610859 y 610861 mostrados por debajo fueron usados con un analizador de ADN AB13130XL (Applied Biosystems, Inc., Foster City, CA, EE.UU.) y la técnica de paseo con cebador con química de terminador de coloración (Giesecke *et al.*, 1992, J. Virol. Methods 38: 47-60).

5 Cebador 610849: 5'-GAACAGGAAGAAATCCAAAA-3' (SEC ID n.º: 13)
 Cebador 610851: 5'-GTCGGCATAGCCACTGCAAT-3' (SEC ID n.º: 14)
 Cebador 610853: 5'-TGTTGCCGCCAAGGGACTTA-3' (SEC ID n.º: 15)
 10 Cebador 610855: 5'-CCGAGAGCGTTGAGTTAATC-3' (SEC ID n.º: 16)
 Cebador 610857: 5'-AGCATTAGGGCTAGCTCCGT-3' (SEC ID n.º: 17)
 Cebador 610859: 5'-CCAAGATGCCATGTCAGGAC-3' (SEC ID n.º: 18)
 Cebador 610861: 5'-TCACAAAAGAGTAGAGGCCA-3' (SEC ID n.º: 19)

15 [0231] El constructo de nucleótidos de la secuencia de ADN genómico (SEC ID n.º: 1), y la secuencia de aminoácidos deducida (SEC ID n.º: 2) del gen *bt1* de *Aspergillus aculeatus* se muestran en las figuras 3A y 3B. La secuencia de codificación genómica de 2503 pares de bases (incluyendo un codón de parada) es interrumpida por tres intrones de 78 pares de bases (465-542), 51 pares de bases (1173-1223) y 61 pares de bases (1747-1807). La secuencia de ADNc correspondiente (secuencia de nucleótidos en negrita mostrada en las Figuras 3A y 3B) es de 2313 pares de bases,
 20 incluyendo un codón de terminación. La proteína codificada predicha es 770 aminoácidos, con una masa molecular predicha de 83,9 kDa y un pH isoelectrónico de 6,9.

Ejemplo 2: clonación de un gen de transportador de bicarbonato de *Aspergillus oryzae* AO090003000798 y construcción de vector de expresión correspondiente

25 [0232] El gen de transportador de bicarbonato *bt2* (AO090003000798) fue clonado a partir de ADN genómico de *Aspergillus oryzae* NRRL3488 por amplificación de PCR usando cebadores homólogos del número de modelo del gen de transportador de bicarbonato predicho de *Aspergillus oryzae* AO090003000798 encontrado en la secuencia genómica publicada de *A. oryzae* ATCC 42149 (Galagan *et al.*, 2005, *supra*).

30 [0233] El ADN genómico de *A. oryzae* NRRL3488 fue aislado y los micelios fueron cosechados y procesados como se describe en el ejemplo 1. El gen *bt2* de *Aspergillus oryzae* fue amplificado utilizando el cebador sentido 0614058 y el cebador antisentido 0614057 mostrado por debajo:

35 Cebador 0614058:

5'-GTGATAGAACATCGTCCATAATGCCGGGCGATCTCAAAACC-3' (SEC ID n.º: 63)

40 Cebador 0614057:

5'-GTGTCAGTCACCTCTAGTTACTATGCATCAAGGACATTC-3' (SEC ID n.º: 64)

45 [0234] Las reacciones de amplificación fueron realizadas usando ADN polimerasa de alta fidelidad Phusion® Hot Start (Finnzymes) según las instrucciones del fabricante. Cada reacción por PCR contenía 47 ng de ADN genómico de *Aspergillus oryzae* NRRL3488, 200 µM dNTPs, 50 pM de cebador sentido, 50 pM de cebador antisentido, 1X tampón de reacción Phusion® GC Buffer (Finnzymes), y 50 unidades de ADN polimerasa de alta fidelidad Phusion® Hot Start en un volumen final de 50 µl. Las reacciones de amplificación fueron incubadas en un MASTERCYCLER® de EPPENDORF® (Eppendorf Scientific Inc.) programadas para 1 ciclo a 98°C durante 2 minutos; 35 ciclos a 98°C durante 15 segundos, 65°C durante 15 segundos y 74°C durante 1 minuto; y 1 ciclo a 74°C durante 1 minuto. El producto de PCR fue purificado por 1% de electroforesis en gel de agarosa en 50 mM Tris base-50 mM acetato-0,1 mM tampón de disodio EDTA (TAE). Un fragmento de aproximadamente 2,7 kb fue extirpado del gel y extraído de la agarosa utilizando un equipo de extracción en gel de QIAQUICK® (QIAGEN Inc.).

55 [0235] El plásmido pShTh77 (figura 16) fue digerido con *Sex AI* y *Pac I*, separado por 0,8% de electroforesis en gel de agarosa en tampón TBE (10,8 g/L Tris Base, 5,5 g/L ácido bórico, 2 mM EDTA, pH 8,0) y purificado utilizando un equipo de extracción en gel de QIAQUICK® (QIAGEN Inc., Valencia, CA, EE.UU.). El producto de PCR purificado anteriormente fue luego insertado en el fragmento pShTh77 digerido utilizando un equipo de reacción Advantage de In-Fusion™ (Clontech) según las instrucciones del fabricante dando como resultado el plásmido pShTh147 (figura 17).

60 [0236] Una parte alícuota de 2,5 µl de la reacción de ligamiento anterior fue transformada en células de *E. coli* químicamente competentes ONE SHOT® TOP10 según las instrucciones del fabricante. Los transformantes fueron colocados en placas sobre placas 2Xit+amp e incubados a 37°C durante toda la noche.

65 [0237] El análisis de secuencia de ADN fue usado en los transformantes resultantes para confirmar la integridad de la secuencia codificante de *bt2*. Los cebadores 0614313, 0614314, 996270 y 0611428, mostrados por debajo fueron

usados con un analizador de ADN ABI3130XL (Applied Biosystems, Inc., Foster City, CA, EE.UU.) y la técnica de paseo con cebador con química de terminador de coloración (Giesecke *et al.*, 1992, J. Virol. Methods 38: 47-60).

Cebador 0614313: 5'- GATTGAGATCGGCATTTACT-3' (SEC ID n.º: 65)

Cebador 0614314: 5'-ACGCGGAACAGCAGAATGGC-3' (SEC ID n.º: 66)

5 Cebador 996270: 5'-CTATAGCGAAATGGATTGATTGTCT-3' (SEC ID n.º: 67)

Cebador 0611428: 5'-TTCACCGTGAAACGTATTGA-3' (SEC ID n.º: 68)

10 [0238] El constructo de nucleótidos de la secuencia de ADN genómico (SEC ID n.º: 3), y la secuencia de aminoácidos deducida (SEC ID n.º: 4) del gen *bt1* del *Aspergillus oryzae* se muestran en las figuras 4A y 4B. La secuencia de codificación genómica de 2657 pares de bases (incluyendo codón de parada) se interrumpe por dos intrones de 64 pares de bases (302-365) y 61 pares de bases (512-572). La secuencia de ADNc correspondiente (secuencia de nucleótidos en negrita mostrada en las Figuras 4A y 4B) es 2532 pares de bases, incluyendo un codón de parada. La proteína codificada predicha es 843 aminoácidos, con una masa molecular predicha de 92,5 kDa y un pH isoelectrónico de 8,4.

15 **Ejemplo 3: clonación de un gen transportador de ácido C4-dicarboxílico de *Aspergillus aculeatus* y construcción del vector de expresión pSaMF36**

20 [0239] ADN genómico de *Aspergillus aculeatus* fue aislado por inoculación de 100 ml de medio YEG en un matraz de agitación con 2×10^6 esporas e incubación del matraz a 34°C durante toda la noche con la agitación a 160 r.p.m. Los micelios fueron cosechados por filtración utilizando un embudo forrado con MIRACLOTH® (Calbiochem) y aproximadamente 2 g de micelios fueron recuperados y congelados en nitrógeno líquido. Los micelios congelados fueron interrumpidos por aplastamiento rápido con un martillo mientras se envolvían dentro del MIRACLOTH®. Los micelios interrumpidos fueron luego transferidos a un tubo centrífugo cónico de polipropileno de 50 ml que contenía 10ml de 1X tampón de lisis (100 mM EDTA, 10 mM Tris pH 8,0, 1% Triton® X-100, 0,5 M Guanidina-HCl, 200 mM NaCl) y 3 µl de RNasa A (QIAGEN Inc.; 100 mg/ml). El tubo fue mezclado por agitación en vórtex suave, y luego incubado a temperatura ambiente durante 5 minutos, después de lo cual se añadió 150 µl de proteinasa K (QIAGEN Inc.; 20 mg/ml). El tubo fue mezclado por inversión e incubado a 50°C durante 1 hora. +l tubo fue luego centrifugado a 7240 x g durante 20 minutos. El sobrenadante fue luego añadido a una Qiagen-tip 100 preequilibrada (QIAGEN Inc.) y los pasos de extracción del ADN restante fueron realizados según las instrucciones del fabricante. El ADN fue resuspendido en 100 µl de tampón TE (10 mM Tris Base, 1 mM EDTA, pH 8,0).

35 [0240] El gen transportador de ácido C4-dicarboxílico de 1257 pares de bases *c4t521* fue amplificado a partir de ADN genómico de *Aspergillus aculeatus* aislado utilizando los cebadores 069700 y 069701 mostrados por debajo.

Cebador 069700:

5'-TGTGATAGAACATCGTCCATAATGCACGACCACAGC-3' (SEC ID n.º: 20)

40 Cebador 069701:

5'-GTGTCAGTCACCTCTAGTTATCATTGGAACAACCTCGGACA-3' (SEC ID n.º: 21)

45 [0241] La reacción por PCR estaba compuesta por 10 µl 5X tampón de reacción, 1 µl modelo de ADN genómico de *A. aculeatus* (105 ng/µl), 1 µl cebador 069700 (100 ng/µl), 1 µl cebador 069701 (100 ng/µl), 1 µl mezcla dNTP (10 mM), 35,5 µl agua desionizada y 0,5 µl ADN polimerasa de alta fidelidad Phusion™ Hot Start (Finnzymes Inc.). La reacción de amplificación fue incubada en un MASTERCYCLER® de EPPENDORF® (Eppendorf Scientific Inc.) programado para 1 ciclo a 98°C durante 30 segundos; 30 ciclos cada uno a 98°C durante 10 segundos, 60°C durante 30 segundos, 72°C durante 1 minuto y un ciclo a 72°C durante 10 minutos. El producto de PCR fue digerido con *Dpn* I durante 1 hora para degradar cualquier modelo de ADN plásmido.

55 [0242] El plásmido pShTh60 (figura 1) fue digerido con *Sex* AI y *Pac* I, separado por 0,8% de electroforesis en gel de agarosa en tampón TBE y purificado utilizando un equipo de extracción en gel de QIAQUICK®. El producto de PCR purificado anteriormente fue luego insertado en el fragmento pShTh60 digerido utilizando un equipo de reacción Advantage de In-Fusion™ compuesto por 2 µl 5X tampón, 3 µl producto de PCR purificado (26 ng/µl), 1,5 µl *Sex* AI y *Pac* I purificado en gel digerido y pShTh60 purificado en gel (132 ng/µl), 1 µl enzima In-Fusion™ y 2,5 µl agua desionizada. La reacción fue incubada a 37°C durante 15 minutos, 50°C durante 15 minutos, colocada en hielo durante 5 minutos y diluida con 40 µl de tampón TE dando como resultado pSaMF36 (figura 5).

60 [0243] Una parte alícuota de 2,5 µl de la reacción de ligamiento anterior fue transformada en células de *E. coli* químicamente competente ONE SHOT® TOP10 según las instrucciones del fabricante. Los transformantes fueron colocados en placas sobre placas 2Xit+amp e incubados a 37°C durante toda la noche. Los transformantes resultantes fueron escogidos y sometidos a secuenciación de ADN para confirmar que el gen *mat521* fuera exitosamente integrado en el vector.

65

[0244] El constructo de nucleótidos de la secuencia de ADN genómico (SEC ID n.º: 5) y secuencia de aminoácidos deducida (SEC ID n.º: 6) del gen *c4t521* de *Aspergillus aculeatus* se muestran en la figura 6. La secuencia codificante genómica de 1257 pares de bases (incluyendo codón de terminación) no contiene intrones. La proteína codificada predicha es 418 aminoácidos, con una masa molecular predicha de 46,8 kDa y un pH isoelectrónico de 6,36. Utilizando el programa SignalP (Nielsen *et al.*, 1997, Protein Engineering 10: 1-6), un péptido señal de 17 residuos fue predicho. Basada en este programa, la proteína madura predicha contiene 401 aminoácidos con una masa molecular predicha de 44,9 kDa y un pH isoelectrónico de 6,89.

Ejemplo 4: clonación de un gen de malato deshidrogenasa de *Aspergillus oryzae* y construcción del vector de expresión pSaMF21

[0245] El plásmido pSaMF21 fue construido para contener la secuencia de génica de la malato deshidrogenasa dependiente de NAD (*mdh3*) (DOGAN: AO090701000013), un fragmento de 1430 pares de bases de *Aspergillus oryzae* como se describe en el número de solicitud PCT. PCT/US10/47002, solicitada el 27 de agosto de 2010. El constructo de nucleótidos de la secuencia de ADN genómico (SEC ID n.º: 7) y la secuencia de aminoácidos deducida (SEC ID n.º: 8) del gen *mdh3* de malato deshidrogenasa de *Aspergillus oryzae* NRRL 3488 se muestran en la figura 7. La secuencia codificante genómica de 1430 pb (incluyendo codón de terminación) es interrumpida por 7 intrones de 57 pb (14-70 pb), 70 pb (103-172 pb), 74 pb (284-357 pb), 68 pb (446-513 pb), 58 pb (892-949 pb), 48 pb (1035-1082 pb) y 62 pb (1228-1289 pb). El contenido G+C de la región de codificación del gen *mdh3* es 50,3%. La secuencia de ADNc correspondiente (secuencia de nucleótidos en negrita mostrada en la figura 7) es 993 pb, incluyendo un codón de terminación. La proteína codificada predicha es 330 aminoácidos con una masa predicha de 34,5 kDa y un pH isoelectrónico de 6,79.

[0246] En resumen, el plásmido fue construido linealizando pShTh60 (figura 1) por digestión de restricción con *Sex AI* y *Pac I*. El vector digerido fue separado por 0,8% de electroforesis en gel de agarosa en tampón TBE y purificado utilizando un equipo de extracción en gel de QIAQUICK®. El gen *mdh3* fue amplificado a partir de pShTh71 (número de solicitud PCT PCT/US10/47002, solicitada el 27 de agosto de 2010) usando los cebadores 067522 y 067525.

Cebador 067522:

5'-AGAACATCGTCCATAATGGTCAAAGCTGGTGAGTTA-3' (SEC ID n.º: 22)

Cebador 067525:

5'-GTGTCAGTCACCTCTAGTTATTACTTTGGTGGTGGGTTCT-3' (SEC ID n.º: 23)

[0247] La reacción por PCR estaba compuesta por 5 µl 10X tampón de reacción, 1 µl modelo pShTh71 (87 ng/µl), 1 µl cebador 067522 (100 ng/µl), 1 µl cebador 067525 (100 ng/µl), 1 µl mezcla dNTP (10 mM), 44,5 µl agua desionizada y 0,5 µl ADN polimerasa HotStart de Herculase® (Stratagene, La Jolla, CA, EE.UU.). La reacción de amplificación fue incubada en un MASTERCYCLER® de EPPENDORF® (Eppendorf Scientific Inc.) programado para 1 ciclo a 95°C durante 2 minutos; 10 ciclos cada uno a 95°C durante 10 segundos, 58°C durante 30 segundos y 72°C durante 1,5 minutos; 20 ciclos cada uno a 95°C durante 10 segundos, 50°C durante 30 segundos y 72°C durante 1,5 minutos más 10 segundos por ciclo. La reacción por PCR fue sometida a una asimilación de restricción con *Dpn I* durante 1 hora para degradar cualquier modelo de ADN plásmido. El producto de PCR fue luego purificado utilizando el equipo de purificación por PCR MinElute® (QIAGEN Inc.). El producto de PCR purificado fue insertado en el vector utilizando una reacción Advantage de In-Fusion™ compuesta por 2 µl 5X tampón, 0,5 µl producto de PCR purificado (110 ng/µl), 1,7 µl restricción de *Sex AI* y *Pac I* purificada de gel digerida pShTh60 (figura 1,78 ng/µl), 1 µl enzima de In-Fusion™ y 4,8 µl agua desionizada. La reacción fue incubada a 37°C durante 15 minutos seguidos de 50°C durante 15 minutos después de lo cual fue colocada en hielo durante 5 minutos y diluida con 40 µl tampón TE dando como resultado pSaMF21 (figura 8). Una parte alícuota de 2 µl de la reacción de ligamiento fue transformada en células de *E. coli* químicamente competentes ONE SHOT® TOP10 (Invitrogen) según las instrucciones del fabricante. Los transformantes fueron colocados sobre placas 2XIT+amp e incubados a 37°C durante toda la noche. Los transformantes resultantes fueron escogidos y sometidos a secuenciación del ADN para confirmar que el gen *mdh3* estaba exitosamente integrado en el vector.

Ejemplo 5: clonación de un gen de piruvato carboxilasa de *Aspergillus oryzae* y construcción del vector de expresión pRyan1

[0248] El plásmido pRyan1 fue construido para contener la secuencia génica de piruvato carboxilasa (*pyc*) (DOGAN: AO090023000801), un fragmento de 3646 pb de *Aspergillus oryzae* (incluyendo dos codones de parada) como se describe en la solicitud PCT número PCT/US10/47002, solicitada el 27 de agosto de 2010. El constructo de nucleótidos de la secuencia de ADN genómica (SEC ID n.º: 9) y secuencia de aminoácidos deducida (SEC ID n.º: 10) del gen de piruvato carboxilasa de *Aspergillus oryzae* se muestran en las figuras 9A y 9B. Tanto el gen de piruvato carboxilasa de *Aspergillus oryzae* NRRL 3488 como el ATCC 56747 tienen la misma secuencia de nucleótidos. El contenido G+C de la región de codificación del gen es 57,1%. La secuencia codificante genómica de 3643 pb (incluyendo un codón de terminación) es interrumpida por 1 intrón de 61 pb (3475-3535 pb). El contenido G+C de la región de codificación del

gen es 57,1%. La secuencia de ADNc correspondiente (secuencia de nucleótidos en negrita mostrada en las Figuras 9A y 9B) es 3582 pb, incluyendo un codón de terminación. La proteína codificada predicha es 1193 aminoácidos con una masa predicha de 131 kDa.

5 [0249] En resumen, el plásmido fue construido linealizando pShTh60 (figura 1) por digestión de restricción con Sex AI y Pac I. El vector digerido fue separado por 0,8% de electroforesis en gel de agarosa en tampón TBE y purificado utilizando un equipo de extracción en gel de QIAQUICK®. El gen *pyc* fue amplificado a partir del ADN genómico de *Aspergillus oryzae* NRRL 3488 usando los cebadores 066549 y 067388 mostrados por debajo.

10 Cebador 066549:

5'-TAGAACATCGTCCATAATGGCGGCTCCGTTTCGTCA-3' (SEC ID n.º: 24)

15 Cebador 067388:

5'-GTGTGTCAGTCACCTCTAGTTATTATTACGCTTTGACGATCT-3' (SEC ID n.º: 25)

[0250] La reacción por PCR estaba compuesta por 5 µl 10X tampón de reacción, 1 µl ADN genómico de *Aspergillus oryzae* NRRL3488 (110 ng/µl), 1 µl cebador 066549 (100 ng/µl), 1 µl cebador 067388 (100 ng/µl), 1 µl mezcla dNTP (10 mM), 45,5 µl agua desionizada y 0,5 µl ADN polimerasa Herculase® HotStart. La reacción de amplificación fue incubada en un MASTERCYCLER® de EPPENDORF® (Eppendorf Scientific Inc.) programado para 1 ciclo a 95°C durante 2 minutos; 10 ciclos cada uno a 95°C durante 10 segundos, 58°C durante 30 segundos y 72°C durante 3,5 minutos; 20 ciclos cada uno a 95°C durante 10 segundos, 58°C durante 30 segundos y 72°C durante 3,5 minutos más 10 segundos por ciclo. El producto de PCR fue luego purificado utilizando un equipo de purificación de PCR MinElute®.

25 [0251] El producto de PCR purificado fue insertado en el vector utilizando una reacción Advantage de In-Fusion™ compuesta por 2 µl 5X tampón, 1 µl producto de PCR purificado (144 ng/µl), 2 µl restricción de Sex AI y Pac I purificada en gel digerida pShTh60 (figura 1,78 ng/µl), 1 µl enzima de In-Fusion™ y 4 µl agua desionizada. La reacción fue incubada a 37°C durante 15 minutos seguido de 50°C durante 15 minutos después de lo cual se colocó en hielo durante 5 minutos y diluyó con 40 µl tampón TE dando como resultado pRYAN1 (figura 10). Un parte alícuota de 2 µl de la reacción de ligamiento fue transformada en células de *E. coli* químicamente competentes ONE SHOT® TOP10 según las instrucciones del fabricante. Transformantes fueron colocados en placas sobre placas 2Xit+amp e incubados a 37°C durante toda la noche. Los transformantes resultantes fueron escogidos y sometidos a secuenciación del ADN para confirmar que el gen *pyc* se había integrado exitosamente en el vector. El nucleótido 1308 fue cambiado de C a T, pero no afectó a la secuencia de la proteína.

Ejemplo 6: preparación de transformantes de *Aspergillus oryzae* ShTh6900

40 [0252] La preparación de protoplasto y la transformación de *Aspergillus oryzae* NRRL3488 fueron realizadas por inoculación de aproximadamente 2×10^7 esporas en 100 ml de medio YEG e incubación en matraz a 27°C durante 16-18 horas a 140 r.p.m. Los micelios fueron recogidos vertiendo el cultivo a través de un embudo estéril forrado con MIRACLOTH® y aclarado con 50 ml de 0,7 M KCl. Los micelios lavados fueron resuspendidos en un matraz de 125 ml con 20 ml de solución protoplástica compuesta por 5 mg de GLUCANEX™ (Novozymes A/S, Bagsvaerd, Dinamarca) y 0,5 mg de quitinasa (Sigma, EE.UU.) por ml de 0,7 M KCl (filtro esterilizado) e incubado a 34°C, durante 30 minutos con mezcla a 80 r.p.m. La solución protoplástica fue vertida a través de un embudo estéril forrado con MIRACLOTH® y enjuagada con 50 ml de STC compuesto por 1 M sorbitol-10 mM Tris-HCl pH 6,5-10 mM de CaCl₂. El flujo pasante fue recogido en dos tubos de polipropileno de 50 ml. Los tubos fueron centrifugados en el centrifugador a 1300 x g durante 10 minutos a temperatura ambiente. El sobrenadante fue descartado y el granulado de protoplasto fue resuspendido en 20 ml de tampón STC. Los protoplastos fueron lavados por resuspensión de granulado de dos rondas en 20 ml de tampón STC y centrifugación a 1300 x g durante 10 minutos a temperatura ambiente. El granulado final fue resuspendido en 2 ml de tampón STC. Los protoplastos fueron contados eliminando una muestra de 10 µl y contando éstos en un hemocitómetro (VWR, West Chester, PA, EE.UU.). El volumen fue ajustado con tampón STC para obtener una concentración de protoplasto de 2×10^7 por ml.

55 [0253] Los vectores de expresión del plásmido pAmFs69 (ejemplo 1), pSaMF36 (ejemplo 3), pSaMF21 (ejemplo 4) y pRyan1 (ejemplo 5) fueron individualmente preparados para transformación por digestión de restricción con *Pme* I durante 4 horas a 37°C. Los casetes de expresión de aproximadamente 5-6 kb de cada constructo fueron separados de las secuencias de vector por 0,8% de electroforesis en gel de agarosa en tampón TBE y purificado utilizando un equipo de extracción en gel de QIAQUICK® según las instrucciones del fabricante.

60 [0254] Cuatro reacciones de transformación fueron preparadas añadiendo 100 µl de la preparación de protoplasto anterior en cuatro tubos de polipropileno de 12 ml. Para cada tubo se añadieron dos microgramos del fragmento de *pyc* digerido pRyan1, y un microgramo cada uno del fragmento de *bt1l* digerido pAmFs69, fragmento de *C4T521* digerido pSaMF36, y el fragmento de *mdh* digerido pSaMF21 a una solución de polietilenglicol 250 µl (PEG) (60% de p/v de polietilenglicol (PEG), 10 mM Tris 6,5, 10 mM de CaCl) seguido de mezcla suave e incubación a 37°C durante 30 minutos. Cada reacción de transformación fue diluida con 6 ml de tampón STC, seguido de colocación en placas de tres

partes alícuotas separadas sobre placas de COVE. Cada placa fue luego incubada a 34°C durante 7-10 días. Sesenta de los transformantes resultantes (designados transformantes ShTh6900) fueron transferidos a placas de COVE individuales e incubados a 34°C durante 5 días. Caldos de esporas fueron preparados recolectando las esporas en 0,1% de TWEEN® 80. Los cultivos fueron almacenados preparando un caldo de glicerol de cada uno (800 µl caldo de espора, 200 µl 0,1% TWEEN® 80) y congelados a -80°C.

[0255] Los transformantes se dejaron crecer en frascos de agitación y ADN genómico aislados según la descripción anterior. Las reacciones de PCR individuales para probar la presencia de cada uno de los cuatro fragmentos de vector de expresión estaban compuestas por 5 µl 10X tampón de reacción; 0,5 µl modelo (80-300 ng/µl); 1,0 µl cebador sentido (50 pM; véase por debajo); 1,0 µl cebador antisentido (50 pM; véase por debajo); 0,5 µl mezcla dNTP (10 mM), 16,75 µl agua desionizada y 0,25 µl ADN polimerasa Phusion®.

Cebador sentido 065067 (para los fragmentos de pRyan1 pyc, pSaMf21 mdh y pSaMf36 C4T521):

15 5'-TGACCTTCCACGCTGACCAC-3' (SEC ID n.º: 46)

Cebador sentido 0610854 (para el fragmento pAmFs69 bt1):

20 5'-GGCTGAGAAAATATGTTGCA-3' (SEC ID n.º: 47)

Cebador antisentido 0611365 (para el fragmento pSaMF36 C4T521):

5'-GATAGACCACTAATCATGGTGGCGATGGAG-3' (SEC ID n.º: 48)

25 Cebador antisentido 061752 (para el fragmento pRyan1 pyc)

5'-TGCGGTCCTGAGTCAGGCCAGTTGCTCGA-3' (SEC ID n.º: 49)

Cebador antisentido 062400 (para el fragmento pSaMF21 mdh)

30 5'-GGGATTTGAACAGCAGAAGG-3' (SEC ID n.º: 50)

Cebador antisentido 996270 (para el fragmento pAmFs69 bt1)

35 5'-TCACAAAAGAGTAGAGGCCA-3' (SEC ID n.º: 51)

[0256] Las reacciones de amplificación fueron incubadas en un MASTERCYCLER® de EPPENDORF® (Eppendorf Scientific Inc.) programadas para 1 ciclo a 98°C durante 30 segundos; 35 ciclos cada una a 98°C durante 10 segundos; 66°C (para el fragmento de pyc pRyan1) o 58°C (para los fragmentos de bt1pAmFs69, mdh pSaMf21 y C4T521 pSaMf36) durante 10 segundos; 72°C durante 15 segundos; y un ciclo de 72°C durante 10 minutos. ADN genómico de *Aspergillus oryzae* NRRL 3488 (110 ng/µl) fue usado como modelo de control negativo y cada plásmido (pRyan1, pAmFs69, pSaMf21 o pSaMf36 diluido a 20 ng/µl) fue usado como modelo de control positivo. Mezclas de reacción de amplificación fueron analizadas por electroforesis en gel utilizando 2 µl de cada mezcla reactiva en un gel de agarosa al 0,8%.

Ejemplo 7: preparación de transformantes de *Aspergillus oryzae* ΔPyrG6900 a partir de ShTh6900

[0257] Preparaciones de protoplasto de transformantes de *Aspergillus oryzae* ShTh6900 (*supra*) fueron realizados por inoculación de aproximadamente 2×10^7 esporas en 100 ml de medio YEG e incubación en el matraz a 34°C durante 16-18 horas a 160 r.p.m. Micelios fueron recogidos vertiendo el cultivo a través de un embudo estéril forrado con MIRACLOTH® y aclarado con 50 ml de 0,7 M KCl. Los micelios lavados fueron resuspendidos en un matraz de 125 ml con 20 ml de solución protoplástica compuesta por 5 mg de GLUCANEX™ (Novozymes A/S) y 0,5 mg de quitinasa (Sigma) por ml de 0,7 M KCl (filtro esterilizado) e incubados a 34°C, durante 45 minutos con mezcla a 80 r.p.m. La solución protoplástica fue vertida a través de un embudo estéril forrado con MIRACLOTH® y enjuagada con 50 ml de tampón STC (1M Sorbitol, 10mM CaCl, 10mM Tris pH 7). El flujo pasante fue recogido en dos tubos de polipropileno de 50 ml. Los tubos fueron centrifugados en el centrifugador a 1900 x g durante 5 minutos a temperatura ambiente. El sobrenadante fue descartado y el granulado de protoplasto fue resuspendido en 20 ml de tampón STC. Los protoplastos fueron lavados por resuspensión de granulado de dos rondas en 20 ml de tampón STC y centrifugación a 1900 x g durante 5 minutos a temperatura ambiente. El granulado final fue resuspendido en 1 ml de tampón STC. Los protoplastos fueron contados eliminando una muestra de 10 µl y recuento de éstos en un hemocitómetro (VWR). El volumen fue ajustado con tampón STC para obtener una concentración de protoplasto de 2×10^7 por ml.

[0258] Un casete de eliminación pyrG fue preparado para transformación por amplificación del fragmento de eliminación a partir de pShTh66 usando los cebadores 62329 y 62199. La reacción por PCR estaba compuesta por 40 µl 10X tampón de reacción, 2 µl modelo pShTh66 (figura 15; SEC ID n.º: 62, 200 ng/µl), 8 µl cebador 62329 (50pM/µl), 8 µl

cebador 62199 (50pM/μl), 8 μl mezcla dNTP (10 mM), 328 μl agua desionizada y 6 μl Expand ADN polimerasa (Roche, EE.UU.) y dividida entre 4 tubos. Las reacciones de amplificación fueron incubadas en un MASTERCYCLER® de EPPENDORF® (Eppendorf Scientific Inc.) programadas para 1 ciclo a 95°C durante 3 minutos; 10 ciclos cada uno a 95°C durante 30 segundos, 57°C durante 30 segundos y 68°C durante 3 minutos; 25 ciclos cada uno a 95°C durante 30 segundos, 57°C durante 30 segundos, y 68°C durante 3 minutos más 5 segundos cada ciclo sucesivo; y un ciclo a 68°C durante 7 minutos. El producto de PCR fue purificado por 1% de electroforesis en gel de agarosa en 50 mM Tris base-50 mM acetato-0,1 mM tampón de disodio EDTA (TAE). Un fragmento de aproximadamente 3,2 kb fue extirpado del gel y extraído de la agarosa utilizando un equipo de extracción en gel QIAQUICK® (QIAGEN Inc.).

