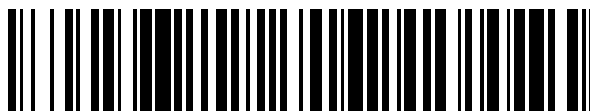


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 730 173**

51 Int. Cl.:

C07K 14/395	(2006.01)
C07K 14/39	(2006.01)
C07K 16/00	(2006.01)
C12N 15/81	(2006.01)
C12P 21/00	(2006.01)
C12N 9/02	(2006.01)
C12N 15/63	(2006.01)
C12N 15/67	(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **05.10.2012 PCT/EP2012/069757**
- 87 Fecha y número de publicación internacional: **11.04.2013 WO13050551**
- 96 Fecha de presentación y número de la solicitud europea: **05.10.2012 E 12768859 (6)**
- 97 Fecha y número de publicación de la concesión europea: **13.03.2019 EP 2764016**

54 Título: **Promotor regulable**

30 Prioridad:

07.10.2011 EP 11184323
07.10.2011 US 201161544451 P
06.06.2012 EP 12171006

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
08.11.2019

73 Titular/es:

LONZA LTD (100.0%)
Lonzastrasse
3930 Visp, CH

72 Inventor/es:

MATTANOVICH, DIETHARD;
GASSER, BRIGITTE;
MAURER, MICHAEL;
PRIELHOFER, ROLAND;
KLEIN, JOACHIM y
WENGER, JANA

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 730 173 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

Promotor regulable

La invención se refiere a promotores regulables y a un método para producir una proteína de interés en un cultivo de células eucariotas bajo el control de un promotor regulable.

5 **Antecedentes**

Se ha logrado la producción con éxito de proteínas recombinantes con hospedantes eucariotas. Los ejemplos más notables son levaduras, tales como *Saccharomyces cerevisiae*, *Pichia pastoris* o *Hansenula polymorpha*, hongos filamentosos, tales como *Aspergillus awamori* o *Trichoderma reesei*, o células de mamífero, tales como, por ejemplo, células CHO. Aunque se logra con facilidad la producción a altas tasas de algunas proteínas, muchas otras proteínas se obtienen solo a unos niveles comparativamente bajos.

10 La expresión heteróloga de un gen en un organismo hospedante habitualmente requiere de un vector que permita la transformación estable del organismo hospedante. Un vector proporcionaría el gen con un promotor funcional adyacente al extremo 5' de la secuencia codificadora. Con ello, la transcripción es regulada e iniciada por esta secuencia de promotor. La mayoría de los promotores usados hasta la fecha se han derivado de genes que
15 codifican proteínas que habitualmente están presentes a altas concentraciones en la célula.

El documento EP0103409A2 describe el uso de promotores de levadura asociados con la expresión de enzimas específicas en la vía glicolítica, concretamente promotores implicados en la expresión de piruvato quinasa, triosafosfato isomerasa, fosfoglucosa isomerasa, fosfoglicerato mutasa, hexoquinasa 1 y 2, glucoquinasa, fosfofructosa quinasa, aldolasa y el gen de la regulación glicolítica.

20 El documento WO 97/44470 describe promotores de levadura procedentes de *Yarrowia lipolytica* para la proteína del factor 1 de alargamiento de la traducción ("translation elongation factor 1", TEF1) y para la proteína ribosómica S7, que son adecuados para la expresión heteróloga de las proteínas en levaduras, y el documento EP1951877A1 describe el uso del promotor de TEF1 de *P. pastoris* para la producción de proteínas heterólogas.

25 El documento WO2005003310 proporciona métodos para la expresión de una secuencia codificadora de interés en levaduras usando un promotor de la gliceraldehído-3-fosfato deshidrogenasa o la fosfoglicerato mutasa procedente de la levadura oleaginoso *Yarrowia lipolytica*.

30 Se describen secuencias de promotores derivadas de genes implicados en la vía metabólica del metanol de *Pichia pastoris* en los documentos US4808537 y US4855231 (alcohol oxidasa AOX1, AOX2) y el documento US6730499B1 (formaldehído deshidrogenasa FLD1). El documento US20080153126A1 incluye secuencias de promotores mutantes basadas en el promotor de AOX1.

El promotor de AOX1 se induce solo en respuesta al metanol y es reprimido por otras fuentes de carbono, tales como glucosa o etanol. El metanol tiene la desventaja de que no resulta adecuado para su uso en la producción de ciertos productos, puesto que es potencialmente peligroso debido a su toxicidad e inflamabilidad. Por tanto, se buscan alternativas al promotor de AOX1.

35 El documento US2008299616A1 describe las secuencias reguladoras del gen de la malato sintasa (MLS1) para la expresión de genes heterólogos en *P. pastoris*, que es reprimido en medios que contienen glucosa y se desreprime bajo condiciones de privación de glucosa o cuando está presente acetato. Sin embargo, este sistema no se considera adecuado para métodos de producción eficaces, puesto que el promotor de MLS1 es débil y presenta poca actividad bajo condiciones de desrepresión.

40 Schöler y Schüller (Mol. Cell. Biol., 1994 14(6):3613-3622) describen la región de control del gen de la isocitrato liasa ICL1, que es desreprimido después de la transferencia de células desde unas condiciones de crecimiento fermentativas a no fermentativas.

45 El documento WO2008063302A2 describe el uso de nuevos promotores inducibles derivados de los genes de ADH1 (alcohol deshidrogenasa), ENO1 (enolasa) y GUT1 de *P. pastoris* para la expresión de proteínas heterólogas, el documento CN1966688A describe la secuencia del promotor de la ácido graso omega 3 deshidrogenasa de *P. pastoris*, y el documento WO002007117062A1 describe el promotor de NPS autoinducible derivado de *P. pastoris*, que es inducido por la limitación de fósforo.

50 El documento WO2008128701A2 describe el uso de nuevos promotores, de los cuales el promotor derivado del gen de THI3 (metabolismo de tiamina) de *P. pastoris* es reprimido en un medio que contiene tiamina, y se desreprime tras la merma de la tiamina.

