

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 716 619**

51 Int. Cl.:

C12P 21/06 (2006.01)

C07H 21/04 (2006.01)

C12N 15/00 (2006.01)

C12N 1/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.03.2014 PCT/US2014/022086**

87 Fecha y número de publicación internacional: **12.09.2014 WO14138679**

96 Fecha de presentación y número de la solicitud europea: **07.03.2014 E 14760869 (9)**

97 Fecha y número de publicación de la concesión europea: **09.01.2019 EP 2964775**

54 Título: **Promotores de levadura para la expresión de proteínas**

30 Prioridad:

08.03.2013 US 201361775029 P

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

13.06.2019

73 Titular/es:

**BIOGRAMMATICS, INC. (100.0%)
2705 Glasgow Drive
Carlsbad, California 92010-6536, US**

72 Inventor/es:

**TOLSTORUKOV, ILYA I.;
CREGG, JAMES M.;
CHAPPELL, THOMAS G. y
MADDEN, KNUT R.**

74 Agente/Representante:

MILTENYI , Peter

ES 2 716 619 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

Promotores de levadura para la expresión de proteínas

5 **Campo de la divulgación**

La presente invención se refiere a ácidos nucleicos aislados, métodos de expresión, células hospedadoras, vectores de expresión y construcciones de ADN para producir proteínas y polipéptidos, y a las proteínas y los polipéptidos producidos usando los métodos de expresión. Más en particular, la invención se refiere a ácidos nucleicos aislados de *Pichia pastoris* donde los ácidos nucleicos tienen actividad promotora. La invención también se refiere a métodos de expresión, células hospedadoras, vectores de expresión y construcciones de ADN, para usar los promotores de *Pichia pastoris* para producir proteínas y polipéptidos, y a las proteínas y los polipéptidos producidos usando los métodos de expresión.

15 **Antecedentes y sumario de la invención**

Los sistemas de expresión de levadura pueden usarse para producir de manera eficaz proteínas, tales como enzimas, hormonas y proteínas de vacunas, en parte, porque algunas levaduras crecen rápidamente hasta altas densidades celulares, se cultivan en medios simples y económicos y son eucariotas de modo que pueden modificar las proteínas de manera similar a las proteínas nativas en los mamíferos. Adicionalmente, con una secuencia de señal adecuada, la proteína expresada puede secretarse en el medio de cultivo para el aislamiento y la purificación convenientes. Algunos sistemas de expresión de levadura también son aceptados en las industrias farmacéutica y alimentaria como seguros para la producción de productos farmacéuticos y productos alimentarios, a diferencia de los sistemas de expresión bacterianos y fúngicos que, en algunos casos, pueden ser peligrosos, por ejemplo, para la fabricación de alimentos humanos.

Por tanto, es beneficioso para una diversidad de industrias, tales como la industria alimentaria y de piensos para animales, las industrias de la salud humana y animal, y similares, desarrollar o mejorar los sistemas de expresión de levadura que pueden usarse para expresar altos niveles de proteínas para aumentar el rendimiento, reducir el gasto de aislamiento y purificación de proteínas y reducir los costes de los productos para la salud humana y animal y los productos alimentarios.

Se ha desarrollado una diversidad de tipos de sistemas de expresión de levadura que implican el uso de la expresión inducible o constitutiva de proteínas usando ácidos nucleicos que codifican proteínas homólogas o heterólogas, con el control de un promotor de levadura. Los promotores son elementos reguladores que están unidos al extremo 5' de un ácido nucleico que codifica una proteína y pueden interactuar con diversos factores reguladores en la célula hospedadora (por ejemplo, una célula hospedadora de levadura) para controlar la transcripción de ARN a partir de ADN. Los promotores también pueden controlar el tiempo de transcripción de ARN a partir de ADN. Por ejemplo, el promotor AOX 1 se ha identificado en la levadura *Pichia pastoris* y se usa habitualmente en sistemas de expresión de levadura porque es un promotor estrechamente regulado y fuerte.

Stadlmayr et al, *J. of Biotechn.* 150: 519-529 (2010), divulgan diferentes promotores de *P. pastoris* inducibles por metanol y/o constitutivos y sus usos para la producción de proteínas heterólogas. Kberl et al., *J. of Biotechn.* 154: 312-320 (2011), divulgan la secuencia del genoma de *Pichia pastoris* (*Komagataella phaffii*) CBS7435, que es la cepa parental de *P. pastoris* GS115.

De Schütter et al., *Nature Biotechn.* 27: 561-566 (2009), divulgan la secuencia del genoma de *P. pastoris* GS115.

Debido a la importancia de los sistemas de expresión de levadura para una diversidad de industrias, incluyendo la industria farmacéutica humana y la industria de alimentos para seres humanos y de piensos para animales, la mejora de los sistemas de expresión de levadura es el foco de mucha investigación y desarrollo. En consecuencia, los presentes inventores han identificado promotores de *Pichia pastoris* que son particularmente eficaces para su uso en la expresión de proteínas en levaduras. Los promotores que se describen en el presente documento pueden usarse, por ejemplo, en sistemas de expresión de levadura inducibles por metanol o en sistemas de expresión para la expresión constitutiva de proteínas.

En una realización ilustrativa de la invención, se proporciona un ácido nucleico aislado donde la secuencia del ácido nucleico aislado comprende una secuencia, por ejemplo, idéntica en al menos el 90 %, el 95 % o el 98 % a una secuencia SEQ ID NO: 2, o idéntica en al menos el 90 %, el 95 % o el 98 % a un fragmento de la misma que tiene actividad promotora, donde el ácido nucleico aislado comprende la secuencia de un promotor de *Pichia pastoris* inducible por metanol. En otras realizaciones, se proporcionan vectores de expresión, células hospedadoras y construcciones de ADN que comprenden estas secuencias promotoras.

En otra realización, se proporciona un método de producción de una proteína usando estas secuencias promotoras. El método comprende las etapas de cultivar en un medio de cultivo una célula hospedadora que comprende un primer casete de expresión que comprende cualquiera de las secuencias promotoras anteriores unidas operativamente a una secuencia codificante heteróloga que codifica una proteína, donde el cultivo se realiza en

condiciones que permiten la expresión de la proteína. En otra realización ilustrativa, se proporciona una proteína aislada producida de acuerdo con este método.

5 En otra realización, se proporciona un método de producción de una proteína usando estas secuencias promotoras. El método comprende las etapas de cultivar en un medio de cultivo una célula hospedadora que comprende un primer casete de expresión que comprende cualquiera de las secuencias promotoras anteriores unidas operativamente a una secuencia codificante heteróloga que codifica una proteína, donde el cultivo se realiza en condiciones que permiten la expresión de la proteína. En otra realización ilustrativa, se proporciona una proteína aislada producida de acuerdo con este método.

10 Todas las realizaciones que se describen en la siguiente lista de cláusulas también se contemplan para su uso de acuerdo con la invención. Para todas las realizaciones que se describen en las siguientes cláusulas, se considera que cualquier combinación aplicable de realizaciones está de acuerdo con la invención.

15 1. Un ácido nucleico aislado donde la secuencia del ácido nucleico aislado comprende una secuencia idéntica en al menos el 90 % a una secuencia de la SEQ ID NO: 2, o idéntica en al menos el 90 % a un fragmento de la misma que tiene actividad promotora, donde el ácido nucleico aislado comprende la secuencia de un promotor de *Pichia pastoris* inducible por metanol.

20 2. El ácido nucleico aislado de la cláusula 1 donde la secuencia del ácido nucleico aislado es idéntica en al menos el 95 % a una secuencia de la SEQ ID NO: 2, o idéntica en al menos el 95 % a un fragmento de la misma que tiene actividad promotora.

25 3. El ácido nucleico aislado de la cláusula 1 donde la secuencia del ácido nucleico aislado es idéntica en al menos el 98 % a una secuencia de la SEQ ID NO: 2, o idéntica en al menos el 98 % a un fragmento de la misma que tiene actividad promotora.

30 4. La secuencia de ácido nucleico aislada de la cláusula 1, donde la secuencia del ácido nucleico aislado es una secuencia de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora.

5. El ácido nucleico aislado de una cualquiera de las cláusulas 1 a 4 unido operativamente a una secuencia codificante heteróloga.

35 6. El ácido nucleico aislado de la cláusula 5 donde la secuencia codificante heteróloga codifica una proteína seleccionada entre el grupo que consiste en una toxina, un anticuerpo, una hormona, una enzima, un factor de crecimiento, una citocina, una proteína estructural, una proteína inmunógena y una proteína de señalización celular.

7. El ácido nucleico aislado de la cláusula 6 donde la proteína es una enzima para su uso en piensos para animales.

40 8. El ácido nucleico aislado de la cláusula 7 donde la proteína se selecciona entre el grupo que consiste en una fitasa, una mananasa, una galactosidasa, una amilasa, una glucanasa, una proteasa, una celulasa y una xilanasas.

9. El ácido nucleico aislado de la cláusula 8 donde la proteína es una fitasa.

45 10. El ácido nucleico aislado de la cláusula 8 donde la proteína es una galactosidasa.

11. Un vector de expresión que comprende el ácido nucleico aislado de una cualquiera de las cláusulas 1 a 10.

50 12. Una célula hospedadora que comprende el vector de expresión de la cláusula 11.

13. Una célula hospedadora que comprende el ácido nucleico aislado de una cualquiera de las cláusulas 1 a 10.

55 14. La célula hospedadora de una cualquiera una de las cláusulas 12 o 13 donde la célula hospedadora es una especie de *Pichia*.

15. La célula hospedadora de la cláusula 14 donde la especie de *Pichia* es *Pichia pastoris*.

16. Una construcción de ADN que comprende el ácido nucleico aislado de una cualquiera de las cláusulas 1 a 10.

60 17. Un método de producción de una proteína, comprendiendo el método la etapa de cultivar en un cultivo una célula hospedadora que comprende un primer casete de expresión que comprende un ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, unido operativamente a una secuencia codificante heteróloga que codifica una proteína

65 donde el cultivo se realiza en condiciones que permiten la expresión de la proteína; donde la proteína se expresa usando el primer casete de expresión en combinación con un segundo casete de expresión, comprendiendo dicho segundo casete de expresión

(a) la secuencia codificante heteróloga que codifica la proteína unida operativamente a un ácido nucleico aislado que tiene una secuencia que comprende la SEQ ID NO: 15 o la SEQ ID NO: 16, donde la SEQ ID NO: 15 o la SEQ ID NO: 16 tienen actividad promotora; o (b) el ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma, que tiene actividad promotora unido operativamente a la secuencia codificante heteróloga que codifica la proteína.

5 18. El método de la cláusula 17 donde la proteína se selecciona entre el grupo que consiste en una toxina, un anticuerpo, una hormona, una enzima, un factor de crecimiento, una citocina, una proteína estructural, una proteína inmunógena y una proteína de señalización celular.

10 19. El método de la cláusula 18 donde la proteína es una enzima para su uso en piensos para animales.

20. El método de la cláusula 19 donde la proteína se selecciona entre el grupo que consiste en una fitasa, una mananasa, una galactosidasa, una amilasa, una glucanasa, una celulasa, una proteasa y una xilanasa.

15 21. El método de la cláusula 20 donde la proteína es una fitasa.

22. El método de la cláusula 20 donde la proteína es una galactosidasa.

20 23. El método de una cualquiera de las cláusulas 17 a 22 donde la proteína se expresa usando el primer casete de expresión en combinación con un segundo casete de expresión.

25 24. El método de la cláusula 23 donde el segundo casete de expresión comprende la secuencia de codificación heteróloga que codifica la proteína unida operativamente a un ácido nucleico aislado que tiene una secuencia que comprende la secuencia de la SEQ ID NO: 15 o la SEQ ID NO: 16 donde la SEQ ID NO: 15 y la SEQ ID NO: 16 tienen actividad promotora o cualquier otra secuencia promotora AOX 1 o AOX 2.

25. El método de la cláusula 24 donde la proteína se expresa usando el primer casete de expresión, el segundo casete de expresión y un tercer casete de expresión.

30 26. El método de la cláusula 25 donde el tercer casete de expresión comprende la secuencia codificante heteróloga que codifica la proteína unida operativamente a un ácido nucleico aislado que tiene una secuencia idéntica en al menos el 90 % a una secuencia SEQ ID NO: 2, o idéntica en al menos el 90 % a un fragmento de la misma que tiene actividad promotora.

35 27. El método de la cláusula 23 donde el segundo casete de expresión comprende la secuencia codificante heteróloga que codifica la proteína unida operativamente a un ácido nucleico aislado que tiene una secuencia idéntica en al menos el 90 % a una secuencia seleccionada de la SEQ ID NO: 2, o idéntica en al menos el 90 % a un fragmento de la misma que tiene actividad promotora.

40 **Breve descripción de los dibujos**

La FIGURA 1 muestra la secuencia de nucleótidos de un promotor inducible por metanol, CAM1_(FR839630_178837..180488)_selección (SEQ ID NO: 1).

45 La FIGURA 2 muestra la secuencia de nucleótidos de un promotor inducible por metanol, PP7435_Chr1-1351_selección (SEQ ID NO: 2).

La FIGURA 3 muestra la secuencia de nucleótidos de un promotor inducible por metanol, THI4_selección (SEQ ID NO: 3).

50 La FIGURA 4 muestra la secuencia de nucleótidos de un promotor inducible por metanol, GPM2_selección (SEQ ID NO: 4).

La FIGURA 5 muestra la secuencia de nucleótidos de un promotor inducible por metanol, PP7435_Chr2-0790_selección (SEQ ID NO: 5).

La FIGURA 6 muestra la secuencia de nucleótidos de un promotor inducible por metanol, PP7435_Chr3-0842_selección (SEQ ID NO: 6).

60 La FIGURA 7 muestra la secuencia de nucleótidos de un promotor constitutivo, PP7435_Chr1-0269_selección (SEQ ID NO: 7).

La FIGURA 8 muestra la secuencia de nucleótidos de un promotor constitutivo, PP7435_Chr2-0207_selección (SEQ ID NO: 8).

65 La FIGURA 9 muestra la secuencia de nucleótidos de un promotor constitutivo, PP7435_Chr2-0208_selección (SEQ ID NO: 9).

ID NO: 9).

La FIGURA 10 muestra la secuencia de nucleótidos de un promotor constitutivo, PP7435_Chr2-0809_selección (SEQ ID NO: 10).

5 La FIGURA 11 muestra la secuencia de nucleótidos de un promotor constitutivo, PP7435_Chr4-0069_selección (SEQ ID NO: 11).

10 La FIGURA 12 muestra la secuencia de nucleótidos de un promotor constitutivo, PP7435_Chr4-0800_selección (SEQ ID NO: 12).

La FIGURA 13 muestra la secuencia de nucleótidos de un promotor constitutivo, cromosoma 1 de *Pichia pastoris* CBS 7435 TEF2_selección, secuencia de replicón completa (SEQ ID NO: 13).

15 La FIGURA 14 muestra la secuencia de nucleótidos de un promotor constitutivo, cromosoma 3 de *Pichia pastoris* CBS 7435 PP7435_Chr3-0476_selección, secuencia de replicón completa (SEQ ID NO: 14).

La FIGURA 15 muestra la secuencia de nucleótidos de la secuencia del promotor AOX1 (SEQ ID NO: 15).

20 La FIGURA 16 muestra la secuencia genómica de 1000 pb corriente arriba del codón de inicio ATG del promotor AOX1 (SEQ ID NO: 16).

La FIGURA 17 muestra el plásmido indicador para someter a ensayo la expresión de alfa-galactosidasa (A-gal) con los promotores.

25 La FIGURA 18 (paneles a - c) muestra las imágenes de la expresión de A-gal de clones de *Pichia* con el plásmido indicador y diferentes promotores.

Descripción detallada de las realizaciones ilustrativas

30 En una realización ilustrativa de la invención, se proporciona un ácido nucleico aislado donde la secuencia del ácido nucleico aislado comprende una secuencia, por ejemplo, idéntica en al menos el 90 %, el 95 % o el 98 % a una secuencia SEQ ID NO: 2, o idéntica en al menos el 90 %, el 95 % o el 98 % a un fragmento de la misma que tiene actividad promotora, donde el ácido nucleico comprende la secuencia de un promotor de *Pichia pastoris* inducible por metanol. En otra realización, la secuencia de ácido nucleico aislada es una secuencia SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora.

35 En otras realizaciones, se proporcionan vectores de expresión, células hospedadoras y construcciones de ADN que comprenden estas secuencias promotoras.

40 En otra realización, se proporciona un método de producción de una proteína usando estas secuencias promotoras. El método comprende las etapas de cultivar en un medio de cultivo una célula hospedadora que comprende un primer casete de expresión que comprende cualquiera de las secuencias promotoras anteriores unidas operativamente a una secuencia codificante heteróloga que codifica una proteína, donde el cultivo se realiza en condiciones que permiten la expresión de la proteína. En otra realización ilustrativa, se proporciona una proteína aislada producida de acuerdo con este método.

45 Los promotores descritos anteriormente se han aislado de una cepa de levadura (es decir, NRRL Y11430) clasificada actualmente como una cepa de levadura *Pichia pastoris*. Sin embargo, la clasificación puede cambiar en algún momento a una especie de *Komagataella* (por ejemplo, *Komagataellaphaffii*).

50 En otra realización, se proporciona un método de producción de una proteína usando cualquiera de las secuencias promotoras en el párrafo anterior. El método comprende las etapas de cultivar en un medio de cultivo una célula hospedadora que comprende un primer casete de expresión que comprende cualquiera de las secuencias promotoras anteriores unidas operativamente a una secuencia codificante heteróloga que codifica una proteína, donde el cultivo se realiza en condiciones que permiten la expresión de la proteína. En otra realización ilustrativa, se proporciona una proteína aislada producida de acuerdo con este método.