10 Cebador 62329:

5'-ATTAATAAGACCACGAGAGGGGAACTAGGGAAATC-3' (SEC ID n.º: 56)

15 Cebador 62199:

5'-TTAATTAACCTTTCCCCCGTAATCTAA-3' (SEC ID n.º: 57)

[0259] Cinco reacciones de transformación cada una fueron preparadas mediante la mezcla de 100 μl de la preparación de protoplasto ShTh6900 de *Aspergillus oryzae* anterior con 1 μg del -3,2 kb fragmento de eliminación pyrG anterior en un tubo de polipropileno de 12 ml. 250 μl de polietilenglicol (PEG) fue añadido y las reacciones suavemente mezcladas, seguido de incubación a 37°C durante 30 minutos. 5 ml YP5% glucosa, 10 mM uridina se añadió a cada reacción y los tubos permitieron incubar a 34°C durante toda la noche. Cada reacción fue colocada en placas sobre dos placas MM 1 M sacarosa + uridina + FOA e incubadas a 34°C durante 7-10 días. Los transformantes resultantes luego fueron transferidos a placas individuales MM + uridina + FOA e incubados a 34°C durante 5 días. Los transformantes fueron sometidos a 2 ciclos de purificación de esporas después de los cuales fueron colocados en placas sobre placas MM + uridina. Caldos de esporas fueron preparados recogiendo las esporas en 0,1% TWEEN® 80. Los cultivos de transformantes resultantes, designados ΔPyrG6900, fueron almacenados a -80°C en un caldo de glicerol (800 μl caldo de espora, 200 μl 0,1% TWEEN® 80).

[0260] Los aislados de esporas fueron confirmados por análisis de Southern y mostraron que carecían del gen pyrG sin integración ectópica. Un segundo análisis de Southern se preparó con los aislados de esporas y se evaluó con el gen amp para asegurar que ninguna secuencia de ADN plásmido estuviera presente en la cepa. No se vio ninguna hibridación entre cualquiera de los transformantes y la sonda de amp.

35 **Ejemplo 8: clonación de un gen de anhidrasa carbónica de *Aspergillus clavatus* y construcción del vector de expresión pSaMF58**

[0261] ADN genómico de *Aspergillus clavatus* NRRL1 fue aislado por la inoculación de 100 ml de medio YEG en un matraz de agitación con todas las esporas cosechadas de una placa PDA de 10 días e incubando el matraz a 34°C durante toda la noche con agitación a 160 r.p.m. Los micelios fueron cosechados por filtración utilizando un embudo forrado con MIRACLOTH® (Calbiochem) y aproximadamente 2 g de micelios fueron recuperados y congelados en nitrógeno líquido. Los micelios congelados fueron interrumpidos por trituración rápida con un martillo mientras estaban envueltos en el Miracloth. Los micelios interrumpidos fueron luego transferidos a un tubo de 50 ml centrífugo cónico de polipropileno que contenía 10ml de 1X tampón de lisis (100 mM EDTA, 10 mM Tris pH 8,0, 1% Tritón X-100, 0,5 M Guanidina-HCl, 200 mM NaCl) y 3 μl de 100 mg/ml RNasa A. El tubo centrífugo fue mezclado por agitación suave en vórtex y luego incubado a temperatura ambiente durante 5 minutos, seguido de la adición de 150 μl de 20 mg/ml de proteinasa K. El tubo centrífugo fue mezclado por inversión, colocado a 50°C para incubar durante 1 hora, y luego centrifugado a 7240 x g durante 20 minutos. El sobrenadante fue luego añadido a una QIAGEN-tip 100 preequilibrada (QIAGEN Inc.) y los pasos de extracción de ADN restantes fueron según las instrucciones del fabricante. El ADN fue resuspendido en 50 μl de tampón TE.

[0262] El gen de anhidrasa carbónica (CA) (ACLA 007930) fue amplificado a partir de ADN genómico de *Aspergillus clavatus* aislado por amplificación de PCR usando los cebadores 0612826 y 0612827 mostrados por debajo.

55 Cebador 0612826:

5'-CCAACAGACACATCTAAACAATGTCCGACAAGGCTC-3' (SEC ID n.º: 52)

Cebador 0612827:

60 5'-GTGTCAGTCACCTCTAGTTATCAGCTCTTGGTGATATTGT-3' (SEC ID n.º: 53)

[0263] La reacción por PCR estaba compuesta por 10 μl 5X tampón de reacción, 0,5 μl modelo de ADN genómico de *A. clavatus* NRRL1 (113 ng/μl), 1 μl cebador 0612826 (100 ng/μl), 1 μl cebador 0612827 (100 ng/μl), 1 μl mezcla dNTP (10 mM), 36 μl agua desionizada y 0,5 μl ADN polimerasa de alta fidelidad Phusion™ Hot Start (Finnzymes Inc.). La reacción de amplificación fue incubada en un MASTERCYCLER® de EPPENDORF® (Eppendorf Scientific Inc.)

programada para 1 ciclo a 98°C durante 30 segundos; 30 ciclos cada uno a 98°C durante 10 segundos, 60°C durante 30 segundos y 72°C durante 30 segundos; y un ciclo a 72°C durante 10 minutos. El producto de PCR fue luego purificado utilizando el equipo de purificación de PCR MinElute® (QIAGEN Inc.) según las instrucciones del fabricante. Un sitio de restricción *Bst* XI fue introducido en el plásmido pAIIo2 (véase WO2004/099228) aguas arriba del promotor NA2-tpi que usando los cebadores de oligómeros 064011 y 064012 con un equipo de mutagénesis dirigida al sitio Quick Change II SL (Agilent Technologies) dando como resultado el plásmido pShTh76 (figura 11).

Cebador 064011:

5'-GGAAACAGCTATGACCATGATTATGGATTGTTAAACGTCGACGC-3' (SEC ID n.º: 58)

Cebador 064012:

5'-GCGTCGACGTTTAAACAATCCATAATCATGGTCATAGCTGTTTCC-3' (SEC ID n.º: 59)

[0264] El plásmido pShTh76 luego fue linealizado por digestión con *Bst* XI y *BGL* II. El vector digerido luego fue separado por 0,8% electroforesis en gel de agarosa en el tampón TBE y purificado en gel utilizando un equipo de purificación de PCR Qiagen MinElute® (QIAGEN Inc.) según las instrucciones del fabricante.

[0265] El promotor *gpd* de *A. oryzae* NRRL3488 fue amplificado usando los cebadores 0612469 y 0612468.

Cebador 0612469:

5'-GGAAACAGCTATGACCATGATTCCAGATTGTAAATTAC-3' (SEC ID n.º: 60)

Cebador 0612468:

5'-AGCACTAGTACGCGTAGATCTGTTTAGATGTGTCTGTTGG-3' (SEC ID n.º: 61)

[0266] La reacción por PCR estaba compuesta por 10 µl 5X tampón de reacción, 0,5 µl modelo *gADN* de *A. oryzae* NRRL3488 (200 ng/µl), 1 µl cebador 0612826 (100 ng/µl), 1 µl cebador 0612827 (100 ng/µl), 1 µl mezcla dNTP (10 mM), 36 µl agua desionizada y 0,5 µl ADN polimerasa de alta fidelidad Phusion™ Hot Start (Finnzymes Inc.). La reacción de amplificación fue incubada en un MASTERCYCLER® de EPPENDORF® (Eppendorf Scientific Inc.) programada para 1 ciclo a 98°C durante 30 segundos; 30 ciclos cada uno a 98°C durante 10 segundos, 65°C durante 30 segundos y 72°C durante 60 segundos; y un ciclo a 72°C durante 10 minutos. El producto de PCR fue luego purificado en columna utilizando el equipo de purificación de PCR MinElute® (QIAGEN Inc.).

[0267] El producto de PCR purificado anterior que contenía el promotor *gpd* luego fue insertado en el vector pShTh76 digerido utilizando una reacción Advantage de In-Fusion™ compuesta por 2 µl 5X tampón, 2 µl producto de PCR purificado (42 ng/µl), 2,7 µl *Bst* XI purificado en gel y restricción digerida de *BGL* II pShTh76 (75 ng/µl), 1 µl enzima In-Fusion™ y 2,3 µl agua desionizada. La reacción fue incubada a 37°C durante 15 minutos luego 50°C durante 15 minutos después de lo cual se colocó sobre hielo durante 5 minutos y diluyó con 40 µl TE tampón dando como resultado pSaMF48. Una parte alícuota de 2,5 µl de la reacción de ligamiento fue transformada en células de *E. coli* químicamente competentes ONE SHOT@ TOP10 (Invitrogen) según las instrucciones del fabricante. Transformantes fueron colocados en placas sobre placas 2Xit+amp e incubados a 37°C durante toda la noche. Los transformantes resultantes, designado pSaMF48 (figura 12), fueron escogidos y sometidos a secuenciación del ADN para confirmar que el promotor *gpd* estaba exitosamente integrado en el vector.

[0268] El plásmido pSaMF48 (figura 12) fue digerido con *BGL* II y *Pac* I, separado por 0,8% electroforesis en gel de agarosa en el tampón TBE, y purificado utilizando un equipo de extracto II Nucleospin® (Macherey-Nagel, Bethlehem, PA, EE.UU.) según las instrucciones del fabricante. El producto de PCR purificado arriba que contenía el gen de anhidrasa carbónica de *Aspergillus clavatus* de SEC ID n.º: 54 fue luego insertado en el vector utilizando una reacción Advantage In-Fusion™ compuesta por 2 µl 5X tampón, 0,25 µl producto de PCR purificado (187 ng/µl), 0,26 µl de la restricción digerida de *BGL* II y *Pac* I y pSaMF48 purificado en gel anterior (557 ng/µl), 1 µl enzima In-Fusion™ y 6,49 µl agua desionizada. La reacción fue incubada a 37°C durante 15 minutos luego 50°C durante 15 minutos después de lo cual fue colocada en hielo durante 5 minutos y diluida con 40 µl tampón TE dando como resultado pSaMF58 (figura 13).

[0269] Una parte alícuota de 2,5 µl de la reacción de ligamiento fue transformada en células de *E. coli* químicamente competentes ONE SHOT@ TOP10 (Invitrogen) según las instrucciones del fabricante. Transformantes fueron colocados en placas sobre placas 2Xit+amp e incubados a 37°C durante toda la noche. Los transformantes resultantes fueron escogidos y sometidos a secuenciación del ADN para confirmar que el gen de CA estaba exitosamente integrados en el vector.

[0270] El constructo de nucleótidos de la secuencia de ADN genómico (SEC ID n.º: 54) y secuencia de aminoácidos deducida (SEC ID n.º: 55) del gen de anhidrasa carbónica de *Aspergillus clavatus* se muestran en la figura 14. La secuencia codificante genómica de 750 pb (incluyendo codón de terminación) se interrumpe por un intrón de 72 pb (153-

224). La secuencia de ADNc correspondiente (secuencia de nucleótidos en negrita mostrada en la figura 13) es 678 pb, incluyendo un codón de terminación. La proteína codificada predicha es 225 aminoácidos, con una masa molecular predicha de 25,7 kDa y un pH isoelectrónico de 6,48.

5 **Ejemplo 9: transformación de un fragmento de vector de expresión de pSaMF58 que contiene un gen de anhidrasa carbónica de *Aspergillus clavatus* en los transformantes de *Aspergillus oryzae* ΔPyrG6900 (SaMF58Q)**

10 [0271] El vector de expresión del plásmido pSaMF58 (figura 12) se preparó para transformación por digestión con Ase I y Xho I durante 4 horas a 37°C. El vector digerido fue separado en un 0,8% de gel TBE de agarosa, la banda de 3292 pb con el casete de expresión fue cortada y purificada utilizando el equipo Nucleospin® Extract II (Macherey-Nagel) según las instrucciones del fabricante.

15 [0272] Dos reacciones de transformación fueron preparadas añadiendo 100 µl de preparación de protoplasto transformante ΔPyrG6900 de *Aspergillus oryzae* (preparada de una manera similar a como se ha descrito anteriormente) en cada uno de dos tubos de polipropileno de 12 ml. A cada tubo fue añadido 5 µg de vector pSaMF58 linealizado, libre de marcador amp (*supra*) y 250 µl de polietilenglicol (PEG) seguido de mezcla suave e incubación a 37°C durante 30 minutos. Cada reacción de transformación fue diluida con 6 ml de tampón STC, separada en 3 ml de partes alícuotas, y luego cada parte alícuota fue colocada en placas sobre una placa de 2 MM + 1 M de sacarosa. Cada placa fue luego incubada a 34°C durante 7-10 días. Los transformantes resultantes (designados SaMF58Q) fueron transferidos a placas individuales MM e incubados a 34°C durante 5 días. Caldos de esporas fueron preparados recogiendo las esporas en 0,1% TWEEN® 80. Cultivos fueron almacenados por preparación de un caldo de glicerol de cada uno (800 µl caldo de espора, 200 µl 0,1 % TWEEN® 80) y congelados a -80°C.

25 **Ejemplo 10: producción de ácido málico en cultivos matraz de agitación de transformantes de *Aspergillus oryzae* que contenían un gen de anhidrasa carbónica de *Aspergillus clavatus* (SaMF58Q)**

30 [0273] Esporas de transformantes SaMF58Q descritas en el ejemplo 9 y *Aspergillus oryzae* NRRL 3488 como un control fueron colocadas en placas sobre placas individuales PDA y se les permitió esporular a 34°C durante 5 a 7 días. Las esporas fueron recogidas en 0,1% TWEEN® 80 y contadas utilizando un hemacitómetro. Cultivos de semillas fueron preparados en matraces de 250 mL que contenían 100 mL de medio B de semilla (suplementados con 4,4 mg/L ZnSO₄·H₂O) e inoculados con 1mL de esporas cosechadas. Los cultivos de semillas se dejaron crecer durante aproximadamente 22 horas a 30°C con agitación a 200 r.p.m. Cultivos de producción ácidos fueron preparados en matraces de 250 mL sin deflectores que contenían 50 mL de medio C de producción ácido (suplementados con 4,4 mg/L ZnSO₄·H₂O) y 3 mL de cultivos de semilla de 22 horas. Los cultivos fueron incubados a 30°C con agitación a 200 r.p.m. durante 3 días.

40 [0274] La cuantificación de ácido málico para los transformantes de cultivo de matraz de agitación fue realizada por cromatografía en fase líquida de alta presión de fase inversa (RP-HPLC) utilizando un 1200 Series Binary LC System y 1200 Series Diode Array Detector (DAD) (Agilent Technologies, Santa Clara, CA, EE.UU.). La separación de fase inversa fue realizada utilizando una columna ID Aqua 5µ C18 125Å 205 x 4,6 mm y AQ C18 4 x 3,0 mm Security Guard Cartridge (Phenomenex, Inc., Torrance, CA, EE.UU.). La fase móvil consistió en 10% metanol (calidad HPLC) y 90% 145 mM tampón de fosfato pH 1,5.

45 [0275] Muestras de cultivo enteras fueron retiradas y diluidas 1:20 en el tampón de ejecución de HPLC compuesto por 900 ml de 145 mM tampón fosfato y 100 ml de metanol pH 1,50. Las muestras fueron luego filtradas a través de una membrana de PVDF Durapore de 96 pocillos de 0,45 micras en una placa de 96 pocillos para análisis ácido.

50 [0276] La RP-HPLC fue realizada utilizando un volumen de inyección de 10 µl a una velocidad de flujo de 0,7 ml/minute (isocrático) y temperatura de columna a 20°C. La detección fue a 210 nm, 4 nm de ancho de banda, con la referencia a 360 nm, 40 nm de ancho de banda. El tiempo de ronda fue 13 minutos. El tiempo cero fue determinado en 3,8 minutos. Las capacidades cuantitativas del método de fase inversa fueron determinadas para ácido málico mediante la realización de inyecciones replicadas de estándares de ácido málico diluidos en serie con concentraciones que variaban de 49,2-3,93 mM. La desviación típica relativa para (RSD) para inyecciones replicadas fue ≤5%. Muestras de ácido málico R² ≥ 0,9999.

60 [0277] Transformantes SaMF58Q de *Aspergillus oryzae* que contienen un gen de anhidrasa carbónica de *Aspergillus clavatus* (SEC ID n.º: 54) mostraron títulos de ácido málico más de dos veces sobre las cepas NRRL 3488 de *A. oryzae*. Adicionalmente, títulos producidos de transformantes SaMF58Q de *A. oryzae* superiores a aquellos observados en experimentos separados con la cepa comparable de *A. oryzae* ShTh6900 (la cepa original para el transformante SaMF58Q, pero carente del gen de anhidrasa carbónica heteróloga de SEC ID n.º: 54).

65 **Ejemplo 11: fermentación de transformantes de *Aspergillus oryzae* que contienen un gen de anhidrasa carbónica de *Aspergillus clavatus* (SaMF58Q)**

[0278] Tres transformantes SaMF58Q de *Aspergillus oryzae* descritos en el ejemplo 9 y el transformante de control de *A. oryzae* ShTh6900 (la cepa original para transformante SaMF58Q, pero carente del gen de anhidrasa carbónica heteróloga de SEC ID n.º: 54) se dejaron crecer durante aproximadamente 7 días a 34°C en placas PDA. Un volumen de 5-6 ml de tampón de fosfato sódico estéril (50 mM, pH 6,8) que contenía 0,2% TWEEN® 80 se añadió a cada placa y las esporas fueron suspendidas por raspado con un bucle de inoculación. Cada suspensión fue transferida por pipeta a un tubo cónico de 50 ml. Para cada tubo, 25 ml de tampón de fosfato de sodio estéril (50 mM, pH 6,8) que contenía 0,2% TWEEN® 80 se añadió a un matraz de 500 ml sin deflectores que contenía 75 ml de medio de semilla, que fue luego inoculado con 2 ml de suspensión de esporas. Los matraces fueron incubados a 34°C y 180 r.p.m. durante aproximadamente 24 horas. Los matraces de semillas fueron combinados para suministrar el inóculo de 144 ml requerido por tanque.

[0279] Fermentadores de tres litros que contenían 1,8 litros de medio de lote fermentador (con o sin 15 µM ZnSO₄ adicional) fueron individualmente inoculados introduciendo 144 ml (8%) del caldo de cultivo de semillas de tres matraces de semillas combinados de bien transformantes SaMF58Q de *Aspergillus oryzae* o transformantes ShTh6900 de *Aspergillus oryzae*. Los fermentadores fueron equilibrados a 34°C ± 0,1 °C y agitados a 700 r.p.m. La corriente de aire de entrada fue mantenida a 1 v/v/m. Una corriente de 30% de glucosa fue administrada a una velocidad de aproximadamente 7,3 g/hr empezando en alrededor de 20 horas de fermentación, aumentando a 9,3 g/hr a 68 horas. 150 G de CaCO₃ estéril fue añadido en el día 3 para mantener el pH de fermentación en el rango de 6 a 7.

[0280] Las muestras fueron retiradas diariamente y analizadas para producción de ácido málico por cromatografía en fase líquida acoplada con espectrometría de masas en serie (LC-MS/MS) utilizando un sistema LC-MS/MS de Agilent con detectores Quad triples y terminal de trabajo Masshunter (Agilent Technologies, Santa Clara, CA, EE.UU.) basado en lo siguiente:

Columna LC: columna Waters XBridge Amide, 3,5 µm, 150X2.1 mm ID (Waters, Milford, MA, USA; P/N: 186004861)
 Volumen de inyección: 2,0 µL
 Tampón de muestra: 584,48 mg EDTA, 154,16 mg NH₄Ac, 200 mL MeOH, 800 mL H₂O, 500 µL 25% NH₃·H₂O, 11,44 mg, estándar interno de ácido málico marcado 13C ((Cambridge Isotope Laboratories, Inc., Andover, MA, EE.UU.) y agua desionizada hasta 1 litro (pH 8,3).
 Solvente A para HPLC: 5 mM NH₄Ac en 80% MeOH
 Solvente B para HPLC: 50 mM (NH₄)₂CO₃ en 20% AcN, ajuste pH a 10 utilizando 25% NH₃·H₂O
 Condición de desarrollo: isocrático 30% B, velocidad de flujo 0,3 mL/min; temperatura de columna 45°C
 Estándares: ácido málico: 60 g/L, 51 g/L, 45 g/L, 30 g/L, 15 g/L, 7,5 g/L, 3,75 g/L, 1,875 g/L cada uno diluido 100 veces con agua destilada doblemente y 10 veces con tampón de muestra (dilución final 1000 veces); ácido succínico, ácido fumárico, ácido cítrico y ácido oxálico: cada uno diluido en 8 niveles a 1000 veces de una manera similar a partir de un caldo de 5-10 g/L.
 Ajustes de MS: gas temp.: 300°C, flujo de gas: 10 L/min, nebulizador: 32 psi, Delta EMV(-): 450
 Ajustes de MRM: tabla 1 por debajo.

Tabla 1.

Compuesto	Precursor	MS1 Res	Producto	MS2 Res	Permanencia	Fragmentador	Energía de colisión	Polaridad
C ₁₃ Ácido málico	137	unidad	92	unidad	50	60	8	negativa
C ₁₃ Ácido málico	137	unidad	74	unidad	50	60	10	negativa
Ácido málico	133	unidad	89	unidad	50	60	8	negativa
Ácido málico	133	unidad	71	unidad	50	60	10	negativa
Ácido succínico	117	unidad	99	unidad	150	50	5	negativa
Ácido succínico	117	unidad	73	unidad	50	50	5	negativa
Ácido fumárico	115	unidad	71	unidad	50	60	3	negativa
Ácido fumárico	115	unidad	27	unidad	50	60	4	negativa
Ácido cítrico	191	unidad	111	unidad	50	80	5	negativa
Ácido cítrico	191	unidad	87	unidad	50	80	11	negativa

ES 2 579 706 T3

Ácido oxálico	89	unidad	61	unidad	50	60	3	negativa
Ácido oxálico	89	unidad	45	unidad	50	60	2	negativa

[0281] Muestras de cultivo entero fueron retiradas diariamente y diluidas 1:10 en el tampón de ejecución de HPLC como se describe para muestras de matraz de agitación en el ejemplo 10. Muestras del día uno fueron además diluidas 1:10 con amonio acuoso (0,0125%), seguido de una dilución 1:10 en el tampón de muestra (dando como resultado una dilución final de 1000 veces). Muestras de los días 2, 3-5 y 6-8, fueron preparadas de manera similar, con agua destilada doblemente adicional para proporcionar diluciones finales de 2000 veces, 4000 veces y 8000 veces, respectivamente. Estándares para ácido málico, ácido succínico, ácido fumárico, ácido cítrico y ácido oxálico fueron preparados como era apropiado para cuantificación y para imitar cualquier efecto matricial en las muestras. Las muestras diluidas fueron luego analizadas por LC-MS/MS utilizando los parámetros anteriormente descritos.

[0282] Los índices de producción de ácido málico para cada transformante SaMF58Q sobre 92 horas de fermentación (normalizado a los controles de ShTh6900 respectivos) se muestran en la tabla 2 por debajo. Dos de los tres transformantes SaMF58Q consiguen índices de producción de ácido málico que están significativamente por encima de los controles respectivos.

Tabla 2.

Cepa	Índice de producción normalizado de ácido C4	
	-Zn	+Zn
ShTh6900-64.3	1,00	1,00
SaMF58Q-2	ND	1,55
SaMF58Q-4	ND	1,15
SaMF58Q-12	1,02	0,99

[0283] Aunque lo anteriormente mencionado se ha descrito con algo de detalle a modo de ilustración y ejemplo para los fines de claridad de la comprensión, es aparente para los expertos en la técnica que cualquier aspecto equivalente o modificación se puede poner en práctica. Por lo tanto, la descripción y los ejemplos no deberían ser interpretados como limitación del ámbito de la invención.

Listado de secuencias

[0284]

- 5 <110> Novozymes, Inc. McFarland, Sarah Brown, Stephen Luttringer, Sheryl
<120> Microorganismos recombinantes para la producción de ácidos C4-dicarboxílicos
<130> 12296-WO-PCT
<150> US 61/525,345
<151> 2011-08-19
<160> 68
10 <170> Versión de PatentIn 3.5
<210> 1
<211> 2503
<212> ADN
<213> *Aspergillus oryzae*
15 <400> 1

ES 2 579 706 T3

atggaatcca gcgctgtaca ggagccgact caacagcgct ctttgcgga tcgcattttt	60
aacctctttc gtacctcttc ctcaaatgat gccccgggtc ttccggcaag actcgttaacc	120
gctgagagcg cagcgcaaaa cgaagggtcg gcgttaatct atccgccacg ggagcctgat	180
gcaaggactc gtcttctcga atcgtacgat cgcggggaac gtggtctgag gaactccggc	240
gttcatggga ctttttcttc acgacctgaa caggaagaaa tccaaaaatg ggatgcaagc	300
tctttgcaga atgctggtaa cgaagaaaga tctcagtccc caggaggagc agacggccat	360
attgggtctc ccggcgacgt ctcaggatac ccacaggac cagagaatat accatcgcta	420
gactcctctt tcacagcatt gcacatgaag aatcataaat ctctgtaggt ttataatcac	480
gttcgccttg ctttctaaca caattgttat ctccatcgtg gagacaacta acgttcatca	540
aggtatatat cttactacat cccatttttc aattggatta ctcaataccg gtggtcgtac	600
attcgaggtg atttggttg cgcgacaacc attgctcca tctatatccc tatggctttg	660
tccttatcct caaatctcgc ccacgcacct cctatcaatg gcctctactc ttttgtgatc	720
aaccctttca tctatgcgat cttcgggagc agcccgtgt taatagtggg ccagaagca	780
gcaggctcct tgcttactgg cacgattgtc aaaactagtg tcagaccagg cccatctggt	840
gaggacgacg aagtagcga tgccatcgtg gtcggcatag ccaactgcaat ggcgggcgcc	900
atgatactga tcgctgggct tacacggctg ggatttctgg acaatgtgct gagccggccc	960
tttcttaggg gtttcattac agcgatcgggt tttgtgattt ttgtggatca actcatcccc	1020
gaagtccgat tgaccgagct agcaaaggaa gctggtgta cccatgggac tacagttgac	1080
aagctcatgt tccttataag aaacatagga ggttgccatg cgcttacaac cgcggtggct	1140
tttgggagct ttgctattat aatggtattt cggttagtgt tggtgactcg gaagcctggt	1200

ES 2 579 706 T3

gcttagactg attaccatta caggactctc aagaaaatgc tccagccgcg gtatcctcag 1260
 gtgatttata ttccggaccg aattctcgtg gttattcttt cagccgtcct gacatggcat 1320
 cttggttggg atgacaaagg gttggagatt cttgggccct tgaacaaaa tgccaatggc 1380
 ctttttgcgt tcaaatggcc tttccagttt agccagatga agcatgtacg cgctgcaatg 1440
 agtacttctt tcgtcatcgc gttacttggc tttttcagag cttctgttgc cgccaaggga 1500
 cttagtggcg aggccagaca agaagggtgc cagggaatgc ctgtcagtgc taacagagag 1560
 atggtggcgc tgggtcttgc taatactgtg gggggctgtt tcatggcgct tcctgcgttt 1620
 ggtggctatg caagaagcaa agtcaacgct tcaactggag ctccggtctcc gatgagcagc 1680
 attttctga gcattattac ctttgtttgt atcatggtgc ttttgccgta cttatactat 1740
 cttccggtga gtctcgacc ccaatacttc cgagcgaagg ctgagaaaat atgttgcaat 1800
 aattcagaaa gccgttcttt cttctatgat atctgtcgtc gcattcagtc tcattgaaga 1860
 atgtcctcac gacgtggctt tctttatccg actgcgcgga tggacggagc tagccctaata 1920
 gcttctcatc tttgtctcga ctatcttcta ttctctagag ctgggaattg cccttggtat 1980
 tggcctttct atcttgatcc ttattcgcca ttctacgcag cctcggatcc aaattctggg 2040
 taagatagca ggcactaccg accgtttcga taacgctgaa ctccaccccc agagcgttga 2100
 gttaatcga ggcgcgctta ttgttaagat cccggaaccg ctcacctttg ccaatactgg 2160
 tgagctcaag aatcgtcttc ggcggttga attatatggc agtagccgag cgcacccttc 2220
 tcttcccccc acgcgcaccc ccgaacataa caagaatatt atatttgatg ttcattggtg 2280
 tactagcatc gatggttccg gtacgcaagt cttatatgag attgtggacg gatatgcaga 2340
 ccagggggtc agcgtcttct tctgccgcgt cgcaactcgc aatgttttcc gcatgtttga 2400
 acgaagtgga attgtggaac gatgcggtgg gataacgcac ttcgttcatg gtgtcgacga 2460
 agccctccgc cttgccgaat cggaagacga gattgaaatc tga 2503

<210> 2

<211> 770

<212> PRT

<213> *Aspergillus oryzae*

<400> 2

5

Met	Glu	Ser	Ser	Ala	Val	Gln	Glu	Pro	Thr	Gln	Gln	Arg	Ser	Leu	Arg
1				5					10					15	
Asp	Arg	Ile	Phe	Asn	Leu	Phe	Arg	Thr	Ser	Ser	Ser	Asn	Asp	Ala	Pro
			20					25					30		
Gly	Leu	Pro	Ala	Arg	Leu	Val	Thr	Ala	Glu	Ser	Ala	Ala	Gln	Asn	Glu
		35					40					45			