El documento US2009325241A1 describe un método de producción de etanol en una célula de levadura que emplea un promotor inducible por xilosa (promotor de FAS2).

Resulta deseable proporcionar líneas de células eucariotas recombinantes mejoradas para producir productos de

fermentación que puedan aislarse con altos rendimientos. Por tanto, un objeto de la presente invención consiste en proporcionar elementos reguladores alternativos adecuados para métodos de producción recombinantes, que sean sencillos y eficaces.

Sumario de la invención

5 El objeto se resuelve por medio de los aspectos reivindicados.

La presente invención proporciona un método para producir una proteína de interés (PDI) mediante el cultivo de una línea de células eucariotas recombinantes que comprenden una construcción de expresión que comprende un promotor regulable y una molécula de ácido nucleico que codifica una PDI bajo el control transcripcional de dicho promotor, que comprende las etapas de:

10 a) cultivar la línea de células con una fuente de carbono basal que reprime al promotor, en el que la fuente de carbono basal es una fuente de carbono adecuada para el crecimiento de las células,

b) cultivar la línea de células sin una fuente de carbono suplementaria, o con una cantidad limitada de esta, que desreprime al promotor para inducir la producción de la PDI a una tasa de transcripción de al menos 20%, comparado con el promotor pGAP nativo de la célula, y

15 c) producir y recuperar la PDI,

en el que el promotor se selecciona del grupo que consiste en:

(i) pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); y

20 (ii) una secuencia que comprende una secuencia de nucleótidos de al menos 200 pb y que tiene al menos 80% de identidad de secuencia con una cualquiera de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); o con una parte de cualquiera de las anteriores con una longitud de al menos 200 pb.

Específicamente, la fuente de carbono basal se selecciona del grupo que consiste en glucosa, glicerol, etanol, una de sus mezclas y un material nutriente complejo.

25 Específicamente, la fuente de carbono suplementaria es una hexosa, tal como glucosa, fructosa, galactosa o manosa, un disacárido, tal como sacarosa, un alcohol, tal como glicerol o etanol, o una de sus mezclas.

Específicamente la etapa b) emplea un medio de alimentación que no proporciona la fuente de carbono suplementaria o la proporciona en una cantidad limitada, preferiblemente a 0-1 g/l en el medio de cultivo.

30 Específicamente, la cantidad limitada de la fuente de carbono suplementaria limita el crecimiento para mantener la tasa de crecimiento específica dentro del intervalo de 0,02 h⁻¹ hasta 0,2 h⁻¹, preferiblemente de 0,02 h⁻¹ hasta 0,15 h⁻¹.

35 Específicamente, el promotor es capaz de controlar la transcripción de un gen de una célula eucariota de tipo salvaje, y dicho gen se selecciona del grupo que consiste en G1 (SEQ ID NO:7), G3 (SEQ ID NO:8), G4 (SEQ ID NO:9), G6 (SEQ ID NO:10), G7 (SEQ ID NO:11) y G8 (SEQ ID NO:12), o uno de sus variantes funcionalmente activos que comprende una secuencia de nucleótidos de al menos 200 pb y una identidad de secuencia de al menos 80% con cualquiera de las secuencias de promotores o con una de sus partes con una longitud de al menos 200 pb.

Específicamente, el promotor es pG1 (SEQ ID NO:1) o uno de sus variantes funcionalmente activos.

40 Específicamente, dichos variantes funcionalmente activos se seleccionan del grupo que consiste en homólogos que pueden obtenerse por medio de la modificación de la secuencia de nucleótidos de origen mediante la inserción, la deleción o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, y análogos derivados de especies distintas de *Pichia pastoris*.

Específicamente, el promotor se selecciona del grupo que consiste en pG1 (SEQ ID NO:1), pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46).

45 Específicamente, la línea de células se selecciona del grupo que consiste en líneas de células de mamífero, insecto, levadura, hongo filamentoso y vegetales, preferiblemente una levadura.

50 Específicamente, la PDI es una proteína heteróloga, preferiblemente seleccionada de proteínas terapéuticas, que incluyen anticuerpos o sus fragmentos, enzimas y péptidos, antibióticos de proteínas, proteínas de fusión de toxinas, conjugados de carbohidrato-proteína, proteínas estructurales, proteínas reguladoras, vacunas y proteínas o partículas similares a vacunas, enzimas de procesos, factores de crecimiento, hormonas y citoquinas, o un

metabolito de una PDI.

Específicamente el promotor es un variante funcionalmente activo de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6), que se selecciona del grupo que consiste en homólogos que pueden obtenerse por medio de la modificación de la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, y análogos derivados de especies distintas de *Pichia pastoris*.

Específicamente, el promotor se selecciona del grupo que consiste en pG1 (SEQ ID NO:1), pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46).

La invención proporciona además un ácido nucleico de promotor aislado que está unido operablemente a una secuencia de nucleótidos que codifica una PDI, y dicho ácido nucleico de promotor no está asociado nativamente con la secuencia de nucleótidos que codifica la PDI, en el que el ácido nucleico de promotor comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en:

a) pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); y

(b) una secuencia que comprende una secuencia de nucleótidos de al menos 200 pb y que tiene al menos 80% de identidad de secuencia con una cualquiera de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); o con una parte de cualquiera de las anteriores con una longitud de al menos 200 pb;

en el que dicho ácido nucleico comprende un promotor funcionalmente activo, que es un promotor regulable por una fuente de carbono capaz de expresar una PDI en una célula eucariota recombinante a una tasa de transcripción de al menos 20%, comparado con el promotor pGAP nativo de la célula.

Específicamente el promotor es un variante funcionalmente activo de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6), que se selecciona del grupo que consiste en homólogos que pueden obtenerse por medio de la modificación de la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, y análogos derivados de especies distintas de *Pichia pastoris*.

Específicamente, el promotor se selecciona del grupo que consiste en pG1 (SEQ ID NO:1), pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46).