55 Todas las realizaciones descritas en la siguiente lista de cláusulas están contempladas para su uso de acuerdo con la invención. Para todas las realizaciones que se describen en las siguientes cláusulas, se considera que cualquier combinación aplicable de realizaciones está de acuerdo con la invención. Cualquier realización que se describa en la siguiente lista de cláusulas también se contempla para su uso con cualquier realización descrita en la sección Sumario de la Invención de la presente solicitud o en la sección Descripción Detallada de la sección de Realizaciones Ilustrativas de la presente solicitud.

60 65 1. Un ácido nucleico aislado donde la secuencia del ácido nucleico aislado comprende un promotor de *Pichia*

pastoris inducible por metanol, teniendo dicho promotor actividad promotora y teniendo dicho promotor una secuencia

- 5 a) idéntica en al menos el 90 % a una secuencia de la SEQ ID NO: 2, o
 b) idéntica en al menos el 90 % a un fragmento de la SEQ ID NO: 2, teniendo dicho fragmento actividad promotora.
2. El ácido nucleico aislado de la realización 1 donde la secuencia es la SEQ ID NO: 2.
- 10 3. El ácido nucleico aislado de una cualquiera de las realizaciones 1 a 2 unido operativamente a una secuencia codificante heteróloga.
- 15 4. El ácido nucleico aislado de la realización 3 donde la secuencia codificante heteróloga codifica una proteína seleccionada entre el grupo que consiste en una toxina, un anticuerpo, una hormona, una enzima, un factor de crecimiento, una citocina, una proteína estructural, una proteína inmunógena y una proteína de señalización celular.
5. El ácido nucleico aislado de la realización 4 donde la proteína es una enzima para su uso en piensos para animales.
- 20 6. El ácido nucleico aislado de la realización 5 donde la proteína se selecciona entre el grupo formado por una mananasa, una amilasa, una glucanasa, una proteasa, una celulasa y una xilanasas.
7. El ácido nucleico aislado de la realización 5 donde la proteína es una fitasa.
- 25 8. El ácido nucleico aislado de la realización 5 donde la proteína es una galactosidasa.
9. Un vector de expresión que comprende el ácido nucleico aislado de una cualquiera de las realizaciones 1 a 8.
- 30 10. Una célula hospedadora que comprende el ácido nucleico aislado de una cualquiera de las realizaciones 1 a 8 o el vector de expresión de la realización 9.
11. La célula hospedadora de la realización 10 donde la célula hospedadora es una especie de *Pichia*.
- 35 12. La célula hospedadora de la realización 11 donde la especie de *Pichia* es *Pichia pastoris*.
13. La célula hospedadora de la realización 10 donde la célula hospedadora se selecciona entre el grupo que consiste en especies de *Hansenula*, especies de *Pichia*, especies de *Saccharomyces*, especies de *Yarrowia*, especies de *Candida*, especies de *Schizosaccharomyces*, especies de *Torulaspota* y especies de *Kluveromyces*.
- 40 14. Un método de producción de una proteína, comprendiendo el método las etapas de cultivar en un medio de cultivo una célula hospedadora que comprende un primer casete de expresión que comprende un ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, unido operativamente a una secuencia de codificación heteróloga que codifica una proteína, donde el cultivo se realiza en condiciones que permiten la expresión de la proteína; donde la proteína se expresa usando el primer casete de expresión en combinación con un segundo casete de expresión, comprendiendo dicho segundo casete de expresión
- 45 (a) la secuencia de codificación heteróloga que codifica la proteína unida operativamente a un ácido nucleico aislado que tiene una secuencia que comprende la SEQ ID NO: 15 o la SEQ ID NO: 16, donde la SEQ ID NO: 15 o la SEQ ID NO: 16 tienen actividad promotora o (b) el ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, unido operativamente a la secuencia codificante heteróloga que codifica la proteína.
- 50 Puede usarse cualquier sistema de expresión de levadura conocido por los expertos en la materia de acuerdo con la presente invención. Por ejemplo, se divulgan diversos sistemas de expresión de levadura en las Patentes de los EE.UU. N.º 6.451.572, 6.841.370, 6.974.690, 7.320.876, 7.078.035, 7.138.260 y en la Publicación PCT N.º WO 2007/112739. En cualquiera de las realizaciones que se describen en el presente documento, puede usarse
- 55 cualquiera de estos sistemas de expresión de levadura. Como alternativa, puede usarse cualquier especie de levadura o sistema de expresión de levadura adecuado para la expresión de una proteína, incluyendo especies de levadura, tales como especies de *Saccharomyces* (por ejemplo, *Saccharomyces cerevisiae*), especies de *Kluveromyces* (por ejemplo, *Kluveromyces lactis*), especies de *Torulaspota*, especies de *Yarrowia* (por ejemplo, *Yarrowia lipolitica*), especies de *Schizosaccharomyces* (por ejemplo, *Schizosaccharomyces pombe*). En otra realización, especies de levaduras metilotróficas tales como especies de *Pichia* (por ejemplo, *Pichia pastoris* o *Pichia methanolica*), especies de *Hansenula* (por ejemplo, *Hansenula polymorpha*), especies de *Torulopsis*, especies de *Komagataella*, especies de *Candida* (por ejemplo, *Candida boidinii*) y pueden usarse especies de *Karwinskia*, en particular cuando el promotor es un promotor inducible por metanol. En una realización, la proteína puede expresarse en la levadura metilotrófica *Pichia pastoris*. Las levaduras metilotróficas son capaces de usar metanol
- 60 como única fuente de carbono para la producción de los recursos energéticos necesarios para mantener la función celular. Las levaduras metilotróficas contienen genes que codifican enzimas para la utilización de metanol, tales
- 65

como los genes que codifican la alcohol oxidasa. Cualquiera de estas células hospedadoras puede ser una cepa de célula hospedadora que sea heteróloga para el promotor que se describe en el presente documento (es decir, la célula hospedadora normalmente no contiene en la naturaleza el promotor que se describe en el presente documento).

5 Puede usarse un sistema de expresión de levadura para producir una cantidad suficiente de la proteína intracelularmente o puede secretarse de las células de levadura de manera que la proteína pueda aislarse y purificarse convenientemente del medio de cultivo. Como se usa en el presente documento, el término "expresión" significa la transcripción y/o traducción de un ácido nucleico en una célula hospedadora. Un sistema de expresión de levadura puede incluir, pero no se limita a, la célula hospedadora de levadura y el vector de expresión (por ejemplo, una construcción de ADN) utilizado para expresar la proteína. El vector de expresión puede contener un promotor que se describe en el presente documento y, como se sabe en la técnica, el promotor es heterólogo para el vector de expresión (es decir, la combinación no se produce en la naturaleza). En una realización, la secreción de la proteína en el medio de cultivo se controla mediante un péptido señal (por ejemplo, el péptido señal del factor α de la levadura, el péptido señal KILM1 de la levadura, el péptido señal PHO1 de la levadura o el péptido señal SUC2 de la levadura) incorporado en el vector de expresión y que es capaz de dirigir la secreción de la proteína expresada fuera de la célula de levadura. En otras realizaciones, los expertos en la materia conocen otros péptidos señal adecuados para facilitar la secreción de la proteína desde las células de levadura. En un aspecto, el péptido señal se escinde normalmente de la proteína después de la secreción.

20 En diversas realizaciones, puede usarse cualquier vector de expresión conocido por el experto en la materia (por ejemplo, un vector que se replique de forma autónoma o se integre en el genoma hospedador) y compatible con un sistema de expresión de levadura. Como se usa en el presente documento, el término "vector" significa cualquier plásmido, u otro vector, en forma bicatenaria o monocatenaria o en forma lineal o circular que puede transformar una célula de levadura mediante la integración en el genoma de la célula de levadura o mediante la existencia extracromosómica (por ejemplo, un plásmido de replicación autónoma). Como se sabe en la técnica, un vector (por ejemplo, vector de expresión o casete de expresión) es una construcción de ácido nucleico utilizada para transformar una célula hospedadora para la expresión de una proteína, polipéptido o péptido y el vector no se encuentra en la naturaleza en la célula hospedadora que transforma.

30 En una realización, el vector de expresión tiene sitios de escisión de endonucleasas de restricción para la inserción de fragmentos de ADN (por ejemplo, uno o más sitios de clonación y/o un sitio de clonación múltiple) y marcadores genéticos para la selección de transformantes. Por ejemplo, los marcadores genéticos para la selección de transformantes pueden incluir un marcador de selección que permita que una levadura transformada crezca en un medio desprovisto de un nutriente necesario que no puede ser producido por una cepa deficiente, un marcador de selección que codifique una enzima para la que se conocen sustratos cromógenos o un marcador de selección que proporcione resistencia a un fármaco, incluyendo, pero no limitado a, G418, Nurseotricina (Nat), Zeocina, Blastidina o Higromicina. En otra realización, el vector de expresión tiene una secuencia de terminación para la terminación de la transcripción (por ejemplo, el terminador AOX 1 o HSP150). En otra realización, el vector de expresión tiene una región 3' no traducida corriente abajo de la secuencia de codificación de la proteína con un sitio de poliadenilación. Como se usa en el presente documento, "región no traducida 3'" significa secuencias de nucleótidos que no se traducen en proteínas y se ubican corriente abajo de una secuencia codificante de una proteína. Normalmente, una región no traducida 3' incluye secuencias reguladoras para el procesamiento de ARNm. En otra realización, el vector de expresión tiene un origen de replicación (por ejemplo, un origen de replicación bacteriano) para su uso en la síntesis y la amplificación del vector, por ejemplo, en un hospedador bacteriano. Se describen diversos vectores de expresión en la Patente de los EE.UU. N.º 6.451.572, 6.841.370, 6.974.690, 7.320.876, 7.078.035, 7.138.260 y la Publicación PCT N.º WO 2007/112739. La construcción y el uso de vectores de expresión se describen en Sambrook et al., "*Molecular Cloning: A Laboratory Manual*", tercera edición, Cold Spring Harbor Laboratory Press, (2001). En otra realización, el vector de expresión o un fragmento del mismo, puede sintetizarse *de novo* o puede usarse PCR para amplificar y unir secciones de vectores de expresión.

50 Como se usa en el presente documento, "secuencias reguladoras" significa secuencias de nucleótidos que se encuentran normalmente corriente arriba o corriente abajo desde el extremo 5' o 3', respectivamente, de una secuencia codificante de proteína. Las secuencias reguladoras no se traducen en proteínas. Las secuencias reguladoras incluyen, pero no se limitan a, secuencias que afectan al procesamiento o a la estabilidad del ARN, tales como secuencias de señal de poliadenilación, potenciadores, sitios de unión a represores y promotores.

60 En una realización, la secuencia codificante de proteínas puede estar unida operativamente en el vector de expresión a la secuencia promotora capaz de dirigir la expresión de la proteína, por ejemplo, en la levadura. Como se usa en el presente documento, "unido operativamente" significa unido funcionalmente. Como se describe en el presente documento, el promotor puede ser un promotor constitutivo o inducible, tal como un promotor inducible por metanol. Como se usa en el presente documento, un "promotor constitutivo" significa un promotor que regula la expresión de un gen de interés. La expresión "promotor constitutivo" es conocida en la técnica. Como se usa en el presente documento, un "promotor inducible" significa un promotor regulado que se activa en una célula mediante un estímulo externo, tal como un producto químico, luz, un cambio en la temperatura, un cambio en la densidad celular o una proteína, tal como una hormona. Un promotor inducible por metanol puede tener cierta actividad constitutiva,

aunque un promotor inducible por metanol tiene actividad máxima cuando se induce en presencia de metanol. Análogamente, un promotor constitutivo puede tener cierta actividad inducible, pero un "promotor constitutivo" como se usa en el presente documento no tiene actividad máxima cuando se induce en presencia de metanol.

5 Como se usa en el presente documento, "promotor" significa una secuencia de nucleótidos que se encuentra normalmente corriente arriba desde el extremo 5' de una secuencia codificante de una proteína que controla la transcripción de ARN a partir de ADN, en parte, mediante la interacción con diversos factores reguladores que controlan la transcripción. En una realización, el promotor puede derivar de la misma especie de levadura que la célula hospedadora de levadura utilizada para la expresión de proteínas. En otra realización, el promotor puede derivar de una especie de levadura diferente de la célula hospedadora de levadura utilizada para la expresión de proteínas. En una realización, un promotor puede incluir una secuencia de caja TATA que actúa como un sitio de reconocimiento para dirigir el inicio de la transcripción, incluyendo, pero no limitado a, uno o más elementos potenciadores de la transcripción. Los elementos potenciadores pueden estar proximales o distales con respecto a la secuencia de la caja TATA y pueden estar en una orientación normal de 5' a 3' o pueden estar en una orientación de 3' a 5'. En otra realización, un elemento potenciador puede ser un elemento potenciador nativo de la secuencia promotora o puede ser un elemento potenciador heterólogo insertado en la construcción del vector de expresión. Un "elemento potenciador" como se usa en el presente documento es un elemento regulador que puede estimular la actividad del promotor.

20 En diversas realizaciones ilustrativas que se describen en el presente documento, el promotor puede ser un ácido nucleico aislado donde la secuencia del ácido nucleico aislado comprende una secuencia idéntica en al menos el 90 %, al menos el 92 %, al menos el 95 %, al menos el 96 %, al menos el 97 %, al menos el 98 % o al menos el 99 % a la secuencia de la SEQ ID NO: 2, o idéntica en al menos el 90 %, al menos el 92 %, al menos el 95 %, al menos el 96 %, al menos el 97 %, al menos el 98 % o al menos el 99 % a un fragmento de la misma que tiene actividad promotora, donde el ácido nucleico aislado comprende la secuencia de un promotor de *Pichia pastoris* inducible por metanol. En otra realización, la secuencia de ácido nucleico aislada es la secuencia de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, donde el ácido nucleico aislado comprende la secuencia de un promotor de *Pichia pastoris* inducible por metanol.

30 Como se usa en el presente documento, "un ácido nucleico aislado" significa un ácido nucleico que está sustancialmente libre de secuencias que flanquean naturalmente el ácido nucleico en el ADN genómico del organismo del que deriva el ácido nucleico. Por ejemplo, en diversas realizaciones, un ácido nucleico aislado de acuerdo con la invención puede contener menos de aproximadamente 2 kb, menos de aproximadamente 1 kb, menos de aproximadamente 0,5 kb, menos de aproximadamente 0,1 kb de secuencias de nucleótidos, menos de aproximadamente 0,05 kb, o ninguna secuencia de nucleótidos que flanquee naturalmente la molécula de ácido nucleico en el ADN genómico del organismo del que deriva el ácido nucleico aislado.

40 Como se usa en el presente documento, "un fragmento de la misma" cuando se refiere a la molécula de ácido nucleico aislada significa un fragmento del ácido nucleico aislado de la SEQ ID NO: 2. En diversas realizaciones ilustrativas, el fragmento puede tener aproximadamente 50 nucleótidos de longitud, aproximadamente 100 nucleótidos de longitud, aproximadamente 200 nucleótidos de longitud, aproximadamente 300 nucleótidos de longitud, aproximadamente 400 nucleótidos de longitud, aproximadamente 500 nucleótidos de longitud, aproximadamente 600 nucleótidos de longitud, aproximadamente 700 nucleótidos de longitud, aproximadamente 800 nucleótidos de longitud o aproximadamente 900 nucleótidos de longitud. En otras realizaciones, el fragmento puede prolongarse aproximadamente 50 nucleótidos, aproximadamente 100 nucleótidos, aproximadamente 200 nucleótidos, aproximadamente 300 nucleótidos, aproximadamente 400 nucleótidos, aproximadamente 500 nucleótidos, aproximadamente 600 nucleótidos, aproximadamente 700 nucleótidos, aproximadamente 800 nucleótidos o aproximadamente 900 nucleótidos corriente arriba desde el extremo 3' del ácido nucleico aislado de la SEQ ID NO: 2. En otras realizaciones más, el fragmento puede incluir aproximadamente 50 nucleótidos, aproximadamente 100 nucleótidos, aproximadamente 200 nucleótidos, aproximadamente 300 nucleótidos o aproximadamente 350 nucleótidos corriente arriba y/o corriente abajo de la secuencia de la caja TATA que se encuentra en la SEQ ID NO: 2.

55 En diversas realizaciones, los ácidos nucleicos aislados que se describen en el presente documento pueden purificarse mediante técnicas para la purificación de ácidos nucleicos (por ejemplo, ADN) que son bien conocidas en la técnica. Por ejemplo, los ácidos nucleicos pueden separarse de los contaminantes mediante métodos físicos incluyendo, pero no limitados a, centrifugación, técnicas de presión o mediante el uso de una sustancia con afinidad por los ácidos nucleicos (por ejemplo, ADN), tal como, por ejemplo, perlas de sílice. Después de un lavado suficiente, los ácidos nucleicos aislados pueden suspenderse en agua o en un tampón. En otras realizaciones, hay kits comerciales disponibles, tales como Qiagen™, Nuclisens™ y Wizard™ (Promega) y Promegam™. Se describen métodos para purificar ácidos nucleicos en Sambrook et al., "Molecular Cloning: A Laboratory Manual", 3ª Edición, Cold Spring Harbor Laboratory Press, (2001). Un ácido nucleico aislado como se describe en el presente documento puede ser un ácido nucleico aislado que también se purifica o el ácido nucleico aislado puede ser impuro. Un "ácido nucleico purificado" está sustancialmente libre de otro material celular o medio de cultivo cuando se produce mediante técnicas recombinantes, o está sustancialmente libre de precursores químicos u otros productos químicos cuando se sintetiza químicamente.