ES 2 579 706 T3

Gly Ser Ala Leu Ile Tyr Pro Pro Arg Glu Pro Asp Ala Arg Thr Arg
 50 55 60

Leu Leu Glu Ser Tyr Asp Arg Gly Glu Arg Gly Leu Arg Asn Ser Gly
 65 70 75 80

Val His Gly Thr Phe Ser Ser Arg Pro Glu Gln Glu Glu Ile Gln Lys
 85 90 95

Trp Asp Ala Ser Ser Leu Gln Asn Ala Gly Asn Glu Glu Arg Ser Gln
 100 105 110

Ser Pro Gly Gly Ala Asp Gly His Ile Gly Ser Pro Gly Asp Val Ser
 115 120 125

Gly Tyr Pro Gln Gly Pro Glu Asn Ile Pro Ser Leu Asp Ser Ser Phe
 130 135 140

Thr Ala Leu His Met Lys Asn His Lys Ser Leu Tyr Ile Ser Tyr Tyr
 145 150 155 160

Ile Pro Phe Phe Asn Trp Ile Thr Gln Tyr Arg Trp Ser Tyr Ile Arg
 165 170 175

Gly Asp Leu Val Ala Ala Thr Thr Ile Ala Ser Ile Tyr Ile Pro Met
 180 185 190

Ala Leu Ser Leu Ser Ser Asn Leu Ala His Ala Pro Pro Ile Asn Gly
 195 200 205

Leu Tyr Ser Phe Val Ile Asn Pro Phe Ile Tyr Ala Ile Phe Gly Ser
 210 215 220

Ser Pro Leu Leu Ile Val Gly Pro Glu Ala Ala Gly Ser Leu Leu Thr
 225 230 235 240

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

Asp Glu Val Ala Asn Ala Ile Val Val Gly Ile Ala Thr Ala Met Ala
 260 265 270

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

Asn Val Leu Ser Arg Pro Phe Leu Arg Gly Phe Ile Thr Ala Ile Gly

ES 2 579 706 T3

Ser Met Ile Ser Val Val Ala Phe Ser Leu Ile Glu Glu Cys Pro His
545 550 555 560

Asp Val Ala Phe Phe Ile Arg Leu Arg Gly Trp Thr Glu Leu Ala Leu
565 570 575

Met Leu Leu Ile Phe Val Ser Thr Ile Phe Tyr Ser Leu Glu Leu Gly
580 585 590

Ile Ala Leu Gly Ile Gly Leu Ser Ile Leu Ile Leu Ile Arg His Ser
595 600 605

Thr Gln Pro Arg Ile Gln Ile Leu Gly Lys Ile Ala Gly Thr Thr Asp
610 615 620

Arg Phe Asp Asn Ala Glu Leu His Pro Glu Ser Val Glu Leu Ile Glu
625 630 635 640

Gly Ala Leu Ile Val Lys Ile Pro Glu Pro Leu Thr Phe Ala Asn Thr
645 650 655

Gly Glu Leu Lys Asn Arg Leu Arg Arg Leu Glu Leu Tyr Gly Ser Ser
660 665 670

Arg Ala His Pro Ser Leu Pro Pro Thr Arg Thr Pro Glu His Asn Lys
675 680 685

Asn Ile Ile Phe Asp Val His Gly Val Thr Ser Ile Asp Gly Ser Gly
690 695 700

Thr Gln Val Leu Tyr Glu Ile Val Asp Gly Tyr Ala Asp Gln Gly Val
705 710 715 720

Ser Val Phe Phe Cys Arg Val Ala Thr Arg Asn Val Phe Arg Met Phe
725 730 735

Glu Arg Ser Gly Ile Val Glu Arg Cys Gly Gly Ile Thr His Phe Val
740 745 750

His Gly Val Asp Glu Ala Leu Arg Leu Ala Glu Ser Glu Asp Glu Ile
755 760 765

Glu Ile
770

<210> 3
<211> 2657
<212> ADN
<213> *Aspergillus oryzae*
5 <400> 3

ES 2 579 706 T3

atgccgggcg atctcaaaac caaaattggt cacggcgcg ccaaggcctt ggggatcaag 60
 atcccctacc gtgatcctct cggagttcat gctgaccag tcacacgagg cgagtcgatg 120
 ttctccgtcg gaacgatcga cacatactcc tatctcgagc ccgaaccac tcccgtgaa 180
 tggctgaagg aagtctgcc tagctggcat caggtgggcc gttatTTTTT caacctttc 240
 ctttctctct cgtggattac gaggtacaac ttgcaatggt tgctgggaga tatgattgcc 300
 ggtaagagcc tttccactgt gtttgatttg atcgacaagt agacaacata ctcatggaa 360
 tgcaggcgtc acggtcgggt ctgtggtcgt tccgcaggga atggcctacg ctaaactggc 420
 aaacctacct gtagagtatg gtctctattc ctcgttcattg ggtgttctca tttattggtt 480
 ttttgccacc tcaaaggata tcaccattgg tgtaagtcat tctgcacca tgtcagcatg 540
 tatcttgcta atatagtatc ttccctgttc agccgggtggc tgtcatgtct acccttacag 600
 gtaagatagt tgccgaggcg caaacgaagc tcccagatgt cgaagggcat gtaatcgctt 660
 cctgtttggc tatcatttgt ggagccgtgg tttgcgctat gggcctgctt cggctgggat 720
 ttatcgtgga tttcattcct ctgccggcaa tttcagcttt catgacgggt tccgccatca 780
 atatctgctc cggacaggtc aaagacatgc tgggagagac ggccgacttc tcgacgaaag 840
 attctaccta tctggttata atcaacaccc tcaagcatct tccctccgca aaaatcgatg 900
 ccgccatggg tgtcagtgtt ttagctatgc tgtacattat ccgttcgggt tgcaattatg 960
 gcgcgaagaa gttccccgt catgcccaagg tttggttctt cgtttcgact ttgcgcacag 1020
 tgttcgtgat cttgttctat acgatgatca gtgccgctgt gaacttgcaac cggcgggtcta 1080
 acccgcggtt caagctcctg ggtaaagttc ctcgtggttt ccaacatgag gctgtccctc 1140
 aggtaaattc gaggatcatc agcgcatttg ctagcgaact tcctgcttcg attattgtcc 1200
 tgcttatcga acacatcgct atctcgaaat cctttggccg tgtcaacaac tacacaattg 1260
 atccctctca ggagctgggt gctattggtg tgtcgaactt gcttgaccg ttccttgggtg 1320
 gttaccagc gactggatcg ttctccgaa ctgcaatcaa atcgaaagcg ggtgtccgca 1380
 ccccacttgc cgggtgttatt actgcggttg ttgtcctcct cgccatttac gctctgcccg 1440
 ctgtcttctt ttacatcccg aaagcttccc ttgctgggtg catcattcat gcagtcgggtg 1500
 acctcattac cccaccaaac accgtttacc agttctggcg cgtgtcccct ctggatgca 1560
 tcattttctt tatcgggtgt atcgtgactg tcttcaccac gattgagatc ggcatttact 1620
 gtaccgtttg tgtgtctggt gccattctgc tgttccgct cgccaaggcc cgcggtcaat 1680
 tcttaggaag agtcactatc cactcgggtga tcggtgacca tctggtacag gatgatggga 1740
 aatatgggtc tgccaactcc cctaagtctg ccagcgatga caaagatgaa ttgagccggt 1800

ES 2 579 706 T3

ctatcttctt gcctatcaac cacacggacg gatcgaatcc cgatgtcgag gtgcagcaac 1860
 cttatcctgg tatcttcac taccgattct cggaaggatt caactacccc aatgccaatc 1920
 actacaccga ttatttggtc cagactatct tcaagcatac acgtcgcaca aatccgttct 1980
 cctacggtaa accgggtgat cggccatgga ataatcctgg ccctcgcagg ggcaagtctg 2040
 aagatgacga gtcgcatttg cccttactgc aggetgtcat tcttgacttc tcatccgtca 2100
 acaatgttga tgtgacctcg gtccagaacc tcatcgatgt ccgcaatcaa ctcgacctct 2160
 acgcttcgcc taagactgtg cagtggcact ttgctcatat taacaaccgc tggacgaaac 2220
 gagcccttgc agcagcaggt ttcggcttcc catctccgga ctcgatgaa ggattccaga 2280
 gatggaagcc aattttcagc gtggctgaga tcgaaggcag tgcctctgcc gcagctcatg 2340
 cagagatggt gaacaacaga cacaccagc ataacatcaa gagcgaagac ctcgagcatg 2400
 gcctcaagca cgattcagag accaccgagc gtgagacaca cggcatcgaa gaatcctccg 2460
 atgccagcag caccggggag gacaagttgc aacgggacct gaaggatagc aaggcttacc 2520
 gcagtcgccg aagggctcgt atggtgcagg gcctcaaccg gccattcttc cacatcgacc 2580
 tgactagtgc actgcagagt gccttgcca acgcgggcca gcagccggac cctaaaatga 2640
 atgtccttga tgcatag 2657

<210> 4
 <211> 843
 <212> PRT
 <213> *Aspergillus oryzae*
 <400> 4

5

Met Pro Gly Asp Leu Lys Thr Lys Ile Gly His Gly Ala Ala Lys Ala
 1 5 10 15

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

Pro Val Thr Arg Gly Glu Ser Met Phe Ser Val Gly Thr Ile Asp Thr
 35 40 45

Tyr Ser Tyr Leu Glu Pro Glu Pro Thr Pro Ala Glu Trp Leu Lys Glu
 50 55 60

Val Cys Pro Ser Trp His Gln Val Gly Arg Tyr Phe Tyr Asn Leu Phe
 65 70 75 80

Pro Phe Leu Ser Trp Ile Thr Arg Tyr Asn Leu Gln Trp Leu Leu Gly
 85 90 95

Asp Met Ile Ala Gly Val Thr Val Gly Ala Val Val Val Pro Gln Gly

ES 2 579 706 T3

Pro Ala Ser Ile Ile Val Leu Leu Ile Glu His Ile Ala Ile Ser Lys
 355 360 365

Ser Phe Gly Arg Val Asn Asn Tyr Thr Ile Asp Pro Ser Gln Glu Leu
 370 375 380

Val Ala Ile Gly Val Ser Asn Leu Leu Gly Pro Phe Leu Gly Gly Tyr
 385 390 395 400

Pro Ala Thr Gly Ser Phe Ser Arg Thr Ala Ile Lys Ser Lys Ala Gly
 405 410 415

Val Arg Thr Pro Leu Ala Gly Val Ile Thr Ala Val Val Val Leu Leu
 420 425 430

Ala Ile Tyr Ala Leu Pro Ala Val Phe Phe Tyr Ile Pro Lys Ala Ser
 435 440 445

Leu Ala Gly Val Ile Ile His Ala Val Gly Asp Leu Ile Thr Pro Pro
 450 455 460

Asn Thr Val Tyr Gln Phe Trp Arg Val Ser Pro Leu Asp Ala Ile Ile
 465 470 475 480

Phe Phe Ile Gly Val Ile Val Thr Val Phe Thr Thr Ile Glu Ile Gly
 485 490 495

Ile Tyr Cys Thr Val Cys Val Ser Val Ala Ile Leu Leu Phe Arg Val
 500 505 510

Ala Lys Ala Arg Gly Gln Phe Leu Gly Arg Val Thr Ile His Ser Val
 515 520 525

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

Ser Pro Asn Ala Ala Ser Asp Asp Lys Asp Glu Leu Ser Arg Ser Ile
 545 550 555 560

Phe Leu Pro Ile Asn His Thr Asp Gly Ser Asn Pro Asp Val Glu Val
 565 570 575

Gln Gln Pro Tyr Pro Gly Ile Phe Ile Tyr Arg Phe Ser Glu Gly Phe
 580 585 590

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

ES 2 579 706 T3

Phe Lys His Thr Arg Arg Thr Asn Pro Phe Ser Tyr Gly Lys Pro Gly
 610 615 620
 Asp Arg Pro Trp Asn Asn Pro Gly Pro Arg Arg Gly Lys Ser Glu Asp
 625 630 635 640
 Asp Glu Ser His Leu Pro Leu Leu Gln Ala Val Ile Leu Asp Phe Ser
 645 650 655
 Ser Val Asn Asn Val Asp Val Thr Ser Val Gln Asn Leu Ile Asp Val
 660 665 670
 Arg Asn Gln Leu Asp Leu Tyr Ala Ser Pro Lys Thr Val Gln Trp His
 675 680 685
 Phe Ala His Ile Asn Asn Arg Trp Thr Lys Arg Ala Leu Ala Ala Ala
 690 695 700
 Gly Phe Gly Phe Pro Ser Pro Asp Ser Asp Glu Gly Phe Gln Arg Trp
 705 710 715 720
 Lys Pro Ile Phe Ser Val Ala Glu Ile Glu Gly Ser Ala Ser Ala Ala
 725 730 735
 Ala His Ala Glu Met Val Asn Asn Arg His Thr Gln His Asn Ile Lys
 740 745 750
 Ser Glu Asp Leu Glu His Gly Leu Lys His Asp Ser Glu Thr Thr Glu
 755 760 765
 Arg Glu Thr His Gly Ile Glu Glu Ser Ser Asp Ala Ser Ser Thr Arg
 770 775 780
 Glu Asp Lys Leu Gln Arg Asp Leu Lys Asp Ser Lys Ala Tyr Arg Ser
 785 790 795 800
 Arg Arg Arg Val Ala Met Val Gln Gly Leu Asn Arg Pro Phe Phe His
 805 810 815
 Ile Asp Leu Thr Ser Ala Leu Gln Ser Ala Leu Ala Asn Ala Gly Glu
 820 825 830
 Gln Pro Asp Pro Lys Met Asn Val Leu Asp Ala
 835 840

ES 2 579 706 T3

<210> 5
 <211> 1257
 <212> ADN
 <213> *Aspergillus aculeatus*
 5 <400> 5
 atgcacgacc acagcactgg atctagtcca tacatctcgg acgtggaaac cttgaaccac 60
 gcctgcgaga agtccgtcaa ccccgagacc aaagtctccc agcctcagga atctcccatt 120
 atcagcaata atgaacatca ggagtttgtt aagctgggca tccgccaacg gctgcgatcat 180
 ttcacctggg cctggtatac cctaaccatg agcgcaggtg gactggccct tcttctccgc 240
 aaccagccgt atcaattcaa ggggttgaag gagataggcc tgggtggtata catagccaat 300
 ctcgtcttct ttactatcat cggctctctt atgatcacca ggtttgttct ttacaacaac 360
 ctgatggact ctctccgcca cgaccgagaa ggtttcttct ttccaacctt ctggctctcc 420
 atcgccacca tgattagtgg tctatctgcc tacttctcta ctgaagacac gcaccgcctc 480
 aattatgctc tcgaggggtct cttctgggcg tactgtatct tcacgtttgc ctcagcagtg 540
 atccagtact cctttgtctt ctccatcac acgttccctc tgcaaaactat gatgccatca 600
 tggatcttac cggcattccc tatcatgctg agcggaaacca ttgcctctgc cgcttccagc 660
 taccagcctg cgggtgtctgc cacgcctatg attgttgccg gcatcacggt ccagggactc 720
 ggattctgca tcagcttcat gatgtacgcc cactacatcg ggcgtctgat ggagacgggc 780
 atcccttcga gcgagcaccg tcctgggatg ttcatctgtg tcggccccc tgccttcacg 840
 ctgctggcta tcatcggcat ggccaacggc cttcccagag gcttcagtat cctgggggat 900
 ggtggcatgg acgaccgtca catcatgcga gtactggccg tctgcgcggg catgttctc 960
 tgggctctga gcatttggtt cttctgtgtc gctctgggct cagttgtgcg ggcgcctccc 1020
 catgatttcc acctcaactg gtgggctatg gtcttcccta acaccggact cactctcgcc 1080
 accatcacc tggccaagtc actggacagt gccgcggtga aatgggtggg cgtgggcatg 1140
 tccctctgcg tgatctgcat gttcatcttc gtcttctgta gcaccattag ggctgttctc 1200
 ttgaagagga tcatgtggcc aggtcgggat gaggatgtgt ccgagttggt cgaatga 1257

<210> 6
 <211> 418
 <212> PRT
 <213> *Aspergillus aculeatus*
 10 <400> 6

ES 2 579 706 T3

Met His Asp His Ser Thr Gly Ser Ser Pro Tyr Ile Ser Asp Val Glu
1 5 10 15

Thr Leu Asn His Ala Cys Glu Lys Ser Val Asn Pro Glu Ala Lys Val
20 25 30

Ser Gln Pro Gln Glu Ser Pro Ile Ile Ser Asn Asn Glu His Gln Glu

ES 2 579 706 T3

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

Asn Gly Leu Pro Glu Gly Phe Ser Ile Leu Gly Asp Gly Gly Met Asp
 290 295 300

Asp Arg His Ile Met Arg Val Leu Ala Val Cys Ala Gly Met Phe Leu
 305 310 315 320

Trp Ala Leu Ser Ile Trp Phe Phe Cys Val Ala Leu Gly Ser Val Val
 325 330 335

Arg Ala Pro Pro His Asp Phe His Leu Asn Trp Trp Ala Met Val Phe
 340 345 350

Pro Asn Thr Gly Leu Thr Leu Ala Thr Ile Thr Leu Ala Lys Ser Leu
 355 360 365

Asp Ser Ala Ala Leu Lys Trp Val Gly Val Gly Met Ser Leu Cys Val
 370 375 380

Ile Cys Met Phe Ile Phe Val Phe Val Ser Thr Val Arg Ala Val Leu
 385 390 395 400

Leu Lys Arg Ile Met Trp Pro Gly Arg Asp Glu Asp Val Ser Glu Leu
 405 410 415

Phe Glu

- 5 <210> 7
- <211> 1430
- <212> ADN
- <213> *Aspergillus oryzae*
- <400> 7

ES 2 579 706 T3

atggtcaaag ctggtgagtt agcaatcctt aacagatgac actctcatag gtactaactc 60
 gaaacgttag cgggtacttgg agcttctggt ggcattggcc aggtatggat atccccacgc 120
 cttacaaccc tggtcacaat atgaccttgt tcgatactga ctatctccca agccactgtc 180
 tctcctggtg aagacctgtc ccttagttga agagcttgct ctctacgatg ttgtgaacac 240
 ccctggtggt gctgctgatc tatcccacat ctcgtctatc gctgtacggt actgccacaa 300
 tgogaattgc ccgatggaag agggcaaaaa tggatatctt cttacctggg cgattagaaa 360
 atctctggtt ttctgcccaa agatgatggg ctgaagcagg cccttactgg tgctaataatt 420
 gttgtcatcc cggctggtat tccccgtaag tccctaccct ttcgcattgc tcctcgtatg 480
 ttcgctggtg gccagttttc tgatagttga taggcaagcc tggtatgacc cgtgacgacc 540
 tcttcaagat caacgccggc atagtgcgag acttgggtcaa gggtatcgcc gagttctgcc 600
 ccaaggcctt tgttctggtt atctcaaacc ccgttaattc tactgttcct attgctgcag 660
 aggtgctcaa agccgctggc gtctttgacc cgaagcgcct ctttgggtgtc accacactgg 720
 acgtcgttcg tgcagagact ttcacccaag agttctcggg ccagaaggat ccttctgctg 780
 ttcaaatacc agttgttggg ggccactctg gagagaccat tgtccccctc ttcagcaaga 840
 ctacccccgc aattcagata cccgaggaga agtatgacgc actgatccac cgtaggttgt 900
 cccaaagaat ctcatgaata tcttgctgta agcactaact atgcttcagg cgtccaattt 960
 ggtggagatg aggtggtcca agctaaggac ggtgctggtt ccgccacctt gtctatggcc 1020
 tatgccggtt acaggtaggg atgctgcgta ccgtgagagc actcgcggct aacatgccat 1080
 aggttcgctg agagtgtaat caaagcttca aagggtcaaa cgggtattgt cgagcctacc 1140
 ttcgtctacc tgccctggaat tcccggcggg gatgagatcg ttaaggcaac tggcgtggaa 1200
 ttcttctcta ctcttgtaac cttaggagta agattcatct cctcacagaa tcttcgttca 1260
 tatcacgcca ggctaacgct attaaacaga ctaatggcgc agagaaggct agcaacgttc 1320
 ttgagggcgt gaccgagaag gaaaagaagc ttctcgaggc ttgcacgaaa ggccttaagg 1380
 gtaatatcga gaaaggcatc gacttcgtta agaaccacc accaaagtaa 1430

5

<210> 8
 <211> 330
 <212> PRT
 <213> *Aspergillus oryzae*
 <400> 8

10

ES 2 579 706 T3

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

Leu Ser Leu Leu Leu Lys Thr Cys Pro Leu Val Glu Glu Leu Ala Leu
 20 25 30

Tyr Asp Val Val Asn Thr Pro Gly Val Ala Ala Asp Leu Ser His Ile
 35 40 45

Ser Ser Ile Ala Lys Ile Ser Gly Phe Leu Pro Lys Asp Asp Gly Leu
 50 55 60

Lys Gln Ala Leu Thr Gly Ala Asn Ile Val Val Ile Pro Ala Gly Ile
 65 70 75 80

Pro Arg Lys Pro Gly Met Thr Arg Asp Asp Leu Phe Lys Ile Asn Ala
 85 90 95

Gly Ile Val Arg Asp Leu Val Lys Gly Ile Ala Glu Phe Cys Pro Lys
 100 105 110

ES 2 579 706 T3

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

Ala Ala Glu Val Leu Lys Ala Ala Gly Val Phe Asp Pro Lys Arg Leu
 130 135 140

Phe Gly Val Thr Thr Leu Asp Val Val Arg Ala Glu Thr Phe Thr Gln
 145 150 155 160

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

Gly Gly His Ser Gly Glu Thr Ile Val Pro Leu Phe Ser Lys Thr Thr
 180 185 190

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

Val Gln Phe Gly Gly Asp Glu Val Val Gln Ala Lys Asp Gly Ala Gly
 210 215 220

Ser Ala Thr Leu Ser Met Ala Tyr Ala Gly Tyr Arg Phe Ala Glu Ser
 225 230 235 240

Val Ile Lys Ala Ser Lys Gly Gln Thr Gly Ile Val Glu Pro Thr Phe
 245 250 255

Val Tyr Leu Pro Gly Ile Pro Gly Gly Asp Glu Ile Val Lys Ala Thr
 260 265 270

Gly Val Glu Phe Phe Ser Thr Leu Val Thr Leu Gly Thr Asn Gly Ala
 275 280 285

Glu Lys Ala Ser Asn Val Leu Glu Gly Val Thr Glu Lys Glu Lys Lys
 290 295 300

Leu Leu Glu Ala Cys Thr Lys Gly Leu Lys Gly Asn Ile Glu Lys Gly
 305 310 315 320

Ile Asp Phe Val Lys Asn Pro Pro Pro Lys
 325 330

5 <210> 9
<211> 3643
<212> ADN
<213> *Aspergillus oryzae*
<400>9

atggcggctc cgtttcgtca gcctgaggag gcggtcgatg acaccgagtt catcgatgac

60

ES 2 579 706 T3

caccatgaac acctccgtga taccgtgcac catcggttgc gcgccaattc ctccattatg 120
 cacttccaga agatcctcgt cgccaaccgt ggtgagatcc ccattcgtat cttcagaacg 180
 gccacgagc tgtcattgca gacggttgct atctactctc atgaggatcg actgtcaatg 240
 caccgtcaaa aggccgatga ggcctacatg attggccacc gcggtcagta caccctgtc 300
 ggtgcgtacc tggcgggca tgagatcatc aagatcgccc tggagcacgg tgtccagctg 360
 atccaccggg gctacggttt cttgtccgag aacgcccact tcgcccgcaa ggttgagaac 420
 gccggcattg tctttgtggg acccactccc gataccattg acagcttggg tgacaagggtg 480
 tcggcccgtc ggctggccat taagtgcgag gtccctgtcg ttccgggtac ggagggcccc 540
 gtcgagcgtc atgaggaggt caaggcgttc acagacacct atggcttccc catcatcatc 600
 aaggctgcct ttggcgggtg tggccgtggt atgcgtgtgg tccgtgacca ggccgagctg 660
 cgtgactcgt tcgagcgagc cacctctgag gcccgctccg ccttcggcaa tggtagcgtc 720
 ttcgtcgagc gcttcctcga caaacccaag cacattgaag tccagcttct gggtagacagc 780
 cacggcaacg ttgtccatct gttgagcgt gactgctccg tgcagcgtcg tcaccagaag 840
 gtcgttgagg ttgctccggc taaggacctg ccagccgatg tccgggaccg catcctggcc 900
 gatgctgtga agctggccaa gtccgtcaac taccgtaacg ccggtacagc tgagttcctg 960
 gtggaccagc agaaccgcca ctacttcatt gaaatcaatc ctcgatcca agtcgagcac 1020
 accatcaccg aagagattac tggtatcgat atcgtggctg cacagatcca gattgctgct 1080
 ggtgcaagcc tcgagcaact gggcctgact caggaccgca tctccgcccg cggatttgcc 1140
 attcaatgtc gtatcaccac ggaagatccc gccaaaggggt tctctccgga tactggtaag 1200
 attgaggttt atcgttccgc tggtggtaac ggtgtccgtc tggatggtg taacggtttc 1260
 gctggtgcta tcatcacccc tctactagac tccatgctgg tcaagtgtac ctgccgtggt 1320
 tcgacctatg aaatcgctcg tcgcaaggtt gtgcgtgcct tggtcgagtt ccgtattcgt 1380
 ggtgtgaaga ccaacattcc cttcctgact tcgcttctga gccacccgac cttcgtcgat 1440
 ggaaactgct ggaccacttt catcgacgac acccctgaat tgttctctct tgtcggcagt 1500
 cagaaccgtg cccagaagct gctcgcatc ctccggcgatg tagctgtcaa cggtagtagc 1560
 atcaagggcc aaattggcga gcccaagctc aaggggtgatg tcatcaagcc gaagcttttc 1620
 gatgccgagg gcaagccgct tgacgtttcc gccccctgca ccaagggttg gaagcagatt 1680
 ctggaccggg agggcccggc tgcctttgcg aaggccgtgc gtgccaacaa gggttgcttg 1740
 atcatggata ctacctggcg tgacgcccac cagtctttgc tggccacccg tgtgcgtacc 1800
 atcgacttgt tgaacatcgc ccatgagacc agctacgct actccaatgc gtacagtttg 1860
 gaatgctggg gtggtgctac cttcgatgtg gccatgcggt tcctctatga ggaccctgg 1920

ES 2 579 706 T3

gaccgcctgc gcaagatgcg taaggctggt cctaacatcc cattccagat gttgctccgt 1980
ggtgccaaacg gtgtcgcccta ctcttccctc ccagacaacg ccatctacca cttctgtaag 2040
caggctaaga agtgcggtgt cgacattttc cgtgttttcg acgccctcaa cgatgtcgat 2100
cagctcgagg tccggtatcaa ggctgttcat gctgccgagg gtgttgtcga ggccaccatg 2160
tgctacagcg gtgacatgct gaacccccac aagaagtaca acctggagta ctacatggcc 2220
ttggtggata agattgtagc catgaagcct cacatccttg gtatcaagga tatggccggg 2280
gtgctgaagc cccaggcccg tcgcctggtg gtgggctcca tccgtcagcg ctaccctgac 2340
cttcccatcc acgtccacac ccacgactcc gctgggtactg gtgtagcttc catgattgcc 2400
tgtgcccagg cgggtgccga cgccgtggac gccgcgaccg acagcatgtc cggtatgacc 2460
tcccagccta gcattggtgc cattctggcc tctcttgagg gcactgagca agaccccggt 2520
ctcaacctcg cccacgtgcg cgctattgat agctactggg cacagctgcg cttgctctac 2580
tctcctttcg aggcgggtct cactggcccc gaccctgagg tctacgagca cgagatccct 2640
ggtggtcagt tgaccaacct tatcttccag gccagtcagc tcggcttggg ccagcagtgg 2700
gccgaaacca agaaggccta tgaggcggct aatgatttac tcggcgacat tgtaaaggtc 2760
actcccacct ccaaggtggt cgggtgacttg gctcagttca tggctctcgaa caaactgact 2820
ccggaggatg ttgttgagcg tgctggtgag ctggacttcc ctggttctgt gctcgaattc 2880
ctcgaaggtc tcatgggaca gcccttcggt ggattccccg agccattgcg ctcccgcgcc 2940
ctgcgcgatc gccgcaagct cgagaagcgt ccaggtctct acctcgagcc tttggatttg 3000
gctaagatca agagccagat ccgtgagaag ttcggtgctg ctactgagta tgacgtggcc 3060
agctatgcc a tgtatcccaa ggtcttcgag gactacaaga agttcgtcca gaagttcggg 3120
gatctctccg tcttgcccac acggtacttc ttggccaagc ctgagattgg cgaggagttc 3180
cacgttgagc tggagaaggg taagggtgctc atcctgaagt tgttggccat cggccctctt 3240
tcagagcaga ctggtcagcg tgaggtcttc tacgaagtca acggtgaggt gcgccaggtc 3300
gctgttgatg acaacaaggc ttccgtggac aacacttcac gccctaaggc cgatgtgggt 3360
gacagcagcc aggtcgggtgc tcctatgagc ggtgtggttg ttgaaatccg tgtccacgat 3420
ggtctggagg ttaagaaggg tgaccactt gccgtcctga gtgccatgaa gatggtaagt 3480
tcattccgaa tcatttttct cactggtcaa ctacagatgc taacagctta tccaggaaat 3540
ggttatctct gctcctcaca gtggaaaggc ctccagcttg ctggtcaagg agggcgattc 3600
tgtggatggc caggatctcg tctgcaagat cgtcaaagcg taa 3643

<210> 10

<211> 1193
<212> PRT
<213> *Aspergillus oryzae*
<400> 10

ES 2 579 706 T3

Met Ala Ala Pro Phe Arg Gln Pro Glu Glu Ala Val Asp Asp Thr Glu
1 5 10 15

Phe Ile Asp Asp His His Glu His Leu Arg Asp Thr Val His His Arg
20 25 30

Leu Arg Ala Asn Ser Ser Ile Met His Phe Gln Lys Ile Leu Val Ala
35 40 45

Asn Arg Gly Glu Ile Pro Ile Arg Ile Phe Arg Thr Ala His Glu Leu
50 55 60

Ser Leu Gln Thr Val Ala Ile Tyr Ser His Glu Asp Arg Leu Ser Met
65 70 75 80

His Arg Gln Lys Ala Asp Glu Ala Tyr Met Ile Gly His Arg Gly Gln
85 90 95

Tyr Thr Pro Val Gly Ala Tyr Leu Ala Gly Asp Glu Ile Ile Lys Ile
100 105 110

Ala Leu Glu His Gly Val Gln Leu Ile His Pro Gly Tyr Gly Phe Leu
115 120 125

Ser Glu Asn Ala Asp Phe Ala Arg Lys Val Glu Asn Ala Gly Ile Val
130 135 140

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

Ser Ala Arg Arg Leu Ala Ile Lys Cys Glu Val Pro Val Val Pro Gly
165 170 175

Thr Glu Gly Pro Val Glu Arg Tyr Glu Glu Val Lys Ala Phe Thr Asp
180 185 190

Thr Tyr Gly Phe Pro Ile Ile Ile Lys Ala Ala Phe Gly Gly Gly Gly
195 200 205

Arg Gly Met Arg Val Val Arg Asp Gln Ala Glu Leu Arg Asp Ser Phe
210 215 220

Glu Arg Ala Thr Ser Glu Ala Arg Ser Ala Phe Gly Asn Gly Thr Val
225 230 235 240

ES 2 579 706 T3

Phe Val Glu Arg Phe Leu Asp Lys Pro Lys His Ile Glu Val Gln Leu
 245 250 255
 Leu Gly Asp Ser His Gly Asn Val Val His Leu Phe Glu Arg Asp Cys
 260 265 270
 Ser Val Gln Arg Arg His Gln Lys Val Val Glu Val Ala Pro Ala Lys
 275 280 285
 Asp Leu Pro Ala Asp Val Arg Asp Arg Ile Leu Ala Asp Ala Val Lys
 290 295 300
 Leu Ala Lys Ser Val Asn Tyr Arg Asn Ala Gly Thr Ala Glu Phe Leu
 305 310 315 320
 Val Asp Gln Gln Asn Arg His Tyr Phe Ile Glu Ile Asn Pro Arg Ile
 325 330 335
 Gln Val Glu His Thr Ile Thr Glu Glu Ile Thr Gly Ile Asp Ile Val
 340 345 350
 Ala Ala Gln Ile Gln Ile Ala Ala Gly Ala Ser Leu Glu Gln Leu Gly
 355 360 365
 Leu Thr Gln Asp Arg Ile Ser Ala Arg Gly Phe Ala Ile Gln Cys Arg
 370 375 380
 Ile Thr Thr Glu Asp Pro Ala Lys Gly Phe Ser Pro Asp Thr Gly Lys
 385 390 395 400
 Ile Glu Val Tyr Arg Ser Ala Gly Gly Asn Gly Val Arg Leu Asp Gly
 405 410 415
 Gly Asn Gly Phe Ala Gly Ala Ile Ile Thr Pro His Tyr Asp Ser Met
 420 425 430
 Leu Val Lys Cys Thr Cys Arg Gly Ser Thr Tyr Glu Ile Ala Arg Arg
 435 440 445
 Lys Val Val Arg Ala Leu Val Glu Phe Arg Ile Arg Gly Val Lys Thr
 450 455 460
 Asn Ile Pro Phe Leu Thr Ser Leu Leu Ser His Pro Thr Phe Val Asp
 465 470 475 480
 Gly Asn Cys Trp Thr Thr Phe Ile Asp Asp Thr Pro Glu Leu Phe Ser
 485 490 495