La invención proporciona además un vector que comprende el ácido nucleico de promotor descrito en la presente, en el que la secuencia de nucleótidos que codifica una PDI está bajo el control transcripcional de dicho promotor, y dicho ácido nucleico de promotor no está nativamente asociado con la secuencia de nucleótidos que codifica la PDI.

La invención proporciona además una célula eucariota recombinante que comprende el vector descrito en la presente.

La invención proporciona además un método para producir una proteína de interés (PDI) usando la célula eucariota descrita en la presente.

En la presente se proporciona un método para producir una proteína de interés (PDI) mediante el cultivo de una línea de células eucariotas recombinantes que comprenden una construcción de expresión que comprende un promotor regulable y una molécula de ácido nucleico que codifica una PDI bajo el control transcripcional de dicho promotor, que comprende las etapas de:

a) cultivar la línea de células con una fuente de carbono basal que reprime al promotor,

b) cultivar la línea de células sin una fuente de carbono suplementaria, o con una cantidad limitada de esta, que desreprime al promotor para inducir la producción de la PDI a una tasa de transcripción de al menos 20%, comparado con el promotor pGAP nativo de la célula, y

c) producir y recuperar la PDI.

Dichas etapas de cultivo comprenden específicamente cultivar la línea de células en presencia de dichas fuentes de carbono, por tanto en un medio de cultivo que comprende dichas fuentes de carbono, o en la etapa b) también en ausencia de una fuente de carbono suplementaria.

Dicha inducción de la producción de PDI se refiere específicamente a la inducción de la transcripción, que incluye específicamente una posterior traducción y una expresión opcional de dicha PDI.

Dicha tasa de transcripción se refiere específicamente a la cantidad de transcritos obtenidos tras inducir totalmente a dicho promotor. Dicho promotor se considera desreprimido y totalmente inducido si las condiciones de cultivo proporcionan aproximadamente una inducción máxima, por ejemplo, a unas concentraciones de glucosa menores que 0,4 g/l, preferiblemente menores que 0,04 g/l, de modo específico menores que 0,02 g/l. El promotor totalmente inducido preferiblemente muestra una tasa de transcripción de al menos 15%, preferiblemente al menos 20%, más preferiblemente al menos 30%, 40%, 50%, 60%, 70%, 80%, 90% y al menos 100% o una tasa de transcripción incluso mayor de al menos 150% o al menos 200%, comparado con el promotor pGAP nativo. La tasa de transcripción puede determinarse, por ejemplo, mediante la cantidad de transcritos de un gen indicador, tal como eGFP, según se describe en la sección de ejemplos a continuación, que muestra una tasa de transcripción relativamente alta del promotor pG1 de al menos 50%, comparado con el promotor pGAP nativo de la célula, tras cultivar un clon en disolución. Como alternativa, la tasa de transcripción puede determinarse por medio de la potencia de transcripción en una micromatriz, en la que los datos de la micromatriz muestran la diferencia en el nivel de expresión entre el estado reprimido y desreprimido y una alta señal de intensidad en el estado totalmente inducido, comparado con el promotor pGAP nativo. Estos datos de la micromatriz muestran específicamente una tasa de transcripción mayor que 200% para pG1, mayor que 30% para pG3 y pG4, mayor que 60% para pG6, mayor que 30% para pG7, mayor que 20% para pG8, y cada valor se compara con el pGAP nativo. Se descubrió que los promotores de MLS1 o ICL1 de la técnica anterior eran demasiado débiles y, por tanto, no resultaban adecuados para el objetivo descrito en la presente.

Dicho promotor pGAP nativo es activo específicamente en dicha célula eucariota recombinante de una manera similar al modo en que lo es en una célula eucariota nativa de la misma especie o cepa, que incluye la célula eucariota no modificada (no recombinante) o recombinante. Se entiende habitualmente que dicho promotor pGAP nativo es un promotor endógeno, por tanto homólogo a la célula eucariota, y actúa como promotor patrón o de referencia para fines comparativos.

Por ejemplo, un promotor pGAP nativo de *P. pastoris* es la secuencia de promotor endógeno no modificada de *P. pastoris*, tal como se emplea para controlar la expresión de GAPDH en *P. pastoris*, por ejemplo que tiene la secuencia mostrada en la figura 13, que es la secuencia del promotor pGAP nativo de *P. pastoris* (GS115) (SEQ ID NO:13). Si se emplea *P. pastoris* como hospedante para producir una PDI descrita en la presente, la tasa o potencia de transcripción del promotor descrito en la presente se compara con dicho promotor pGAP nativo de *P. pastoris*.

Como otro ejemplo, un promotor pGAP nativo de *S. cerevisiae* es la secuencia de promotor endógeno no modificada de *S. cerevisiae*, tal como se emplea para controlar la expresión de GAPDH en *S. cerevisiae*. Si se emplea *S. cerevisiae* como hospedante para producir una PDI descrita en la presente, la tasa o potencia de transcripción del promotor descrito en la presente se compara con dicho promotor pGAP nativo de *S. cerevisiae*.

Por tanto, la tasa o potencia de transcripción relativa de un promotor proporcionado en la presente normalmente se compara con el promotor pGAP nativo de una célula de la misma especie o cepa que se emplea como hospedante para producir una PDI.

Según una realización específica, la fuente de carbono basal es diferente de la fuente de carbono suplementaria, por ejemplo, es diferente desde el punto de vista cuantitativo y/o cualitativo. La diferencia cuantitativa puede proporcionar diferentes condiciones para reprimir o desreprimir la actividad del promotor.

Según otra realización específica, las fuentes de carbono basal y suplementaria comprenden el mismo tipo de moléculas o carbohidratos, preferiblemente en diferentes concentraciones. Según otra realización específica, la fuente de carbono es una mezcla de dos o más fuentes de carbono diferentes.

Puede usarse cualquier tipo de carbono orgánico adecuado utilizado para el cultivo de células eucariotas. Según una realización específica, la fuente de carbono es una hexosa, tal como glucosa, fructosa, galactosa o manosa, un disacárido, tal como sacarosa, un alcohol, tal como glicerol o etanol, o una de sus mezclas.