Los ácidos nucleicos aislados que se describen en el presente documento son susceptibles de hibridación específica, en condiciones de hibridación apropiadas (por ejemplo, tampón apropiado, fuerza iónica, temperatura, formamida y concentraciones de $MgCl_2$), con un ácido nucleico complementario. Los ácidos nucleicos aislados que se describen en el presente documento pueden modificarse mediante sustitución, supresión, truncamiento y/o pueden fusionarse con otras moléculas de ácido nucleico donde los ácidos nucleicos aislados resultantes se hibridan específicamente con los ácidos nucleicos complementarios.

Dentro del alcance de la invención también están los ácidos nucleicos complementarios a los ácidos nucleicos aislados, o fragmentos de los mismos, que se describen en el presente documento, y aquellos que se hibridan con los ácidos nucleicos aislados que se describen en el presente documento o aquellos que se hibridan con sus complementos en condiciones de alta rigurosidad. Como se usa en el presente documento, el término "complementario" se refiere a la capacidad de las secuencias de nucleótidos púricos y pirimidínicos para asociarse a través de enlaces de hidrógeno para formar moléculas de ácido nucleico bicatenario. La guanina y la citosina, la adenina y la timina, y la adenina y el uracilo son complementarios y pueden asociarse a través de enlaces de hidrógeno, dando como resultado la formación de moléculas de ácido nucleico bicatenario cuando dos moléculas de ácido nucleico tienen secuencias "complementarias". Las secuencias de ADN complementarias se denominan "complemento".

De acuerdo con la invención, "condiciones de alta rigurosidad" significa hibridación a 65 °C en SSPE 5X y formamida al 50 %, y lavado a 65 °C en SSPE0,5X. Se describen condiciones de alta rigurosidad para la hibridación en Sambrook et al., "*Molecular Cloning: A Laboratory Manual*", 3ª edición, Cold Spring Harbor Laboratory Press, (2001). En algunos aspectos ilustrativos, la hibridación puede producirse en toda la longitud del ácido nucleico aislado o a lo largo de parte de su longitud o en un fragmento del mismo.

También se incluyen moléculas de ácido nucleico aisladas que tienen una identidad de aproximadamente el 90 %, aproximadamente el 92 %, aproximadamente el 95 %, el 96 %, el 97 %, el 98 % y el 99 % con los ácidos nucleicos aislados que se describen en el presente documento, o con un fragmento de los mismos que tienen actividad promotora. La determinación del porcentaje de identidad o similitud entre secuencias puede realizarse, por ejemplo, usando el programa GAP (Genetics Computer Group, software; ahora disponible a través de Accelrys en <http://www.accelrys.com>) y las alineaciones pueden hacerse usando, por ejemplo, el algoritmo ClustalW (software VNTI, InforMax Inc.). Se puede buscar en una base de datos de secuencias usando la secuencia del ácido nucleico aislado de interés o la secuencia de un fragmento del mismo. Los algoritmos para la búsqueda en bases de datos se basan normalmente en el software BLAST (Altschul et al., 1990). En algunas realizaciones, el porcentaje de identidad puede determinarse en toda la longitud del ácido nucleico aislado.

Las técnicas para sintetizar los ácidos nucleicos aislados que se describen en el presente documento son bien conocidas en la técnica e incluyen síntesis química y métodos recombinantes. Dichas técnicas se describen en Sambrook et al., "*Molecular Cloning: A Laboratory Manual*", 3ª edición, Cold Spring Harbor Laboratory Press, (2001). Las moléculas de ácido nucleico aisladas también pueden prepararse comercialmente. Las técnicas para sintetizar los ácidos nucleicos aislados que se describen en el presente documento son bien conocidas en la técnica. Los ácidos nucleicos aislados que se describen en el presente documento pueden analizarse mediante técnicas conocidas en la técnica, tales como análisis de enzimas de restricción o secuenciación, para determinar la secuencia de los ácidos nucleicos aislados. Por tanto, los ácidos nucleicos aislados que se describen en el presente documento pueden ser sintéticos.

En aspectos ilustrativos, las células de levadura se transforman con un vector de expresión que comprende un ácido nucleico heterólogo que codifica la proteína de interés unida operativamente a uno de los promotores que se describen en el presente documento usando procedimientos bien conocidos por los expertos en la materia. El término "transformación" significa la transferencia de un ácido nucleico o un fragmento de ácido nucleico, a una célula hospedadora. En realizaciones ilustrativas, dichos protocolos de transformación incluyen electroporación, métodos de acetato de litio y uso de esferoplastos. En aspectos ilustrativos, la secuencia codificante de ácido nucleico expresada puede ser una secuencia codificante de ácido nucleico heteróloga. Como se usa en el presente documento, una secuencia codificante heteróloga se define como un ácido nucleico artificial o sintético o un ácido nucleico que se origina a partir de una especie diferente a la especie de la que derivó la secuencia promotora. Por tanto, una secuencia codificante heteróloga unida a un promotor que se describe en el presente documento no se produce en la naturaleza.

En diversas realizaciones, las células de levadura transformadas pueden cultivarse mediante técnicas incluyendo la fermentación discontinua y continua en un medio líquido o en un medio semisólido, o un medio sólido. Normalmente, "las condiciones que permiten la expresión de la proteína" como se usa en el presente documento significan condiciones para la fermentación discontinua o continua de levadura en un medio líquido, pero no se excluye el crecimiento en un medio semisólido, tal como el agar. Los medios de cultivo para células de levadura son conocidos en la técnica y se complementan normalmente con una fuente de carbono (por ejemplo, glucosa). Un medio de cultivo de levadura típico es el caldo YPD (Sunrise Science Products, Inc.) que comprende extracto de levadura (10 gramos), peptona Bacto (20 gramos) y dextrosa (20 gramos). En un aspecto ilustrativo, las células de levadura transformadas pueden cultivarse aeróbicamente a 30 °C en un entorno de pH controlado (un pH de

aproximadamente 6) y con la fuente de carbono (por ejemplo, glucosa) mantenida continuamente a un nivel predeterminado conocido para respaldar el crecimiento de las células de levadura a una densidad deseada dentro de un período específico de tiempo.

5 En una realización ilustrativa, se proporciona un método de producción de una proteína. El método comprende la etapa de cultivar en un medio de cultivo una célula hospedadora que comprende un primer casete de expresión que comprende un ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, unido operativamente a una secuencia codificante heteróloga que codifica una proteína, donde el cultivo se realiza en condiciones que permiten la expresión de la proteína. El método puede comprender adicionalmente la
10 etapa de purificar la proteína del medio de la célula hospedadora cultivada. Como se usa en el presente documento, un "casete de expresión" significa los elementos de un vector de expresión que dirigen a la célula de levadura a producir ARN. Un casete de expresión comprende al menos secuencias reguladoras (por ejemplo, un promotor) y una secuencia codificante del ARN y la proteína (es decir, un marco abierto de lectura). El ácido nucleico aislado puede ser homólogo en al menos el 90 %, el 92 %, el 95 %, el 96 %, el 97 %, el 98 % o el 99 % al ácido nucleico
15 aislado de la SEQ ID NO: 2 o un fragmento de la misma que tenga actividad promotora. En diversos aspectos ilustrativos, la secuencia codificante de proteína puede ser de una bacteria, una levadura, un hongo o un virus.

En esta realización del método, la proteína puede expresarse usando el primer casete de expresión en combinación con un segundo casete de expresión. En otra realización, el segundo casete de expresión puede comprender 1) la
20 secuencia de codificación heteróloga que codifica la proteína unida operativamente a un ácido nucleico aislado que tiene una secuencia que comprende la secuencia de la SEQ ID NO: 15 o la SEQ ID NO: 16 donde la SEQ ID NO: 15 y la SEQ ID NO: 16 tienen actividad promotora, o cualquier otro promotor regulado por metanol conocido, tal como las secuencias promotoras AOX 1, AOX 2, FLD o DAS, o 2) el ácido nucleico aislado de la SEQ ID NO: 2, o un
25 fragmento de la misma que tenga actividad promotora, unido operativamente a la secuencia codificante heteróloga que codifica la proteína.

En otra realización más, la proteína puede expresarse usando el primer casete de expresión, el segundo casete de expresión y un tercer casete de expresión. En otro aspecto ilustrativo, el tercer casete de expresión puede comprender 1) la secuencia codificante heteróloga que codifica la proteína unida operativamente a un ácido nucleico
30 aislado que tiene una secuencia que comprende la secuencia de la SEQ ID NO: 15 o la SEQ ID NO: 16 donde la SEQ ID NO: 15 y la SEQ ID NO: 16 tienen actividad promotora, o cualquier otro promotor conocido regulado por metanol, tal como las secuencias promotoras AOX 1, AOX 2, FLD o DAS, o 2) el ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, unido operativamente a la secuencia codificante heteróloga que codifica la proteína.
35

En otra realización, puede usarse cualquier número de casetes de expresión adicionales y los casetes de expresión pueden comprender 1) la secuencia codificante heteróloga que codifica la proteína unida operativamente a un ácido nucleico aislado que tiene una secuencia que comprende la secuencia de la SEQ ID NO: 15 o la SEQ ID NO: 16 donde la SEQ ID NO: 15 y la SEQ ID NO: 16 tienen actividad promotora, o cualquier otro promotor conocido regulado por metanol, tal como las secuencias promotoras AOX 1, AOX 2, FLD o DAS, o 2) el ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, unido operativamente a la
40 secuencia codificante heteróloga que codifica la proteína. En otra realización, el primer casete de expresión, el segundo casete de expresión y el tercer casete de expresión como se han descrito anteriormente se usan en el método, junto con un cuarto casete de expresión, un quinto casete de expresión y un sexto casete de expresión donde todos los casetes de expresión comprenden el ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, unido operativamente a la secuencia de codificación heteróloga que
45 codifica la proteína.

En otra realización más, se proporciona un método de producción de una o más proteínas. El método comprende la etapa de cultivar en un medio de cultivo una célula hospedadora que comprende un primer casete de expresión, un segundo casete de expresión y, opcionalmente, uno o más casetes de expresión adicionales, donde cada uno de los casetes de expresión comprende 1) una secuencia de codificación heteróloga que codifica la una o más proteínas unidas operativamente a un ácido nucleico aislado que tiene una secuencia que comprende la secuencia de la SEQ ID NO: 15 o la SEQ ID NO: 16 donde la SEQ ID NO: 15 y la SEQ ID NO: 16 tienen actividad promotora, o cualquier
50 otro promotor regulado por metanol, tal como las secuencias promotoras AOX 1, AOX 2, FLD o DAS, o 2) el ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, unido operativamente a una secuencia codificante heteróloga que codifica una o más proteínas, donde el cultivo se realiza en condiciones que permiten la expresión de una o más proteínas. El método puede comprender adicionalmente la etapa de purificar una de las una o más proteínas del medio de la célula hospedadora cultivada.
55
60

En cualquiera de las realizaciones descritas en los dos párrafos anteriores, el ácido nucleico aislado puede ser homólogo en al menos el 90 %, el 92 %, el 95 %, el 96 %, el 97 %, el 98 % o el 99 % al ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora. En cualquiera de las realizaciones descritas en los dos párrafos anteriores, los casetes de expresión pueden incluirse en un vector de expresión o en
65 múltiples vectores de expresión. En cualquiera de las realizaciones descritas en los dos párrafos anteriores, los vectores de expresión en los que se incorporan los casetes de expresión pueden ser vectores que se replican de

forma autónoma o que se integran en el genoma de la célula hospedadora.

En diversas realizaciones ilustrativas, la proteína codificada por cualquiera de las secuencias codificantes heterólogas que se describen en el presente documento unidas operativamente al promotor puede ser una proteína seleccionada entre el grupo que consiste en una toxina, un anticuerpo, una hormona, una enzima, un factor de crecimiento, una citocina, una proteína estructural, una proteína inmunógena (por ejemplo, un antígeno de vacuna) y una proteína de señalización celular. En una realización, la proteína es una enzima para su uso en piensos para animales. En la presente realización, la proteína puede seleccionarse entre el grupo que consiste en una fitasa, una mananasa, una galactosidasa, una amilasa, una glucanasa, una celulasa, una proteasa y una xilanasa. En otra realización, la proteína puede ser una enzima útil para la glucosilación. En otro aspecto, puede expresarse más de una proteína usando múltiples casetes de expresión, cada uno con una secuencia de codificación heteróloga, unida operativamente a un promotor. Sin embargo, estos ejemplos de proteínas no son limitantes y puede expresarse cualquier proteína, polipéptido o péptido capaz de expresarse en levadura de acuerdo con los ácidos nucleicos aislados, vectores de expresión, células hospedadoras, construcciones de ADN, métodos y proteínas aisladas descritos en el presente documento. Las enzimas descritas anteriormente pueden ser de cualquier especie (por ejemplo, especies de hongos, tales como una especie de levadura).

Las proteínas expresadas en levadura para su uso de acuerdo con la presente invención pueden producirse en forma purificada mediante técnicas convencionales. Como se usa en el presente documento, "proteína aislada" significa una proteína purificada. Una proteína purificada está sustancialmente libre de otros contaminantes de células de levadura o contaminantes del medio de cultivo. Por ejemplo, "sustancialmente libre" de otros contaminantes de células de levadura o contaminantes del medio de cultivo significa que la proteína es pura en al menos aproximadamente un 60 %, pura en al menos aproximadamente un 70 %, pura en al menos aproximadamente un 80 %, pura en al menos aproximadamente un 90 %, pura en al menos aproximadamente el 95 %, pura en al menos aproximadamente el 98 %, pura en aproximadamente el 60 %, pura en aproximadamente el 70 %, pura en aproximadamente el 80 %, pura en aproximadamente el 90 %, pura en aproximadamente el 95 % o pura en aproximadamente el 98 % (todo basado en seco peso). Normalmente, la proteína se secreta en el medio de cultivo de levadura y se recoge del medio de cultivo.

En una realización ilustrativa, para la purificación a partir del medio de cultivo, la proteína puede, por ejemplo, someterse a precipitación con sulfato de amonio seguida de cromatografía en columna con DEAE-Sefarosa. En otras realizaciones, pueden usarse técnicas convencionales conocidas por los expertos en la materia, tales como precipitación con sulfato de amonio o etanol, extracción con ácido, filtración en gel, cromatografía de intercambio aniónico o catiónico, cromatografía en columna de DEAE-Sefarosa, cromatografía en hidroxapatita, cromatografía en lectina, cromatografía de afinidad, extracción disolvente-disolvente, ultrafiltración y HPLC.

Como alternativa, las etapas de purificación pueden no ser necesarias porque la proteína puede estar presente en concentraciones tan altas en el medio de cultivo que la proteína esté esencialmente pura en el medio de cultivo (por ejemplo, del 70 al 80 % de pureza). En una realización, la proteína se recoge del medio de cultivo sin etapas de purificación adicionales enfriando el cultivo de levadura (por ejemplo, a aproximadamente 4 °C a aproximadamente 8 °C) y retirando las células de levadura usando técnicas tales como centrifugación, microfiltración y filtración al vacío rotativa. La proteína en el medio libre de células puede entonces concentrarse mediante técnicas tales como, pero no limitadas a, ultrafiltración y filtración por flujo tangencial.

En algunas realizaciones en las que la proteína no se secreta en el medio de cultivo, las células de levadura pueden lisarse, por ejemplo, mediante tratamiento químico, térmico o por ultrasonidos, y el homogeneizado puede centrifugarse para retirar los restos celulares. El sobrenadante puede después someterse a precipitación con sulfato de amonio y técnicas de fraccionamiento adicionales según sea necesario, tales como filtración en gel, cromatografía de intercambio iónico, cromatografía en columna de DEAE-Sefarosa, cromatografía de afinidad, extracción con disolvente-disolvente, ultrafiltración y HPLC para purificar la proteína. Debe entenderse que los métodos de purificación descritos anteriormente para la purificación de proteínas del medio de cultivo o de células de levadura lisadas no son limitantes y puede usarse técnica de purificación conocida por los expertos en la materia para purificar la proteína expresada en levadura si se requieren dichas técnicas para obtener una proteína sustancialmente pura.

Pueden prepararse diversas formulaciones de las preparaciones de proteína purificada de acuerdo con la invención. En algunas realizaciones, las proteínas pueden estabilizarse a través de la adición de otras proteínas (por ejemplo, gelatina y leche en polvo desnatada), agentes químicos (por ejemplo, glicerol, polietilenglicol, EDTA, sorbato de potasio, benzoato de sodio y agentes reductores y aldehídos), polisacáridos, monosacáridos, lípidos (aceites vegetales hidrogenados) y similares. En una realización, las proteínas para la adición a productos alimentarios o mezclas de piensos para animales pueden secarse (por ejemplo, secado por pulverización, secado en tambor y liofilización) y pueden formularse en forma de polvos, gránulos, píldoras, bloques minerales, líquidos y geles a través de procesos conocidos. En una realización, pueden usarse agentes gelificantes tales como gelatina, alginato, colágeno, agar, pectina y carragenina.

En realizaciones alternativas, la expresión de la proteína puede ser para la expresión intracelular, tal como para la

acción enzimática en la levadura en un proceso de biotransformación o para la presentación en la superficie de la célula de levadura. Para dichas realizaciones, la proteína, expresada como se describe en el presente documento, no se purifica.

5 En diversas realizaciones, las proteínas descritas anteriormente se seleccionan entre el grupo que consiste en una toxina, un anticuerpo, una hormona, una enzima, un factor de crecimiento, una citocina, una proteína estructural, una proteína inmunógena (por ejemplo, un antígeno de vacuna) y una proteína de señalización celular. En otra
10 realización, las proteínas descritas anteriormente se seleccionan entre el grupo que consiste en un anticuerpo, una hormona, una enzima, un factor de crecimiento, una citocina, una proteína estructural, una proteína inmunógena (por ejemplo, un antígeno de vacuna) y una proteína de señalización celular. En otra realización más, de acuerdo con la invención, puede usarse la secuencia codificante de una proteína, un fragmento de la misma, una proteína de fusión (por ejemplo, una proteína quimérica) o un péptido. En otra realización, puede expresarse una proteína modificada, tal como una proteína mutada o una proteína con aminoácidos no naturales.