ES 2 579 706 T3

Leu Val Gly Ser Gln Asn Arg Ala Gln Lys Leu Leu Ala Tyr Leu Gly
 500 505 510

Asp Val Ala Val Asn Gly Ser Ser Ile Lys Gly Gln Ile Gly Glu Pro
 515 520 525

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

Lys Pro Leu Asp Val Ser Ala Pro Cys Thr Lys Gly Trp Lys Gln Ile
 545 550 555 560

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

Lys Gly Cys Leu Ile Met Asp Thr Thr Trp Arg Asp Ala His Gln Ser
 580 585 590

Leu Leu Ala Thr Arg Val Arg Thr Ile Asp Leu Leu Asn Ile Ala His
 595 600 605

Glu Thr Ser Tyr Ala Tyr Ser Asn Ala Tyr Ser Leu Glu Cys Trp Gly
 610 615 620

Gly Ala Thr Phe Asp Val Ala Met Arg Phe Leu Tyr Glu Asp Pro Trp
 625 630 635 640

Asp Arg Leu Arg Lys Met Arg Lys Ala Val Pro Asn Ile Pro Phe Gln
 645 650 655

Met Leu Leu Arg Gly Ala Asn Gly Val Ala Tyr Ser Ser Leu Pro Asp
 660 665 670

Asn Ala Ile Tyr His Phe Cys Lys Gln Ala Lys Lys Cys Gly Val Asp
 675 680 685

Ile Phe Arg Val Phe Asp Ala Leu Asn Asp Val Asp Gln Leu Glu Val
 690 695 700

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

Cys Tyr Ser Gly Asp Met Leu Asn Pro His Lys Lys Tyr Asn Leu Glu
 725 730 735

Tyr Tyr Met Ala Leu Val Asp Lys Ile Val Ala Met Lys Pro His Ile
 740 745 750

ES 2 579 706 T3

Leu Gly Ile Lys Asp Met Ala Gly Val Leu Lys Pro Gln Ala Ala Arg
755 760 765

Leu Leu Val Gly Ser Ile Arg Gln Arg Tyr Pro Asp Leu Pro Ile His
770 775 780

Val His Thr His Asp Ser Ala Gly Thr Gly Val Ala Ser Met Ile Ala
785 790 795 800

Cys Ala Gln Ala Gly Ala Asp Ala Val Asp Ala Ala Thr Asp Ser Met
805 810 815

Ser Gly Met Thr Ser Gln Pro Ser Ile Gly Ala Ile Leu Ala Ser Leu
820 825 830

Glu Gly Thr Glu Gln Asp Pro Gly Leu Asn Leu Ala His Val Arg Ala
835 840 845

Ile Asp Ser Tyr Trp Ala Gln Leu Arg Leu Leu Tyr Ser Pro Phe Glu
850 855 860

Ala Gly Leu Thr Gly Pro Asp Pro Glu Val Tyr Glu His Glu Ile Pro
865 870 875 880

Gly Gly Gln Leu Thr Asn Leu Ile Phe Gln Ala Ser Gln Leu Gly Leu
885 890 895

Gly Gln Gln Trp Ala Glu Thr Lys Lys Ala Tyr Glu Ala Ala Asn Asp
900 905 910

Leu Leu Gly Asp Ile Val Lys Val Thr Pro Thr Ser Lys Val Val Gly
915 920 925

Asp Leu Ala Gln Phe Met Val Ser Asn Lys Leu Thr Pro Glu Asp Val
930 935 940

Val Glu Arg Ala Gly Glu Leu Asp Phe Pro Gly Ser Val Leu Glu Phe
945 950 955 960

Leu Glu Gly Leu Met Gly Gln Pro Phe Gly Gly Phe Pro Glu Pro Leu
965 970 975

Arg Ser Arg Ala Leu Arg Asp Arg Arg Lys Leu Glu Lys Arg Pro Gly
980 985 990

Leu Tyr Leu Glu Pro Leu Asp Leu Ala Lys Ile Lys Ser Gln Ile Arg

ES 2 579 706 T3

995					1000					1005				
Glu	Lys	Phe	Gly	Ala	Ala	Thr	Glu	Tyr	Asp	Val	Ala	Ser	Tyr	Ala
	1010					1015					1020			
Met	Tyr	Pro	Lys	Val	Phe	Glu	Asp	Tyr	Lys	Lys	Phe	Val	Gln	Lys
	1025					1030					1035			
Phe	Gly	Asp	Leu	Ser	Val	Leu	Pro	Thr	Arg	Tyr	Phe	Leu	Ala	Lys
	1040					1045					1050			
Pro	Glu	Ile	Gly	Glu	Glu	Phe	His	Val	Glu	Leu	Glu	Lys	Gly	Lys
	1055					1060					1065			
Val	Leu	Ile	Leu	Lys	Leu	Leu	Ala	Ile	Gly	Pro	Leu	Ser	Glu	Gln
	1070					1075					1080			
Thr	Gly	Gln	Arg	Glu	Val	Phe	Tyr	Glu	Val	Asn	Gly	Glu	Val	Arg
	1085					1090					1095			
Gln	Val	Ala	Val	Asp	Asp	Asn	Lys	Ala	Ser	Val	Asp	Asn	Thr	Ser
	1100					1105					1110			
Arg	Pro	Lys	Ala	Asp	Val	Gly	Asp	Ser	Ser	Gln	Val	Gly	Ala	Pro
	1115					1120					1125			
Met	Ser	Gly	Val	Val	Val	Glu	Ile	Arg	Val	His	Asp	Gly	Leu	Glu
	1130					1135					1140			
Val	Lys	Lys	Gly	Asp	Pro	Leu	Ala	Val	Leu	Ser	Ala	Met	Lys	Met
	1145					1150					1155			
Glu	Met	Val	Ile	Ser	Ala	Pro	His	Ser	Gly	Lys	Val	Ser	Ser	Leu
	1160					1165					1170			
Leu	Val	Lys	Glu	Gly	Asp	Ser	Val	Asp	Gly	Gln	Asp	Leu	Val	Cys
	1175					1180					1185			
Lys	Ile	Val	Lys	Ala										
	1190													

<210> 11
 <211> 40
 <212> ADN
 <213> *Aspergillus oryzae*
 5 <400> 11
 gtgatagaac atcgtccata atggaatcca gcgctgtaca 40

<210> 12
 <211> 40
 <212> ADN
 <213> *Aspergillus oryzae*
 10 <400> 12
 15 gtgtcagtca cctctagtta tcagattca atctcgtctt 40

<210> 13
 <211> 20
 <212> ADN
 <213> *Aspergillus oryzae*
 20 <400> 13
 gaacaggaag aaatccaaaa 20

<210> 14
 <211> 20
 <212> ADN
 <213> *Aspergillus oryzae*
 25 <400> 14
 30 gtcggcatag ccactgcaat 20

<210> 15
 <211> 20
 <212> ADN
 <213> *Aspergillus oryzae*
 35 <400> 15
 40 tgttgccgcc aagggactta 20

<210> 16
 <211> 20
 <212> ADN
 <213> *Aspergillus oryzae*
 45 <400> 16
 ccgagagcgt tgagttaatc 20

<210> 17
 <211> 20
 <212> ADN
 <213> *Aspergillus oryzae*
 50 <400> 17
 55 agcattaggg ctagtccgt 20

<210> 18
 <211> 20
 <212> ADN
 <213> *Aspergillus oryzae*
 60 <400> 18
 ccaagatgcc atgtcaggac 20

<210> 19
 65

<211> 20
 <212> ADN
 <213> *Aspergillus oryzae*

5 <400> 19
 tcacaaaaga gtagaggcca 20

<210> 20
 <211> 36
 10 <212> ADN
 <213> *Aspergillus oryzae*

<400> 20
 15 tgtgatagaa catcgtccat aatgcacgac cacagc 36

<210> 21
 <211> 40
 <212> ADN
 <213> *Aspergillus oryzae*

20 <400> 21
 gtgtcagtca cctctagta tcattcgaac aactcggaca 40

<210> 22
 <211> 36
 25 <212> ADN
 <213> *Aspergillus oryzae*

<400> 22
 30 agaacatcgt ccataatggt caaagctggt gagtta 36

<210> 23
 <211> 40
 <212> ADN
 35 <213> *Aspergillus oryzae*

<400> 23
 gtgtcagtca cctctagta ttactttggt ggtgggttct 40

<210> 24
 <211> 36
 <212> ADN
 <213> *Aspergillus oryzae*

45 <400> 24
 tagaacatcg tccataatgg cggtccggt togtca 36

<210> 25
 <211> 40
 50 <212> ADN
 <213> *Aspergillus oryzae*

<400> 25
 55 gtgtcagtca cctctagta ttattacgct ttgacgatct 40

<210> 26
 <211> 1143
 <212> ADN
 <213> *Aspergillus oryzae*

60 <400> 26

ES 2 579 706 T3

atgctgacac ctcccaagtt tgaggatgag aagcagctgg gccccgtggg tatccgggag 60
 aggcttcgcc atttcaactg ggcctggtac acattaacga tgagtggagg agggctggcc 120
 gtcctcatca tcagccagcc ctttgggttc cgcggattga gagagatcgg catcgcgtgc 180
 tatatcctca acctgatcct ctctgccctt gtctgctcta ccatggctat aaggttcatc 240
 ctgcacggca acottctgga gtccctccgt catgaccgcg agggctctctt cttcccgacc 300
 ttctggctct ccgtcgcaac catcatctgc ggcttgtctc gctacttcgg tgaagaatcg 360
 aatgagtcct tccaactagc cctogaagcc ctcttctgga tctactgctt ctgcacotta 420
 ctcgtcgcaa tcatccaata ctcgttcgtc ttctcatccc acaagtacgg ccttcaaacc 480
 atgatgcctt catggatcct tccagccttc cccatcatgc tcagcggcac catcgcctcc 540
 gtcacggtg aacaacaacc cgctcgcgca gccctcccca tcatcggcgc cggcgtcacc 600
 ttccagggcc tcggcttctc catcagcttc atgatgtacg cccactacat cggccgactg 660
 atggagtccg gcctcccca cagcgaccac agaccaggca tgttcatctg cgtcggaccc 720
 cccgccttca cagccctcgc cctcgtcggc atgagcaaag gcctccccga agacttcaag 780
 ctgctccacg acgcccacgc cctggaagat ggccgcatca tcgagctgct ggccatctcc 840
 gccggcgtct tcctctgggc cctgagtctc tggttcttct gcatcgccat tgtcgcctc 900
 atccgctcgc cccccgaggc cttccacctc aactggtggg ccatggtctt ccccaacacc 960
 ggcttcaccc tggccacccat caccctgggc aaggctctca acagtaacgg cgtgaagggc 1020
 gtcggctccg ccatgtctat ctgcatcgtg tgcattgaca tcttcgtctt tgtcaacaat 1080
 gtccgcgccg ttatccggaa ggatatcatg taccogggta aagatgagga tgtatctgat 1140
 tag 1143

<210> 27

<211> 380

<212> PRT

<213> *Aspergillus oryzae*

<400> 27

5

ES 2 579 706 T3

Met Leu Thr Pro Pro Lys Phe Glu Asp Glu Lys Gln Leu Gly Pro Val
1 5 10 15

Gly Ile Arg Glu Arg Leu Arg His Phe Thr Trp Ala Trp Tyr Thr Leu
20 25 30

ES 2 579 706 T3

Thr Met Ser Gly Gly Gly Leu Ala Val Leu Ile Ile Ser Gln Pro Phe
35 40 45

Gly Phe Arg Gly Leu Arg Glu Ile Gly Ile Ala Val Tyr Ile Leu Asn
50 55 60

Leu Ile Leu Phe Ala Leu Val Cys Ser Thr Met Ala Ile Arg Phe Ile
65 70 75 80

Leu His Gly Asn Leu Leu Glu Ser Leu Arg His Asp Arg Glu Gly Leu
85 90 95

Phe Phe Pro Thr Phe Trp Leu Ser Val Ala Thr Ile Ile Cys Gly Leu
100 105 110

Ser Arg Tyr Phe Gly Glu Glu Ser Asn Glu Ser Phe Gln Leu Ala Leu
115 120 125

Glu Ala Leu Phe Trp Ile Tyr Cys Val Cys Thr Leu Leu Val Ala Ile
130 135 140

Ile Gln Tyr Ser Phe Val Phe Ser Ser His Lys Tyr Gly Leu Gln Thr
145 150 155 160

Met Met Pro Ser Trp Ile Leu Pro Ala Phe Pro Ile Met Leu Ser Gly
165 170 175

Thr Ile Ala Ser Val Ile Gly Glu Gln Gln Pro Ala Arg Ala Ala Leu
180 185 190

Pro Ile Ile Gly Ala Gly Val Thr Phe Gln Gly Leu Gly Phe Ser Ile
195 200 205

Ser Phe Met Met Tyr Ala His Tyr Ile Gly Arg Leu Met Glu Ser Gly
210 215 220

Leu Pro His Ser Asp His Arg Pro Gly Met Phe Ile Cys Val Gly Pro
225 230 235 240

Pro Ala Phe Thr Ala Leu Ala Leu Val Gly Met Ser Lys Gly Leu Pro
245 250 255

Glu Asp Phe Lys Leu Leu His Asp Ala His Ala Leu Glu Asp Gly Arg
260 265 270

Ile Ile Glu Leu Leu Ala Ile Ser Ala Gly Val Phe Leu Trp Ala Leu
275 280 285

ES 2 579 706 T3

Ser Leu Trp Phe Phe Cys Ile Ala Ile Val Ala Val Ile Arg Ser Pro
 290 295 300

Pro Glu Ala Phe His Leu Asn Trp Trp Ala Met Val Phe Pro Asn Thr
 305 310 315 320

Gly Phe Thr Leu Ala Thr Ile Thr Leu Gly Lys Ala Leu Asn Ser Asn
 325 330 335

Gly Val Lys Gly Val Gly Ser Ala Met Ser Ile Cys Ile Val Cys Met
 340 345 350

Tyr Ile Phe Val Phe Val Asn Asn Val Arg Ala Val Ile Arg Lys Asp
 355 360 365

Ile Met Tyr Pro Gly Lys Asp Glu Asp Val Ser Asp
 370 375 380

<210> 28

5 <211> 1182

<212> ADN

<213> *Aspergillus terreus*

<400> 28

10

ES 2 579 706 T3

atgtttgaga aactgcccc tccagggagc tcccgctccg actctggcat cctggaccat 60
 gaattcgaga agcagccggg ttccgtgggc atgctggaac gcatccgcca ttttacctgg 120
 gcctggtata ctctcacaat gagtgcctgg ggcttggccc tcctccttgg gagccagcca 180
 aacaccttca ccggcctgag ggagattgga ctgcgctgt acctgctcaa cctgctcttc 240
 tttgccctgg tctgctcgac catggccggc cggttcatcc tgcacggagg gctggctgac 300
 tctctccggc acgaacgcga gggcatcttc ttcccaacct tctggctctc gatcgccacc 360
 atcatcacag gcctgtaccg ctacttcggc gaagacgccc gacgcccctt cgtgctcgcc 420
 ctogaagccc tcttctggat ctactgcgct tgcaccctcc tcgtcgccgt catccaatac 480
 tcctggctct tctccggccc caaataccgc ctccaaaccg ccatgcccg ctggatcttc 540
 cccgccttcc ctgtcatgct ctctggcacc atcgctccg tcatcgccga gcagcagccg 600
 gcccgcccg ccatcccat catcgctgcc ggcaccacct tccagggcct gggcttctcc 660
 atcagcatga tcatgtacgc cactacgtc ggccgcctca tggagtccgg cctgccgtgc 720
 cgcgagcacc gcccgggcat gttcatcgcc gtcggcccgc cggctttcac ggcgctggcc 780
 ctcgctcgga tgaccaaggg gctcccgcac gacttcagc tcatcgccga tgacttcgcc 840
 ttcgaggatg cccgcatcct gcagctgctg gcgatcgccg tcggcgtgtt tctctgggcg 900
 ctgagtctgt ggttcttttg cattgcggcc attgccgtcg tgcgctcccc gccaacggcc 960
 ttccacctga gctgggtggc catggtcttc cccaacacgg gcttcaccct cgccacgatc 1020
 aacctgggta cggccctcaa gagcgagggt atccagggtg tggggacggc catgtcgatt 1080
 ggaattgtgt ctatcttctt gtttgtgttt atcagccatg tgcgggctgt catcaggaaa 1140
 gacattatgt atcctgggaa agacgaggat gtggtggagt aa 1182

5 <210> 29
 <211> 393
 <212> PRT
 <213> *Aspergillus terreus*

10 <400> 29

ES 2 579 706 T3

Met Phe Glu Asn Thr Ala Pro Pro Gly Ser Ser Arg Ser Asp Ser Gly
 1 5 10 15

Ile Leu Asp His Glu Phe Glu Lys Gln Pro Gly Ser Val Gly Met Arg
 20 25 30

Glu Arg Ile Arg His Phe Thr Trp Ala Trp Tyr Thr Leu Thr Met Ser
 35 40 45

Ala Gly Gly Leu Ala Leu Leu Leu Gly Ser Gln Pro Asn Thr Phe Thr
 50 55 60

Gly Leu Arg Glu Ile Gly Leu Ala Val Tyr Leu Leu Asn Leu Leu Phe
 65 70 75 80

Phe Ala Leu Val Cys Ser Thr Met Ala Gly Arg Phe Ile Leu His Gly
 85 90 95

Gly Leu Val Asp Ser Leu Arg His Glu Arg Glu Gly Ile Phe Phe Pro
 100 105 110

Thr Phe Trp Leu Ser Ile Ala Thr Ile Ile Thr Gly Leu Tyr Arg Tyr
 115 120 125

Phe Gly Glu Asp Ala Gly Arg Pro Phe Val Leu Ala Leu Glu Ala Leu
 130 135 140

Phe Trp Ile Tyr Cys Ala Cys Thr Leu Leu Val Ala Val Ile Gln Tyr
 145 150 155 160

Ser Trp Leu Phe Ser Gly Pro Lys Tyr Arg Leu Gln Thr Ala Met Pro
 165 170 175

Gly Trp Ile Leu Pro Ala Phe Pro Val Met Leu Ser Gly Thr Ile Ala
 180 185 190

ES 2 579 706 T3

Ser Val Ile Ala Glu Gln Gln Pro Ala Arg Ala Ala Ile Pro Ile Ile
 195 200 205

Val Ala Gly Thr Thr Phe Gln Gly Leu Gly Phe Ser Ile Ser Met Ile
 210 215 220

Met Tyr Ala His Tyr Val Gly Arg Leu Met Glu Ser Gly Leu Pro Cys
 225 230 235 240

Arg Glu His Arg Pro Gly Met Phe Ile Ala Val Gly Pro Pro Ala Phe
 245 250 255

Thr Ala Leu Ala Leu Val Gly Met Thr Lys Gly Leu Pro His Asp Phe
 260 265 270

Gln Leu Ile Gly Asp Asp Phe Ala Phe Glu Asp Ala Arg Ile Leu Gln
 275 280 285

Leu Leu Ala Ile Ala Val Gly Val Phe Leu Trp Ala Leu Ser Leu Trp
 290 295 300

Phe Phe Cys Ile Ala Ala Ile Ala Val Val Arg Ser Pro Pro Thr Ala
 305 310 315 320

Phe His Leu Ser Trp Trp Ala Met Val Phe Pro Asn Thr Gly Phe Thr
 325 330 335

Leu Ala Thr Ile Asn Leu Gly Thr Ala Leu Lys Ser Glu Gly Ile Gln
 340 345 350

Gly Val Gly Thr Ala Met Ser Ile Gly Ile Val Ser Ile Phe Leu Phe
 355 360 365

Val Phe Ile Ser His Val Arg Ala Val Ile Arg Lys Asp Ile Met Tyr
 370 375 380

Pro Gly Lys Asp Glu Asp Val Val Glu
 385 390

ES 2 579 706 T3

<210> 30

<211> 1317

<212> ADN

5 <213> *Schizosaccharomyces pombe*

<400> 30

atgggagaat tgaaggaaat tctcaagcag cgctaccatg aattgctcga ctggaacgtc 60
aaagcaccac acgtccctct ctgcgagagg ttgaagcatt tcacatggtc gtggttcgcg 120
10 tgtacgatgg caaccggtgg cgtcggactc atcatcggat ccttcccttt ccgattctac 180
ggactcaaca cgatcggcaa gattgtgtac atcctccaga ttttcctctt ctccctgttc 240
ggctcgtgta tgctcttcag gttcatcaag tatccgtcca caatcaagga ctccctggaac 300
catcatctcg agaaactctt cattgcgact tgtctcctct cgatttcgac attcatcgat 360
atggtggcga tctacgccta ccccgacaca ggcgagtgga tgggtgtgggt catccgaatc 420
ctctactaca tctacgtcgc ggtctccttc atttactgtg tgatggcggt cttcacgatc 480
ttcaacaacc acgtctatac cattgaaacc gcctcgcctg catggatcct ccctatcttc 540
cctccgatga tctgtgggtg cattgccggt gcggtgaaact ccaccagcc tgcgcaccag 600
ctcaaaaaca tgggtgatttt cggaatcctc ttccagggat tgggtttctg ggtctacttg 660
ctcttggtcg cagtcaacgt gctccgggtc ttcaagggtc gcttggcaaa gccccaggac 720
cgacctggca tgttcatggt cgtgggacct cctgcggtct ccggccttggc actcatcaac 780
atcgcgaggg gtgccatggg ctgcgagccg tacatcttcg tgggagcaaa ctccctcggaa 840
tacttggggt tcgtgtcgcac gttcatggcg attttcatct ggggcttggc agcatggtgt 900
tattgtctcg ccatggtgtc cttcctcgcg ggcttcttca cacgcgcacc tttgaagttc 960
gcgtgtgggt ggttcgcatt catcttcccc aacgtgggct tcgtgaactg tacgattgag 1020
atcggcaaga tgatcgactc caaagccttc cagatgttcg gccacattat cgggtgtcatc 1080
ctctgtatcc agtggatfff gctcatgtat ttgatgggtc gtgcggttctt ggtcaacgac 1140
ttgtgttatc ccggtaaaga cgaggacgcc catccgcctc ccaaaccxaa cacaggcgtc 1200
ctcaacccca ccttccctcc cgaaaaagca cctgcctccc tcgaaaaagt cgatacacat 1260
gtcacttcca ctggcggaga gtcggatcct ccgtcctccg aacacgagtc ggtctaa 1317

<210> 31

<211> 1317

<212> ADN

15 <213> *Schizosaccharomyces pombe*

<400> 31

ES 2 579 706 T3

atgggtgaac tcaaggaaat cttgaaacag aggtatcatg agttgcttga ctggaatgtc 60
aaagcccctc atgtccctct cagtcaacga ctgaagcatt ttacatggtc ttggtttgca 120
tgtactatgg caactgggtg tgttggtttg attattgggt ctttcccctt tcgattttat 180
ggtcttaata caattggcaa aattgtttat attcttcaaa tctttttggt ttctctcttt 240
ggatcatgca tgctttttcg ctttattaaa tctcttcaa ctatcaagga ttctctggaac 300
catcatttgg aaaagctttt cattgctact tgtcttcttt caatatccac gttcatcgac 360
atgcttgcca tatacgctta tcctgatacc ggcgagtgga tgggtgtgggt cattcgaatc 420
ctttattaca tttacgttgc agtatccttt atatactgcg taatggcttt ttttacaatt 480
ttcaacaacc atgtatatac cattgaaacc gcatctcctg cttggattct tcctattttc 540
cctcctatga tttgtggtgt cattgctggc gccgtcaatt ctacacaacc cgctcatcaa 600
ttaaaaaata tggttatctt tggtatcctc tttcaaggac ttggtttttg ggtttatctt 660
ttactgtttg ccgtcaatgt cttacggttt tttactgtag gcctggcaaa accccaagat 720
cgacctggtg tgtttatggt tgcgggtcca ccagctttct caggtttggc cttaattaat 780
attgcgctg gtgctatggg cagtcgccct tatatttttg ttggcgccaa ctcatccgag 840
tatcttggtt ttgtttctac ctttatggct atttttattt ggggtcttgc tgcttggtgt 900
tactgtctcg ccatggttag ctttttagcg ggctttttca ctcgagcccc tctcaagttt 960
gcttggtgat ggtttgcatt ctttttccc aacgtgggtt ttgttaattg taccattgag 1020
ataggtaaaa tgatagattc caaagctttc caaatgtttg gacatatcat tggggtcatt 1080
ctttgtattc agtggatcct cctaattgat ttaatggtcc gtgcgtttct cgtcaatgat 1140
ctttgctatc ctggcaaaga cgaagatgcc catcctccac caaaaccaa tacaggtgtc 1200
cttaacccta ctttcccacc tgaaaaagca cctgcatctt tggaaaaagt cgatacacat 1260
gtcacatcta ctggtggtga atcggatcct cctagtagtg aacatgaaag cgtttaa 1317

5 <210> 32
<211> 438
<212> PRT
<213> *Schizosaccharomyces pombe*

10 <400> 32

ES 2 579 706 T3

Met Gly Glu Leu Lys Glu Ile Leu Lys Gln Arg Tyr His Glu Leu Leu
 1 5 10 15

Asp Trp Asn Val Lys Ala Pro His Val Pro Leu Ser Gln Arg Leu Lys
 20 25 30

His Phe Thr Trp Ser Trp Phe Ala Cys Thr Met Ala Thr Gly Gly Val
 35 40 45

Gly Leu Ile Ile Gly Ser Phe Pro Phe Arg Phe Tyr Gly Leu Asn Thr
 50 55 60

Ile Gly Lys Ile Val Tyr Ile Leu Gln Ile Phe Leu Phe Ser Leu Phe
 65 70 75 80

Gly Ser Cys Met Leu Phe Arg Phe Ile Lys Tyr Pro Ser Thr Ile Lys
 85 90 95

Asp Ser Trp Asn His His Leu Glu Lys Leu Phe Ile Ala Thr Cys Leu
 100 105 110

ES 2 579 706 T3

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

Met Tyr Leu Met Val Arg Ala Phe Leu Val Asn Asp Leu Cys Tyr Pro
 370 375 380

Gly Lys Asp Glu Asp Ala His Pro Pro Pro Lys Pro Asn Thr Gly Val
 385 390 395 400

Leu Asn Pro Thr Phe Pro Pro Glu Lys Ala Pro Ala Ser Leu Glu Lys
 405 410 415

Val Asp Thr His Val Thr Ser Thr Gly Gly Glu Ser Asp Pro Pro Ser
 420 425 430

Ser Glu His Glu Ser Val
 435

- <210> 33
- <211> 1194
- <212> ADN
- <213> *Aspergillus aculeatus*
- <400> 33

ES 2 579 706 T3

atgctcgggc aacatccgcc tcccgcacacc tcctgctcgg accttacaac ataccagcat 60
gagctcaaag cctccaaata ctctagttcc accaatgtgt ctctacggga ccgtctgcgt 120
cattttacct gggcgtggta tactctgact atgagcaccg gcgggtctagc cctcctgctg 180
gccagccagc cctactcctt ctccggactg caacagatcg ggcttgcagt ctacatcatc 240
aacctggcct tctttgcggt gctgtgtagc ctcatggccg cacgcttcat tctccaaggc 300
aacttcctcg actccctccg acacgaccgc gagggctctt tctttcctac tttctggctt 360
tctattgcaa ctatcatcac cggcctgtac cgctacttcg gcgacaccac acagcctgca 420
ttcatttacg ctcttgaggt gctcttctgg ctctactgtg ccttcactct gatgaccgct 480
attatccaat actcctttgt ctttaccgcc caccactacc ctctacaaac gatgatgcc 540
tcatggatcc tccccgcatt ccctatcatg ctacagcggca cgatcgctc cgtcattgcc 600
gaacagcagc ccgcgcgctc tgctattccc atgatcgtcg ccggcaccac cttccaaggc 660
cttggcttct ccatcagttt cctcatgtac gcgcactata tcgggcggct catggagacg 720
ggccttccgt cccgggaaca ccgacccggg atgttcatct gcggttgccc cccggctttc 780
acggcccttg ccctaacgg catgaccaac ggccttcctg aggatthtca agtccttcaa 840
gacccgcacc cctttcaaga cgcgcacatc ctccgactcc ttgccatcgc cacggggcgc 900
ttcctctggg ccctcagtct ctggttcttt agcattgcc aatcgcaccac catccgcctc 960
ccacctacag ccttccacct caactgggtg gccatggttt ttccaaacac gggttttact 1020
ctcgcgacca tcacgctggg caaagccttc gatagccctg gagtcaaggg cgtcggatct 1080
gccatgtcca tttgcatcgt ggggatgtgg ctggtcgtgt ttgcgagcaa tatccgtgcc 1140
gttgtcaaac gggatattgt tttccctggg aaggacgagg atgtatcgga gtaa 1194