Según una realización específicamente preferida, la fuente de carbono basal se selecciona del grupo que consiste en glucosa, glicerol, etanol, una de sus mezclas y un material nutriente complejo. Según una realización preferida, la fuente de carbono basal es glicerol.

Según otra realización específica, la fuente de carbono suplementaria es una hexosa, tal como glucosa, fructosa, galactosa y manosa, un disacárido, tal como sacarosa, un alcohol, tal como glicerol o etanol, o una de sus mezclas. Según una realización preferida, la fuente de carbono suplementaria es glucosa.

Específicamente, el método puede emplear glicerol como la fuente de carbono basal, y glucosa como la fuente de carbono suplementaria.

Las condiciones de desrepresión pueden lograrse de modo adecuado por medios específicos. La etapa b) opcionalmente emplea un medio de alimentación que no proporciona la fuente de carbono suplementaria o la proporciona en una cantidad limitada.

Específicamente, el medio de alimentación está químicamente definido y no contiene metanol.

- El medio de alimentación puede añadirse al medio de cultivo en forma líquida o en otra forma alternativa, tal como un sólido, por ejemplo, como un comprimido u otro medio de liberación sostenida, o un gas, por ejemplo, dióxido de carbono. Según una realización preferida, la cantidad limitada de la fuente de carbono suplementaria añadida al medio de cultivo de células incluso puede ser cero. Preferiblemente, la concentración de la fuente de carbono suplementaria en el medio de cultivo es de 0-1 g/l, preferiblemente menor que 0,6 g/l, más preferiblemente menor que 0,3 g/l, más preferiblemente menor que 0,1 g/l, preferiblemente 1-50 mg/l, más preferiblemente 1-10 mg/l, y se prefiere específicamente 1 mg/l o incluso menor, tal como por debajo del límite de detección, según se mide con un ensayo convencional adecuado, por ejemplo determinado como la concentración residual en el medio de cultivo tras el consumo por el cultivo de las células en crecimiento.
- 5
- En un método preferido, la cantidad limitada de la fuente suplementaria proporciona una cantidad residual en el cultivo de células que está por debajo del límite de detección, según se determina en el caldo de fermentación al final de una fase de producción o en la salida de un proceso de fermentación, preferiblemente tras recolectar el producto de la fermentación.
- 10
- Preferiblemente, la cantidad limitada de la fuente de carbono suplementaria limita el crecimiento para mantener la tasa de crecimiento específica dentro del intervalo de 0,02 h⁻¹ hasta 0,2 h⁻¹ preferiblemente de 0,02 h⁻¹ hasta 0,15 h⁻¹.
- 15
- Según una realización específica, el promotor es un promotor de *Pichia pastoris* o uno de sus variantes funcionalmente activos.
- Según una realización específica, el promotor siempre se refiere a las secuencias descritas en la presente y sus variantes funcionalmente activos. Tal como se explica en detalle a continuación, dichos variantes incluyen homólogos y análogos derivados de especies distintas de *Pichia pastoris*.
- 20
- El método proporcionado en la presente puede emplear un promotor que es un promotor de tipo salvaje de *P. pastoris* o uno de sus variantes funcionalmente activos, por ejemplo, capaz de controlar la transcripción de un gen específico en una célula eucariota recombinante o de tipo salvaje, por ejemplo, un promotor de tipo salvaje para genes seleccionados, y dicho gen se selecciona del grupo que consiste en G1 (SEQ ID NO:7), tal como un gen que codifica un transportador de glucosa (de alta afinidad), G3 (SEQ ID NO:8), G4 (SEQ ID NO:9), tal como un gen que codifica una aldehído deshidrogenasa mitocondrial, G6 (SEQ ID NO:10), G7 (SEQ ID NO:11), tal como un gen que codifica un miembro de la familia de facilitadores principales de transportadores de azúcares, o G8 (SEQ ID NO:12), tal como un gen que codifica un miembro de la superfamilia Gti1_Pac2, o uno de sus variantes funcionalmente activos.
- 25
- 30
- Específicamente, se proporciona un promotor, o uno de sus variantes funcionalmente activos, que estaría asociado nativamente con uno de dichos genes en una célula de levadura de tipo salvaje.
- Según una realización específica, la línea de células se selecciona del grupo que consiste en líneas de células de mamífero, insecto, levadura, hongo filamentoso y vegetales, preferiblemente una levadura.
- 35
- Específicamente la levadura se selecciona del grupo que consiste en *Pichia*, *Candida*, *Torulopsis*, *Arxula*, *Hansenula*, *Yarrowia*, *Kluyveromyces*, *Saccharomyces*, *Komagataella*, preferiblemente una levadura metilotrófica.
- Una levadura específicamente preferida es *Pichia pastoris*, *Komagataella pastoris*, *K. phaffii*, o *K. pseudopastoris*.
- Según otra realización específica, el promotor no está asociado nativamente con la secuencia de nucleótidos que codifica la PDI.
- 40
- Específicamente, la PDI es una proteína eucariótica, preferiblemente una proteína de mamífero.
- Una PDI producida según el método proporcionado en la presente puede ser una proteína multimérica, preferiblemente un dímero o un tetrámero.
- Según una realización específica, la PDI es una proteína recombinante o heteróloga, preferiblemente seleccionada de proteínas terapéuticas, que incluyen anticuerpos o sus fragmentos, enzimas y péptidos, antibióticos de proteínas, proteínas de fusión de toxinas, conjugados de carbohidrato-proteína, proteínas estructurales, proteínas reguladoras, vacunas y proteínas o partículas similares a vacunas, enzimas de procesos, factores de crecimiento, hormonas y citoquinas, o un metabolito de una PDI.
- 45
- 50
- Una PDI específica es una molécula de unión a un antígeno, tal como un anticuerpo, o uno de sus fragmentos. Entre las PDI específicas se encuentran anticuerpos, tales como anticuerpos monoclonales (mAb), inmunoglobulinas (Ig) o inmunoglobulinas de clase G (IgG), anticuerpos de cadena pesada (HcAb), o sus fragmentos, tales como el fragmento de unión al antígeno (Fab), Fd, fragmentos variables monocatenarios (mcFv), o sus variantes modificados, tales como, por ejemplo, dímeros de Fv (diacuerpos), trímeros de Fv (triacuerpos), tetrámeros de Fv, o minicuerpos y anticuerpos de dominio único, tales como VH o VHH o V-NAR.
- Según una realización específica, se fabrica un producto de fermentación usando la PDI, un metabolito o un

derivado de esta.