15 En una realización, las toxinas pueden ser proteínas tales como, pero no limitadas a, toxina botulínica o verotoxina-1 y, después de la preparación usando los métodos, ácidos nucleicos aislados, vectores de expresión, células hospedadoras y construcciones de ADN descritos en el presente documento, las toxinas pueden modificarse usando un agente de direccionamiento de manera que se dirijan específicamente a las células enfermas. En otro aspecto ilustrativo, el anticuerpo puede ser un anticuerpo humanizado, un anticuerpo que no está humanizado, un
20 nanocuerpo o un fragmento de anticuerpo, tal como un fragmento Fab de un anticuerpo o un anticuerpo monocatenario. En otra realización, la hormona puede ser, por ejemplo, una gonadotropina, una hormona adrenocorticotrófica, una hormona de crecimiento, vasopresina, oxitocina, somatostatina, gastrina o leptina. En otra realización ilustrativa, el factor de crecimiento puede ser insulina, factor de crecimiento epidérmico, factor de crecimiento de fibroblastos, factor de crecimiento endotelial vascular, eritropoyetina, factor de crecimiento derivado de plaquetas, trombopoyetina o una proteína morfogénica ósea. En un aspecto, la citocina puede ser IL-2, IFN- α , IFN- γ o GM-CSF. En otro aspecto ilustrativo, las proteínas de la vacuna pueden ser cualquier proteína de vacuna adecuada que sea inmunógena en un paciente o un animal, incluyendo, entre otras, proteínas de VPH (por ejemplo, VPH 16 y VPH 18) y proteínas de vacuna contra el tétanos, como ejemplos. En otra realización ilustrativa, las
25 enzimas pueden ser, por ejemplo, enzimas para piensos para animales como se analiza en el presente documento, acetilcolinesterasa, o ciclooxigenasa, o cualquier otra enzima útil que pueda expresarse en levadura. En otra realización, pueden expresarse proteínas estructurales, por ejemplo, netrin, proteínas de unión a actina o miosina y, en otra realización, pueden expresarse proteínas de señalización celular tales como proteínas *ras*, cinasas, la proteína ErbB2 (el receptor Her-2) usando los métodos, ácidos nucleicos aislados, vectores de expresión, células hospedadoras y construcciones de ADN descritos en el presente documento.

35 En una realización, la proteína es una enzima para su uso en piensos para animales. En la presente realización, la proteína puede seleccionarse entre el grupo que consiste en una fitasa, una mananasa, una galactosidasa, una amilasa, una glucanasa, una celulasa, una proteasa y una xilanasas o una combinación de las mismas. Por ejemplo, puede expresarse una diversidad de fitasas de acuerdo con los métodos que se describen en el presente
40 documento. Son ejemplos de genes de fitasa (es decir, una secuencia codificante de fitasa) que pueden expresarse de acuerdo con la invención genes de fitasa derivados de bacterias, hongos filamentosos, plantas y levaduras, tales como los genes *appA* (número de acceso de Gene Bank M58708) y *appA2* (número de acceso de Gene Bank 250016) derivados de *Escherichia coli* y los genes *phyA* y *phyB* derivados del hongo *Aspergillus niger*, o cualquier mutante de estos genes que conserve o tenga actividad mioinositol hexaquisfosfato fosfohidrolasa mejorada (véase, por ejemplo, Rodríguez et al., *Arch. of Biochem y Biophys.* 382: 105-112 (2000)). También pueden expresarse de
45 acuerdo con la invención genes de fitasa sustituidos, suprimidos y truncados o un fragmento de los mismos.

En una realización, la proteína expresada usando los métodos que se describen en el presente documento puede usarse en piensos para animales que comprenden una mezcla de pienso para animales. En diversas realizaciones,
50 puede usarse cualquier mezcla de pienso para animales conocida en la técnica, tal como harina de colza, harina de semilla de algodón, harina de soja y harina. Los ingredientes opcionales de la mezcla de pienso para animales incluyen azúcares e hidratos de carbono complejos tales como monosacáridos, disacáridos y polisacáridos hidrosolubles y no hidrosolubles. Son ingredientes aminoácidos opcionales que pueden añadirse a la mezcla de pienso arginina, histidina, isoleucina, leucina, lisina, metionina, fenilalanina, treonina, triptófano, valina, etil HCl de tirosina, alanina, ácido aspártico, glutamato de sodio, glicina, prolina, serina, etil HCl de cisteína y análogos y sales de los mismos. Son vitaminas que pueden añadirse opcionalmente HCl de tiamina, riboflavina, HCl de piridoxina, niacina, niacinamida, inositol, cloruro de colina, pantotenato de calcio, biotina, ácido fólico, ácido ascórbico y
55 vitaminas A, B, K, D, E y similares. También pueden añadirse minerales, ingredientes proteínicos, incluyendo proteína obtenida a partir de harina de carne o harina de pescado, huevo líquido o en polvo, solubles de pescado, concentrado de proteína de suero, aceites (por ejemplo, aceite de soja), almidón de maíz, calcio, fosfato inorgánico,
60 sulfato de cobre, sal y piedra caliza. También pueden añadirse antioxidantes.

Ejemplos

Ejemplo 1

5 NIVELES DE EXPRESION DE PROMOTOR

Tabla 1. Muestras de fermentación

Gen	Inanición de glicerol	6 h de alimentación con glicerol	2 h de inducción con metanol	6 h de inducción con metanol	18 h de inducción con metanol	48 h de inducción con metanol	66 h de inducción con metanol
AOX1	958	54	21806	17607	10035	581	911
AOX2	185	28	10035	1374	1697	64	119
CAM1	316	166	13201	13201	24457	36959	36959
FLD1	3405	1293	11173	10632	13473	21806	14995
GPM2	884	549	7280	8614	8357	14995	17607
PP7435_Chr1-1351	2102	149	12171	13473	14995	17607	11173
PP7435_Chr1-0269	7766	8636	3252	3772	6615	9483	5969
PP7435_Chr2-0207	13473	17607	1520	858	4496	3914	3742
PP7435_Chr2-0208	21806	28506	2298	1136	7155	7280	8949
PP7435_Chr2-0790	2457	356	7829	7502	6433	2064	822
PP7435_Chr2-0809	5566	4305	1668	4205	4799	6758	2899
PP7435_Chr3-0476	6758	9425	1648	5566	6104	10632	10035
PP7435_Chr3-0842	243	134	8614	6987	5684	7155	12171
PP7435_Chr4-0069	7502	12171	1746	1162	7829	9425	13473
PP7435_Chr4-0800	5064	5916	2382	5415	5415	6987	3964
SPI1	275	203	673	376	267	164	195
TDH1	8949	841	2102	951	2965	813	104
TEF2	7155	2248	4305	7829	10632	3246	757
THI4	1668	68	14995	12171	2542	429	419

* Todos los números como transcritos por millón

10 **Tabla 2. Muestras en matraz de agitación**

Gen	Cepa 1: matraz de agitación con glicerol	Cepa 2: matraz de agitación con glicerol	Cepa 3: matraz de agitación con glicerol	Cepa 4: matraz de agitación con glicerol	Cepa 5: inducción con metanol
AOX1	7	8	24	22	3167
AOX2	12	16	9	8	109
CAM1	537	513	661	789	24110
FLD1	762	743	357	1202	17607
GPM2	303	352	699	561	9425
PP7435_Chr1-1351	9	10	15	19	10035
PP7435_Chr1-0269	5587	5535	11983	10362	6375
PP7435_Chr2-0207	2226	2152	3466	3246	680
PP7435_Chr2-0208	3167	3121	7502	4402	864

Gen	Cepa 1: matraz de agitación con glicerol	Cepa 2: matraz de agitación con glicerol	Cepa 3: matraz de agitación con glicerol	Cepa 4: matraz de agitación con glicerol	Cepa 5: inducción con metanol
PP7435_Chr2-0790	41	48	31	26	1877
PP7435_Chr2-0809	6542	6433	2965	4305	6758
PP7435_Chr3-0476	28506	21806	28506	28506	7829
PP7435_Chr3-0842	14	16	48	44	1171
PP7435_Chr4-0069	4887	5684	8141	9425	6542
PP7435_Chr4-0800	21806	17607	10035	21806	7502
SPI1	2500	2035	4799	1529	136
TDH1	1770	5415	1177	1095	2654
TEF2	17607	28506	3914	3803	12171
THI4	10	58	40	30	452

* Todos los números como transcritos por millón

Ejemplo 2

5 CRECIMIENTO DE LA CEPA PARA EL AISLAMIENTO DE ARN PARA EXAMINAR LOS NIVELES DE EXPRESIÓN DEL PROMOTOR

La cepa BG10 se mantuvo en forma de parches sobre placas de agar YPD. Para el análisis del transcriptoma, se inocularon 50 ml de cultivos de la cepa de un parche y se cultivaron en BMGY a 30 °C (200 rpm) durante aproximadamente 16 horas. El cultivo estacionario se diluyó 100 veces en medio BMGY fresco y se cultivó a 30 °C (200 rpm) durante 6 horas. Este punto temporal se consideró crecimiento exponencial con glicerol como fuente de carbono. Las alícuotas se hicieron girar en tubos de 15 ml. Después de esto, los sobrenadantes se descartaron y los sedimentos de células se congelaron rápidamente en nitrógeno líquido. Los sedimentos de células se almacenaron a -80 °C para el aislamiento posterior del ARN total.

Para el análisis por secuenciación de ARN durante la fermentación, un cultivo en matraz de agitación cultivado como se ha descrito anteriormente se expandió en un fermentador de 1 litro usando condiciones convencionales de inducción por metanol. Se tomaron muestras celulares en diversos puntos temporales antes, durante y al final de la inducción por metanol. Los sedimentos de células se recogieron y se congelaron en nitrógeno líquido. Los sedimentos de células se almacenaron a -80 °C para el aislamiento posterior del ARN total.

Ejemplo 3

25 AISLAMIENTO DE ARN TOTAL

Se usaron kits FastRNA SPIN (MP Bio) para aislar el ARN total. La lisis celular se realizó usando un BioSpec Mini-Beadbeater 96. El ARN total se eluyó de la columna de centrifugación en 15 µl de agua sin ARNasa/ADNasa, se congeló en nitrógeno líquido y se almacenó a -80 °C. Las muestras de ARN se enviaron en hielo seco para el análisis por secuenciación de ARN en una máquina Illumina HiSeq. Las muestras de ARN se analizaron con un Agilent BioAnalyzer y todas mostraron picos de ARN ribosómico de levadura intacto.

Ejemplo 4

35 GENERACIÓN Y SECUENCIACIÓN DE BIBLIOTECAS DE ARN

Se prepararon bibliotecas de ARNm usando reactivos Illumina. Se usó un kit de preparación de muestras de ARN TruSeq para generar selectivamente ADNc codificado en barras a partir de poliA ARN. Después de la codificación en barras y la amplificación, se agruparon un total de 12 muestras (8 muestras de fermentación, 4 muestras de matraz de agitación) para el análisis. Se realizaron cincuenta lecturas de bases, de un solo extremo. Los datos se suministraron a BioGrammatics en formato FASTQ convencional. Las lecturas se recortaron según las bases ambiguas y la puntuación de calidad y, después, se filtraron para eliminar todas las lecturas recortadas que tuvieran menos de 40 bases de longitud. Aproximadamente se retiró el 0,3 % de las lecturas de cada conjunto de datos.

Ejemplo 5

ANÁLISIS POR SECUENCIACIÓN DE ARN

5 Se importaron conjuntos de datos de una máquina Illumina HiSeq a CLC Genomics Workbench (versión 6). Se usaron las herramientas de software convencionales de CLC Genomics Workbench para el análisis por secuenciación de ARN. Las lecturas de ARN se asignaron a un genoma de *Pichia pastoris* de referencia. Se generaron perfiles de transcripción de 5202 genes anotados en los 12 conjuntos de datos de muestra. Basándose en los perfiles de transcripción, se identificaron genes con patrones de transcripción constitutivos o regulados por metanol.

Ejemplo 6

ANÁLISIS DE EXPRESIÓN DE PROTEÍNAS

15 Se insertaron múltiples promotores en un vector comprobador del promotor, (plásmido indicador, pJ-G-Agal, figura 17) para someter a ensayo la expresión de la proteína. Brevemente, se usa un marco abierto de lectura (ORF, por sus siglas en inglés) de alfa-galactosidasa (A-gal) en el pJ-G-A-gal para medir el nivel de expresión de un promotor insertado solo 5' del ORF A-gal. Se usaron cebadores de ADN seleccionados (IDT-DNA, tabla 3) para amplificar por PCR el ADN aislado del ADN genómico de *Pichia* (ADNg, cepa Bg10 de *Pichia pastoris* de tipo silvestre de BioGrammatics).

Tabla 3. Cebadores de PCR para promotores.

Promotor	SEQ ID NO:	Cebador directo	Secuencia del cebador	Posición del cebador en la secuencia del promotor
SAM1	1	11231-F685	NNNGGTCTCNATCCACGAGTTTCTGGACCGTATC	-685
SAM1	1	11231-F447	NNNGGTCTCTATCCGAAAACGTTAAGAGATG	-447
SAM1	1	11231-F3	NNNGGTCTCTATCCCTCTTACTAAATTGCCCAAGTG	-532
Chr1-1351	2	1351-F1	NNNGGTCTCNATCCGATGGAGACTCAGTATGATGGGGC	-606
Chr1-1351	2	1351-F2	NNNGGTCTCNATCCAGTATGATGGGGCAAGGAAAACG	-595
THI4	3	THI4-F1	NNNGGTCTCNATCCTGGAGACCCTTAACAGGTCCG	-404
GPM2	4	GPM2-F1	NNNGGTCTCNATCCGTTGGAACTGTGCCTG	-637
Chr2-0790	5	790-F1	NNNGGTCTCNATCCACAGTGGTAGGTCCAACCTGG	-543
Chr3-0842	6	842-F1	NNNGGTCTCNATCCGTAGTAGCCTCTCCAGCCTG	-635
Chr1-0269	7	269-F1	NNNGGTCTCNATCCTGAAGCCCCTGCAACTACAGAG	-611
Chr2-0207	8	207-F1	NNNGGTCTCNATCCGTAGACGACATCCAGAGAAGTAACAG	-632
Chr2-0208	9	208-F1	NNNGGTCTCNATCCTCAGGTCAGTCTTGAAGTCCTGAG	-618
Chr2-0809	10	809-F1	NNNGGTCTCNATCCTGTGGAATTCAAAGAAGGGG	-653
Chr4-0069	11	069-F1	NNNGGTCTCNATCCGTCCGTGATGTAAAATGAGACTAC	-467
Chr4-0800	12	800-F1	NNNGGTCTCNATCCAGTCAACTGGGAGCTACGGT	-490
TEF2	13	TEF2-FF1	NNNGGTCTCNATCCGATGTGAGGATGCGCTC	-729
Chr3-0476	14	476-F1	NNNGGTCTCNATCCTCAATGACCACGGTAACATGAAAAC	-650
GAP		GAP-F1	NNNGGTCTCNATCCAATGGACCAAATTGTTGCAAGGT	-619
DAS		07226F620	NNNGGTCTCCATCCCTTTGTTGAGCAACA	-382
DAS		07226F518	NNNGGTCTCGATCCGCCAAACGAACAG	-483
AOX		AOXF1	NNNGGTCTCCATCAAAGACGAAAGGTTGAATGA	-931

Se aisló ADNg de Bg10 rompiendo las células mediante agitación vigorosa con 0,5 mm de perlas de circona/sílice (BioSpec Products, Inc.) en presencia de fenol-cloroformo, antes de precipitación con EtOH y suspensión en una solución de Tris-EDTA (Tris 10 mM, EDTA 1 mM). La PCR se realizó durante 35 ciclos (98 °C durante 10 segundos, ~55 °C durante 10 segundos y 72 °C durante ~30 segundos) con una polimerasa correctora de errores mediante métodos convencionales y las recomendaciones de los fabricantes (New England BioLabs, NEB). Después de la purificación de los amplicones de PCR (aislamiento y extracción en gel, DNA Clean y Concentrator, Zymo-Research), después se usaron los sitios de enzimas de restricción ubicados estratégicamente en los cebadores para cortar y ligar cada amplicón de PCR purificado, independientemente, en el vector indicador pJ-G-A-gal (figura 17) mediante ligadura según lo recomendado por el fabricante (NEB). Se usaron métodos convencionales para aislar, amplificar y purificar cada uno de los plásmidos pJ-G-A-gal-promotor resultantes. En la mayoría de los casos, la "A" del codón de inicio del indicador es la posición +1 con respecto a la posición de los pares de bases aproximadamente -1 a -1500 del indicador (tal como se indica en la tabla 3) con los cebadores utilizados para amplificar los promotores.

En todos los casos, la secuencia de los promotores en estos plásmidos indicadores se determinó usando los cebadores 5' y 3' del sitio de inserción en el vector; la secuencia se obtuvo a través de todo el promotor y las uniones de clonación (Genewiz, Inc.) y se comparó con las secuencias de ARNm respectivas para confirmar la identidad del clon promotor. En todos los casos, la unión de clonación entre el promotor y el ORF A-gal se diseñó para ubicar el codón de inicio ATG de A-gal, de manera que se encontrara en la misma posición que el ATG previsto del codón de inicio del ORF nativo del promotor.