5

<210> 34
<211> 397
<212> PRT
<213> *Aspergillus aculeatus*

10

<400> 34

ES 2 579 706 T3

Met Leu Gly Gln His Pro Pro Pro Asp Thr Ser Cys Ser Asp Leu Thr
1 5 10 15

Thr Tyr Gln His Glu Leu Lys Ala Ser Lys Tyr Ser Ser Ser Thr Asn
20 25 30

Val Ser Leu Arg Asp Arg Leu Arg His Phe Thr Trp Ala Trp Tyr Thr
35 40 45

Leu Thr Met Ser Thr Gly Gly Leu Ala Leu Leu Leu Ala Ser Gln Pro
50 55 60

Tyr Ser Phe Ser Gly Leu Gln Gln Ile Gly Leu Ala Val Tyr Ile Ile
65 70 75 80

Asn Leu Ala Phe Phe Ala Leu Leu Cys Ser Leu Met Ala Ala Arg Phe
85 90 95

Ile Leu His Gly Asn Phe Leu Asp Ser Leu Arg His Asp Arg Glu Gly
100 105 110

Leu Phe Phe Pro Thr Phe Trp Leu Ser Ile Ala Thr Ile Ile Thr Gly
115 120 125

Leu Tyr Arg Tyr Phe Gly Asp Thr Thr Gln Pro Ala Phe Ile Tyr Ala
130 135 140

Leu Glu Val Leu Phe Trp Leu Tyr Cys Ala Phe Thr Leu Met Thr Ala
145 150 155 160

Ile Ile Gln Tyr Ser Phe Val Phe Thr Ala His His Tyr Pro Leu Gln
165 170 175

Thr Met Met Pro Ser Trp Ile Leu Pro Ala Phe Pro Ile Met Leu Ser
180 185 190

Gly Thr Ile Ala Ser Val Ile Ala Glu Gln Gln Pro Ala Arg Ser Ala
195 200 205

Ile Pro Met Ile Val Ala Gly Thr Thr Phe Gln Gly Leu Gly Phe Ser
 210 215 220

Ile Ser Phe Leu Met Tyr Ala His Tyr Ile Gly Arg Leu Met Glu Thr
 225 230 235 240

Gly Leu Pro Ser Arg Glu His Arg Pro Gly Met Phe Ile Cys Val Gly
 245 250 255

Pro Pro Ala Phe Thr Ala Leu Ala Leu Ile Gly Met Thr Asn Gly Leu
 260 265 270

Pro Glu Asp Phe Gln Val Leu Gln Asp Pro His Pro Phe Gln Asp Ala
 275 280 285

His Ile Leu Arg Leu Leu Ala Ile Ala Thr Gly Ala Phe Leu Trp Ala
 290 295 300

Leu Ser Leu Trp Phe Phe Ser Ile Ala Ile Ile Ala Thr Ile Arg Leu
 305 310 315 320

Pro Pro Thr Ala Phe His Leu Asn Trp Trp Ala Met Val Phe Pro Asn
 325 330 335

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

Pro Gly Val Lys Gly Val Gly Ser Ala Met Ser Ile Cys Ile Val Gly
 355 360 365

Met Trp Leu Phe Val Phe Ala Ser Asn Ile Arg Ala Val Val Lys Arg
 370 375 380

Asp Ile Val Phe Pro Gly Lys Asp Glu Asp Val Ser Glu
 385 390 395

<210> 35
 <211> 1194
 <212> ADN
 <213> *Aspergillus aculeatus*

ES 2 579 706 T3

<400> 35

atgctcgggc aacactcgcc tcccggcacc tcctgctcgg accttacaac ataccaacat	60
gagcttaaag cctccaaata ctctagttcc accaatgtgt ctctacggga ccgtctgcgt	120
cattttacct gggcctggta tactctgact atgagcaccg gcggcctagc gcttctgctg	180
gccagccagc cctacacctt ctccggactg caacagatcg ggcttgcagt ctatatcatc	240
aacctggtct tctttgcttt gctgtgcagc ctcatggcca cgcgcttcat tctccacggc	300
aacttcctcg actccctccg acacgaccgc gagggctctt tctttccac tttctggctt	360
tccattgcaa ctatcatcac cggactctac cgctacttcg gcgacaccac acagcctgca	420
ttcatttacg cccttgaggt gcttttctgg ctctactgtg ccttcacact gatgaccgct	480
atcatccaat actcttttgt ctttactgcc caccactacc ctctacaaac gatgatgcc	540
tcgtggatcc tccccgcatt ccccatcatg ctaagcggca cgatcgctc tgtcattgcc	600
gaacagcagc ccgcgcgctc tgctattccc atgatcgtcg ccggcaccac cttccaaggc	660
cttggcttct ccatcagttt cctcatgtac gcgcactata tcggacgcct catggagacg	720
ggccttccgt cccgggaaca ccgacccggg atgttcatct gcggttgccc ccctgctttc	780
acggcccttg ccctaatcgg catgaccaac ggccttcctg aggattttca agtccttcaa	840
gacccgcacc cctttcaaga cgcgcatatc ctccgactcc ttgccatcgc cacgggcgcc	900
ttcctctggg ccctcagtct ctggttcttc agcattgcca ttatcgccac catccgcctc	960
ccacctacgg ccttccacct caactggtgg gccatggttt ttccaaacac gggttttact	1020
ctcgcgacca tcacgctggg caaagccttc gatagccctg gagtcaaggg cgtcggatct	1080
gccatgtcca tttgcatcgt ggggatgtgg ctgttcgtgt ttgcgagcaa tatccgcgcc	1140
5 gttgtcaaac gggatattgt gtttcctggc aaggacgagg atgtatcgga gtaa	1194

<210> 36

<211> 397

<212> PRT

10 <213> *Aspergillus aculeatus*

<400> 36

ES 2 579 706 T3

Met Leu Gly Gln His Ser Pro Pro Gly Thr Ser Cys Ser Asp Leu Thr
 1 5 10 15

Thr Tyr Gln His Glu Leu Lys Ala Ser Lys Tyr Ser Ser Ser Thr Asn
 20 25 30

Val Ser Leu Arg Asp Arg Leu Arg His Phe Thr Trp Ala Trp Tyr Thr
 35 40 45

Leu Thr Met Ser Thr Gly Gly Leu Ala Leu Leu Leu Ala Ser Gln Pro
 50 55 60

Tyr Thr Phe Ser Gly Leu Gln Gln Ile Gly Leu Ala Val Tyr Ile Ile
 65 70 75 80

Asn Leu Val Phe Phe Ala Leu Leu Cys Ser Leu Met Ala Thr Arg Phe
 85 90 95

ES 2 579 706 T3

Ile Leu His Gly Asn Phe Leu Asp Ser Leu Arg His Asp Arg Glu Gly
100 105 110

Leu Phe Phe Pro Thr Phe Trp Leu Ser Ile Ala Thr Ile Ile Thr Gly
115 120 125

Leu Tyr Arg Tyr Phe Gly Asp Thr Thr Gln Pro Ala Phe Ile Tyr Ala
130 135 140

Leu Glu Val Leu Phe Trp Leu Tyr Cys Ala Phe Thr Leu Met Thr Ala
145 150 155 160

Ile Ile Gln Tyr Ser Phe Val Phe Thr Ala His His Tyr Pro Leu Gln
165 170 175

Thr Met Met Pro Ser Trp Ile Leu Pro Ala Phe Pro Ile Met Leu Ser
180 185 190

Gly Thr Ile Ala Ser Val Ile Ala Glu Gln Gln Pro Ala Arg Ser Ala
195 200 205

Ile Pro Met Ile Val Ala Gly Thr Thr Phe Gln Gly Leu Gly Phe Ser
210 215 220

Ile Ser Phe Leu Met Tyr Ala His Tyr Ile Gly Arg Leu Met Glu Thr
225 230 235 240

Gly Leu Pro Ser Arg Glu His Arg Pro Gly Met Phe Ile Cys Val Gly
245 250 255

Pro Pro Ala Phe Thr Ala Leu Ala Leu Ile Gly Met Thr Asn Gly Leu
260 265 270

Pro Glu Asp Phe Gln Val Leu Gln Asp Pro His Pro Phe Gln Asp Ala
275 280 285

His Ile Leu Arg Leu Leu Ala Ile Ala Thr Gly Ala Phe Leu Trp Ala
290 295 300

Leu Ser Leu Trp Phe Phe Ser Ile Ala Ile Ile Ala Thr Ile Arg Leu
305 310 315 320

Pro Pro Thr Ala Phe His Leu Asn Trp Trp Ala Met Val Phe Pro Asn
325 330 335

Thr Gly Phe Thr Leu Ala Thr Ile Thr Leu Gly Lys Ala Phe Asp Ser

ES 2 579 706 T3

340

345

350

Pro Gly Val Lys Gly Val Gly Ser Ala Met Ser Ile Cys Ile Val Gly
355 360 365

Met Trp Leu Phe Val Phe Ala Ser Asn Ile Arg Ala Val Val Lys Arg
370 375 380

Asp Ile Val Phe Pro Gly Lys Asp Glu Asp Val Ser Glu
385 390 395

<210> 37

<211> 1353

<212> ADN

<213> *Schizosaccharomyces japonicus*

<400> 37

5

ES 2 579 706 T3

atgtcgtcgg agacaccgac atactcctcc tgggcatccc agaggtacaa cgaattgatt 60
 gcatggaacg tcaagggtcc gaggttgccc atcgcacaga ggctcaagca cttcacgtgg 120
 tcgtggttca cgtgtactat ggcaacaggc ggtgtcggaa tgatcctcgc gtcgctcccc 180
 tataggttca cgggcttgaa caccatcgga aaggctcgtct tcattttcca ggtggtcctg 240
 ttggccatct tctgttcggc catggccttc aggttcattc gctacccgga gactttcaag 300
 aagtcgatct atcaccactt ggagaaattg ttcatcggta cattccttgc ctcgatgtcg 360
 accttcacg atatgctcgc agcctacggc tacccttcca ccggtgaatg gatggtgtac 420
 ttgatccgaa tcttctactg gatgtacttc gccgtctcct tcgtctacgc gatcttcgca 480
 ttcgcaacta ctttccatat gcatccttat accctcgaaa cggcatcgcc tgcctggatc 540
 ctcccgattt tccctgcgat gatctccgga gcagtggcag gaaccgtggc attcactcag 600
 cctccccatc agctcaaaaa cctcgtggtg tgtggcatta tgttccaggg tttgggcttc 660
 tgggtctaca tcatgttggt cgcgggtcaac atgctcaaat tgttcacaaa gggcatgatg 720
 ggagcctcgg aacgaccggg tttgttcatg ttcgtcggac ctccggcata cacaggcctc 780
 gccctcatcg gtatgggcaa gaccgccatg gattccaaaa tctccatggt ctccgccact 840
 cccgtctcct ccgaacacct cgcattcatg tgtaccttca tggcactctt catgtggggt 900
 ctcgcagcgt ggtgttattg tgtggcgatg gtctgtttcg cagcaggttt catgtccagg 960
 gcacctatcc agttcaagtt gggatggttc gcgttcatct tccctgtcgt gggcttcgtg 1020
 aacgtcacca tgaagatcgg cgagatgatt gactcggcag cttcaaaaat cttcggccac 1080
 gtcatcggag ccatgttggc catccagtgg atgttcgtga tgttcttcat ggtgcgagcg 1140
 gtcttggtgc aggaaatcat gtatcctgga cgggacgagg acgtcaaaac accgcctgga 1200
 gccacacctc ctccgaccct cgtgacctcc cctctctcct tcgcatccct ccaggatgtc 1260
 aaggatggac accccatcca ggtgacggtc tcccgcacta gggatcggtc gaaacagcac 1320
 atgtcccagg gctcggacga ggaaaagatt taa 1353

5 <210> 38
 <211> 1353
 <212> ADN
 <213> *Schizosaccharomyces japonicus*

10 <400> 38

ES 2 579 706 T3

atgtcttcgg aaacgccgac ttacagctct tgggcgagcc aacggtaaca cgaattgatc 60
gcttggaaacg tcaagggtcc tcgcttacct attgcccagc gtctaaaaca cttcacttgg 120
tcctggttta cctgtaccat ggccaccggt ggtgtcggta tgattctggc atcgctgccc 180
taccgattca caggcttaaa cacgattggt aaagtctgtg tcattttcca ggctcgtttg 240
ctggcgattt tctgttcggc gatggccttt cggttcattc gttaccccga aaccttcaaa 300
aagtctatth accatcattt ggagaagctc ttcattggta ccttcctgct ttccatgtcg 360
acgttcatcg atatgctcgc cgcctacgga taccocagca ctggcgagtg gatgggtgtac 420
ctaattcgca ttttttactg gatgtacttt gccgtctcgt tcgtatacgc catcttcgca 480
tttgctacca cttttcacat gcatccctac accctggaga cggcttcccc agcatggatt 540
ctgcctatth tcccagctat gattagcggc gctgtcggcg gtactgtggc cttcacacaa 600
ccgccgcacc aattgaagaa tttggtcgtg tgcggtatca tgttccaggg cttgggtttc 660
tgggtgtaca tcatgctgth cgcctgaac atgctcaagc tgtttacgaa gggtatgatg 720
ggtgcctctg aacgccctgg tctttttatg ttcggtggc ctccggccta taccggcttg 780
gctttaatcg gtatgggtaa aactgctatg gactccaaga tctccatgth ttctgcaacc 840
cccgtttctt ctgaacacct tgcctttatg tgtacctta tggccttgth tatgtggggt 900
cttgctgctt ggtgctattg tgtggccatg gtctgctthg ctgctggtht catgtctcgt 960
gctcctatth aattcaaact cggctggthc gcatttattt tcccagtcgt tggthttgth 1020
aacgttacta tgaagattgg tgagatgatt gattcggccg cgttcaagat ctttggthcat 1080
gtcattggtg caatgctthc cattcagthg atgthttgth tgthctthcat ggtccgcgcc 1140
gtcttactgc aagagatcat gtaccocggc cgcgacgaa atgtcaagac acctcccggt 1200
gccactcctc ctcccactth ggtgacgagth cccttgthct ttgcttcgct gcaagacgth 1260
aaagatggcc atcccattca ggtcacctgth tcccgcactc gagacagaag caaacagcac 1320
atgtcgcagg gctctgatga agaaaaatc tag 1353

<210> 39

<211> 450

<212> PRT

<213> *Schizosaccharomyces japonicus*

5

<400> 39

ES 2 579 706 T3

Met Ser Ser Glu Thr Pro Thr Tyr Ser Ser Trp Ala Ser Gln Arg Tyr
1 5 10 15

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

Gln Arg Leu Lys His Phe Thr Trp Ser Trp Phe Thr Cys Thr Met Ala
35 40 45

Thr Gly Gly Val Gly Met Ile Leu Ala Ser Leu Pro Tyr Arg Phe Thr
50 55 60

Gly Leu Asn Thr Ile Gly Lys Val Val Phe Ile Phe Gln Val Val Leu
65 70 75 80

Leu Ala Ile Phe Cys Ser Ala Met Ala Phe Arg Phe Ile Arg Tyr Pro
85 90 95

Glu Thr Phe Lys Lys Ser Ile Tyr His His Leu Glu Lys Leu Phe Ile
100 105 110

Gly Thr Phe Leu Leu Ser Met Ser Thr Phe Ile Asp Met Leu Ala Ala
115 120 125

Tyr Gly Tyr Pro Ser Thr Gly Glu Trp Met Val Tyr Leu Ile Arg Ile
130 135 140

Phe Tyr Trp Met Tyr Phe Ala Val Ser Phe Val Tyr Ala Ile Phe Ala
145 150 155 160

Phe Ala Thr Thr Phe His Met His Pro Tyr Thr Leu Glu Thr Ala Ser
165 170 175

Pro Ala Trp Ile Leu Pro Ile Phe Pro Ala Met Ile Ser Gly Ala Val
180 185 190

Ala Gly Thr Val Ala Phe Thr Gln Pro Pro His Gln Leu Lys Asn Leu
195 200 205

Val Val Cys Gly Ile Met Phe Gln Gly Leu Gly Phe Trp Val Tyr Ile
210 215 220

Met Leu Phe Ala Val Asn Met Leu Lys Leu Phe Thr Lys Gly Met Met
225 230 235 240

Gly Ala Ser Glu Arg Pro Gly Leu Phe Met Phe Val Gly Pro Pro Ala

ES 2 579 706 T3

				245					250					255		
Tyr	Thr	Gly	Leu	Ala	Leu	Ile	Gly	Met	Gly	Lys	Thr	Ala	Met	Asp	Ser	
			260					265					270			
Lys	Ile	Ser	Met	Phe	Ser	Ala	Thr	Pro	Val	Ser	Ser	Glu	His	Leu	Ala	
		275					280					285				
Phe	Met	Cys	Thr	Phe	Met	Ala	Leu	Phe	Met	Trp	Gly	Leu	Ala	Ala	Trp	
	290					295					300					
Cys	Tyr	Cys	Val	Ala	Met	Val	Cys	Phe	Ala	Ala	Gly	Phe	Met	Ser	Arg	
305					310					315					320	
Ala	Pro	Ile	Gln	Phe	Lys	Leu	Gly	Trp	Phe	Ala	Phe	Ile	Phe	Pro	Val	
				325					330					335		
Val	Gly	Phe	Val	Asn	Val	Thr	Met	Lys	Ile	Gly	Glu	Met	Ile	Asp	Ser	
			340					345					350			
Ala	Ala	Phe	Lys	Ile	Phe	Gly	His	Val	Ile	Gly	Ala	Met	Leu	Ala	Ile	
		355					360					365				
Gln	Trp	Met	Phe	Val	Met	Phe	Phe	Met	Val	Arg	Ala	Val	Leu	Leu	Gln	
	370					375					380					
Glu	Ile	Met	Tyr	Pro	Gly	Arg	Asp	Glu	Asp	Val	Lys	Thr	Pro	Pro	Gly	
385					390					395					400	
Ala	Thr	Pro	Pro	Pro	Thr	Leu	Val	Thr	Ser	Pro	Leu	Ser	Phe	Ala	Ser	
				405					410					415		
Leu	Gln	Asp	Val	Lys	Asp	Gly	His	Pro	Ile	Gln	Val	Thr	Val	Ser	Arg	
			420					425					430			
Thr	Arg	Asp	Arg	Ser	Lys	Gln	His	Met	Ser	Gln	Gly	Ser	Asp	Glu	Glu	
		435					440						445			
Lys	Ile															
																450

ES 2 579 706 T3

<210> 40

<211> 1179

<212> ADN

5 <213> *Aspergillus clavatus*

<400> 40

```

atgttcgaaa atcgtataacc gccgacctcg tctcagtcag actctggctt cctcgagAAC      60
cagctggaaa aacaacatcg actcagcctc cgtgagaggt taaggcactt tacctgggcc      120

tggtacacat tgaccatgag cacaggtggg ttggctctcc tgatagcgag ccagccatac      180

accttcaagg ggttgaagac cattggactg gtggctctaca tcgtgaactt gatcttgttt      240

ggtcttgtct gttcccttat ggccactagg ttcatcctcc acggtggctt cctcgactcc      300

cttcgccatg agcgcgaggg tcttttcttt cctaccttct ggctatccgt agcaaccatc      360

atcaccggct tgcacgccta cttcggctcc gatgctcgag aatcgtacct gattgcactc      420

gaagtactct tctgggtcta ctgtgcctgt aactggcca cagcagtgat ccagtactcc      480

ttcatcttct ctgcgcacag atacggcctc cagacatga tgcctcctg gattctccca      540

gccttcccca tcatgctcag tggcacgatt gcctccgtca tcggcgaagc tcaaccgcga      600

cggtcacga tccccgtcat catggccgga gtcaccttc agggcctggg gttctcgatc      660

agcttcatga tgtacgcca ctatatcggc cggctgatgg aatcagggct cccctgccgc      720

gagcacagac ccggcatggt catctgcgtt ggtcccccg ctttcacagc cctcgtctta      780

gtcgggatgg ccaagggcct gcccgccgag ttcaagctca tcaacgacgc acagccctc      840

gaagacgcgc ggatcctcga gctgctcgca atcaccgcgg gcatcttct ctgggccctg      900

agtctgtggt tcttcttcat cgccgtcatc gccgtcctcc ggtccccgcc tacttcttc      960

catctcaact ggtgggcctt ggtcttccc aacacgggct tcactttggc caccatcacg     1020

cttgaaagg cattgggcag tcccgggatc ttgggcgttg gttctgccat gtccttggc     1080

atcgttggca tgtggctggt tgtttttgtc agccatatcc gtgccatcat caaccaggat     1140
10 atcatgtatc cgggcaaaga tgaggatgct gcagactag                               1179

```

<210> 41

<211> 392

<212> PRT

15 <213> *Aspergillus clavatus*

<400> 41

ES 2 579 706 T3

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

ES 2 579 706 T3

His Leu Asn Trp Trp Ala Leu Val Phe Pro Asn Thr Gly Phe Thr Leu
325 330 335

Ala Thr Ile Thr Leu Gly Lys Ala Leu Gly Ser Pro Gly Ile Leu Gly
340 345 350

Val Gly Ser Ala Met Ser Leu Gly Ile Val Gly Met Trp Leu Phe Val
355 360 365

Phe Val Ser His Ile Arg Ala Ile Ile Asn Gln Asp Ile Met Tyr Pro
370 375 380

Gly Lys Asp Glu Asp Ala Ala Asp
385 390

- 5 <210> 42
<211> 1182
<212> ADN
<213> *Aspergillus fumigatus*
- 10 <400> 42

ES 2 579 706 T3

atgttcaacg atcatgatca tgttccacca acatcatcac agtcggattc tggctttttt 60
 gaacaagaaa tgaagaaatc tcctcgacta agccttcgtg agcgcctacg gcacttcacc 120
 tgggcggtgt ataccttgac gatgagtacc ggtggactgg ctcttctgat tgctagtcag 180
 ccgtatacct tcaatggcat gaagggcatc gggatggctg tttatatcct caatcttctg 240
 ttattcgctc ttgtctgttc tttgatgggt ctgagattcg ttttgcattg cggtttctct 300
 gacagcttgc gccaccctcg cgagggcttc ttcttcccta ccttctggct atccattgca 360
 acgatcatca ctggcttgca tcgttacttc ggctccgacg acctagagtc gtacctcatc 420
 gcactcgaag tcctcttctg ggtctactgt agttgcaccc tcgccacagc tgtgatccag 480
 tactcattcc tctttgccgc ccactcctac ggcctgcaga caatgatgcc atcatggatc 540
 ctaccagcct tccccatcat gctcagcggg accatcgcct cggatcatcag cgaatcccag 600
 cccgcgcgat ccgcgatccc catcatcact gccggcggtt ccttccaggg cctcggcttc 660
 tcaatcagct tcataatgta cgcccactac atcgcccgac tcatgcagtc agggcttccc 720
 tgccgcgaac acagaccagc catgttcatt tgcgtggggc ctccgtcttt caccgcgttg 780
 gcgctagtag ggatggccaa gggcctgccc gacgaattca agataatcaa agacgcacac 840
 gtcgaggacg cccggatcct cgagctgatg gctattatcg tcggcgtggt cctgtggggc 900
 ctgagtctct ggttcttctt cattgccttt gttgctgtcg tccggtgccg gccactgcg 960
 ttccacctta gctgggtggg catggctctc cccaacactg ggttcacgct ggccactatt 1020
 accctgggga gggcattggg gagccctggc gtcttgggcg tcggctcggc catgtcggtc 1080
 ggtgttgtct gcatgtgggt ctctgctttc gtctaccaca ttcgtgctgt catcaggcaa 1140
 gacatcatgt acccgggcaa agacgaggat gtgctagatt aa 1182

5

<210> 43
 <211> 393
 <212> PRT
 <213> *Aspergillus fumigatus*

10

<400> 43

ES 2 579 706 T3

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

Ser Gly Phe Phe Glu Gln Glu Met Lys Lys Ser Pro Arg Leu Ser Leu
20 25 30

Arg Glu Arg Leu Arg His Phe Thr Trp Ala Trp Tyr Thr Leu Thr Met
35 40 45

Ser Thr Gly Gly Leu Ala Leu Leu Ile Ala Ser Gln Pro Tyr Thr Phe
50 55 60

Asn Gly Met Lys Gly Ile Gly Met Val Val Tyr Ile Leu Asn Leu Leu
65 70 75 80

Leu Phe Ala Leu Val Cys Ser Leu Met Val Leu Arg Phe Val Leu His
85 90 95

Gly Gly Phe Leu Asp Ser Leu Arg His Pro Arg Glu Gly Leu Phe Phe
100 105 110

Pro Thr Phe Trp Leu Ser Ile Ala Thr Ile Ile Thr Gly Leu His Arg
115 120 125

Tyr Phe Gly Ser Asp Asp Leu Glu Ser Tyr Leu Ile Ala Leu Glu Val
130 135 140

Leu Phe Trp Val Tyr Cys Ser Cys Thr Leu Ala Thr Ala Val Ile Gln
145 150 155 160

Tyr Ser Phe Leu Phe Ala Ala His Ser Tyr Gly Leu Gln Thr Met Met
165 170 175

Pro Ser Trp Ile Leu Pro Ala Phe Pro Ile Met Leu Ser Gly Thr Ile
180 185 190

Ala Ser Val Ile Ser Glu Ser Gln Pro Ala Arg Ser Ala Ile Pro Ile
195 200 205

Ile Thr Ala Gly Val Thr Phe Gln Gly Leu Gly Phe Ser Ile Ser Phe
 210 215 220

Ile Met Tyr Ala His Tyr Ile Gly Arg Leu Met Gln Ser Gly Leu Pro
 225 230 235 240

Cys Arg Glu His Arg Pro Ala Met Phe Ile Cys Val Gly Pro Pro Ser
 245 250 255

Phe Thr Ala Leu Ala Leu Val Gly Met Ala Lys Gly Leu Pro Asp Glu
 260 265 270

Phe Lys Ile Ile Lys Asp Ala His Val Glu Asp Ala Arg Ile Leu Glu
 275 280 285

Leu Met Ala Ile Ile Val Gly Val Phe Leu Trp Ala Leu Ser Leu Trp
 290 295 300

Phe Phe Phe Ile Ala Phe Val Ala Val Val Arg Cys Arg Pro Thr Ala
 305 310 315 320

Phe His Leu Ser Trp Trp Ala Met Val Phe Pro Asn Thr Gly Phe Thr
 325 330 335

Leu Ala Thr Ile Thr Leu Gly Arg Ala Leu Gly Ser Pro Gly Val Leu
 340 345 350

Gly Val Gly Ser Ala Met Ser Val Gly Val Val Cys Met Trp Val Phe
 355 360 365

Val Phe Val Tyr His Ile Arg Ala Val Ile Arg Gln Asp Ile Met Tyr
 370 375 380

Pro Gly Lys Asp Glu Asp Val Leu Asp
 385 390

<210> 44
 <211> 1294
 <212> ADN

ES 2 579 706 T3

<213> *Aspergillus oryzae*

<400> 44

atgttcgctg ctcgccagtc tttcaacctc ctccagaagc ggccttctc cgcctctgcc 60
agccaggtgt gtgattgaat ggatccattg gacctcggag ctagctctgc aacatcaaca 120
aaactaacat actaacttat cttcttcata ggcttccaag gttgccgttc ttggtgccgc 180
tggtggcatt ggccagcctc tctcccttct cctcaagctc aacccccgtg tttctgagct 240
5 tgcctctac gatatccgcg gtggccctgg tatgtttttg cacagcttgc aacatctccg 300
acttcggtga ttcaagacag ggctaacata aggatacaat aggtgttgcc gctgacctga 360
gccacatcaa caccaacagc accgtctctg gctacgaggc taccctctct ggcctccgtg 420
atgctctcaa gggctccgag atcgtcctca tccctgccg tgttcctcgc aagcccggca 480
tgacctctga cggtatgaac cgttaacttg tcaatggcac tgggaattga atactaatta 540
taatatcgcc agacctgttc aacaccaacg cctccattgt ccgcgacctt gctaaggccg 600
ccgccgaggc ttcccccgag gccaacatcc tcgtcatctc caaccctgta tgacgctttc 660
caccactgc taccagttat ctccgctaa ttgcaatcag gtcaactcca ccgtccccat 720
cgtctctgag gtcttcaagt ccaagggtgt ctacaacccc aagcgtctct tcggtgtcac 780
tacccttgac gttgtccgtg cctctcgtt catctcccag gtccagaaga ccgaccctc 840
caacgaggcc gtcactgtcg tcggtggta ctccggtgtg accattgtcc ctcttctctc 900
ccagtccagc caccocagca ttgagggtaa gaccocgat gagctcgtca accgcatcca 960
gttcggtggt gatgaggtg tcaaggccaa ggatggtgct ggctctgcca ccctctccat 1020
ggccatggct ggtgctcgca tggctgagtc cctcctgaag gccgcccagg gtgagaaggg 1080
tgtcgttgag cccactttcg tcgacagccc tctctacaag gaccagggtg ttgacttctt 1140
cgcctccaag gtcgagctcg gcccacagc tgttgagaag atcctccccg ttggccaggt 1200
caacgcctac gaggagaagc tcctcgaggc ctgccttggg gacctcaaga agaacatcca 1260
gaagggtatt gacttcgtca aggccaaccc ttaa 1294