Según una realización específica, se proporciona un método para controlar la expresión de una PDI en una célula eucariota recombinante bajo el control transcripcional de un promotor regulable por una fuente de carbono que tiene una potencia de transcripción de al menos 20%, comparado con el promotor pGAP nativo de la célula, en el que la expresión se induce bajo condiciones que limitan la fuente de carbono. El promotor regulable por una fuente de carbono preferiblemente tiene una potencia de transcripción de al menos 20%, comparado con el promotor pGAP nativo, y específicamente tiene una potencia de transcripción tal como se describió anteriormente con respecto a la tasa de transcripción, comparado con el promotor pGAP nativo. Por tanto, el promotor totalmente inducido preferiblemente tiene una potencia de transcripción de al menos 20%, preferiblemente al menos 30%, 40%, 50%, 60%, 70%, 80%, 90% y al menos 100% o una potencia de transcripción incluso mayor de al menos 150% o al menos 200%, comparado con el promotor pGAP nativo, según se determina en la célula eucariota seleccionada para producir la PDI.

En una realización preferida, dicho promotor empleado tiene una actividad transcripcional o una potencia transcripcional en el estado desreprimido que es al menos 2 veces, más preferiblemente al menos 5 veces, aún más preferiblemente al menos 10 veces, más preferiblemente al menos 20 veces, más preferiblemente al menos 30, 40, 50, o 100 veces mayor en el estado desreprimido, comparado con el estado reprimido.

Según una realización específica, se proporciona un método para producir una PDI en una célula eucariota recombinante bajo el control transcripcional de un promotor regulable por una fuente de carbono, en el que dicho promotor tiene una potencia de transcripción tal como se describió anteriormente, es decir, al menos 20%, comparado con el promotor pGAP nativo de la célula. El promotor regulable por una fuente de carbono preferiblemente tiene una potencia de transcripción de al menos 20%, comparado con el promotor pGAP de referencia, más preferiblemente al menos 30%, 40%, 50%, 60%, 70%, 80%, 90% y al menos 100% o una potencia de transcripción incluso mayor de al menos 150% o al menos 200%, comparado con el promotor pGAP nativo. En una realización preferida, dicho promotor empleado tiene una actividad transcripcional en el estado desreprimido que es al menos 2 veces, más preferiblemente al menos 5 veces, aún más preferiblemente al menos 10 veces, más preferiblemente al menos 20 veces, más preferiblemente al menos 30, 40, 50, o 100 veces mayor en el estado desreprimido, comparado con el estado reprimido. De modo adecuado, se emplea un promotor específico, según se describe en la presente, en dicho método.

En una realización específicamente preferida, el promotor es un promotor regulable que comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en:

a) pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6);

b) una secuencia que tiene al menos 80% de homología con pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6);

c) una secuencia que se hibrida bajo condiciones rigurosas con pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); y

d) un fragmento o variante derivado de a), b) o c),

en el que dicho promotor es un promotor funcionalmente activo, que es un promotor regulable por una fuente de carbono capaz de expresar una PDI en una célula eucariota recombinante a una tasa de transcripción de al menos 20%, comparado con el promotor pGAP nativo de la célula.

Específicamente, el variante de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6) es un variante funcionalmente activo seleccionado del grupo que consiste en homólogos con al menos aproximadamente 80% de identidad de secuencia de nucleótidos, homólogos que pueden obtenerse modificando la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, preferiblemente con una secuencia de nucleótidos de al menos 200 pb, preferiblemente al menos 250 pb, preferiblemente al menos 300 pb, más preferiblemente al menos 400 pb, al menos 500 pb, al menos 600 pb, al menos 700 pb, al menos 800 pb, al menos 900 pb, o al menos 1000 pb, y análogos derivados de especies distintas de *Pichia pastoris*.

Algunos de los variantes funcionalmente activos preferidos del promotor descrito en la presente son fragmentos de cualquiera de las secuencias de nucleótidos de los promotores pG1, pG3, pG6, pG7 o pG8, preferiblemente fragmentos que incluyen el extremo 3' de una secuencia de nucleótidos de promotor, por ejemplo, una secuencia de nucleótidos derivada de una de las secuencias de nucleótidos de promotor que tiene una longitud específica y una delección de la región 5' terminal, por ejemplo, una escisión de la secuencia de nucleótidos en el extremo 5', para obtener una longitud específica con una amplitud desde el extremo 3' hasta un extremo 5' variable, tal como con una longitud de la secuencia de nucleótidos de al menos 200 pb, preferiblemente al menos 250 pb, preferiblemente al menos 300 pb, más preferiblemente al menos 400 pb, al menos 500 pb, al menos 600 pb, al menos 700 pb, al

menos 800 pb, al menos 900 pb, o al menos 1000 pb.

Los ejemplos de variantes que han demostrado ser funcionalmente activos comprenden o consisten en dichos fragmentos, por ejemplo, fragmentos con una longitud específica con una amplitud de 200 a 1000 pb, preferiblemente con una amplitud de 250 a 1000 pb, más preferiblemente con una amplitud de 300 a 1000 pb, por ejemplo, que incluyen la secuencia 3' terminal. Por ejemplo, un variante funcionalmente activo de pG1 se selecciona del grupo que consiste en pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46), por tanto, una secuencia de nucleótidos con una amplitud de 300-1000 pb, que incluye la secuencia 3' terminal hasta el nucleótido 1001.