Las construcciones de promotor-indicador purificadas, verificadas de secuencia, se linealizaron para la transformación e integración en el genoma de Pichia de células de Pichia Bg10 por electroporación. Después de la digestión con enzimas de restricción, el ADN se limpió y se concentró a aproximadamente ~200 ng/ul antes de que cada construcción se transformara independientemente en Pichia electro-competente (Bg10, BioGrammatics, Inc.). Las células Bg10 se hicieron competentes mediante la incubación de células en fase logarítmica con DTT y lavados con sorbitol posteriores (Protocolos de Pichia). Después de la electroporación, se seleccionaron transformantes en YPD con G418 800 ug/ml y se incubaron a 30 °C durante ~3 días. Se parchearon células de colonias aisladas en placas de YPD-G418 similares y las células de estos parches se usaron para medir el nivel de expresión de los genes indicadores.

La expresión del ORF de A-gal, regulada por los diferentes promotores, se determinó midiendo la actividad enzimática de A-gal. La actividad indicadora se puntuó a partir de placas de agar YPD fabricadas con tampón fosfato 100 mM, pH 6,5 y alfa-X-GAL como sustrato cromógeno. La actividad indicadora se detectó mediante un producto de color azul en, o alrededor de las células (figura 18). Las placas 12 y 23 en la Fig. 18, panel A, se replicaron en diferentes fuentes de carbono en las placas A y B, respectivamente, que se muestran en la Fig. 18, paneles B y C, respectivamente. Por tanto, las designaciones en la Fig. 18, panel A de las placas 12 y 23 se refieren a la Fig. 18, paneles B y C, respectivamente. Las notaciones de NO: 7, NO: 8, etc. se refieren a las SEQ ID NO (es decir, al promotor utilizado para expresar la α -galactosidasa). El nivel de expresión de la α -galactosidasa con los diversos promotores se indica mediante 0, +1, +2, +3, +4 y +5 en las ocho tablas a continuación, con el nivel de expresión que aumenta de 0 a +5. Las tablas etiquetadas como "12A" corresponden a los resultados que se muestran en la Fig. 18, panel B, para las diversas fuentes de carbono. Las tablas etiquetadas como "23B" corresponden a los resultados que se muestran en la Fig. 18, panel C, para las diversas fuentes de carbono.

Expresión relativa de alfa galactosidasa, placa 12/A, glucosa

Promotor (SEQ ID/nombre)	Fig. 18 (Panel/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 7	A/A/2	Glucosa	3
No: 7	A/A/3	Glucosa	3
No: 7	A/A/4	Glucosa	3
No: 7	A/B/1	Glucosa	3
No: 8	A/B/2	Glucosa	2
No: 8	A/B/3	Glucosa	4
No: 8	A/B/4	Glucosa	2
No: 8	A/B/5	Glucosa	4
Control	A/C/1	Glucosa	0
No: 8	A/C/2	Glucosa	2
No: 5	A/C/3	Glucosa	3
No: 10	A/C/4	Glucosa	2
No: 10	A/C/5	Glucosa	3
No: 10	A/D/1	Glucosa	3
No: 14	A/D/2	Glucosa	2

ES 2 716 619 T3

Promotor (SEQ ID/nombre)	Fig. 18 (Panel/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 14	A/D/3	Glucosa	4
No: 14	A/D/4	Glucosa	0
No: 14	A/D/5	Glucosa	5
Control	A/E/1	Glucosa	0
Control	A/E/2	Glucosa	0
No: 12	A/E/3	Glucosa	2
No: 12	A/E/4	Glucosa	2
No: 12	A/E/5	Glucosa	1
09479	A/F/1	Glucosa	3
09479	A/F/2	Glucosa	2
09479	A/F/3	Glucosa	3
No: 13	A/F/4	Glucosa	4
UPP-513	A/F/5	Glucosa	4
UPP-354	A/G/1	Glucosa	4
DAS	A/G/2	Glucosa	0
1-0469	A/G/3	Glucosa	4
GAP	A/G/4	Glucosa	3
No: 3	A/H/1	Glucosa	2
No: 2	A/H/2	Glucosa	0
No: 2	A/H/3	Glucosa	0
No: 11	A/H/4	Glucosa	3
No: 11	A/H/5	Glucosa	3
No: 3	A/I/1	Glucosa	2
No: 3	A/I/2	Glucosa	2
No: 3	A/I/3	Glucosa	1
No: 3	A/I/4	Glucosa	2
No: 3	A/J/2	Glucosa	1
No: 3	A/J/3	Glucosa	2

Expresión relativa de alfa galactosidasa, placa 12/A, glicerol

Promotor (SEQ ID/nombre)	Fig. 18 (Panel/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 7	A/A/2	Glicerol	3
No: 7	A/A/3	Glicerol	3
No: 7	A/A/4	Glicerol	3
No: 7	A/B/1	Glicerol	3
No: 8	A/B/2	Glicerol	3
No: 8	A/B/3	Glicerol	4
No: 8	A/B/4	Glicerol	3
No: 8	A/B/5	Glicerol	4
Control	A/C/1	Glicerol	0
No: 8	A/C/2	Glicerol	2
No: 5	A/C/3	Glicerol	2
No: 10	A/C/4	Glicerol	2
No: 10	A/C/5	Glicerol	2
No: 10	A/D/1	Glicerol	2
No: 14	A/D/2	Glicerol	2
No: 14	A/D/3	Glicerol	4
No: 14	A/D/4	Glicerol	0
No: 14	A/D/5	Glicerol	5
Control	A/E/1	Glicerol	0
Control	A/E/2	Glicerol	0

ES 2 716 619 T3

Promotor (SEQ ID/nombre)	Fig. 18 (Panel/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 12	A/E/3	Glicerol	2
No: 12	A/E/4	Glicerol	2
No: 12	A/E/5	Glicerol	1
09479	A/F/1	Glicerol	3
09479	A/F/2	Glicerol	2
09479	A/F/3	Glicerol	3
No: 13	A/F/4	Glicerol	3
UPP-513	A/F/5	Glicerol	3
UPP-354	A/G/1	Glicerol	4
DAS	A/G/2	Glicerol	0
1-0469	A/G/3	Glicerol	3
GAP	A/G/4	Glicerol	2
No: 3	A/H/1	Glicerol	3
No: 2	A/H/2	Glicerol	0
No: 2	A/H/3	Glicerol	0
No: 11	A/H/4	Glicerol	2
No: 11	A/H/5	Glicerol	3
No: 3	A/I/1	Glicerol	2
No: 3	A/I/2	Glicerol	2
No: 3	A/I/3	Glicerol	1
No: 3	A/I/4	Glicerol	2
No: 3	A/J/2	Glicerol	1
No: 3	A/J/3	Glicerol	2

Expresión relativa de alfa galactosidasa, placa 12/A, placas de metanol.

Promotor (SEQ ID/nombre)	Fig. 18 (Panel/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 7	A/A/2	Metanol	2
No: 7	A/A/3	Metanol	3
No: 7	A/A/4	Metanol	2
No: 7	A/B/1	Metanol	2
No: 8	A/B/2	Metanol	3
No: 8	A/B/3	Metanol	4
No: 8	A/B/4	Metanol	2
No: 8	A/B/5	Metanol	4
Control	A/C/1	Metanol	1
No: 8	A/C/2	Metanol	2
No: 5	A/C/3	Metanol	3
No: 10	A/C/4	Metanol	2
No: 10	A/C/5	Metanol	3
No: 10	A/D/1	Metanol	3
No: 14	A/D/2	Metanol	2
No: 14	A/D/3	Metanol	5
No: 14	A/D/4	Metanol	0
No: 14	A/D/5	Metanol	5
Control	A/E/1	Metanol	0
Control	A/E/2	Metanol	1
No: 12	A/E/3	Metanol	2
No: 12	A/E/4	Metanol	4
No: 12	A/E/5	Metanol	1
09479	A/F/1	Metanol	4
09479	A/F/2	Metanol	2

ES 2 716 619 T3

Promotor (SEQ ID/nombre)	Fig. 18 (Panel/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
09479	A/F/3	Metanol	4
No: 13	A/F/4	Metanol	3
UPP-513	A/F/5	Metanol	3
UPP-354	A/G/1	Metanol	4
DAS	A/G/2	Metanol	2
1-0469	A/G/3	Metanol	3
GAP	A/G/4	Metanol	3
No: 3	A/H/1	Metanol	2
No: 2	A/H/2	Metanol	2
No: 2	A/H/3	Metanol	2
No: 11	A/H/4	Metanol	1
No: 11	A/H/5	Metanol	3
No: 3	A/I/1	Metanol	2
No: 3	A/I/2	Metanol	2
No: 3	A/I/3	Metanol	0
No: 3	A/I/4	Metanol	2
No: 3	A/J/2	Metanol	1
No: 3	A/J/3	Metanol	2

Expresión relativa de alfa galactosidasa, placa 12-A, Etanol.

Promotor (SEQ ID/nombre)	Fig. 18 (Panel/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 7	A/A/2	Etanol	2
No: 7	A/A/3	Etanol	2
No: 7	A/A/4	Etanol	2
No: 7	A/B/1	Etanol	1
No: 8	A/B/2	Etanol	0
No: 8	A/B/3	Etanol	3
No: 8	A/B/4	Etanol	2
No: 8	A/B/5	Etanol	4
Control	A/C/1	Etanol	0
No: 8	A/C/2	Etanol	1
No: 5	A/C/3	Etanol	2
No: 10	A/C/4	Etanol	1
No: 10	A/C/5	Etanol	2
No: 10	A/D/1	Etanol	2
No: 14	A/D/2	Etanol	1
No: 14	A/D/3	Etanol	5
No: 14	A/D/4	Etanol	0
No: 14	A/D/5	Etanol	5
Control	A/E/1	Etanol	0
Control	A/E/2	Etanol	0
No: 12	A/E/3	Etanol	1
No: 12	A/E/4	Etanol	1
No: 12	A/E/5	Etanol	0
09479	A/F/1	Etanol	3
09479	A/F/2	Etanol	2
09479	A/F/3	Etanol	3
No: 13	A/F/4	Etanol	3
UPP-513	A/F/5	Etanol	3
UPP 354	A/G/1	Etanol	4
DAS	A/G/2	Etanol	0

ES 2 716 619 T3

Promotor (SEQ ID/nombre)	Fig. 18 (Panel/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
1-0469	A/G/3	Etanol	3
GAP	A/G/4	Etanol	2
No: 3	A/H/1	Etanol	2
No: 2	A/H/2	Etanol	0
No: 2	A/H/3	Etanol	1
No: 11	A/H/4	Etanol	2
No: 11	A/H/5	Etanol	3
No: 3	A/I/1	Etanol	2
No: 3	A/I/2	Etanol	1
No: 3	A/I/3	Etanol	0
No: 3	A/I/4	Etanol	2
No: 3	A/J/2	Etanol	1
No: 3	A/J/3	Etanol	2

Expresión relativa de alfa galactosidasa, placa 23/B, glucosa

Promotor (SEQ ID/nombre)	Fig. 18 (Placa B/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 4	B/A/1	Glucosa	0
No: 4	B/A/2	Glucosa	0
No: 4	B/A/3	Glucosa	0
No: 4	B/A/4	Glucosa	0
No: 4	B/B/1	Glucosa	0
11231	B/B/2	Glucosa	0
11231	B/B/3	Glucosa	0
No: 1	B/B/4	Glucosa	0
No: 1	B/B/5	Glucosa	0
No: 1	B/C/1	Glucosa	0
No: 14	B/C/2	Glucosa	0
No: 14	B/C/3	Glucosa	0
No: 14	B/C/4	Glucosa	0
No: 14	B/C/5	Glucosa	0
No: 14	B/D/1	Glucosa	2
No: 14	B/D/2	Glucosa	1
No: 2	B/D/3	Glucosa	0
No: 2	B/D/4	Glucosa	0
No: 2	B/D/5	Glucosa	0
No: 2	B/E/1	Glucosa	0
No: 2	B/E/2	Glucosa	0
No: 2	B/E/3	Glucosa	0
No: 2	B/E/4	Glucosa	0
No: 2	B/E/5	Glucosa	0
UPP-513	B/F/1	Glucosa	4
UPP-345	B/F/2	Glucosa	4
DAS	B/F/3	Glucosa	0
1-0469	B/F/4	Glucosa	3
GAP	B/F/5	Glucosa	4
09476	B/G/1	Glucosa	2
UPP 222	B/G/2	Glucosa	3
No: 13	B/G/3	Glucosa	4
No: 13	B/G/4	Glucosa	4
No: 2	B/H/1	Glucosa	0
No: 2	B/H/2	Glucosa	0

ES 2 716 619 T3

Promotor (SEQ ID/nombre)	Fig. 18 (Placa B/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 13	B/H/3	Glucosa	4
No: 2	B/H/4	Glucosa	0
No: 3	B/I/2	Glucosa	0
No: 2	B/I/3	Glucosa	0
No: 3	B/I/4	Glucosa	0
No: 3	B/J/3	Glucosa	0

Expresión relativa de alfa galactosidasa, placa 23/B, glicerol

Promotor (SEQ ID/nombre)	Fig. 18 (Placa B/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 4	B/A/1	Glicerol	0
No: 4	B/A/2	Glicerol	0
No: 4	B/A/3	Glicerol	0
No: 4	B/A/4	Glicerol	0
No: 4	B/B/1	Glicerol	0
11231	B/B/2	Glicerol	0
11231	B/B/3	Glicerol	0
No: 1	B/B/4	Glicerol	0
No: 1	B/B/5	Glicerol	0
No: 1	B/C/1	Glicerol	0
No: 14	B/C/2	Glicerol	0
No: 14	B/C/3	Glicerol	0
No: 14	B/C/4	Glicerol	0
No: 14	B/C/5	Glicerol	0
No: 14	B/D/1	Glicerol	2
No: 14	B/D/2	Glicerol	1
No: 2	B/D/3	Glicerol	0
No: 2	B/D/4	Glicerol	0
No: 2	B/D/5	Glicerol	0
No: 2	B/E/1	Glicerol	0
No: 2	B/E/2	Glicerol	0
No: 2	B/E/3	Glicerol	0
No: 2	B/E/4	Glicerol	0
No: 2	B/E/5	Glicerol	0
UPP-513	B/F/1	Glicerol	4
UPP-345	B/F/2	Glicerol	4
DAS	B/F/3	Glicerol	0
1-0469	B/F/4	Glicerol	3
GAP	B/F/5	Glicerol	4
09476	B/G/1	Glicerol	2
UPP 222	B/G/2	Glicerol	3
No: 13	B/G/3	Glicerol	4
No: 13	B/G/4	Glicerol	4
No: 2	B/H/1	Glicerol	0
No: 2	B/H/2	Glicerol	0
No: 13	B/H/3	Glicerol	4
No: 2	B/H/4	Glicerol	0
No: 3	B/I/2	Glicerol	0
No: 2	B/I/3	Glicerol	0
No: 3	B/I/4	Glicerol	0
No: 3	B/J/3	Glicerol	0

ES 2 716 619 T3

Expresión relativa de alfa galactosidasa, placa 23-B, metanol

Promotor (SEQ ID/nombre)	Fig. 18 (Placa/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 4	B/A/1	Metanol	4
No: 4	B/A/2	Metanol	4
No: 4	B/A/3	Metanol	4
No: 4	B/A/4	Metanol	4
No: 4	B/B/1	Metanol	4
11231	B/B/2	Metanol	2
11231	B/B/3	Metanol	3
No: 1	B/B/4	Metanol	4
No: 1	B/B/5	Metanol	5
No: 1	B/C/1	Metanol	4
No: 14	B/C/2	Metanol	3
No: 14	B/C/3	Metanol	2
No: 14	B/C/4	Metanol	1
No: 14	B/C/5	Metanol	0
No: 14	B/D/1	Metanol	2
No: 14	B/D/2	Metanol	1
No: 2	B/D/3	Metanol	3
No: 2	B/D/4	Metanol	3
No: 2	B/D/5	Metanol	3
No: 2	B/E/1	Metanol	4
No: 2	B/E/2	Metanol	2
No: 2	B/E/3	Metanol	3
No: 2	B/E/4	Metanol	3
No: 2	B/E/5	Metanol	3
UPP-513	B/F/1	Metanol	5
UPP-345	B/F/2	Metanol	5
DAS	B/F/3	Metanol	3
1-0469	B/F/4	Metanol	3
GAP	B/F/5	Metanol	3
09476	B/G/1	Metanol	2
UPP 222	B/G/2	Metanol	4
No: 13	B/G/3	Metanol	4
No: 13	B/G/4	Metanol	4
No: 2	B/H/1	Metanol	2
No: 2	B/H/2	Metanol	1
No: 13	B/H/3	Metanol	4
No: 2	B/H/4	Metanol	1
No: 3	B/I/2	Metanol	3
No: 2	B/I/3	Metanol	2
No: 3	B/I/4	Metanol	3
No: 3	B/J/3	Metanol	3

Expresión relativa de alfa galactosidasa, placa 23-B, etanol

Promotor (SEQ ID/nombre)	Fig. 18 (Placa B/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
No: 4	B/A/1	Etanol	0
No: 4	B/A/2	Etanol	0
No: 4	B/A/3	Etanol	0
No: 4	B/A/4	Etanol	0
No: 4	B/B/1	Etanol	0
11231	B/B/2	Etanol	0