<210> 45

<211> 340

<212> PRT

<213> *Aspergillus oryzae*

<400> 45

10

15

ES 2 579 706 T3

Ala Thr Pro Ser Gly Leu Arg Asp Ala Leu Lys Gly Ser Glu Ile Val
85 90 95

Leu Ile Pro Ala Gly Val Pro Arg Lys Pro Gly Met Thr Arg Asp Asp
100 105 110

Leu Phe Asn Thr Asn Ala Ser Ile Val Arg Asp Leu Ala Lys Ala Ala
115 120 125

Ala Glu Ala Ser Pro Glu Ala Asn Ile Leu Val Ile Ser Asn Pro Val
130 135 140

Asn Ser Thr Val Pro Ile Val Ser Glu Val Phe Lys Ser Lys Gly Val
145 150 155 160

Tyr Asn Pro Lys Arg Leu Phe Gly Val Thr Thr Leu Asp Val Val Arg
165 170 175

Ala Ser Arg Phe Ile Ser Gln Val Gln Lys Thr Asp Pro Ser Asn Glu
180 185 190

Ala Val Thr Val Val Gly Gly His Ser Gly Val Thr Ile Val Pro Leu
195 200 205

Leu Ser Gln Ser Ser His Pro Ser Ile Glu Gly Lys Thr Arg Asp Glu
210 215 220

Leu Val Asn Arg Ile Gln Phe Gly Gly Asp Glu Val Val Lys Ala Lys
225 230 235 240

Asp Gly Ala Gly Ser Ala Thr Leu Ser Met Ala Met Ala Gly Ala Arg
245 250 255

Met Ala Glu Ser Leu Leu Lys Ala Ala Gln Gly Glu Lys Gly Val Val
260 265 270

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

Phe Phe Ala Ser Lys Val Glu Leu Gly Pro Asn Gly Val Glu Lys Ile
290 295 300

Leu Pro Val Gly Gln Val Asn Ala Tyr Glu Glu Lys Leu Leu Glu Ala
305 310 315 320

Cys Leu Gly Asp Leu Lys Lys Asn Ile Gln Lys Gly Ile Asp Phe Val
325 330 335

Lys Ala Asn Pro
340

5 <210> 46
<211> 20
<212> ADN
<213> *Aspergillus oryzae*

10 <400> 46
gacctcca cgctgaccac 20

15 <210> 47
<211> 20
<212> ADN
<213> *Aspergillus oryzae*

20 <400> 47
ggctgagaaa atatgtgca 20

25 <210> 48
<211> 30
<212> ADN
<213> *Aspergillus oryzae*

30 <400> 48 30
gatagaccac taatcatggt gccgatggag 30

35 <210> 49
<211> 30
<212> ADN
<213> *Aspergillus oryzae*

40 <400> 49
tgcggtcctg agtcaggccc agttgctcga 30

45 <210> 50
<211> 20
<212> ADN
<213> *Aspergillus oryzae*

50 <400> 50
gggatttga cagcagaagg 20

55 <210> 51
<211> 20
<212> ADN
<213> *Aspergillus oryzae*

60 <400> 51
tcacaaaaga gtagaggcca 20

65 <210> 52
<211> 36
<212> ADN
<213> *Aspergillus clavatus*

<400> 52
ccaacagaca catctaaaca atgtccgaca aggctc 36

70 <210> 53
<211> 40
<212> ADN
<213> *Aspergillus clavatus*

75 <400> 53

ES 2 579 706 T3

gtgtcagtca cctctagtta tcagctcttg gtgatattgt 40

<210> 54

<211> 750

5 <212> ADN

<213> *Aspergillus clavatus*

<400> 54

atgtccgaca aggctcgcga ggcgggtctcg caatatctga agcagttctca cgagcgcattc 60

ttcgagaaca acccgcgctg ggtcgcggcc aagaaggaag aggacccccgc gtttttcgag 120

aaactgggcg ccggacagac gccgcaatat ctgtaaagac cagcatcagc gatttgtcat 180

tggagttgtc ggatgacttt gccatactga ccgtggctgt gcaggtacat cggatgcagt 240

gacagtcgcg tgcccgcaa tgacattatg ggtctcacgg ccggcggagg ttttgtgcac 300

cgcaacatcg ccaatctggt gcccaacacc gacctcaatg tcatgtcggc catcaactac 360

gccgtccggc atctgaaggt caagcacatc gttgtctgcg gccactaaa ctgtggcggc 420

gtcaaggctg cgctgacgcc ctccgacctg gggctgctga acccctggct gcgcaatgtc 480

cgggatgtgt atcggttgca cgagcgcgag ctggacgcca tcgaagacga agaggcgaag 540

tataatcgcc tgggtggagct gaatgttgtt gagtcctgcc gcaacgtcat caagacggcg 600

gcgggtgcagc agagctacca cgacaaccag ttccccgtgg tccacggatg gatctttgat 660

gtgocggacgg gtctgcttcg ggatctcaac attgatttcg aggagacgct gcgggatatc 720

10 aagaagatct acaatatcac caagagctga 750

<210> 55

<211> 225

<212> PRT

15

<213> *Aspergillus clavatus*

<400> 55

ES 2 579 706 T3

Met Ser Asp Lys Ala Arg Glu Ala Val Ser Gln Tyr Leu Lys Gln Ser
1 5 10 15

His Glu Arg Ile Phe Glu Asn Asn Arg Ala Trp Val Ala Ala Lys Lys
20 25 30

Glu Glu Asp Pro Ala Phe Phe Glu Lys Leu Gly Ala Gly Gln Thr Pro

ES 2 579 706 T3

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

Ser
225
<210> 56

<211> 37
 <212> ADN
 <213> *Aspergillus oryzae*

5 <400> 56
 attaactaaa gaccacgaga ggggaactag ggaaatc 37

<210> 57
 <211> 28
 10 <212> ADN
 <213> *Aspergillus oryzae*

<400> 57
 15 ttaattaact ttcccccccg taatctaa 28

<210> 58
 <211> 45
 <212> ADN
 <213> *Aspergillus oryzae*

20 <400> 58
 ggaaacagct atgacatga ttatggattg ttaaactgc gacgc 45

<210> 59
 <211> 45
 <212> ADN
 <213> *Aspergillus oryzae*

25 <400> 59
 30 gcgtcgacgt taaacaatc cataatcatg gtcatactg ttcc 45

<210> 60
 <211> 38
 <212> ADN
 35 <213> *Aspergillus oryzae*

<400> 60
 ggaaacagct atgacatga ttccagattg taaattac 38

40 <210> 61
 <211> 40
 <212> ADN
 <213> *Aspergillus oryzae*

45 <400> 61
 agcactagta cgcgtagatc tgtttagatg tgtctgttg 40

<210> 62
 <211> 3228
 50 <212> ADN
 <213> Secuencia artificial

<220>
 <223> SECUENCIA DE ADN ARTIFICIAL

55 <400> 62

ES 2 579 706 T3

attaactaaa gaccacgaga ggggaactag ggaaatccgt gatagaatga agtatacagt	60
tgatthttgtc agcaagggat tcaaacctaa taccagaggt gactttatcc aagattcctt	120
cacgacactg gaagaactgc ttgaagagct gcctgagtca atcagcttca atatcgagat	180
aagtaggta tatgctgcca ctggtggttc tccacttgcg ggagacaaag ctaacaacgt	240
cccaatgaag agtaccaccag gcttcatgaa gctatagaag caggtgtagc accagtggct	300
attgaaatca acaccttcat cgacaaagcg cttgagagac tcttttctta cggcaacaaa	360

ES 2 579 706 T3

aaacggacca ttatcctatc ctcatttact cccgagatct gcattttatt ggccatcaaa 420
 caacagacgt accctgtgat gttcatcact aatgccggca agcctccagt tacggatcga 480
 gagatgaggg ctgccagcat acagtccgct gttcgatttg ccaagaggtg gaatttatct 540
 ggccttgtct ttgcatctga ggcgctggta atgtgccccca ggcttgtcag atatgttcaa 600
 cgatcaggat tgatctgtgg atcctatgga tctcagaaca atataccaga aaatgcgaag 660
 gtaagtgctt ctatattgat ccttagtgct ttcaaaactgt gatgtagaag ttgctcggta 720
 gctgattaaa tattctagac ccaagccgct gctggaattg acattattat ggccgatagg 780
 gttgggctta ttgctatgtc cctgaaagga tatcaaaagc aggcaaaaag ccaggcataa 840
 tccccgcgtg gacggtagcc taaggatagg ccctaattct atctacatgt gactgcatcg 900
 atgtgtttgg tcaaaatgag gcatgtggct caccacacag gcggagaaac gtgtggctag 960
 tgcatgacag tcccctccat agattcaatt taatttttcg cggcaattgt cgtgcagttt 1020
 gtatctacat ttcattccat atatcaagag ttagtagttg gacatcctga ttattttgtc 1080
 taattactga aaactcgaag tactaaccta ctaataagcc agtttcaacc actaagtgct 1140
 catttataca atatttgtag aaccccgcg c taccctcca tcgccaacat gtaagtagtg 1200
 gtggatacgt actcctttta tggcagtagc tcgcaagtag gatgtagttt ataaattcag 1260
 cactcgaat gactactact atgtgtctac gacagatacc ctctccgtac gaataagaca 1320
 cctgcctcga tatatggaca aattcaaaat cagggtcaag ggtcatgttt caaagtcaca 1380
 acaatctcca acatagacga gaatttgtac cggagtgctc gaaggtgcag ctggagattg 1440
 gtctattttc ttagagtggg gtatcactaa tgtacagtcg gtcactatcg tacaacaat 1500
 cacaattata tacaagattt cccaccaccc cctactctaa cacggcaca ttatccatcg 1560
 agtcagagcc tagccaccat ttggtgctct cgtagagacc aaagtataat cctgatccga 1620
 cagcggccat aaacgtgttg atagcacacc ctcggaatag tcctctcggg ccatctgttc 1680
 gtacaatctc ccgtacggta ttgatcatcc ttttcttctg aggtgcagtt gtatctgcag 1740
 catcgagcat gattcgtgtc cggaccatat ccatgggtgc tgtcaagaca ctagctatac 1800
 cgcccagac cgcagcactt attgcggtg tcgctgcagc ctctccgatt gtcgaatggg 1860
 cctctttctt tccatactct cttggtcttt ctagcacctt ctctcgatct ccgaatctat 1920
 attcaaaaat tcgataccga aaagactcgt acagaggcat ctgaatcgcc gacactggca 1980
 agctatgogc cacaagagcc gggatccgc tccaaagctg tctaggggtg ataaacttct 2040
 tgaaagctag ccgtgtcgtc ttctgggcta caccacctac ccttccccca gctacaggtg 2100
 ctgatgcgtc tggatgggtg gattggatca tctgcgcggt gtgttttaat gcatcagccg 2160
 gagcaaagac tccgcaagca gcaagatccg caacggaggg tcgcgcaaaaa tcggaaaaga 2220
 gccgagctga gctagactca tgcgttccaa gtttttgatg tatgacttgg agtcctgact 2280

ES 2 579 706 T3

gtgcatactc gtatgtgatg aagaatgcg cgcctttgac cgttatttt ttgctcaatg 2340
aatatgagtt gggtaaggat aatgttgggt atcaccaacc tgtgggaaat gaggcagcgg 2400
ttacgctcgc gataccttgg taaaggccgc ggaatatccc tgggtgtctc catatgctgg 2460
ttcctgtggt ggttctcagg aactgcgaat attcgcgaga ctgaatgcgg gttttgatcg 2520
tatctaaggg atatgtgaat agatctagt agagggatgc ggttgtgctt gcctagttcc 2580
ctcaggtaa tgtatgctgt ttgtccaagg gggtagaaa atataccagc aatatactaa 2640
gtggctcggc ggccatatta tttgtcgata gaggcgggag gacgaggatg ttcaatcgtg 2700
gattttgtaa ccgtcctggc acacatcacc ttacaaggct ttgagttttc gttctataat 2760
ttgagacagg tcacagtcga ttgccatatt agattacagc aacgtaggta gacattagga 2820
ttgatgccag gcagcgaagt ttttttact attggatgag tttattccgg gctccgcggg 2880
aaccgattgg ggtctgacat ttctggaaa tccccaaaca tatttcattt tgtctttgcg 2940
gacagataaa acgctttcga ttccgatagt agaggtatat acacatcgag tgtaaacagt 3000
ataagcatat agtctctcgg tttgtagcg gttttgacgc atactaataa atcatcgata 3060
actagtactg ctatcgataa gctccttgac ggggaaagg agcaggacc actgcagtca 3120
ctttagtaag tggatgtcag gatgataagg agatatagag atatgcagct gataagctta 3180
gaacgtgacc tgaaggaaga cattaaaatt agattacggg ggggaaag 3228

5 <210> 63
<211> 41
<212> ADN
<213> Secuencia artificial

10 <220>
<223> CEBADOR DE ADN ARTIFICIAL

<400> 63
gtgatagaac atcgccata atgcccggcg atctcaaac c 41

15 <210> 64
<211> 39
<212> ADN
<213> Secuencia artificial

20 <220>
<223> CEBADOR DE ADN ARTIFICIAL

<400> 64
25 gtgcagtca cctctagta ctatgcatca aggacattc 39

<210> 65
<211> 20
<212> ADN
30 <213> Secuencia artificial

<220>
 <223> CEBADOR DE ADN ARTIFICIAL

 <400> 65
 5 gattgagatc ggcatttact 20

 <210> 66
 <211> 20
 <212> ADN
 10 <213> Secuencia artificial

 <220>
 <223> CEBADOR DE ADN ARTIFICIAL

 <400> 66
 15 acgcggaaca gcagaatggc 20

 <210> 67
 <211> 25
 20 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> CEBADOR DE ADN ARTIFICIAL
 25
 <400> 67
 ctatagcgaa atggattgat tgtct 25

 <210> 68
 <211> 20
 <212> ADN
 <213> Secuencia artificial

 <220>
 <223> CEBADOR DE ADN ARTIFICIAL
 35
 <400> 68
 ttcaccgtga aacgtattga 20

 40

REIVINDICACIONES

1. Célula huésped recombinante que comprende un polinucleótido heterólogo que codifica un polipéptido que tiene actividad de anhidrasa carbónica, donde la célula huésped es capaz de producir una cantidad superior de ácido C4-dicarboxílico en comparación con la célula huésped sin el polinucleótido heterólogo cuando se cultiva bajo las mismas condiciones; donde la célula huésped recombinante es una célula huésped fúngica; y donde el polinucleótido heterólogo:
- 5 (a) codifica una anhidrasa carbónica que tiene al menos 80% de identidad de secuencia con SEC ID n.º: 55;
- 10 (b) hibridiza bajo condiciones de astringencia media con (i) la cadena complementaria en toda su longitud de SEC ID n.º: 54, o (ii) la cadena complementaria en toda su longitud de la secuencia de ADNc de SEC ID n.º: 54; o
- (c) tiene al menos 80% de identidad de secuencia con SEC ID n.º: 54 o la secuencia de ADNc de SEC ID n.º: 54.
2. Célula huésped recombinante según la reivindicación 1, donde el polinucleótido heterólogo codifica una anhidrasa carbónica que tiene al menos 90% de identidad de secuencia con SEC ID n.º: 55.
- 15 3. Célula huésped recombinante según la reivindicación 1 o 2, donde el polinucleótido heterólogo codifica una variante de anhidrasa carbónica de SEC ID n.º: 55, donde la variante comprende una sustitución, eliminación y/o inserción en no más de 10 posiciones.
- 20 4. Célula huésped recombinante de cualquiera de las reivindicaciones 1-3, donde el polinucleótido heterólogo codifica una anhidrasa carbónica que comprende o consiste en SEC ID n.º: 55.
5. Célula huésped recombinante de cualquiera de las reivindicaciones 1-4, que comprende además un polinucleótido heterólogo que codifica un transportador de bicarbonato.
- 25 6. Célula huésped recombinante de cualquiera de las reivindicaciones 1-5, que comprende además un polinucleótido heterólogo que codifica un transportador de ácido C4-dicarboxílico.
7. Célula huésped recombinante de cualquiera de las reivindicaciones 1-6, que comprende además un polinucleótido heterólogo que codifica una malato deshidrogenasa.
- 30 8. Célula huésped recombinante de cualquiera de las reivindicaciones 1-7, que comprende además un polinucleótido heterólogo que codifica una piruvato carboxilasa.
- 35 9. Célula huésped recombinante de cualquiera de las reivindicaciones 1-8, donde la célula huésped es una célula huésped fúngica filamentosa.
10. Célula huésped recombinante según la reivindicación 9, donde la célula huésped se selecciona del grupo que consiste en una *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Rhizopus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes* y *Trichoderma*.
- 40 11. Célula huésped recombinante según la reivindicación 10, donde la célula huésped es una célula huésped de *Aspergillus*, tal como una célula huésped de *Aspergillus oryzae*.
12. Célula huésped recombinante de cualquiera de las reivindicaciones 1-11, donde el ácido C4-dicarboxílico es ácido málico.
- 50 13. Célula huésped recombinante de cualquiera de las reivindicaciones 1-12, donde la célula huésped es capaz de producir una cantidad superior de ácido C4-dicarboxílico de por lo menos el 10%, preferiblemente por lo menos el 50%, en comparación con la célula huésped sin el polinucleótido heterólogo que codifica la anhidrasa carbónica, cuando se cultiva bajo las mismas condiciones.
- 55 14. Método para producir un ácido C4-dicarboxílico, que comprende:
- (a) cultivo de la célula huésped recombinante de cualquiera de las reivindicaciones 1-13 en un medio bajo condiciones adecuadas para producir el ácido C4-dicarboxílico; y
- (b) recuperación del ácido C4-dicarboxílico.
- 60 15. Método según la reivindicación 14, donde el ácido C4-dicarboxílico es ácido málico.

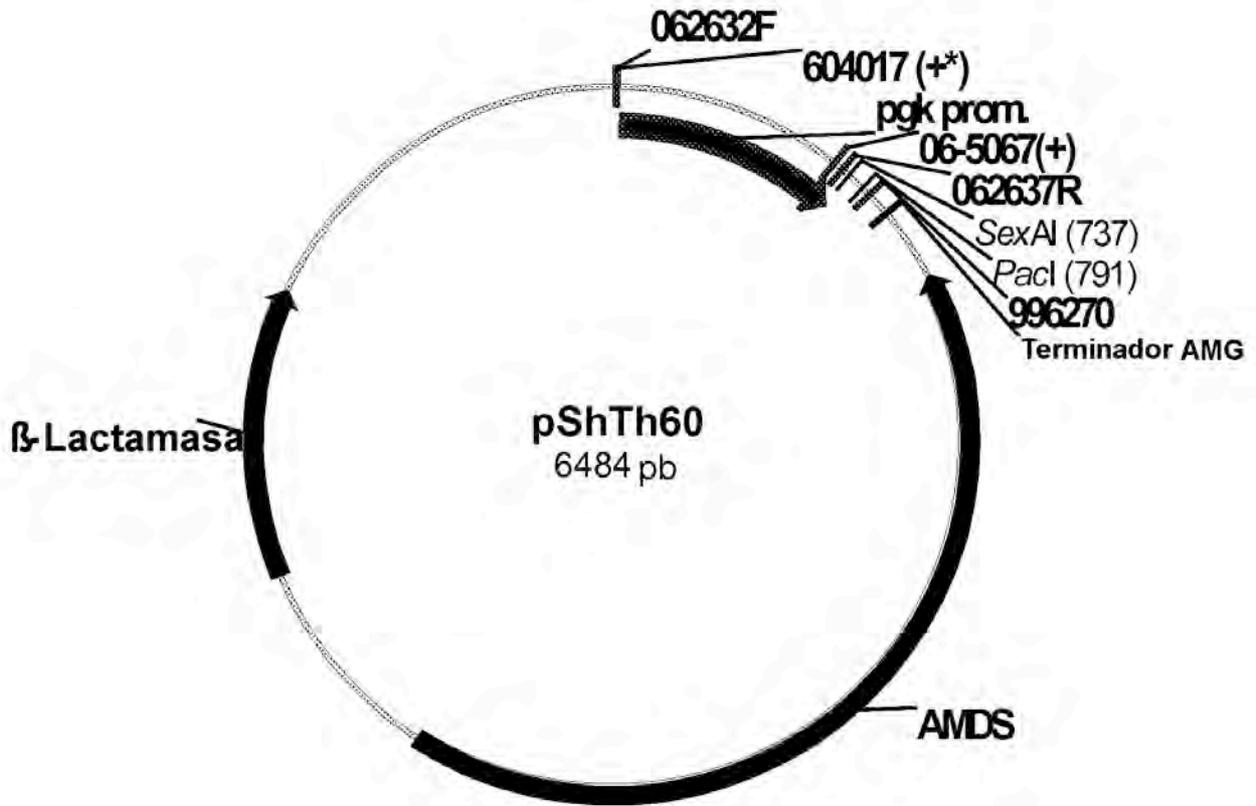


Fig. 1

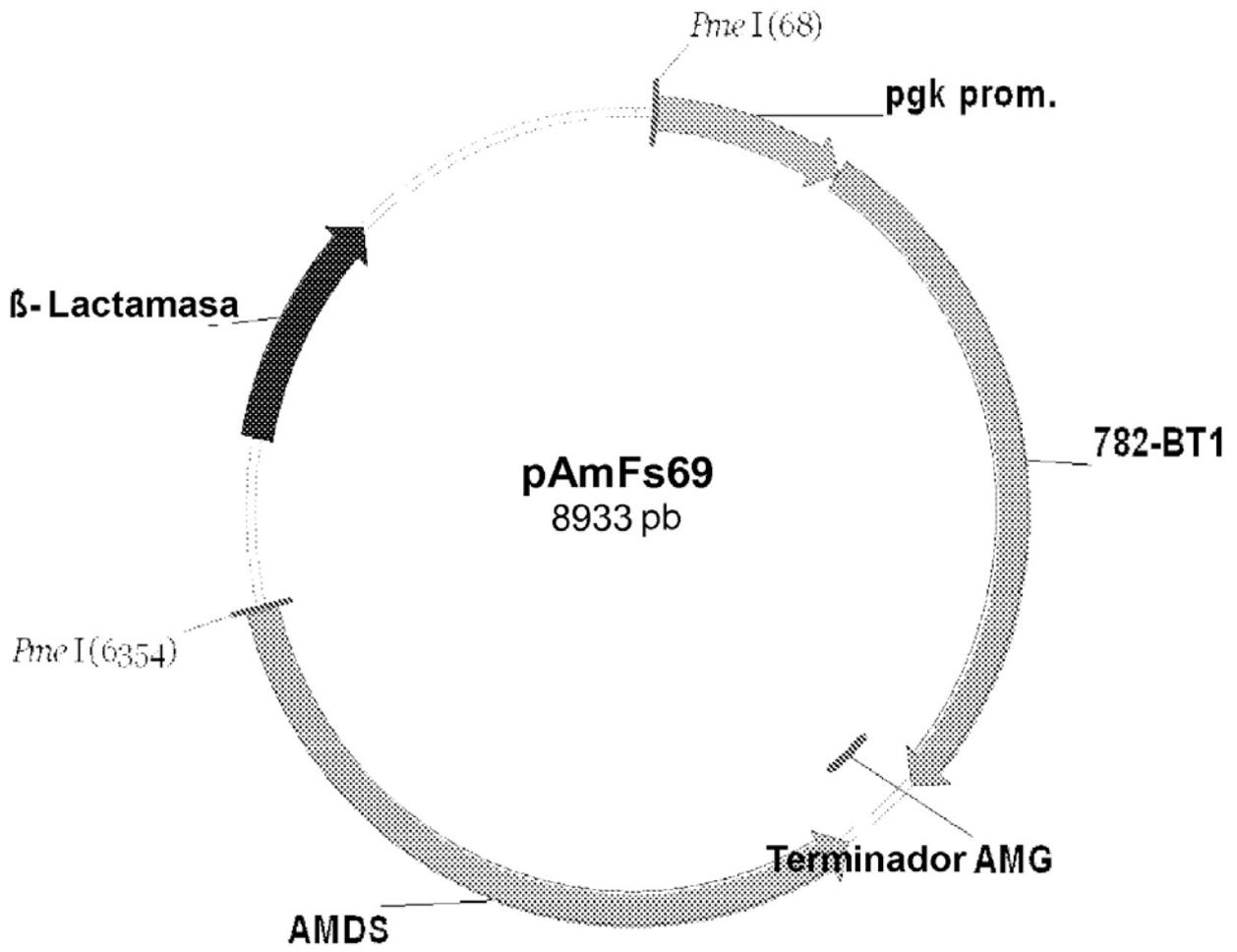


Fig. 2

M E S S A V Q E P T Q Q R S L R D R I F
 1 ATGGAATCCA GCGCTGTACA GGAGCCGACT CAACAGCGCT CTTTGCGGGA TCGCATTTTT
 N L F R T S S S N D A P G L P A R L V T
 61 AACCTCTTTC GTACCTCTTC CTCAAATGAT GCCCCGGGTC TTCCGGCAAG ACTCGTAACC
 A E S A A Q N E G S A L I Y P P R E P D
 121 GCTGAGAGCG CAGCGCAAAA CGAAGGGTCG GCGTTAATCT ATCCGCCACG GGAGCCTGAT
 A R T R L L E S Y D R G E R G L R N S G
 181 GCAAGGACTC GTCTTCTCGA ATCGTACGAT CGCGGGGAAC GTGGTCTGAG GAACTCCGGC
 V H G T F S S R P E Q E E I Q K W D A S
 241 GTTCATGGGA CTTTTTCTTC ACGACCTGAA CAGGAAGAAA TCCAAAAATG GGATGCAAGC
 S L Q N A G N E E R S Q S P G G A D G H
 301 TCTTTGCAGA ATGCTGGTAA CGAAGAAAGA TCTCAGTCCC CAGGAGGAGC AGACGGCCAT
 I G S P G D V S G Y P Q G P E N I P S L
 361 ATTGGGTCTC CCGGCGACGT CTCAGGATAC CCACAGGGAC CAGAGAATAT ACCATCGCTA
 D S S F T A L H M K N H K S L
 421 GACTCCTCTT TCACAGCATT GCACATGAAG AATCATAAAT CTCTGTAGGT TTATAATCAC

 481 GTTCGCCCTG CTTTCTAACA CAATTGTTAT CTCCATCGTG GAGACAAC TA ACGTTCATCA
 Y I S Y Y I P F F N W I T Q Y R W S Y
 541 AGGTATATAT CTTACTACAT CCCATTTTTT AATTGGATTA CTCAATACCG GTGGTCGTAC
 I R G D L V A A T T I A S I Y I P M A L
 601 ATTCGAGGTG ATTTGGTTGC TGCACAACC ATTGCGTCCA TCTATATCCC TATGGCTTTG
 S L S S N L A H A P P I N G L Y S F V I
 661 TCCTTATCCT CAAATCTCGC CCACGCACCT CCTATCAATG GCCTCTACTC TTTTGTGATC
 N P F I Y A I F G S S P L L I V G P E A
 721 AACCCTTTCA TCTATGCGAT CTTCCGGGAGC AGCCCGCTGT TAATAGTGGG CCCAGAAGCA
 A G S L L T G T I V K T S V R P G P S G
 781 GCAGGCTCCT TGCTTACTGG CACGATTGTC AAAACTAGTG TCAGACCAGG CCCATCTGGT
 E D D E V A N A I V V G I A T A M A G A
 841 GAGGACGACG AAGTAGCGAA TGCCATCGTG GTCGGCATAG CCACTGCAAT GCGGGGCGCC
 M I L I A G L T R L G F L D N V L S R P
 901 ATGATACTGA TCGCTGGGCT TACACGGCTG GGATTCTGG ACAATGTGCT GAGCCGGCCC
 F L R G F I T A I G F V I F V D Q L I P
 961 TTTCTTAGGG GTTTCATTAC AGCGATCGGT TTTGTGATTT TTGTGGATCA ACTCATCCCC
 E V G L T E L A K E A G V T H G T T V D
 1021 GAAGTCGGAT TGACCGAGCT AGCAAAGGAA GCTGGTGTTA CCCATGGGAC TACAGTTGAC
 K L M F L I R N I G G C H A L T T A V A
 1081 AAGCTCATGT TCCTTATAAG AAACATAGGA GGTGCCATG CGCTTACAAC CGCGGTGGCT
 F G S F A I I M V F R
 1141 TTTGGGAGCT TTGCTATTAT AATGGTATTT CGGTTAGTGT TGGTGACTCG GAAGCCTGGT
 T L K K M L Q P R Y P Q
 1201 GCTTAGACTG ATTACCATTA CAGGACTCTC AAGAAAATGC TCCAGCCGCG GTATCCTCAG
 V I Y L P D R I L V V I L S A V L T W H
 1261 GTGATTTATC TTCCGGACCG AATTCTCGTA GTTATTCTTT CAGCCGTCCT GACATGGCAT
 L G W D D K G L E I L G P L K Q N A N G
 1321 CTTGGTTGGG ATGACAAAGG GTTGGAGATT CTTGGGCCCT TGAAACAAAA TGCCAATGGC
 L F A F K W P F Q F S Q M K H V R A A M
 1381 CTTTTTGCCT TCAAATGGCC TTTCCAGTTT AGCCAGATGA AGCATGTACG CGCTGCAATG
 S T S F V I A L L G F F E S S V A A K G
 1441 AGTACTTCTT TCGTCATCGC GTTACTTGGC TTTTTCGAGT CTTCTGTTGC CGCCAAGGGA
 L S G E A R Q E G V Q G M P V S A N R E
 1501 CTTAGTGGCG AGGCCAGACA AGAAGGTGTC CAGGGAATGC CTGTCAAGTGC TAACAGAGAG
 M V A L G L A N T V G G C F M A L P A F
 1561 ATGGTGGCGC TGGGTCTTGC TAATACTGTG GGGGCTGTT TCATGGCGCT TCCTGCGTTT

Fig. 3A

G G Y A R S K V N A S T G A R S P M S S
 1621 GGTGGCTATG CAAGAAGCAA AGTCAACGCT TCAACTGGAG CTCGGTCTCC GATGAGCAGC
 I F L S I I T F V C I M V L L P Y L Y Y
 1681 ATTTTCCTGA GCATTATTAC CTTTGTTTGT ATCATGGTGC TTTTGCCGTA CTTATACTAT
 L P
 1741 **CTTCCG**GTGA *GTCTCGACCC CAAATACTTC CGAGCGAAGG CTGAGAAAAT ATGTTGCAAT*
 K A V L S S M I S V V A F S L I E E
 1801 *AATTCAGAAA* GCCGTTCTTT CTTCTATGAT ATCTGTCGTC GCATTGAGTC TCATTGAAGA
 C P H D V A F F I R L R G W T E L A L M
 1861 ATGTCCTCAC GACGTGGCTT TCTTTATCCG ACTGCGCGGA TGGACGGAGC TAGCCCTAAT
 L L I F V S T I F Y S L E L G I A L G I
 1921 GCTTCTCATC TTTGTCTCGA CTATTTTCTA TTCTCTAGAG CTGGGAATTG CCCTTGATAT
 G L S I L I L I R H S T Q P R I Q I L G
 1981 TGGCCTTTCT ATCTTGATCC TTATTCGCCA TTCTACGCAG CCTCGGATCC AAATTCTGGG
 K I A G T T D R F D N A E L H P E S V E
 2041 TAAGATAGCA GGCACCTACCG ACCGTTTCGA TAACGCTGAA CTCCACCCCG AGAGCGTTGA
 L I E G A L I V K I P E P L T F A N T G
 2101 GTTAATCGAA GGCGCGCTTA TTGTTAAGAT CCCGGAACCG CTCACCTTTG CCAATACTGG
 E L K N R L R R L E L Y G S S R A H P S
 2161 TGAGCTCAAG AATCGTCTTC GGCGGTTGGA ATTATATGGC AGTAGCCGAG CGCACCTTTC
 L P P T R T P E H N K N I I F D V H G V
 2221 TCTTCCCCCC ACGCGCACCC CCGAACATAA CAAGAATATT ATATTTGATG TTCATGGTGT
 T S I D G S G T Q V L Y E I V D G Y A D
 2281 TACTAGCATC GATGGTCCG GTACGCAAGT CTTATATGAG ATTGTGGACG GATATGCAGA
 Q G V S V F F C R V A T R N V F R M F E
 2341 CCAGGGGGTC AGCGTCTTCT TCTGCCGCGT CGCAACTCGC AATGTTTTCC GCATGTTTGA
 R S G I V E R C G G I T H F V H G V D E
 2401 ACGAAGTGGA ATTGTTGGAAC GATGCGGTGG GATAACGCAC TTCGTTTCATG GTGTCGACGA
 A L R L A E S E D E I E I *
 2461 AGCCCTCCGC CTTGCCGAAT CGGAAGACGA GATTGAAATC TGA