Según otra realización específica, se proporciona un ácido nucleico aislado que comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en:

a) pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6),

b) una secuencia que tiene al menos 80% de homología con pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6),

c) una secuencia que se hibrida bajo condiciones rigurosas con pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6), y

d) un fragmento o variante derivado de a), b) o c),

en el que dicho ácido nucleico comprende un promotor funcionalmente activo, que es un promotor regulable por una fuente de carbono capaz de expresar una PDI en una célula eucariota recombinante a una tasa de transcripción de al menos 20%, comparado con el promotor pGAP nativo de la célula.

Específicamente, el variante de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6) es un variante funcionalmente activo seleccionado del grupo que consiste en homólogos con al menos aproximadamente 80% de identidad de secuencia de nucleótidos, homólogos que pueden obtenerse modificando la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, preferiblemente con una secuencia de nucleótidos de al menos 200 pb, preferiblemente al menos 250 pb, preferiblemente al menos 300 pb, más preferiblemente al menos 400 pb, al menos 500 pb, al menos 600 pb, al menos 700 pb, al menos 800 pb, al menos 900 pb, o al menos 1000 pb, y análogos derivados de especies distintas de *Pichia pastoris*.

Algunos de los variantes funcionalmente activos preferidos del promotor descrito en la presente son fragmentos de cualquiera de las secuencias de nucleótidos de los promotores pG1, pG3, pG4, pG6, pG7 o pG8, preferiblemente fragmentos que incluyen el extremo 3' de una secuencia de nucleótidos de promotor, por ejemplo, una secuencia de nucleótidos derivada de una de las secuencias de nucleótidos de promotor que tiene una longitud específica y una delección de la región 5' terminal, por ejemplo, una escisión de la secuencia de nucleótidos en el extremo 5', para obtener una longitud específica con una amplitud desde el extremo 3' hasta un extremo 5' variable, tal como con una longitud de la secuencia de nucleótidos de al menos 200 pb, preferiblemente al menos 250 pb, preferiblemente al menos 300 pb, más preferiblemente al menos 400 pb, al menos 500 pb, al menos 600 pb, al menos 700 pb, al menos 800 pb, al menos 900 pb, o al menos 1000 pb.

Los ejemplos de variantes que han demostrado ser funcionalmente activos comprenden o consisten en dichos fragmentos, por ejemplo, fragmentos con una longitud específica con una amplitud de 200 a 1000 pb, preferiblemente con una amplitud de 250 a 1000 pb, más preferiblemente con una amplitud de 300 a 1000 pb, por ejemplo, que incluyen la secuencia 3' terminal. Por ejemplo, un variante funcionalmente activo de pG1 se selecciona del grupo que consiste en pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46), por tanto, una secuencia de nucleótidos con una amplitud de 300-1000 pb, que incluye la secuencia 3' terminal hasta el nucleótido 1001.

El promotor regulable por una fuente de carbono preferiblemente tiene una potencia de transcripción como se describió anteriormente, preferiblemente de al menos 20%, comparado con el promotor pGAP de referencia, más preferiblemente al menos 30%, 40%, 50%, 60%, 70%, 80%, 90% y al menos 100% o una potencia de transcripción incluso mayor de al menos 150% o al menos 200%, comparado con el promotor pGAP nativo. En una realización preferida, dicho promotor empleado tiene una actividad transcripcional en el estado desreprimido que es al menos 2 veces, más preferiblemente al menos 5 veces, aún más preferiblemente al menos 10 veces, más preferiblemente al menos 20 veces, más preferiblemente al menos 30, 40, 50, o 100 veces mayor en el estado desreprimido, comparado con el estado reprimido. De modo adecuado, se emplea un promotor específico, según se describe en la presente, en dicho método.

Además, según otra realización específica, se proporciona una construcción de expresión que comprende un promotor descrito en la presente, unido operablemente a una secuencia de nucleótidos que codifica una PDI bajo el

control transcripcional de dicho promotor, y dicho promotor no está asociado nativamente con la secuencia codificadora de la PDI.

Otra realización específica se refiere a un vector que comprende la construcción descrita en la presente.

5 Otra realización específica se refiere a una célula eucariota recombinante que comprende la construcción o el vector descritos en la presente.

Específicamente, la célula se selecciona del grupo que consiste en líneas de células de mamífero, insecto, levadura, hongo filamentosos y vegetales, preferiblemente una levadura.

La levadura puede seleccionarse de modo adecuado del grupo que consiste en *Pichia*, *Candida*, *Torulopsis*, *Arxula*, *Hansenula*, *Yarrowia*, *Kluyveromyces*, *Saccharomyces*, *Komagataella*, preferiblemente una levadura metilotrófica.

10 Preferiblemente, la levadura es *Pichia pastoris*, *Komagataella pastoris*, *K. phaffii*, o *K. pseudopastoris*.

Según una realización específica, se emplea una célula que tiene una tasa de crecimiento específica mayor en presencia de un exceso de fuente de carbono, con relación a unas condiciones de fuente de carbono limitada.

Otra realización específica se refiere al uso de la célula eucariota recombinante descrita en la presente para la producción de la PDI.

15 Según otra realización específica, se proporciona un método para seleccionar o identificar un promotor regulable por una fuente de carbono procedente de células eucariotas, que comprende las etapas de:

a) cultivar células eucariotas en presencia de una fuente de carbono en un cultivo discontinuo bajo condiciones de crecimiento de las células,

20 b) cultivar después las células en un cultivo de alimentación discontinua en presencia de una cantidad limitada de una fuente de carbono suplementaria,

c) proporcionar muestras del cultivo de células de la etapa a) y b), y

d) realizar un análisis de la transcripción en dichas muestras para identificar un promotor regulable que muestra una potencia transcripcional mayor en las células de la etapa b) que en las células de la etapa a).

25 Dicha potencia transcripcional mayor puede determinarse mediante la potencia de transcripción en el estado totalmente inducido, que se obtiene, por ejemplo, bajo condiciones de cultivos en quimiostato con limitación de glucosa, que es al menos 2 veces, más preferiblemente al menos 5 veces, aún más preferiblemente al menos 10 veces, más preferiblemente al menos 20 veces, más preferiblemente al menos 30, 40, 50, o 100 veces mayor en el estado desreprimido, comparado con el estado reprimido.