ES 2 716 619 T3

Promotor (SEQ ID/nombre)	Fig. 18 (Placa B/fila/columna)	Fuente de carbono	Expresión relativa de A-gal (1-5)
11231	B/B/3	Etanol	0
No: 1	B/B/4	Etanol	0
No: 1	B/B/5	Etanol	0
No: 1	B/C/1	Etanol	0
No: 14	B/C/2	Etanol	0
No: 14	B/C/3	Etanol	0
No: 14	B/C/4	Etanol	0
No: 14	B/C/5	Etanol	0
No: 14	B/D/1	Etanol	1
No: 14	B/D/2	Etanol	0
No: 2	B/D/3	Etanol	0
No: 2	B/D/4	Etanol	0
No: 2	B/D/5	Etanol	0
No: 2	B/E/1	Etanol	0
No: 2	B/E/2	Etanol	0
No: 2	B/E/3	Etanol	0
No: 2	B/E/4	Etanol	0
No: 2	B/E/5	Etanol	0
UPP-513	B/F/1	Etanol	4
UPP-345	B/F/2	Etanol	4
DAS	B/F/3	Etanol	0
1-0469	B/F/4	Etanol	3
GAP	B/F/5	Etanol	4
09476	B/G/1	Etanol	2
UPP 222	B/G/2	Etanol	3
No: 13	B/G/3	Etanol	4
No: 13	B/G/4	Etanol	4
No: 2	B/H/1	Etanol	0
No: 2	B/H/2	Etanol	0
No: 13	B/H/3	Etanol	4
No: 2	B/H/4	Etanol	0
No: 3	B/I/2	Etanol	0
No: 2	B/I/3	Etanol	0
No: 3	B/I/4	Etanol	0
No: 3	B/J/3	Etanol	0

REIVINDICACIONES

- 5 1. Un ácido nucleico aislado donde la secuencia del ácido nucleico aislado comprende un promotor de *Pichia pastoris* inducible por metanol, teniendo dicho promotor actividad promotora y teniendo dicho promotor una secuencia
- 10 a) idéntica en al menos el 90 % a una secuencia de la SEQ ID NO: 2, o
b) idéntica en al menos el 90 % a un fragmento de la SEQ ID NO: 2, teniendo dicho fragmento actividad promotora.
- 15 2. El ácido nucleico aislado de la reivindicación 1 donde la secuencia es la SEQ ID NO: 2.
3. El ácido nucleico aislado de una cualquiera de las reivindicaciones 1 a 2 unido operativamente a una secuencia codificante heteróloga.
- 20 4. El ácido nucleico aislado de la reivindicación 3 donde la secuencia codificante heteróloga codifica una proteína seleccionada entre el grupo que consiste en una toxina, un anticuerpo, una hormona, una enzima, un factor de crecimiento, una citocina, una proteína estructural, una proteína inmunógena y una proteína de señalización celular.
- 25 5. El ácido nucleico aislado de la reivindicación 4 donde la proteína es una enzima para su uso en piensos para animales.
6. El ácido nucleico aislado de la reivindicación 5 donde la proteína se selecciona entre el grupo que consiste en una mananasa, una amilasa, una glucanasa, una proteasa, una celulasa y una xilanasas.
7. El ácido nucleico aislado de la reivindicación 5 donde la proteína es una fitasa.
8. El ácido nucleico aislado de la reivindicación 5 donde la proteína es una galactosidasa.
- 30 9. Un vector de expresión que comprende el ácido nucleico aislado de una cualquiera de las reivindicaciones 1 a 8.
10. Una célula hospedadora que comprende el ácido nucleico aislado de una cualquiera de las reivindicaciones 1 a 8 o el vector de expresión de la reivindicación 9.
- 35 11. La célula hospedadora de la reivindicación 10 donde la célula hospedadora es una especie de *Pichia*.
12. La célula hospedadora de la reivindicación 11 donde la especie de *Pichia* es *Pichia pastoris*.
- 40 13. La célula hospedadora de la reivindicación 10 donde la célula hospedadora se selecciona entre el grupo que consiste en especies de *Hansenula*, especies de *Pichia*, especies de *Saccharomyces*, especies de *Yarrowia*, especies de *Candida*, especies de *Schizosaccharomyces*, especies de *Torulaspota* y especies de *Kluveromyces*.
- 45 14. Un método de producción de una proteína, comprendiendo el método la etapa de cultivar en un medio de cultivo una célula hospedadora que comprende un primer casete de expresión que comprende un ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora, unido operativamente a una secuencia de codificación heteróloga que codifica una proteína,
donde el cultivo se realiza en condiciones que permiten la expresión de la proteína;
donde la proteína se expresa usando el primer casete de expresión en combinación con un segundo casete de expresión,
comprendiendo dicho segundo casete de expresión
- 50 (a) la secuencia de codificación heteróloga que codifica la proteína unida operativamente a un ácido nucleico aislado que tiene una secuencia que comprende la SEQ ID NO: 15 o la SEQ ID NO: 16, donde la SEQ ID NO: 15 o la SEQ ID NO: 16 tienen actividad promotora; o
- 55 (b) el ácido nucleico aislado de la SEQ ID NO: 2, o un fragmento de la misma que tenga actividad promotora unido operativamente a la secuencia codificante heteróloga que codifica la proteína.

SECUENCIA DE PROMOTOR INDUCIBLE POR METANOL
CAM1 (FR839630_178837..180488)_SELECCIÓN (SEQ ID NO: 1)

ATTGTTGTGAATACTCTCCTTCATTTGGATTTCTTGGACTTCGGACTCTCTTGATCTC
TCTTCGAAAGTTTTAACTCTGTTCATGTATAATTTTACCCGCTGTAGGTTCGCTCATA
ATACCATGAGTATGCACATCTTTTACTCCATTAACTTTCAGGTATGCAAAATACAAT
GAAGATAGTATATAGCTCAAAGAATTTAGCATTTCATTGATCTAATTGTGACATT
TTCTCTATGATATCATCTAGCTTCTTAAACTCGAGAATCTCGTCCAACGAGGCAGAA
ACATTGTCCAGTCTTACGTCAAGATTATTCACGAGTTTCTGGACCGTATCAACGTTT
TCCATCTTAAGATTACAGTAAGTATCGTCCTTTTGAAGTGC AAAAGGTAGAAAAGTT
AATTTTTGATTTGGTAGTACACTATGAACTTGCTCACCCCAATCTTTCCTCCTGAC
AGGTTGATCTTTATCCCTCTACTAAATTGCCCAAGTGTATCAAGTAGACTAGATCT
CGCGAAAGAACAGCCTAATAAACTCCGAAGCATGATGGCCTCTATCCGGAAAACG
TTAAGAGATGTGGCAACAGGAGGGCACATAGAATTTTTAAAGACGCTGAAGAATG
CTATCATAGTCCGTAAAAATGTGATAGTACTTTGTTTAGTGCGTACGCCACTTATTC
GGGGCCAATAGCTAAACCCAGGTTTGCTGGCAGCAAATTCAACTGTAGATTGAATC
TCTCTAACAATAATGGTGTTC AATCCCCTGGCTGGTCACGGGGAGGACTATCTTGC
GTGATCCGCTTGAAAAATGTTGTGTATCCCTTTCTCAATTGCGGAAAGCATCTGCTA
CTTCCCATAGGCACCAGTTACCCAATTGATATTTCCAAAAAAGATTACCATATGTTT
ATCTAGAAGTATAAATACAAGTGGACATTCAATGAATATTTCAATTC AATTAGTCAT
TGACACTTTCATCAACTTACTACGTCTTATTCAACAAT

FIG. 1

SECUENCIA DE PROMOTOR INDUCIBLE POR METANOL
PP7435_CHRI-1351_SELECCIÓN (SEQ ID NO: 2)

GTCAACTGCGTACTCTTTTGTGCGAATGGACTACTGAATCTGCCTCGATAGCCACTAT
AGGAAGGTCCATAGAGGCCAGTTTTTCAACTAGTCTTGGTGGAAAGAAACCGACA
AAGCCTTTCATGGAGTCACCGATACTGAAAGGTTCAAACAAAGAATGCTTGGGTAG
TCTCTTAATACCCATGGCAACGAAAAAGGGGTCTTCATTGTTCAACATGAATTCGT
ATCCACCTTTAATGTAGTCATAAAGCTGCTGAAGTTCCGAATCAGTGATGGAACTG
TCTACAGTGACAATATAGGAGTTCTCAATCACCTTATATCCAGTCGAATATATCTGG
ATAGGGTCGGGTCTCACTGTGGAAGATTCAAATGGGTTAGATCCCTGTAATTTTCAG
CGATGGAGACTCAGTATGATGGGGCAAGGAAAACGGCAATTGGATATTCAATTGG
TCAAGAGATGGTATCAAAGCGAGTGTGCCAGGGTAGCCACGGTAGCCACTGATG
CTAATCTGATAATTTTCATTTCTGGAGTGTCAAACAGTAGTGATAAAAGGCTATG
AAGGAGTTGTCTAGGGGCTCGCGGAGGAAAGTGATTCAAACAGACCTGCCAAAA
AGAGAAAAAAGAGGGAATCCCTGTTCTTTCCAATGGAAATGACGTAACTTTAACTT
GAAAAATACCCCAACCAGAAGGGTTCAAACCAAGGATTGCGTAATTCCTAC
AAGTAGCTTAGAGCTGGGGGAGAGACAACTGAAGGCAGCTTAACGATAACGCGGG
GGGATTGGTGCACGACTCGAAAGGAGGTATCTTAGTCTTGTAACCTCTTTTTTCCAG
AGGCTATTCAAGATTCATAGGCGATATCGATGTGGAGAAGGGTGAACAATATAAA
AGGCTGGAGAGATGTCAATGAAGCAGCTGGATAGATTTCAAATTTTCTAGATTTCA
GAGTAATCGCACAAAACGAAGGAATCCCACCAAGCAAAAAAAAAAAAAATCTAA

FIG. 2

SECUENCIA DE PROMOTOR INDUCIBLE POR METANOL
THI4_SELECCIÓN (SEQ ID NO: 3)

TGCTGTTTTGGGCTCGTACGGATGTTTCTTAGGTCCGATGATTGGTGTTATGACTTG
TGACTACTACTTCGTACGTCATCAGAACTCAAGCTGACAGACCTCTACAAAGCCG
ACAAGAGTTCTATTTACTGGTTCTACAAAGGATTCAACTGGAGAGGTTTCGTTGCCT
GGATTTGCGGTTTTACTCCAGGTATTACAGGGTTTCCTAGCGTCAACCCCAACTTGA
CTGGGGTTCCTACAGCCTGTATCAAGATGTTCTACATTTTCGTTTATCATTGGTTACC
CGATCGGATTCTTAGTTCATCTGGCACTCAATAAGCTATTCCCTCCACCAGGTCTTG
GTGAAGTCGATGAGTATGACTACTACCACTCTTTCACCGAAAAGGAAGCACTGAAA
TTAGGAATGGCCCCTAGTTCCGAGTTGGACAGAGTCAGCACCGATGACCCGATCAA
TATTCCTTACGACGAGAAGTCTTTAGGCTAATGTAGTTAAATAGTTAATCGAAACA
ATCGTGTATCCTCTTTATCGTACCAGCGGGATTCGCTGCTTGGATGGGTGACTCCTG
TCCAGTTGACTCAAAGTAGTCAAAATAGGCCTGGAGACCCTTAACAGGTCGATGAG
TAGCCTACTATGAGAAAACCCCTCACCACAACCTGGACTATAAAAGGGCACGTCAAT
CCCCAAAGCAACTCTTTTCTTTCATCCCTACTTTATTACTTTATCCTTTGATCTTCAT
TGAAGAAAATCTGAAACAATTGTAAGCAGCAATCCACACCTCCCCAGCAATGACTC
AATTTACTAACCCCAATTGACAGAAAATGTGAGCATCTTTTTTAGATGTCATGATGAT
AGGTGGAGTATTCTTAATTATTGCTTTCAGCAAACCGGTGCCATAAAGTGTTTCCC
ATTAAATCAATGAGAGGCATTAAGGCTGAGATTAACGGTTGAACTTGAAGTAGAT
AATTCTAGCGGAAAGAATTGCTCTTTTTATTACGTCGT

FIG. 3

SECUENCIA DE PROMOTOR INDUCIBLE POR METANOL
GPM2_SELECCIÓN (SEQ ID NO:4)

AGGCTCTGCGCAAGGCAACTGAGAAATTGAATAGTGGTTTCAAGCCCGCTGACTTT
TTGTATTATCTCAATGTCGGTGTTTCACAGTCCCCAGAAGGGGGCTTGCCTTCAAG
GGAGACGGAAGAGACATCGTCAACCCTGGGGAGAAGTATTTCAAATGGCGCAAGT
TCGCTAATTTTTACGATTAAGCAGTGCTGTATGGGGTAGTTAATAAATCGGGAATA
TCCTTCTGACGTGACTGTAACAAATCTCTTTTTACGTGGTGCGCATACTGGACAGAG
GCAGAGTCTCAATTTCTTCTTTTGAGACAGGCTACTACAGCCTGTGATTCCTCTTGG
TACTTGGATTTGCTTTTATCTGGCTCCGTTGGGAAGTGTGCCTGGGTTTTGAAGTAT
CTTGTGGATGTGTTTCTAACACTTTTTCAATCTTCTTGGAGTGAGAATGCAGGACTT
TGAACATCGTCTAGCTCGTTGGTAGGTGAACCGTTTTACCTTGCATGTGGTTAGGAG
TTTTCTGGAGTAACCAAGACCGTCTTATCATCGCCGTAATAATCGCTCTTACTGTCCG
TAATAATCCCGCTGGAAGAGAAGTTCGAACAGAAGTAGCACGCAAAGCTCTTGTC
AAATGAGAATTGTTAATCGTTTGACAGGTCACACTCGTGGGCTATGTACGATCAAC
TTGCCGGCTGTTGCTGGAGAGATGACACCAGTTGTGGCATGGCCAATTGGTATTCA
GCCGTACCACTGTATGGAAAATGAGATTATCTTGTTCCTTGATCTAGTTTCTTGCCAT
TTTAGAGTTGCCACATTCGTAGGTTTCAGTACCAATAATGGTAACTTCCAAACTTCC
AACGCAGATACCAGAGATCTGCCGATCCTTCCCCAACAATAGGAGCTTACTACGCC
ATACATATAGCCTATCTATTTTCACTTTCGCGTGGGTGCTTCTATATAAACGGTCC
CCATCTTCCGTTTCATACTACTTGAATTTAAGCACTAAA

FIG. 4

SECUENCIA DE PROMOTOR INDUCIBLE POR METANOL
PP7435_CHR2-0790_SELECCIÓN (SEQ ID NO: 5)

CAAACCTTAACCGACCGTTCTTCCATCCGTTTATTAATATACACACCTATGAACTGAG
CCAGGTTTTTCAGGTCTCTGTGACTCTCTATACATTGACGGAACAACATCCGTTTCAGT
CTCATCCAATTGCAGCCCAAACCTCTGAGTTTAGCAATTGCAAATGGTTATTATCTGA
CGAGTAATCGTTGATGGCACATGCCCTCTGTTTGAACATCTCTTGAACAATAGCCAT
CAGTTCTGTGTCATTAACATGCTTCCCCATTTCACTGACAGTTTGTAGAAATAGGG
CAACAATTGATGCAAATCGATTTTCAACGCATTGGTTTTGATAGCATTGATGATCTT
GGAGCTGTAAAAGTCCGGCTGGATAAGCTCAATGAAATAGGTTGGTTGATCTGGAT
CTTCTTTTGGGTCATTTTGTTCGCTCTGTATTTACAAAATTGCCAGAATCTCTGCCAA
CCACAGTGGTAGGTCCAACCTGGTGTCTGAATCACAGGCTTCCCCGGGTTGTTCTC
TAAATAACCGAGGCCCGGCACAGAAATCGTAAACCGACACGGTATCTTTTGTCCGT
CCGCCAGTATCTCATCAAGGTCGTAGTAGCCCATGATGAGTATCAAAGGGGATTG
GTTATGCGATGCAACGAGAGATTGTTTATCCCAGATGCTGATGTAAAAACCTTAAC
CAGCGTGACAGTAGAAATAAGACACGTAAAATTACCCGCGCTTCCCTAACAATTG
GCTCTGCCTTTCGGCAAGTTTCTAACTGCCCTCCCCTCTCACATGCACCACGAACTT
ACCGTTCGCTCCTAGCAGAACCACCCCAAAGTTTAATCAGGACCGCATTTTAGCCT
ATTGCTGTAGAACCCACAACATAACCTGGTCCAGAGCCAGCCCTTTATATATGGT
AAATCCCGTTTGAACCTCGAAGTGAATCGGAATTTTTACATCAAAGAACTGATA
CTGAAACTTTTGGCTTCGACTTGGACTTCTCTTAATC

FIG. 5

SECUENCIA DE PROMOTOR INDUCIBLE POR METANOL
PP7435_CHR3-0842_SELECCIÓN (SEQ ID NO: 6)

GTAAATAAGTTAAAGTTTAAAAAGGAAAGGATGCAAAAAATATCCTTGAAGGCAA
CGAATATTTTGAAATCCCCGATGCCAAATAAATGCTATCACTTAAAGAGCACATA
GAGGTGGAAAAAGAAAACTTGGTCAAGCTAAGGGTTAGCAAGTTTCTGTTTGTGA
TAATCAGGGAGAAGGTGTCAGAAAAACATGATTATGTAAGTGGTGTAGGAGCC
GTAAAGCATTCTGTTCGGCCAATAGCAAGCCCGCCTTTTGTTCATCTTTTTCGCGTTT
CTCTGTGTGAGACACTAATCACCTTTGTAAGACATCGGGAAAACCGTTGCGCAAAA
TGAGATAGAGATTGTTCTCGATAGAGGAGCGTAGTAGCCTCTCCAGCCTGCTTTAG
CAACATAATAGAAAAGAAATATGCGTTGCCTAGGGAGGCTACGTATGCCAGCAT
AAACGAGTGTTTACCTTACTTCGCACGAGCAGTAGCCACTAAGATCATTATAAACT
CACCTATTGTCTTCATGCTGTGCTCCGCGTATTTGTCTGTTTCAGGGTGTCAATTTCTCG
TCATGAGAATCTGATTGATGACTATGCGAGATTACCCCTGGATTTTTTTTGATCCCG
TAACGCGAACTTGAACATTGACTTTGATATGGCAATGGGCCCTAATATGCCCTAAT
ATGCCCTAAGCTTAACAATTGACTTCTGTTCTCTGGCAGACTCCACAGAAAACCTGG
TTGACAGGTCTAATTTCTTTTTGAATCATTTCCGGTGATTCATTTTGATGCTTAGAGT
GAGTCATGGGTTCTTTATCCGCATTCTTCTTCGCGTCTGCTGTGCTTAATAATAGCC
TACTAAAAATGTGGGGAGCCTCTTACCTTATGTCTATAAAAAACAAGCACATGACTA
TGCCATCGCCTTCATAGTTGTTCTGCGGGTTTTTGCTTGTTTTATGACCGTAGAGAC
CAACCAATTTACATATCTACAGGGTAGCACATTTCGATAAGAAA