Fig. 3B

M P G D L K T K I G H G A A K A L G I K
 1 ATGCCGGGCG ATCTCAAAAC CAAAATTGGT CACGGCGCGG CCAAGGCCTT GGGGATCAAG
 I P Y R D P L G V H A D P V T R G E S M
 61 ATCCCCTACC GTGATCCTCT CGGAGTTCAT GCTGACCCAG TCACACGAGG CGAGTCGATG
 F S V G T I D T Y S Y L E P E P T P A E
 121 TTCTCCGTCG GAACGATCGA CACATACTCC TATCTCGAGC CCGAACCCAC TCCCCTGAA
 W L K E V C P S W H Q V G R Y F Y N L F
 181 TGGCTGAAGG AAGTCTGCCC TAGCTGGCAT CAGGTGGGCC GTTATTTTTTA CAACCTTTTC
 P F L S W I T R Y N L Q W L L G D M I A
 241 CCTTTCCTCT CGTGGATTAC GAGGTACAAC TTGCAATGGT TGCTGGGAGA TATGATTGCC

 301 GGTAAGAGCC TTTCCTACTGT GTTTGATTGG ATCGACAAGT AGACAACATA CTCATTGGAA
 G V T V G A V V V P Q G M A Y A K L A
 361 TGCAGGCGTC ACGGTCCGGTG CTGTGGTCGT TCCGAGGGA ATGGCCTACG CTAAACTGGC
 N L P V E Y G L Y S S F M G V L I Y W F
 421 AAACCTACCT GTAGAGTATG GTCTCTATTG CTCGTTCATG GGTGTTCTCA TTTATTGGTT
 F A T S K D I T I G
 481 TTTTGCCACC TCAAAGGATA TCACCATTGG TGTAAGTCAT TCTGCACCCA TGTCAGCATG
 P V A V M S T L T
 541 TATCTTGCTA ATATAGTATC TTCCCTGTTC AGCCGGTGGC TGTCATGTCT ACCCTTACAG
 G K I V A E A Q T K L P D V E G H V I A
 601 GTAAGATAGT TGCCGAGGCG CAAACGAAGC TCCCAGATGT CGAAGGGCAT GTAATCGCCT
 S C L A I I C G A V V C A M G L L R L G
 661 CCTGTTTTGGC TATCATTTGT GGAGCCGTGG TTTGCGCTAT GGCCTGCTT CGGCTGGGAT
 F I V D F I P L P A I S A F M T G S A I
 721 TTATCGTGGG TTTTATTCTT CTGCCGGCAA TTTTCTGTTT CATGACGGGT TCCGCCATCA
 N I C S G Q V K D M L G E T A D F S T K
 781 ATATCTGCTC CGGACAGGTC AAAGACATGC TGGGAGAGAC GGCCGACTTC TCGACGAAAG
 D S T Y L V I I N T L K H L P S A K I D
 841 ATCTTACATA TCTGTTTATC ATCAACACCC TCAAGCATCT TCCCTCCGCA AAAATCGATG
 A A M G V S A L A M L Y I I R S G C N Y
 901 CCGCCATGGG TGTCAGTGCT TTAGCTATGC TGTACATTAT CCGTTCGGGT TGCAATTATG
 G A K K F P R H A K V W F F V S T L R T
 961 GCGCGAAGAA GTTCCCCCGT CATGCCAAGG TTTGGTTCTT CGTTTCGACT TTGCGCACAG
 V F V I L F Y T M I S A A V N L H R R S
 1021 TGTTTCGTGAT CTTGTTCTAT ACGATGATCA GTGCCGTGT GAACTTGCAC CGGCGGTCTA
 N P R F K L L G K V P R G F Q H A A V P
 1081 ACCCGCGGTT CAAGCTCCTG GGTAAGTTC CTCGTGGTTT CCAACATGCG GCTGTCCCTC
 Q V N S R I I S A F A S E L P A S I I V
 1141 AGGTAAATTC GAGGATCATC AGCGCATTTC CTAGCGAACT TCCTGCTTCG ATTATGTCC
 L L I E H I A I S K S F G R V N N Y T I
 1201 TGCTTATCGA ACACATCGCT ATCTCGAAAT CCTTTGGCCG TGTCACAAC TACACAATTG
 D P S Q E L V A I G V S N L L G P F L G
 1261 ATCCCTCTCA GGAGCTGGTT GCTATTGGTG TGTGAACTT GCTTGGACCG TTCCTTGGTG
 G Y P A T G S F S R T A I K S K A G V R
 1321 GTTACCCAGC GACTGGATCG TTCTCCGAA CTGCAATCAA ATCGAAAGCG GGTGTCCGCA
 T P L A G V I T A V V V L L A I Y A L P
 1381 CCCCACTTGC CGGTGTTATT ACTGCGGTTG TTGTCTCCT CGCCATTTAC GCTCTGCCCG
 A V F F Y I P K A S L A G V I I H A V G
 1441 CTGTCTTCTT TTACATCCCG AAAGCTTCCC TTGCTGGTGT CATCATTCAT GCAGTCCGGT
 D L I T P P N T V Y Q F W R V S P L D A
 1501 ACCTCATTAC CCCACCAAAC ACCGTTTACC AGTTCGGCG CGTGTCCCT CTGGATGCGA
 I I F F I G V I V T V F T T I E I G I Y
 1561 TCATTTTCTT TATCGGTGTT ATCGTGACTG TCTTACCAC GATTGAGATC GGCATTTACT
 C T V C V S V A I L L F R V A K A R G Q
 1621 GTACCGTTTG TGTGTCTGTT GCCATTCTGC TGTTCGGCGT CGCCAAGGCC CGCGTCAAT

Fig. 4A

```

      F L G R   V T I   H S V   I G D H   L V Q   D D G
1741 TCTTAGGAAG AGTCACTATC CACTCGGTGA TCGGTGACCA TCTGGTACAG GATGATGGGA
      K Y G S   A N S   P N A   A S D D   K D E   L S R
1801 AATATGGGTC TGCCAACCTCC CCTAATGCTG CCAGCGATGA CAAAGATGAA TTGAGCCGGT
      S I F L   P I N   H T D   G S N P   D V E   V Q Q
1861 CTATCTTCTT GCCTATCAAC CACACGGACG GATCGAATCC CGATGTCGAG GTGCAGCAAC
      P Y P G   I F I   Y R F   S E G F   N Y P   N A N
1921 CTTATCCTGG TATCTTCATC TACCGATTCT CGGAAGGATT CAACTACCCC AATGCCAATC
      H Y T D   Y L V   Q T I   F K H T   R R T   N P F
1981 ACTACACCGA TTATTTGGTC CAGACTATCT TCAAGCATAAC ACGTGCGACA AATCCGTTCT
      S Y G K   P G D   R P W   N N P G   P R R   G K S
2041 CCTACGGTAA ACCGGGTGAT CGGCCATGGA ATAATCCTGG CCCTCGCAGG GGCAAGTCTG
      E D D E   S H L   P L L   Q A V I   L D F   S S V
2101 AAGATGACGA GTCGCATTTG CCCTTACTGC AGGCTGTTCAT TCTTGACTTC TCATCCGTCA
      N N V D   V T S   V Q N   L I D V   R N Q   L D L
2161 ACAATGTTGA TGTGACCTCG GTCCAGAACC TCATCGATGT CCGCAATCAA CTCGACCTCT
      Y A S P   K T V   Q W H   F A H I   N N R   W T K
2221 ACGCTTCGCC TAAGACTGTG CAGTGGCACT TTGCTCATAT TAACAACCGC TGGACGAAAC
      R A L A   A A G   F G F   P S P D   S D E   G F Q
2281 GAGCCCTTGC AGCAGCAGGT TTCGGCTTCC CATCTCCGGA CTCGGATGAA GGATTCCAGA
      R W K P   I F S   V A E   I E G S   A S A   A A H
2341 GATGGAAGCC AATTTTCAGC GTGGCTGAGA TCGAAGGCAG TGCCTCTGCC GCAGCTCATG
      A E M V   N N R   H T Q   H N I K   S E D   L E H
2401 CAGAGATGGT GAACAACAGA CACACCCAGC ATAACATCAA GAGCGAAGAC CTCGAGCATG
      G L K H   D S E   T T E   R E T H   G I E   E S S
2461 GCCTCAAGCA CGATTCAGAG ACCACCGAGC GTGAGACACA CGGCATCGAA GAATCCTCCG
      D A S S   T R E   D K L   Q R D L   K D S   K A Y
2521 ATGCCAGCAG CACCCGGGAG GACAAGTTGC AACGGGACCT GAAGGATAGC AAGGCTTACC
      R S R R   R V A   M V Q   G L N R   P F F   H I D
2581 GCAGTCGCCG AAGGGTCGCT ATGGTGCAGG GCCTCAACCG GCCATTCTTC CACATCGACC
      L T S A   L Q S   A L A   N A G E   Q P D   P K M
2641 TGACTAGTGC ACTGCAGAGT GCCTTGGCCA ACGCGGGCGA GCAGCCGGAC CCTAAAATGA
      N V L D   A *
2701 ATGTCCTTGA TGCATAG

```

Fig. 4B

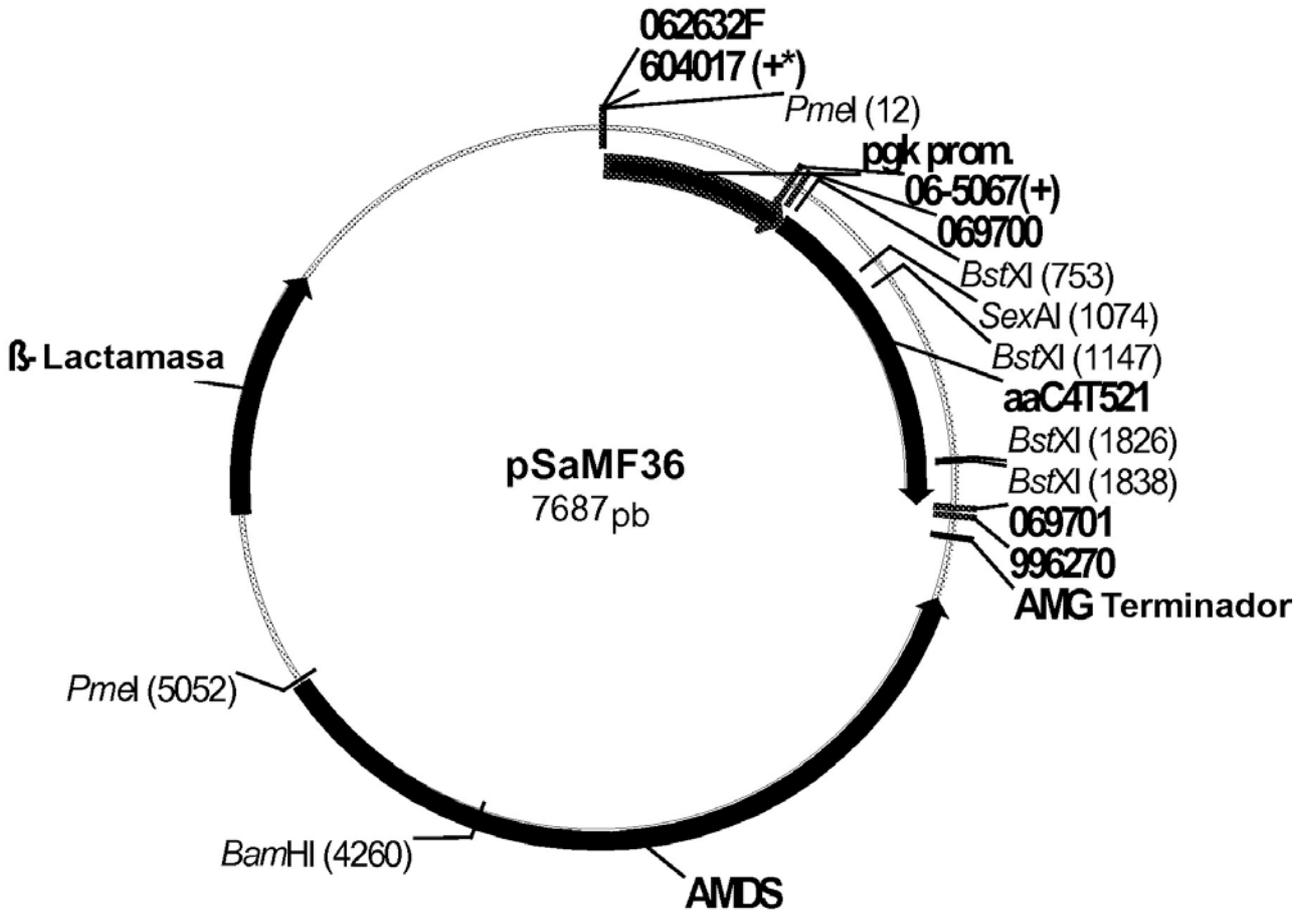


Fig. 5

	M H D	H S T G	S S P	Y I S	D V E T	L N H
1	ATGCACGACC	ACAGCACTGG	ATCTAGTCCA	TACATCTCGG	ACGTGGAAAC	CTTGAACCAC
	A C E	K S V N	P E T	K V S	Q P Q E	S P I
61	GCCTGCGAGA	AGTCCGTCAA	CCCCGAGACC	AAAGTCTCCC	AGCCTCAGGA	ATCTCCCATT
	I S N	N E H Q	E F V	K L G	I R Q R	L R H
121	ATCAGCAATA	ATGAACATCA	GGAGTTTGTT	AAGCTGGGCA	TCCGCCAACG	GCTGCGTCAT
	F T W	A W Y T	L T M	S A G	G L A L	L L R
181	TTCACCTGGG	CCTGGTATAC	CCTAACCATG	AGCGCAGGTG	GACTGGCCCT	TCTTCTCCGC
	N Q P	Y Q F K	G L K	E I G	L V V Y	I A N
241	AACCAGCCGT	ATCAATTCAA	GGGGTTGAAG	GAGATAGGCC	TGGTGGTATA	CATAGCCAAT
	L V F	F T I I	G S L	M I T	R F V L	Y N N
301	CTCGTCTTCT	TTACTATCAT	CGGCTCTCTT	ATGATCACCA	GGTTTGTCTT	TTACAACAAC
	L M D	S L R H	D R E	G F F	F P T F	W L S
361	CTGATGGACT	CTCTCCGCCA	CGACCGAGAA	GGTTTCTTCT	TTCCAACCTT	CTGGCTCTCC
	I A T	M I S G	L S A	Y F S	T E D T	H R L
421	ATCGCCACCA	TGATTAGTGG	TCTATCTGCC	TACTTCTCTA	CTGAAGACAC	GCACCGCCTC
	N Y A	L E G L	F W A	Y C I	F T F A	S A V
481	AATTATGCTC	TCGAGGGTCT	CTTCTGGGCG	TACTGTATCT	TCACGTTTGC	CTCAGCAGTG
	I Q Y	S F V F	S Y H	T F P	L Q T M	M P S
541	ATCCAGTACT	CCTTTGTCTT	CTCCTATCAC	ACGTTCCCTC	TGCAAACAT	GATGCCATCA
	W I L	P A F P	I M L	S G T	I A S A	A S S
601	TGGATCTTAC	CGGCATTCCC	TATCATGCTG	AGCGGAACCA	TTGCCTCTGC	CGTTCCAGC
	Y Q P	A V S A	T P M	I V A	G I T F	Q G L
661	TACCAGCCTG	CGGTGTCTGC	CACGCCTATG	ATTGTTGCCG	GCATCACGTT	CCAGGGACTC
	G F C	I S F M	M Y A	H Y I	G R L M	E T G
721	GGATTCTGCA	TCAGCTTCAT	GATGTACGCC	CACTACATCG	GGCGTCTGAT	GGAGACGGGC
	I P S	S E H R	P G M	F I C	V G P P	A F T
781	ATCCCTTCGA	GCGAGCACCG	TCCTGGTATG	TTCATCTGTG	TCGGCCCCCC	TGCCTTCACG
	L L A	I I G M	A N G	L P E	G F S I	L G D
841	CTGCTGGCTA	TCATCGGCAT	GGCCAACGGC	CTTCCCGAGG	GCTTCAGTAT	CCTGGGCGAT
	G G M	D D R H	I M R	V L A	V C A G	M F L
901	GGTGGCATGG	ACGACCGTCA	CATCATGCGA	GTACTGGCCG	TCTGCGCGGG	CATGTTCCCTC
	W A L	S I W F	F C V	A L G	S V V R	A P P
961	TGGGCTCTGA	GCATTTGGTT	CTTCTGTGTC	GCTCTGGGCT	CAGTTGTGCG	GGCGCCTCCC
	H D F	H L N W	W A M	V F P	N T G L	T L A
1021	CATGATTTCC	ACCTCAACTG	GTGGGCTATG	GTCTTCCCTA	ACACCGGACT	CACTCTCGCC
	T I T	L A K S	L D S	A A L	K W V G	V G M
1081	ACCATCACCC	TGGCCAAGTC	ACTGGACAGT	GCCGCGTTGA	AATGGGTGGG	CGTGGGCATG
	S L C	V I C M	F I F	V F V	S T I R	A V L
1141	TCCCTCTGCG	TGATCTGCAT	GTTTCATCTT	GTCTTCGTGA	GCACCATTAG	GGCTGTTCTC
	L K R	I M W P	G R D	E D V	S E L F	E *
1201	TTGAAGAGGA	TCATGTGGCC	AGGTCGGGAT	GAGGATGTGT	CCGAGTTGTT	CGAATGA

Fig. 6

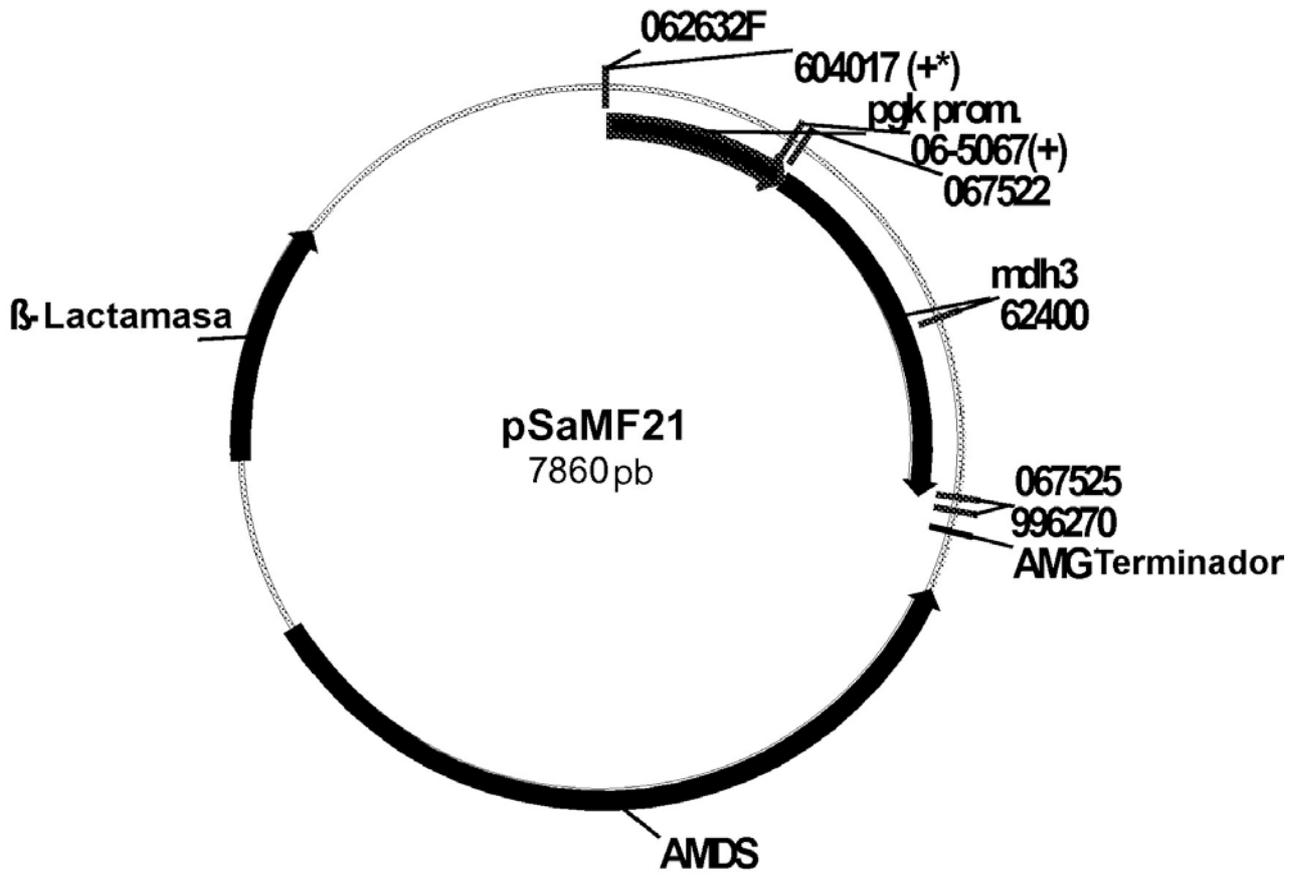


Fig. 8

M A A P F R Q P E E A V D D T E F I D D
 1 ATGGCGGCTC CGTTTCGTCA GCCTGAGGAG GCGGTGATG ACACCGAGTT CATCGATGAC
 H H E H L R D T V H H R L R A N S S I M
 61 CACCATGAAC ACCTCCGTGA TACCGTGAC CATCGGTTGC GCGCCAATTC CTCCATTATG
 H F Q K I L V A N R G E I P I R I F R T
 121 CACTTCCAGA AGATCCTCGT CGCCAACCGT GGTGAGATCC CCATTCGTAT CTTCAGAACG
 A H E L S L Q T V A I Y S H E D R L S M
 181 GCCCAGGAGC TGTCTTGCA GACGGTTGCT ATCTACTCTC ATGAGGATCG ACTGTCAATG
 H R Q K A D E A Y M I G H R G Q Y T P V
 241 CACCGTCAAA AGGCCGATGA GGCCTACATG ATTGGCCACC GCGGTACAGTA CACCCCTGTC
 G A Y L A G G D E I I K I A L E H G V Q L
 301 GGTGCGTACC TGGCGGGCGA TGAGATCATC AAGATCGCCC TGGAGCACGG TGTCCAGTGC
 I H P G Y G F L S E N A D F A R K V E N
 361 ATCCACCCGG GCTACGGTTT CTTGTCCGAG AACGCCGACT TCGCCCGCAA GGTGAGAAC
 A G I V F V G P T P D T I D S L G D K V
 421 GCCGGCATTG TCTTTGTGGG ACCCACTCCC GATACCATTG ACAGCTTGGG TGACAAGGTG
 S A R R L A I K C E V P V V P G T E G P
 481 TCGGCCCGTC GGCTGGCCAT TAAGTGCAGG GTCCCTGTGC TTCCGGGTAC GGAGGGCCCC
 V E R Y E E V K A F T D T Y G F P I I I
 541 GTCGAGCGCT ATGAGGAGGT CAAGGCCTTC ACAGACACCT ATGGCTTCCC CATCATCATC
 K A A F G G G G R G M R V V R D Q A E L
 601 AAGGTCCTT TTGGCCGTGG TGGCCGTGGT ATGCGTGTGG TCCGTGACCA GGCCGAGCTG
 R D S F E R A T S E A R S A F G N G T V
 661 CGTGACTCGT TCGAGCGAGC CACCTCTGAG GCCCGCTCCG CCTTCGGCAA TGGTACCCTG
 F V E R F L D K P K H I E V Q L L G D S
 721 TTCGTCGAGC GCTTCCTCGA CAAACCCAAG CACATGAAG TCCAGCTTCT GGGTGACAGC
 H G N V V H L F E R D C S V Q R R H Q K
 781 CACGGCAACG TTGTCCATCT GTTTGAGCGT GACTGCTCCG TGCAGCGTCG TCACCAGAAG
 V V E V A P A K D L P A D V R D R I L A
 841 GTCGTTGAGG TTGCTCCGGC TAAGGACCTG CCAGCCGATG TCCGGGACCG CATCCTGGCC
 D A V K L A K S V N Y R N A G T A E F L
 901 GATGCTGTGA AGCTGGCCAA GTCCGTCAAC TACCGTAACG CCGGTACAGC TGAGTTCCTG
 V D Q Q N R H Y F I E I N P R I Q V E H
 961 GTGACCAGC AGAACCGCCA CTACTTCATT GAAATCAATC CTCGTATCCA AGTCGAGCAC
 T I T E E I T G I D I V A A Q I Q I A A
 1021 ACCATCACCG AAGAGATTAC TGGTATCGAT ATCGTGGCTG CACAGATCCA GATTGCTGCT
 G A S L E Q L G L T Q D R I S A R G F A
 1081 GGTGCAAGCC TCGAGCAACT GGGCCTGACT CAGGACCGCA TCTCCGCCCG CGGATTTGCC
 I Q C R I T T E D P A K G F S P D T G K
 1141 ATTCAATGTC GTATCACCAC GGAAGATCCC GCCAAGGGGT TCTCTCCGGA TACTGGTAAG
 I E V Y R S A G G N G V R L D G G N G F
 1201 ATTGAGGTTT ATCGTTCCGC TGGTGGTAAC GGTGTCCGTC TGGATGGTGG TAACGGTTTC
 A G A I I T P H Y D S M L V K C T C R G
 1261 GCTGGTGCTA TCATACCCCC TCACTACGAC TCCATGCTGG TCAAGTGCAC CTGCCGTGGT
 S T Y E I A R R K V V R A L V E F R I R
 1321 TCGACCTATG AAATCGCTCG TCGCAAGGTT GTGCGTGCCT TGGTCCGAGT CCGTATTCTG
 G V K T N I P F L T S L L S H P T F V D
 1381 GGTGTGAAGA CCAACATTCC CTTCTGACT TCGTCTTGA GCCACCCGAC CTTCTGCTGAT
 G N C W T T F I D D T P E L F S L V G S
 1441 GGAACTGCT GGACCACTTT CATCGACGAC ACCCCTGAAT TGTCTCTCT TGTCCGGCAGT
 Q N R A Q K L L A Y L G D V A V N G S S
 1501 CAGAACCGTG CCCAGAAGCT GCTCGCATACT CCGGCGATG TAGCTGTCAA CGGTAGTAGC
 I K G Q I G E P K L K G D V I K P K L F
 1561 ATCAAGGGCC AAATTGGCGA GCCCAAGCTC AAGGGTGATG TCATCAAGCC GAAGCTTTTC
 D A E G G K P L D V S A P C T K G W K Q I
 1621 GATCCCGAGG GCAAGCCGCT TGACGTTTCC GCCCCCTGCA CCAAGGGTTG GAAGCAGATT
 L D R E G P A A F A K A V R A N K G C L
 1681 CTGGACCGGG AGGGTCCGGC TGCCTTTGCG AAGGCCGTGC GTGCCAACAA GGGTTGCTTG
 I M D T T W R D A H Q S L L A T R V R T
 1741 ATCATGGATA CTACCTGGCG TGACGCCAC CAGTCTTTCG TGCCACCCG TGTGCGTACC
 I D L L N I A H E T S Y A Y S N A Y S L
 1801 ATCGACTTGT TGAACATCGC CCATGAGACC AGCTACGCCT ACTCCAATGC GTACAGTTTG

Fig. 9A

E C W G G A T F D V A M R F L Y E D P W
 1861 GAATGCTGGG GTGGTGCTAC CTTGATGTG GCCATGCGTT TCCTCTATGA GGACCCCTGG
 D R L R K M R K A V P N I P F Q M L L R
 1921 GACCGCTGCG GCAAGATGCG TAAGGCTGTT CCTAACATCC CATTCCAGAT GTTGCTCCGT
 G A N G V A Y S S L P D N A I Y H F C K
 1981 GGTGCCAACG GTGTGCGCTA CTCTTCCCTC CCAGACAACG CCATCTACCA CTTCTGTAAG
 Q A K K C G V D I F R V F D A L N D V D
 2041 CAGGCTAAGA AGTGCCTGTT CGACATTTTC CGTGTTTTCG ACGCCCTCAA CGATGTCGAT
 Q L E V G I K A V H A A E G V V E A T M
 2101 CAGCTCGAGG TCGGTATCAA GGCTGTTCAT GCTGCCGAGG GTGTGTGCGA GGCCACCATG
 C Y S G D M L N P H K K Y N L E Y Y M A
 2161 TGCTACGAGG GTGACATGCT GAACCCCCAC AAGAAGTACA ACCTGGAGTA CTACATGGCC
 L V D K I V A M K P H I L G I K D M A G
 2221 TTGGTGGATA AGATTGTAGC CATGAAGCCT CACATCCTTG GTATCAAGGA TATGGCCCGT
 V L K P Q A A R L L V G S I R Q R Y P D
 2281 GTGCTGAAGC CCCAGGCCGC TCGCCTGTTG GTGGGCTCCA TCCGTCAGCG CTACCCTGAC
 L P I H V H T H D S A G T G V A S M I A
 2341 CTTCCCATCC ACGTCCACAC CCACGATDCC GCTGGTACTG GTGTAGCTTC CATGATTGCC
 C A Q C A D A A V D A A T D S M S G M T
 2401 TGTGCCCAGG CGGGTGCCGA CGCCGTGGAC GCCGCGACCG ACAGCATGTC CGGTATGACC
 S Q P S I G A I L A S L E G T E Q D P G
 2461 TCCCAGCCTA GCATTGGTGC CATTCTGGCC TCTCTGAGG GCACTGAGCA AGACCCCGT
 L N L A H V R A I D S Y W A Q L R L L Y
 2521 CTC AACCTCG CCCACGTGCG CGTATTGAT AGCTACTGGG CACAGCTGCG CTTGCTCTAC
 S P F E A G L T G P D P E V Y E H E I P
 2581 TCTCCTTTTCG AGGCGGGTCT CACTGGCCCC GACCCTGAGG TCTACGAGCA CGAGATCCCT
 G G Q L T N L I F Q A S Q L G L G Q Q W
 2641 GGTGTCAGT TGACCAACCT TATCTTCCAG GCCAGTCAGC TCGGCTTGGG CCAGCAGTGG
 A E T K K A Y E A A N D L L G D I V K V
 2701 GCCGAAACCA AGAAGCCCTA TGAGGCGGCT AATGATTTAC TCGGCGACAT TGTAAAGTTC
 T P T S K V V G D L A Q F M V S N K L T
 2761 ACTCCACCT CCAAGGTGGT CGGTGACTTG GCTCAGTTCA TGGTCTCGAA CAACTGACT
 P E D V V E R A G E L D F P G S V L E F
 2821 CCAGAGGATG TTGTTGAGCG TGCTGGTGAG CTGGACTTCC CTGGTTCTGT GCTCGAATTC
 L E G L M G Q P F G G F P E P L R S R A
 2881 CTCGAAGGTC TCATGGGACA GCCCTTCCGT GGATTCCCCG AGCCATTGCG CTCCCGCCCC
 L R D R R K L E K R P G L Y L E P L D L
 2941 CTGCGGATC GCCGCAAGCT CGAGAAGCGT CCAGGTCTCT ACCTCGAGCC TTTGGATTTG
 A K I K S Q I R E K F G A A T E Y D V A
 3001 GCTAAGATCA AGAGCCAGAT CCGTGAGAAG TTCGGTGCTG CTAAGTGTGTA TGACGTGGCC
 S Y A M Y P K V F E D Y K K F V Q K F G
 3061 AGCTATGCCA TGTATCCCAA GGTCTTCCGAG GACTACAAGA AGTTCGTCCA GAAGTTCGGT
 D L S V L P T R Y F L A K P E I G E E F
 3121 GATCTCTCCG TCTTGCCAC ACGGTACTTC TTGGCCAAGC CTGAGATTGG CGAGGAGTTC
 H V E L E K G K V L I L K L L A I G P L
 3181 CACGTTGAGC TGGAGAAGGG TAAGGTGCTC ATCCTGAAGT TGTGGCCAT CGGCCCTCTT
 S E Q T G Q R E V F Y E V N G E V R Q V
 3241 TCAGAGCAGA CTGGTCAGCG TGAGGTCTTC TACGAAGTCA ACGGTGAGGT GCGCCAGGTC
 A V D D N K A S V D N T S R P K A D V G
 3301 GCTGTGATG ACAACAAGGC TTCCGTGGAC AACACTTCAC GCCCTAAGGC CGATGTGGGT
 D S S Q V G A P M S G V V V E I R V H D
 3361 GACAGCAGCC AGGTGCGGTGC TCCTATGAGC GGTGTGGTTG TTGAAATCCG TGTCCACGAT
 G L E V K K G D P L A V L S A M K M
 3421 GGTCTGGAGG TTAAGAAGGG TGACCCACTT GCCGTCCTGA GTGCCATGAA GATGTAAGT
 E M
 3481 TCATTCCGAA TCATTTTTCT CACTGGTCAA CTACAGATGC TAACAGCTTA TCCAGGAAAT
 V I S A P H S G K V S S L L V K E G D S
 3541 GGTATCTCT GCTCCTCACA GTGGAAAGGT CTCCAGCTTG CTGGTCAAGG AGGGCAGATC
 V D G Q D L V C K I V K A *
 3601 TGTGATGGC CAGGATCTCG TCTGCAAGAT CGTCAAAGCG TAA