30 Preferiblemente, el análisis de la transcripción es cuantitativo o semicuantitativo, y emplea preferiblemente micromatrices de ADN, secuenciación de ARN y análisis del transcriptoma.

Figuras

Figura 1: Secuencia del promotor pG1 (SEQ ID NO:1) de *P. pastoris*.

Figura 2: Secuencia del promotor pG3 (SEQ ID NO:2) de *P. pastoris*.

Figura 3: Secuencia del promotor pG4 (SEQ ID NO:4) de *P. pastoris*.

35 Figura 4: Secuencia del promotor pG6 (SEQ ID NO:3) de *P. pastoris*.

Figura 5: Secuencia del promotor pG7 (SEQ ID NO:5) de *P. pastoris*.

Figura 6: Secuencia del promotor pG8 (SEQ ID NO:6) de *P. pastoris*.

Figura 7: Secuencias codificadoras del gen G1 del genoma de GS115 (SEQ ID NO:7) de *P. pastoris*.

Figura 8: Secuencias codificadoras del gen G3 del genoma de GS115 (SEQ ID NO:8) de *P. pastoris*.

40 Figura 9: Secuencias codificadoras del gen G4 del genoma de GS115 (SEQ ID NO:9) de *P. pastoris*.

Figura 10: Secuencias codificadoras del gen G6 del genoma de GS115 (SEQ ID NO:10) de *P. pastoris*.

Figura 11: Secuencias codificadoras del gen G7 del genoma de GS115 (SEQ ID NO:11) de *P. pastoris*.

Figura 12: Secuencias codificadoras del gen G8 del genoma de GS115 (SEQ ID NO:12) de *P. pastoris*.

Figura 13: Secuencia de promotor pGAP nativo de *P. pastoris* (GS115) (SEQ ID NO:13)

n.º	Nombre	PAS*	PIPA*	Descripción de GS115
pGAP	TDH3	PAS_chr2-1_0437	PIPA02510	Gliceraldehído-3-fosfato deshidrogenasa
*PAS: nombre de la ORF de <i>P. pastoris</i> GS115; PIPA: nombre de la ORF del tipo de <i>P. pastoris</i> cepa DSMZ70382				

Figura 14: Propiedades de desrepresión del promotor pG1 (círculo), pG3 (triángulo), pG4 (rombo) y pG6 (cuadrado): la máxima actividad de transcripción se alcanza para pG1 a aproximadamente 0,04 g de glucosa/l o menor, mientras que los demás promotores pG la alcanzan a aproximadamente 4 g/l o menor. Para comparar los comportamiento de inducción relativa de los diferentes promotores, los datos se normalizaron dividiendo cada valor entre el valor de D20 de la respectiva construcción del promotor. Por tanto, los datos son valores de fluorescencia relativa, y los puntos de datos en D20 son 1,0.

Figura 15: Variantes funcionalmente activos de la secuencia de promotor pG1; pG1a-f (SEQ ID NO:41-46) de *P. pastoris*.

10 Descripción detallada de la invención

Los términos y expresiones específicos empleados a lo largo de la memoria descriptiva tienen el siguiente significado.

La expresión “fuente de carbono”, tal como se emplea en la presente, significa un sustrato de carbono fermentable, generalmente un carbohidrato, adecuado como fuente de energía para los microorganismos, tales como los que son capaces de ser metabolizados por organismos hospedantes o líneas celulares de producción, en particular fuentes seleccionadas del grupo que consiste en monosacáridos, oligosacáridos, polisacáridos, alcoholes, que incluyen glicerol, en forma purificada o proporcionados como materia prima, tal como un material nutriente complejo. La fuente de carbono puede usarse como se describe en la presente como una única fuente de carbono o como una mezcla de diferentes fuentes de carbono.

Una “fuente de carbono basal”, tal como se emplea en la presente, generalmente es una fuente de carbono adecuada para el crecimiento celular, tal como un nutriente para células eucariotas. La fuente de carbono basal puede proporcionarse en un medio, tal como un medio basal o un medio complejo, pero también en un medio químicamente definido que contiene una fuente de carbono purificada. La fuente de carbono basal generalmente se proporciona en una cantidad para proporcionar el crecimiento celular, en particular durante la fase de crecimiento en un proceso de cultivo, por ejemplo para obtener unas densidades de células de al menos 5 g/l de masa seca de células, preferiblemente al menos 10 g/l de masa seca de células, o al menos 15 g/l de masa seca de células, por ejemplo, que muestran una viabilidades mayores que 90% durante las etapas de subcultivo convencionales, preferiblemente mayor que 95%.

Tal como se describe en la presente, la fuente de carbono basal generalmente se emplea en una cantidad en exceso, lo cual se entiende como un exceso que proporciona energía para aumentar la biomasa, por ejemplo, durante la fase de crecimiento de una línea celular en un proceso de cultivo de alimentación discontinua. Esta cantidad en exceso está, en particular, en exceso de la cantidad limitada de una fuente de carbono suplementaria para lograr una concentración residual en el caldo de fermentación que sea mensurable y generalmente al menos 10 veces mayor, preferiblemente al menos 50 veces o al menos 100 veces mayor que durante el suministro de la cantidad limitada de la fuente de carbono suplementaria.

La expresión “químicamente definido” con respecto a un medio de cultivo de células, tal como un medio de alimentación en un proceso de alimentación discontinua, significa un medio de crecimiento adecuado para el cultivo de células *in vitro* de una línea celular de producción, en el que todos los componentes químicos y péptidos son conocidos. Generalmente, un medio químicamente definido está totalmente exento de componentes derivados de animales y representa un entorno de cultivo celular puro y constante.