FIG. 6

SECUENCIA DE PROMOTOR CONSTITUTIVO
PP7435_CHR1-0269_SELECCIÓN (SEQ ID NO: 7)

GAAGGGCTACAGGTTGTTCGAGTCAACCAAGAAAAAATGTCTCCATTCCATTACC
TTCAGAAGTCTGTGATGTGAGATTTCAGTAGCAACGAACTACACCCATAGTCAGAC
CAAATGAAACCAAGTCAAGACTTTTCAATCTGGCTCAAAGACCAGCCAGGAATCAG
AGAGTTTTTACAAAATTCGACCATCGCACTTGGATCTAGACTCAATCTCTTAAGCA
ATGTAAATATTGAGATACATGGGGAGACAGTGAATTATACCTTCACAGAAATCGAA
CAGAGGAGACAATTAGAAGTGAATTACAGACAACACCTTCTTCAATATTCAGCAGT
GGAAGGAGGACTACTGGGTGGGAAAAGAACAGAAGTCAATCTGGTAAATATGGCT
GAAGCCCCTGCAACTACAGAGTCATTTGCCAAGTTTTCCAAGGACGTTCTGGATCT
TGTTTCGTATACTAAGTTGTACATAGAGTGAATATAAGAACAGGGGCCAGTTAAGA
CTAGATCTATCAATACTGTAATCTAGACTGAATAAACCTGAATGGCAAAGAAACAC
CAAATGGAGATTTCCGAGGTAGTATCGACCGCACTTACGAGGGCGGATATGTGTCC
TCCTGATGTATAGTTACGTCATAAAGTATTGATGCCTAGACCATTTCAATTTTCGTAG
TCATCTATTCGTGAAAAGAGGCAAGGCTGATTGTTCAATTTGGCAATGATGAGCCA
ACAGTGCTTAAAACCTGGACAAAACCTTAGATCAAATATTCACCGTTGTAAAACACAG
TTCTCTCGTCTTCATTTTAGCCCATTTGCACAAAACCCCACTTTTACATTCCGGA
AAAAAATGCTCCTCCACCTAGATGCAGAGTGTACCGCTCAAAGAACCCTCGATGA
GTTTCGACCATCTTCACCAAGTGAATAAATATTTATAGTGAGGGAAAGAACATTTT
TTCCCTGTTCGTTTCTTTACCAGTAACTACTACAGAATCTTCTAAA

FIG. 7

SECUENCIA DE PROMOTOR CONSTITUTIVO
PP7435_CHR2-0207_SELECCIÓN (SEQ ID NO: 8)

TGGATCTGGATGTAGATGCAACTGCCCGAAAAAATACAAAGAATTAGAAGATAGT
CCCAGCTGCTGTATTGGCAGATGGATGGTATTAACCTAACGATAAGTATAAAGGCGA
AAAATACCTAGATAAATACTCAATGGCCGAAGAGGTCAAGCAGACTCTAAGCAAG
GGTGAAAACTAAACGTTAAAGAGATATTGAAGAGGCACCACAAGTATCCAACAT
GAGAAGAGAACCCAGAAAGAGTCAAAATACAACTCGAAAACATACCGATTGGCA
GAATACCAGCCATACATTACATTTGATACTAATATACATAAAAATACACTAACCTCC
GAAACCATACAAGGTACGTCCTTGTCTTTTCAAAGCGTAGACGACATCCAGAGAAG
TAACAGTCTTTCTCTTGGCATGCTCAGTGTAAGTAACAGCATCTCTAATGACGTTCT
CTAAGAAAGTCTTCAGCACAGCTCTGACTTCTTCGTAGATCAAAGCGGAAATACGC
TTAACACCACCTCTTCTGGCCAATCTTCTGATAGCTGGCTTTGTGATACCTTGAATG
TTGTCTCTAAGAATCTTTCTGTGACGCTTAGCACCGGATTTTCCTAGACCTTTTCCA
CCTTTTCCTCTACCAGACATATTTATTGATTATTTGTTTATGGGTGAGTCTAGAAAA
GGACGCACTCGTCTTGTATTTATAGATGAAAAGAGTTAAGGTGGACAATGCAGTGC
CAAACCGTAATGTTGATACGACACGGTACGCGATGAACTGAGCCGTA CTCTAGAG
GCAGAGCGTGCGTAAATTTGAAACTGTAAGAACATGTCCCATTTCTATGCTACCAG
AACTCATAATAAACTTGCATCCCATTTCAGTCGGGAGTGTCCGAGTGC GAAATACT
GTTGGCACTGATGGCAAAAAGTGCCAAATCGTTATCTATGGGAGACTCTTATAAAT
ATCCAGATCCTCCCCCTCCTGCTTCTTCTTGTGTCTATCGTAGTAAAA

FIG. 8

SECUENCIA DE PROMOTOR CONSTITUTIVO
PP7435_CHR2-0208_SELECCIÓN (SEQ ID NO: 9)

GATCGAGGACATGGTGTTTTTTTTTTTATAGAAGTATCATACTAAGTGTACTTTTC
ACTTGCTTATCCTCTTTTCTACCTCACCTCATAAGCAGGAAATTCCTTGACAATCGC
AAGAGTAAGCGTGCCTACACAGCCATCGGCAACACCGAATTCACCTTACAGCCA
GTTTGTATCTCCCTTAAATATCGTATTCGATTAATATACTACTACATTACACTTAGGA
TCTTTCACCTCTCAGTCTTCTAGCTAACAGGATATCCTTCTTTTGGATAGTAACTCTC
TTAGCGTGAATAGCACACAAGTTAGTATCTTCGAACAAGGAAACCAAGTAAGCTTC
GACGGATTCTTGCAAAGCACCGATGGCAGAAGATTGGAATCTCAGGTCAGTCTTGA
AGTCCTGAGCAATTTCTCTGACCAATCTTTGGAAAGGCAGCTTTCTGATAAGCAAC
TCAGTAGACTTTTGGAAATCTTCTGATTTCTCTCAGAGCAACGGTACCTGGCTTATAT
CTGTGGGGCTTCTTGACACCTCCAGCAGCAGATGGAGCGGATTTTCTGGCAGCCTT
GGAAGCCAATTGCTTTCTTGGGGCTTACCACCAGTGGATTTTCTTGCTGTTTGTTT
AGTTCTAGCCATTTTACTACGATAGACACAAGAAGAAGCAGGAGGGGGAGGATC
TGGATATTTATAAGAGTCTCCCATAGATAACGATTTGGCACTTTTGGCCATCAGTGC
CAACAGTATTTTCGCACTGCGACACTCCCGACTGAAATGGGATGCAAGTTTATTATG
AGTTCTGGTAGCATAGAAATGGGACATGTTCTTACAGTTTCAAATTTACGCACGCT
CTGCCTCTAGGAGTACGGCTCAGTTCATCGCGTACCGTGTCTGATCAACATTACGGT
TTGGCACTGCATTGTCCACCTTAACTCTTTTCATCTATAAATACAAGACGAGTGCCT
CCTTTTCTAGACTCACCCATAAACAAATAATCAATAAAT

FIG. 9

SECUENCIA DE PROMOTOR CONSTITUTIVO
PP7435_CHR2-0809_SELECCIÓN (SEQ ID NO: 10)

TTCCATATTACTTCAGCAAATACAAGTACAAATCACTTGACTCGTATCAGTTCCTGG
ATACCCTTTATGAGTTTTTCTCAGATAAACATGAAATTTTAGACAAAGTTGACTGGG
AAACTTGGTTGTACAAGTTTGGATTACCACCAAAGCCAAAGTTTGATACCTCGCTA
GTTGATGAATGCTATGATTTAGCTGCCAAATGGGTTGATGTCACCAAGAAGGATAG
CAAAGATCTTTTGAAGAAAACTTTCAGCTCAGATGATATTTCTAATTTACGGGTA
ACCAAACCTAACGTCTTCCTTGATACTTTAGTATCTTATCAAGGAGTGGAAGGTTTCC
TGTGGAATTCCAAAGAAGGGGAACAAGCTTTGACGTTTATGAGAGAGAGCTATAG
CGTTTATGACGACTCCAAGAATGCAGAAGTCATTTTCCGCTGGTACAGACTACAAT
TAACAGGGAGAAGCAAGCAGTACTACCAGCGTCTGGCCGATTGGCTGGGAACAAT
CGGACGAATGAAATTTGTTTCGACCATCTTATCGAATGCTCAACGATGTGGACCCTC
AACTGGCCAAGGACACATTCTTGAAGTTTGAAGCAATCTACCATCCGATCTGTCGA
TCAATGATCCGCAAGGACTTACACTTGTAGAGAAAAAAGGTATGTGGGGTATACGT
AAGTACGTTAAGTATGCCAAAAACATCACAATAGACAATACGCCCTACAACAGAA
CGTCAAGTTGTACGCGAGCACCGTTAAATCACACGTAACATCCAACACCTTTCTTT
GGTAGGGCATAGCCCCAGGTGGCACCACGTGCATAAACCATTTTACACCCCAACAC
CCACCAATCCTCGCCCTCTGGCATTGCTCAAATTTTTTAACCAGCTTCTGATTACA
TAGGTAACCAGTTCAATTCATAGGTAAGTCAACGCCGAACAACAAGGGAAATCAC
CCATCCGCTCACATCAGTCCAGTTGAAGACTAACATTAAGACAAAA

FIG. 10

SECUENCIA DE PROMOTOR CONSTITUTIVO
PP7435_CHR4-0069_SELECCIÓN (SEQ ID NO: 11)

TGGTAATCAACTCACTATTTTCATTCATGTCAACAGCCCCGAAAACGATAATTAATA
GACAACCCCAATCATATTTAGTGAGTTTTTCGGATTAAAGCTTGTTCGCGTTCCGGTT
GATGATGGTTTTCAAGAAAATGCTGAGCTCGCTCTTATACTGGGTCCCTTGAAATGA
TATGCATATGTTGGGAATACGCTAATGTACTTTGACAGATATTGCTTCTCCTTGTAG
CAAATTGACAGTTTCGTTTGCAACTGTAACACTTGGTGATCGATATTTGGATAAAC
ACAATTACTATTTTGGAGTTACAACCAGAATTCGTCTTAATGTTTCGCATCCTTGTT
CAAAGACTAAAACCTTCTTCGATTTCGAATTTGTCTTAAGTATTGACTAGTCGGGCTC
AAAGAGCAACACTAGCTGGAAACGGCTCAGGATCACTTTCGTTTTTTTTCTAGAAA
ATGAGGCAATTGTAGGACTGGGTCTCTTTATCTGCAATTTGTGTAAGCTGGTTTACT
GTTCTCAACAAAATAGAAGTGAACCGTCCGTGATGTAATAAGACTACTCCGTTG
AATCGCAACAAGTATTATATCCTAGCTTAAGGTGATTCTTAACCTTGGAGCAAAAA
TATCGTGTAACAACGCCGGCTCAGTCAATCTTCTTTGAATTAACACTGCACCTCTAT
CTTGAACACTATCTACATGAGTGTGCAACCAGAAAGAATGCAAAGATTAATTTTCT
TACCCACAATGTGGTCCGATCCACAATTCTAGTGAATTCCTAACATGTTTAGTACCA
ACGTTTGAACAACAGATACATCACATTGAAAAAAGGATTGAGAATAATAATCAA
ATTTTCAGCATTATATAGTACCACGAACTAAATCGCTTTGGAAGGAGAAACATTT
TCAAGCTGTTAAACTGGATTTATCCAAAATAGTTAGTTTGCACAAGTTTAAAAGGA
GGTGTCTAACAGGTGGGGAAAAATCAGTTAATTGAGTTTTTA

FIG. 11

SECUENCIA DE PROMOTOR CONSTITUTIVO
PP7435_CHR4-0800_SELECCIÓN (SEQ ID NO: 12)

CAGGGGGCAGTTACTTCATTATATGTAGTAAGTTGTCTAATAACACTTTAGAATTG
AGGGTTGTAATAACCAAACAGACACTGTTGCAAGTTATGGTGACAAGGTTGCTTTC
AAGCCCTGATACTAAAGTGGTTCATATTTGCATGAGTGGTTCGTATCCGGAGGATC
ATCTACAGAACGTGCTAAAATCATTATCTTTTGGTGACGAAAAAAGAAGTTACCA
TTGGATGACAAGACAAAGAATACTATTTTAGATAGTATTTGGCTGGCTTACCCTGA
CGATCTGAATGAGTCGTACGAACAATTGGAGAACTTCAAGAAGAGAATTCAGCG
GGTCTCTCAAACACTGATGATTGTTATTCAAGATTTGGATGATATGGCTGAGTTATAC
GCCTTCAGAGATTTACATCAGGTAAGTTTACTCTGAAATTTTGATTACAAAATA
TATACTAACACATTTAGCCTGTGCTACATTGATGAATTTTCATGCTTAAATTAAGAAA
AATCAGTCAACTGGGAGCTACGGTCATAGTTACTAATATCAATCATGACTCTGGAG
TGTCTATGATATTCTCCTACATAGCCAAGTTTTACGACAATAAACTAACGTTTCAGCT
TTCACAAAGGCTCAATACTTTTAGAGATACGTCCTCTAACGATATCAACCGAACTT
CAGCTCAGCACTACTACTCTCTATCAACCCACAAACGATGTACCACATGCTTTTC
CCCAGCTCTGAAAATGCCCTTATCAGCACGTGATTAAAGATGAGGGCTCGCTGACA
AGATGTGACTAGACTGCCACCTTTACACCCTCAGAGAAGACAAACGTCTCATTTCG
CGGTTAGTAAGGTAAACCCACAAATTTTCAACGTGTACTACCAAGAGCTCTCTGT
CTGCCATGGGCGCGCACACTCCCTGGTGTGCGAATCCAACCAAGTATGAAAAA
AAAAACCATTACTTCAAAGTTTTCTTCAACACACAAAAGAAG

FIG. 12

SECUENCIA DE PROMOTOR CONSTITUTIVO
CROMOSOMA 1 DE PICHIA PASTORIS CBS 7435 TEF2 SELECCIÓN,
SECUENCIA DE REPLICÓN COMPLETA (SEQ ID NO: 13).

GTATTTGACAGGTTGGGGAGCAAATAAGTGATGATGTCCCATGAAAGTAGAAAAT
GGCTAGTAGAAGGCAAAAATTTGAAATTCTTAGAGTCAAATAGTTAGACTCCAAGT
TCTAATCCACATTTGGTCAGTTTCATAGCATCCAGAGCTTTTGCCACTGGTGAACAT
ATCTACCCATTGCGATGCAACAAGTCACTGAAAGCCTAAAACGGAGATCCCCTAT
CTTACAGCCTCGTTCAAAAAAAGTCTACCGTTTATCTGCTATGGCCGATGTGAGG
ATGCGCTCATGCCAAGAGTCCAACCTTATCAAAAACTTGACCCGTCATACAGGCT
CTAGATCAAGAAGCAAACCTAATCTCAGCATCTGGTTACGTAACCTCTGGCAACCAG
TAACACGCTTAAGGTTTGGAAACAACACTAACTACCTTGCGGTAACCTATTGACA
CTACACATCCTTAATTCCAATCCTGTCTGGCCTCCTTCACCTTTTAACCATCTTGCCC
ATTCCAACCTCGTGTGTCAGATTGCGTATCAAGTGAAAAAATAATTTTAAATCTTT
AACCCAATCAGGTAATAACTGTCGCCTCTTTTATCTGCCGCACTGCATGAGGTGTCC
CCTTAGTGGGAAAGAGTACTGAGCCAACCCTGGAGGACAGCAAGGGAAAAAATACC
TACAACTTGCTTCATAATGGTCGTAAAAACAATCCTTGTCGGATATAAGTGTTGTA
GACTGTCCCTTATCCTCTGCGATGTTCTTCTCTCAAAGTTTGCGATTTCTCTCTATC
AGAATTGCCATCAAGAGACTCAGGACTAATTCGCAGTCCCACACGCACTCGTACA
TGATTGGCTGAAATTTCCCTAAAGAATTTCTTTTTCACGAAAATTTTTTTTTTACACA
AGATTTTCAGCAGATATAAAATGGAGAGCAGGACCTCCGCTGTGACTCTTCTTTTTT
TTCTTTTATTCTCACTACATACATTTTAGTTATTCGCCAAC

FIG. 13

SECUENCIA DE PROMOTOR CONSTITUTIVO
CROMOSOMA 3 DE PICHIA PASTORIS CBS 7435 PP7435 CHR3-0476 SELECCIÓN.
SECUENCIA DE REPLICÓN COMPLETA (SEQ ID NO: 14)