Fig. 9B

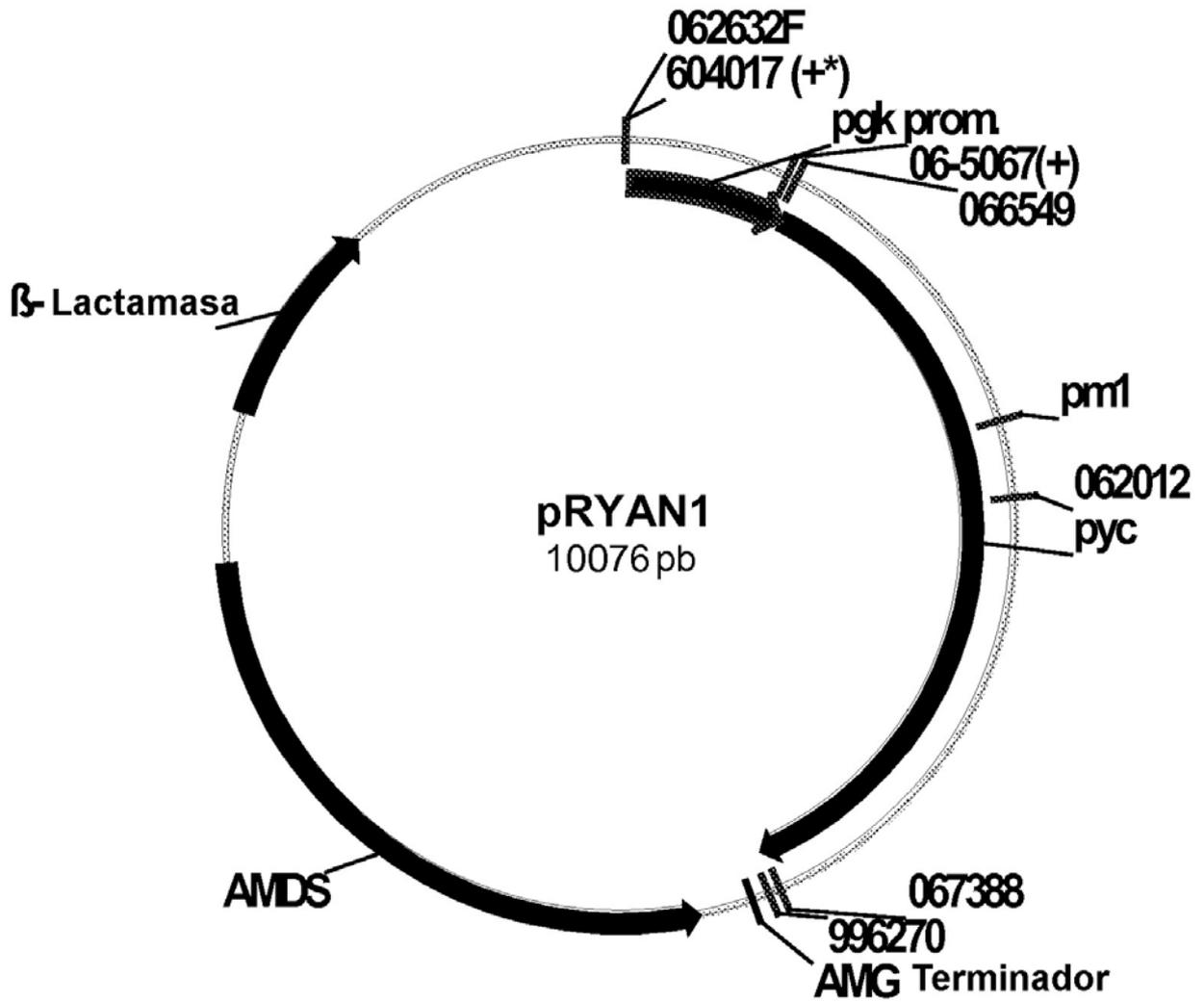


Fig. 10

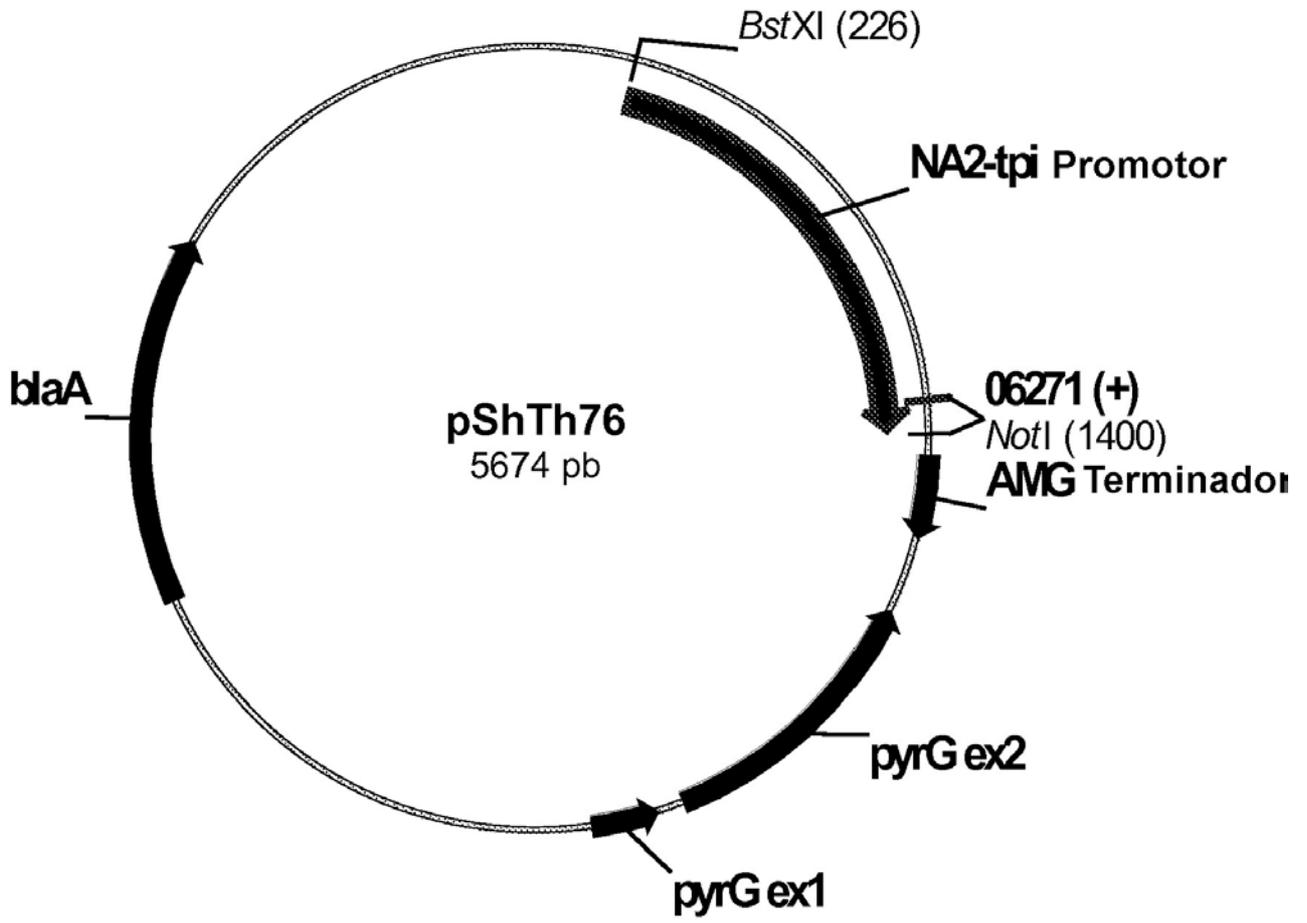


Fig. 11

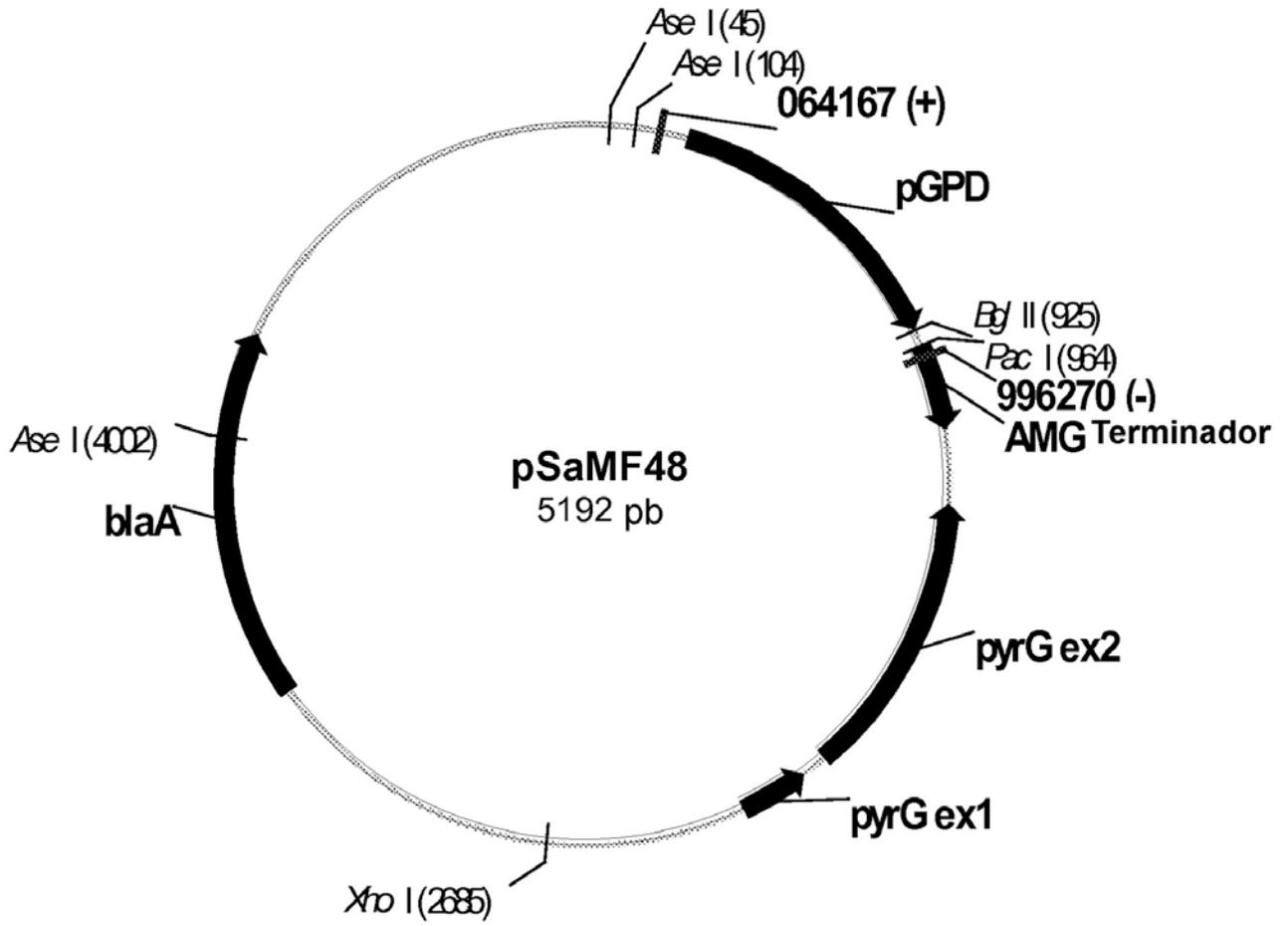


Fig. 12

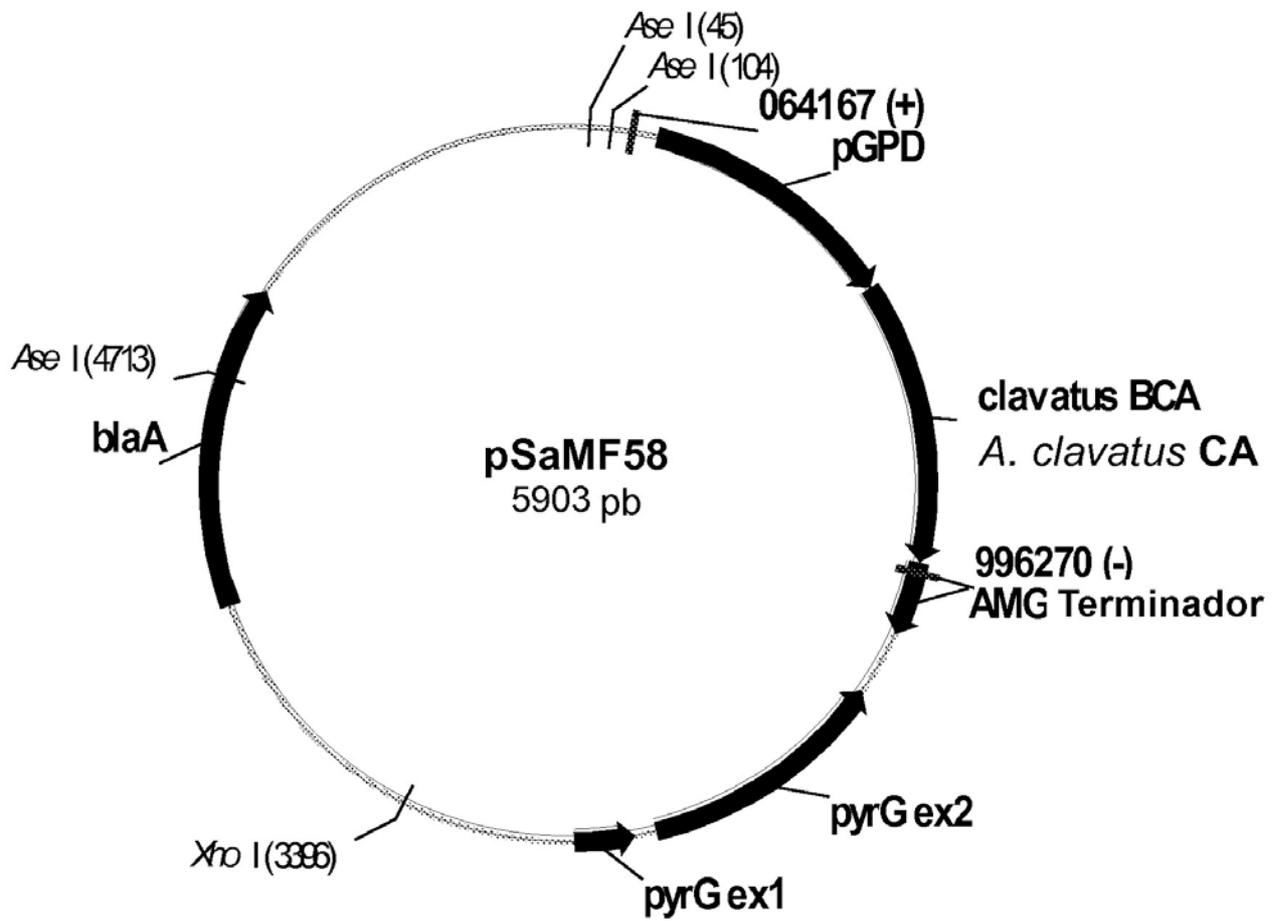


Fig. 13

M S D K A R E A V S Q Y L K Q S H E R I
 1 ATGTCCGACA AGGCTCGCGA GGCGGTCTCG CAATATCTGA AGCAGTCTCA CGAGCGCATC
 F E N N R A W V A A K K E E D P A F F E
 61 TTCGAGAACA ACCGCGCCTG GGTCGCGGCC AAGAAGGAAG AGGACCCCGC GTTTTTCGAG
 K L G A G Q T P Q Y L
 121 AAActGGGCG CCGGACAGAC GCCGCAATAT CTGTAAAGAC CAGCATCAGC GATTTGTCAT
 Y I G C S
 181 TGGAGTTGTC GGATGACTTT GCCATACTGA CCGTGGCTGT GCAGGTACAT CGGATGCAGT
 D S R V P A N D I M G L T A G E V F V H
 241 GACAGTCGCG TGCCCGCCAA TGACATTATG GGTCTCACGG CCGGCGAGGT CTTTGTGCAC
 R N I A N L V P N T D L N V M S V I N Y
 301 CGCAACATCG CCAATCTGGT GCCCAACACC GACCTCAATG TCATGTCGGT CATCAACTAC
 A V R H L K V K H I V V C G H Y N C G G
 361 GCCGTCCGGC ATCTGAAGGT CAAGCACATC GTTGTCTGCG GCCACTACAA CTGTGGCGGT
 V K A A L T P S D L G L L N P W L R N V
 421 GTCAAGGCTG CGCTGACGCC CTCCGACCTG GGGCTGCTGA ACCCCTGGCT GCGCAATGTC
 R D V Y R L H E R E L D A I E D E E A K
 481 CGGGATGTGT ATCGGTTGCA CGAGCGCGAG CTGGACGCCA TCGAAGACGA AGAGGCGAAG
 Y N R L V E L N V V E S C R N V I K T A
 541 TATAATCGCC TGGTGGAGCT GAATGTTGTT GAGTCCTGCC GCAACGTCAT CAAGACGGCG
 A V Q Q S Y H D N Q F P V V H G W I F D
 601 GCGGTGCAGC AGAGCTACCA CGACAACCAG TTCCCCGTGG TCCACGGATG GATCTTTGAT
 V R T G L L R D L N I D F E E T L R D I
 661 GTGCGGACGG GTCTGCTTCG GGATCTCAAC ATTGATTTCG AGGAGACGCT GCGGGATATC
 K K I Y N I T K S *
 721 AAGAAGATCT ACAATATCAC CAAGAGCTGA

Fig. 14

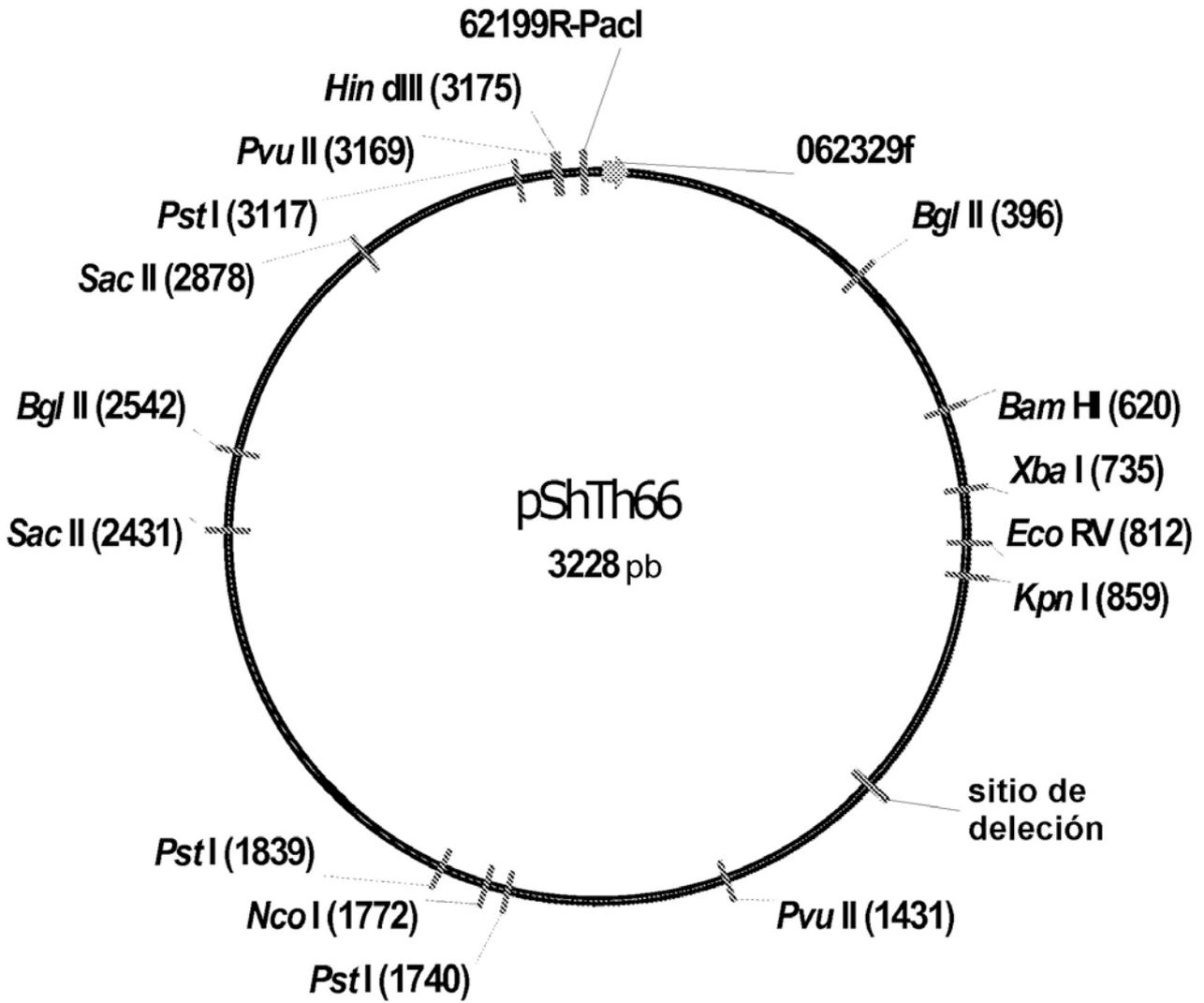


Fig. 15

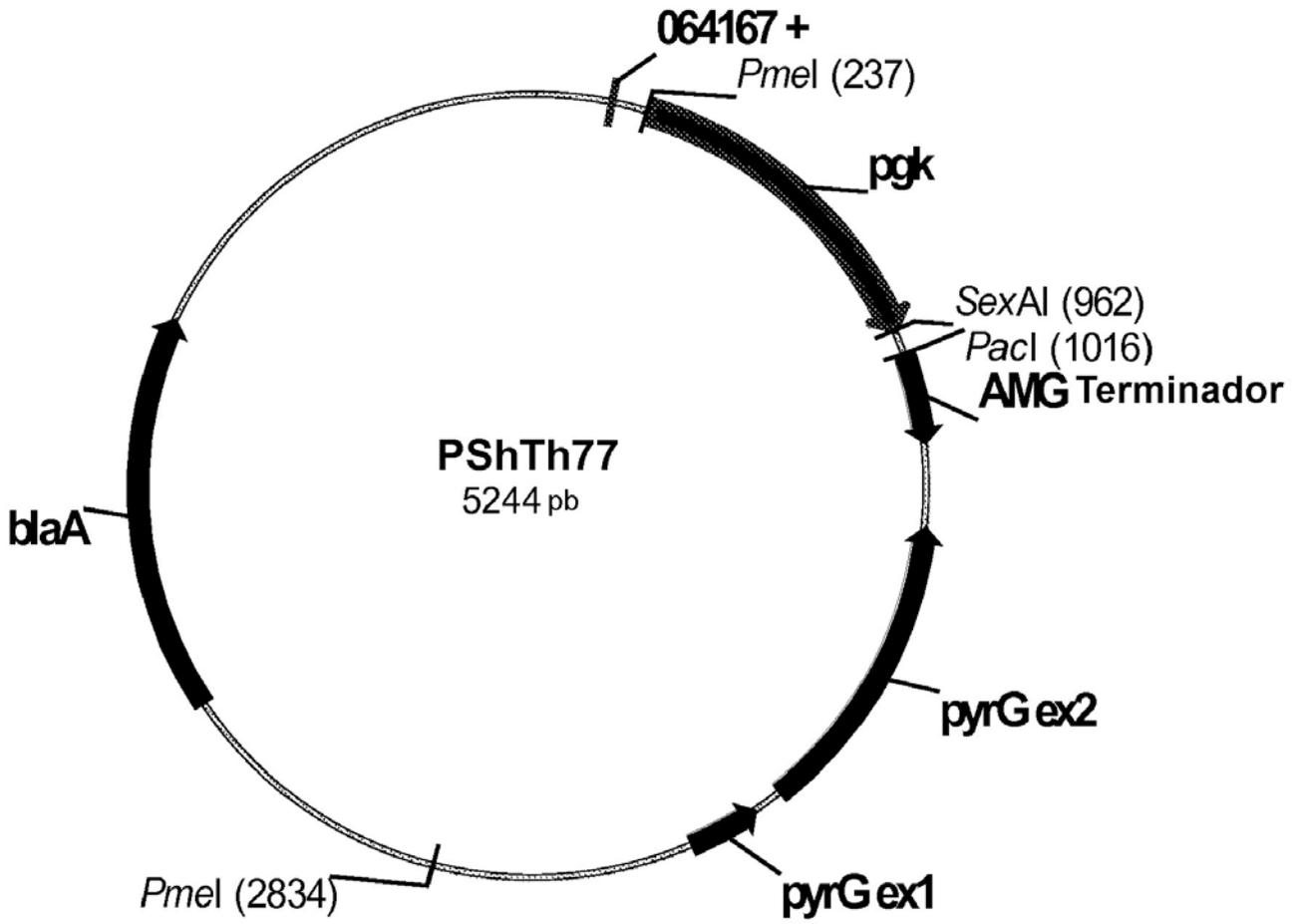


Fig. 16

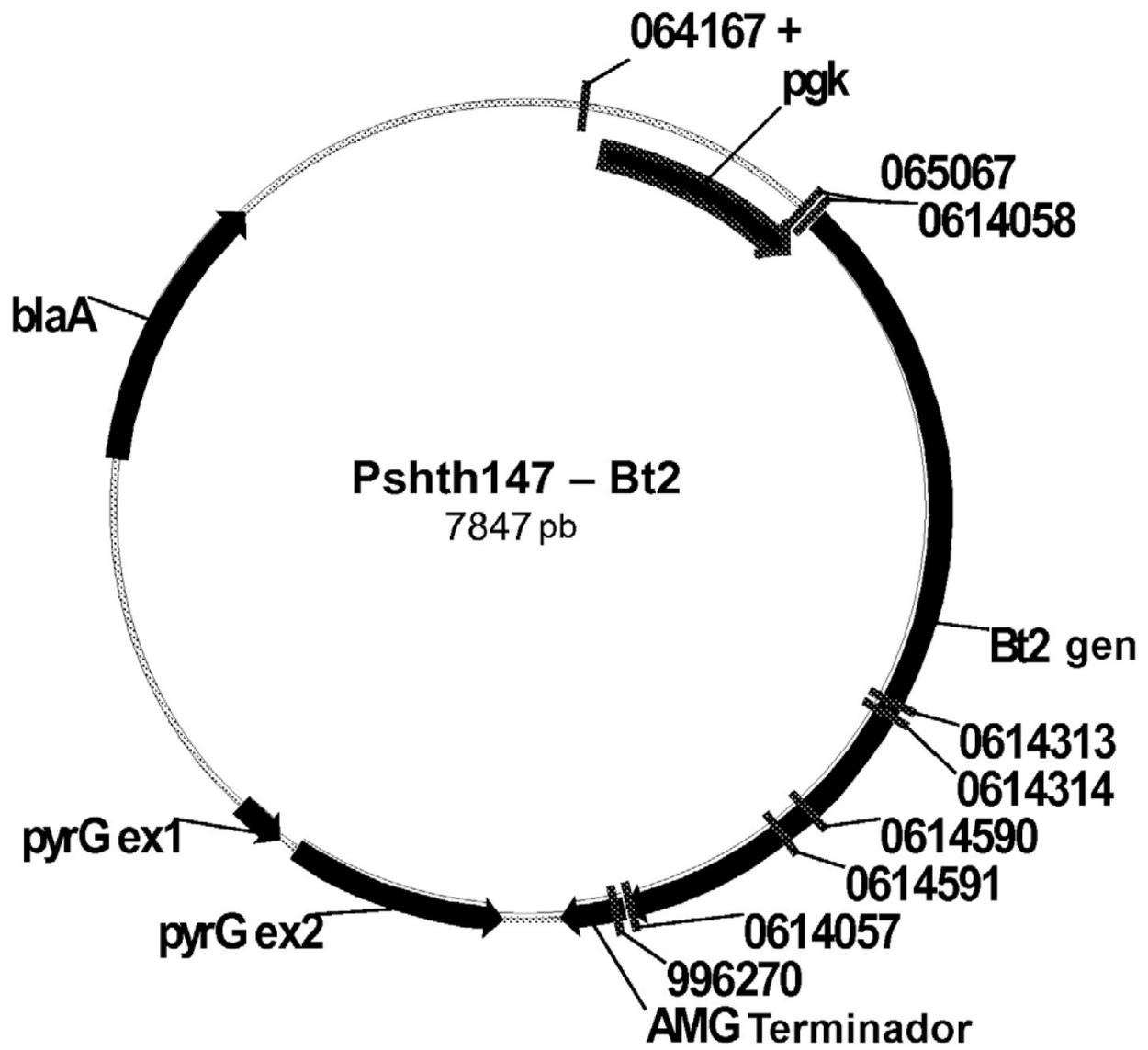


Fig. 17