ATTGTTTATAGCCTATAATCGCAGACTTTGTTAACCCGTAGGAATAATCCTACCACT
AGAATTAGATCTTCCAACCATACATAGCGTGTACTACAACCTAAGCTGGTCCTCTTC
TTTTATCATTTCGCGCGGCTCTATCTCATCTACACCTCCATTGAACTAATGGAAAGCA
AATACGCCGAAAAATATCAAGATCGATTGATTCTTGGAGATGATGGCGATGAGGA
CGAATTGTTTGACGAGCTGGAAAAAGATATTGAGGATCAATTCTTGGCCAAATACA
GAGCAGAGAGAATCCAACAGTTAAAACAGGAGATTACCAAGATCAAGGACCATAG
TTCAAACATCAACCTCAATGACCACGGTAACATGAAAACAATAGATACTGATGACG
AACTATTGAAAGAACTGTTGATAGCGAACGTGTTGTGATCCATTTCTTTAACCCAT
CGTTTAGCACTTGCCGTATCATGGATGAGAAGCTGTCTATAATCAGCACCAAACAT
ATTGGAACGCGTTTTTTTCAGAATTGAAGCACATAGAGCTCCATTTTTAGTTGCAAA
GCTTGGTATCAAAGTGCTTCCATGTGTTGTATTGTACTACAAAGGATTAGAAAGGG
ATAGAATTGTCGGATTTGACAGATTAAGTAATTCTCAGACCAATTTTGAGCTAGAA
GCTTTAGAGGAGTTACTCTTAGATAGTGGAATTGTGGAACGAAGAACTGTGCGATTT
TAGCAACCTGAGAAACAAGGTCCAAAACAAGGTGGATCAGTCAAATCAGACTCA
GAAAGTGATCTAGATATGTGATAGATGGCGGATGGCAGGTTTATTCTAGTGTTTCA
CGTGACACACGTGAGCGTTTAAAGGGCACACACCCTGACTGACGCGCGAACATCTAA
TCTGTTCCGCATGAAAAAAAAAAAACTACCTCGACGAAATTCTTCTTAGACAGTTT
TTACATTGGTAAGAAAGAAGCATTACGTATTGCCGACGAAGCCAAA

FIG. 14

SECUENCIA DEL PROMOTOR A0X1 (SEQ ID NO: 15)

GATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGACATCCA
CAGGTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACACTAGCAG
CAGACCGTTGCAAACGCAGGACCTCCACTCCTCTTCTCCTCAACACCCACTTTTGCC
ATCGAAAACCAGCCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCAATTCCTT
CTATTAGGCTACTAACACCATGACTTTATTAGCCTGTCTATCCTGGCCCCCCTGGCG
AGGTTTCATGTTTGTTTATTTCCGAATGCAACAAGCTCCGCATTACACCCGAACATCA
CTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCCCAAATGG
CCCAAACTGACAGTTTAAACGCTGTCTTGGAACCTAATATGACAAAAGCGTGATC
TCATCCAAGATGAACTAAGTTTGGTTCGTTGAAATGCTAACGGCCAGTTGGTCAA
AAGAACTTCCAAAAGTCGGCATAACGTTTGTCTTGTGTTGGTATTGATTGACGAAT
GCTCAAAAATAATCTCATTAAATGCTTAGCGCAGTCTCTCTATCGCTTCTGAACCCCG
GTGCACCTGTGCCGAAACGCAAATGGGGAAACACCCGCTTTTTGGATGATTATGCA
TTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATAGCCTAAC
GTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATATATAAACAG
AAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTATCATCATTATTAGCTTACTTTT
AATTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAACGACTTTTAACGACA
TGAGAAGATCAAAAACAATAATTATTCGAAACG

FIG. 15

SECUENCIA GENÓMICA, 1000 PB CORRIENTE ARRIBA DEL ATG DEL PROMOTOR AOX1
(SEQ ID NO: 16)

CATGTTGGTATTGTGAAATAGACGCAGATCGGGAACACTGAAAAATAACAGTTATT
ATTCGAGATCTAACATCCAAAGACGAAAGGTTGAATGAAACCTTTTTGCCATCCGA
CATCCACAGGTCCATTCTCACACATAAGTGCCAAACGCAACAGGAGGGGATACT
AGCAGCAGACCGTTGCAAACGCAGGACCTCCACTCCTCTTCTCCTCAACACCCACT
TTTGCCATCGAAAAACCAGCCCAGTTATTGGGCTTGATTGGAGCTCGCTCATTCCA
ATTCCTTCTATTAGGCTACTAACACCATGACTTTATTAGCCTGTCTATCCTGGCCCC
CCTGGCGAGGTTTCATGTTTGTTTATTTCCGAATGCAACAAGCTCCGCATTACACCCG
AACATCACTCCAGATGAGGGCTTTCTGAGTGTGGGGTCAAATAGTTTCATGTTCCC
CAAATGGCCCAAACACTGACAGTTTAAACGCTGTCTTGGAACCTAATATGACAAAAG
CGTGATCTCATCCAAGATGAACTAAGTTTGGTTCGTTGAAATGCTAACGGCCAGTT
GGTCAAAAAGAAACTTCCAAAAGTCGGCATAACCGTTTGTCTTGTGTTGGTATTGATT
GACGAATGCTCAAAAATAATCTCATTAAATGCTTAGCGCAGTCTCTCTATCGCTTCTG
AACCCCGGTGCACCTGTGCCGAAACGCAAATGGGGAAACACCCGCTTTTTGGATGA
TTATGCATTGTCTCCACATTGTATGCTTCCAAGATTCTGGTGGGAATACTGCTGATA
GCCTAACGTTTCATGATCAAAATTTAACTGTTCTAACCCCTACTTGACAGCAATATAT
AAACAGAAGGAAGCTGCCCTGTCTTAAACCTTTTTTTTTATCATCATTATTAGCTTA
CTTTCATAATTGCGACTGGTTCCAATTGACAAGCTTTTGATTTTAAACGACTTTTAAAC
GACAACTTGAGAAGATCAAAAAACAACCTAATTATTCGAAACG

FIG. 16

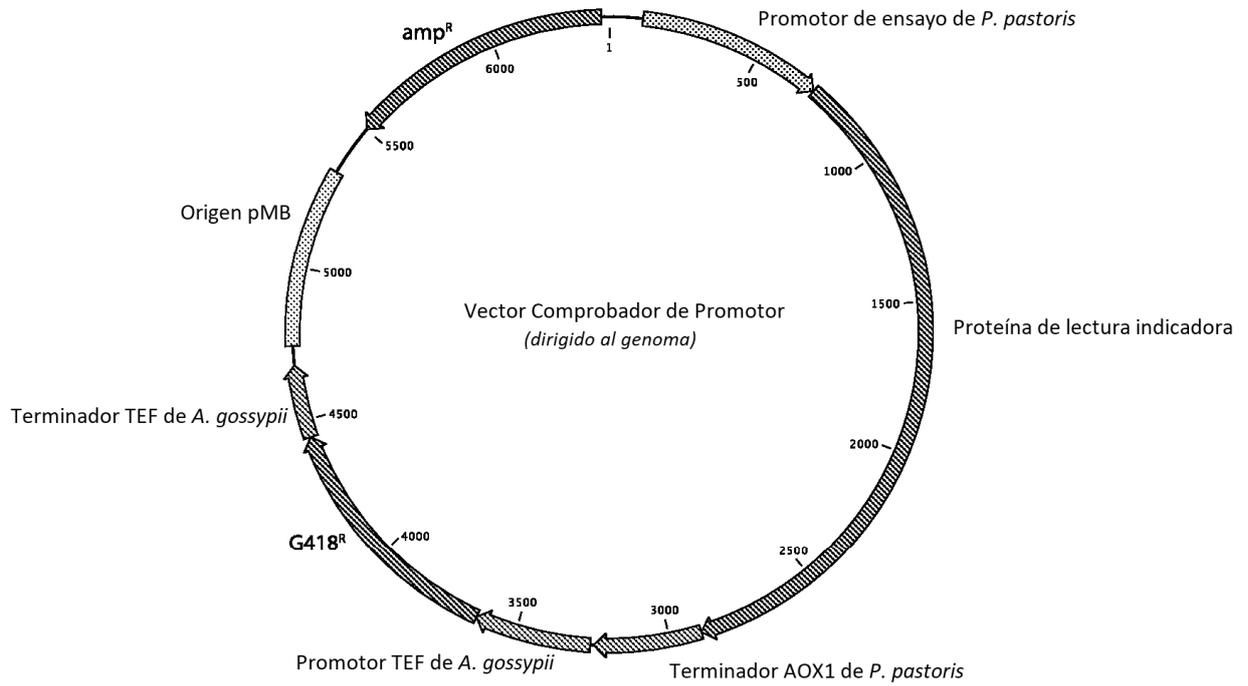


FIG. 17

ES 2 716 619 T3

Ensayo en placa de alfa-X-Gal para determinar la actividad alfa-galactosidasa de transformantes de *Pichia pastoris* que albergan una construcción de un ORF de alfa-gal sintético fusionado con secuencias de promotor aislado SEQ ID NO:). No se ha insertado ningún promotor en un plásmido de control. Todas las construcciones están integradas en el genoma de los transformantes.

Posición	1	2	3	4	5
Fila	Placa 12				
A		NO:7	NO:7	NO:7	
B	NO:7	NO:8	NO:8	NO:8	NO:8
C	Control	NO:8	NO:5	NO:10	NO:10
D	NO:10	NO:14	NO:14	NO:14	NO:14
E	Control	Control	NO:12	NO:12	NO:12
F	09479	09479	09479	NO:13	UPP-513
G	UPP-354	DAS	1-0469	GAP	
H	NO:3	NO:2	NO:2	NO:11	NO:11
I	NO:3	NO:3	NO:3	NO:3	
J		NO:3	NO:3		

Posición	1	2	3	4	5
Fila	Placa 23				
A	NO:4	NO:4	NO:4	NO:4	
B	NO:4	11231	11231	NO:1	NO:1
C	NO:1	NO:14	NO:14	NO:14	NO:14
D	NO:14	NO:14	NO:2	NO:2	NO:2
E	NO:2	NO:2	NO:2	NO:2	NO:2
F	UPP-513	UPP-354	DAS	1-0469	GAP
G	09479	UPP 222	NO:13	NO:13	
H	NO:2	NO:2	NO:13	NO:2	
I		NO:3	NO:2	NO:3	
J			NO:3		

FIGURA 18, panel A.

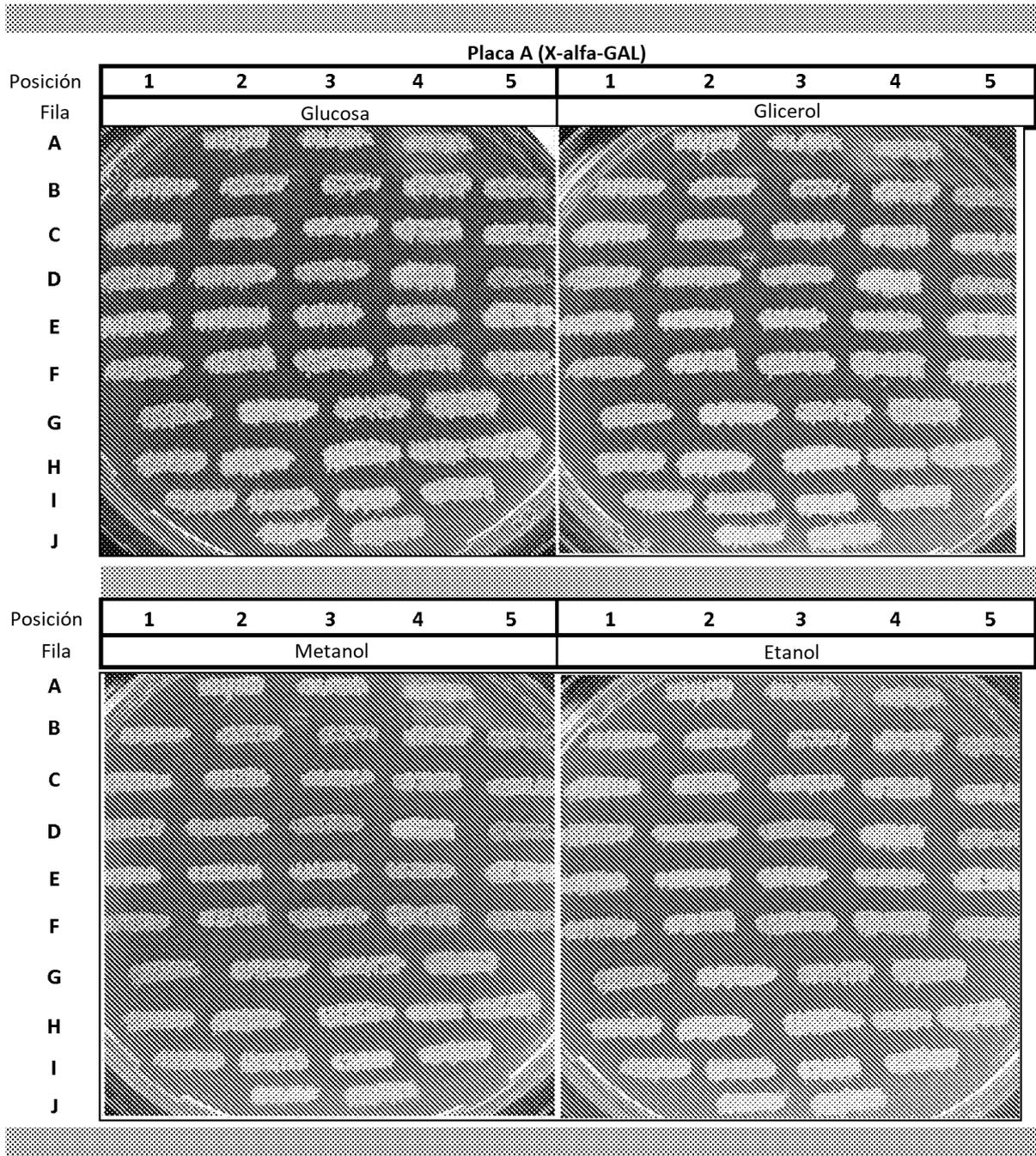


FIGURA 18, panel B

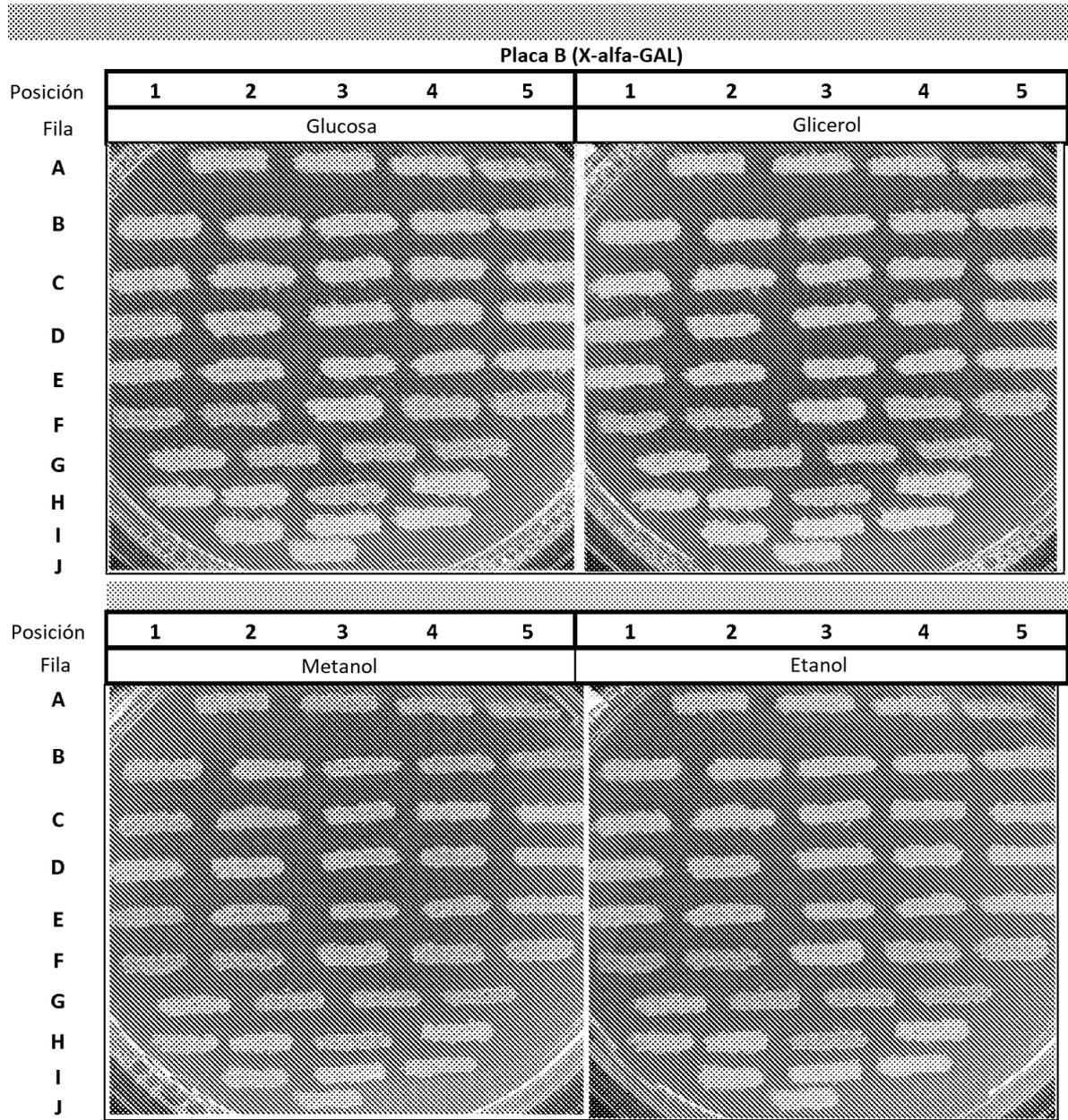


FIGURA 18, panel C