Una “fuente de carbono suplementaria”, tal como se emplea en la presente, generalmente es un sustrato suplementario que facilita la producción de productos de fermentación por las líneas celulares de producción, en particular en la fase de producción de un proceso de cultivo. La fase de producción específicamente sigue a la fase de crecimiento, por ejemplo, en un proceso de cultivo continuo, de alimentación discontinua y discontinuo. La fuente de carbono suplementaria específicamente puede estar contenida en la alimentación de un proceso de alimentación discontinua.

Una “cantidad limitada” de una fuente de carbono o una “fuente de carbono limitada” se entiende en la presente como la cantidad de una fuente de carbono necesaria para mantener una línea celular de producción en una fase de producción o un modo de producción. Dicha cantidad limitada puede emplearse en un proceso de alimentación discontinua, en el que la fuente de carbono está contenida en un medio de alimentación y se suministra al cultivo a bajas tasas de alimentación para el suministro sostenido de energía para producir una PDI, al mismo tiempo que se

mantiene la biomasa a unas tasas de crecimiento bajas. Generalmente, se añade un medio de alimentación al caldo de fermentación durante la fase de producción de un cultivo de células.

5 La cantidad limitada de la fuente de carbono suplementaria puede determinarse, por ejemplo, por medio de la cantidad residual de la fuente de carbono suplementaria en el caldo de cultivo celular, que está por debajo de un umbral predeterminado o incluso por debajo del límite de detección, tal como se mide en un ensayo convencional (carbohidratos). La cantidad residual generalmente se determina en el caldo de fermentación tras recolectar un producto de fermentación.

10 La cantidad limitada de una fuente de carbono suplementaria puede determinarse definiendo la tasa promedio de alimentación de la fuente de carbono suplementaria al fermentador, por ejemplo, según se determina mediante la cantidad añadida a lo largo del proceso de cultivo completo, por ejemplo, la fase de alimentación discontinua, por tiempo de cultivo, para determinar una cantidad promedio calculada por tiempo. Esta tasa promedio de alimentación se mantiene baja para asegurar el uso completo de la fuente de carbono suplementaria por el cultivo celular, por ejemplo, entre $0,6 \text{ g l}^{-1} \text{ h}^{-1}$ (g de fuente de carbono por l de volumen de fermentación inicial y h tiempo) y $25 \text{ g l}^{-1} \text{ h}^{-1}$, preferiblemente entre $1,6 \text{ g l}^{-1} \text{ h}^{-1}$ y $20 \text{ g l}^{-1} \text{ h}^{-1}$.

15 La cantidad limitada de una fuente de carbono suplementaria también puede determinarse midiendo la tasa de crecimiento específica antes y durante el proceso de producción, en el que la tasa de crecimiento específica se mantiene baja durante el proceso de producción, por ejemplo, dentro de un intervalo predeterminado, tal como en el intervalo de $0,02 \text{ h}^{-1}$ a $0,20 \text{ h}^{-1}$, preferiblemente entre $0,02 \text{ h}^{-1}$ y $0,15 \text{ h}^{-1}$.

20 La expresión "línea celular" o "línea de células", tal como se emplea en la presente, se refiere a un clon establecido de un tipo concreto de célula que ha adquirido la capacidad de proliferar a lo largo de un periodo de tiempo prolongado. La expresión "línea de células hospedantes" se refiere a una línea de células tal como se emplea para expresar un gen endógeno o recombinante o productos de una vía metabólica para producir polipéptidos o metabolitos de células mediados por dichos polipéptidos. Una "línea de células hospedantes de producción" o "línea celular de producción" se entiende habitualmente como una línea de células listas para usar para el cultivo en un biorreactor para obtener el producto de un proceso de producción, tal como una PDI. La expresión "hospedante eucariota" o "línea de células eucariotas" significa cualquier organismo o célula eucariota que puede cultivarse para producir una PDI o un metabolito de la célula hospedante. Se entiende que la expresión no incluye a los seres humanos.

30 El término "expresión" o las expresiones "sistema de expresión" o "módulo de expresión" se refieren a moléculas de ácidos nucleicos que contienen una secuencia codificadora deseada y secuencias de control en una unión operable, de modo que los hospedantes transformados o transfectados con estas secuencias son capaces de producir las proteínas codificadas o los metabolitos de la célula hospedante. Para realizar la transformación, el sistema de expresión puede incluirse en un vector; sin embargo, el ADN pertinente también puede integrarse en el cromosoma del hospedante. La expresión puede referirse a productos de expresión segregados o no segregados, que incluyen polipéptidos o metabolitos.

35 Las "construcciones de expresión" o "vectores" empleados en la presente se definen como secuencias de ADN que son necesarias para la transcripción de secuencias de nucleótidos recombinantes clonadas, es decir, de genes recombinantes, y para la traducción de su ARNm en un organismo hospedante adecuado. Los vectores de expresión habitualmente comprenden un origen para la replicación autónoma en las células hospedantes, marcadores seleccionables (por ejemplo, un gen de síntesis de aminoácidos o un gen que confiere resistencia a antibióticos, tales como zeocina, kanamicina, G418 o higromicina), una serie de sitios de ruptura de enzimas de restricción, una secuencia de promotor adecuada y un terminador de la transcripción, y estos componentes están unidos operablemente entre sí. Los términos "plásmido" y "vector", tal como se emplean en la presente, incluyen secuencias de nucleótidos de replicación autónoma, así como secuencias de nucleótidos que se integran en el genoma.

45 El término "variante", tal como se emplea en la presente, se refiere a cualquier secuencia con una homología o analogía específica. El promotor variante puede derivarse, por ejemplo, de la secuencia de promotor pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6) mediante mutagénesis para producir secuencias adecuadas para su uso como promotor en líneas de células recombinantes. Estos promotores variantes pueden obtenerse a partir de un banco de secuencias mutantes seleccionando los miembros del banco con propiedades predeterminadas. Los promotores variantes pueden tener las mismas propiedades o propiedades incluso mejoradas, por ejemplo, mejoradas para inducir la producción de una PDI, con un mayor efecto diferencial bajo condiciones de represión y desrepresión. El promotor variante también puede derivarse de secuencias análogas, por ejemplo, de especies eucariotas distintas de *Pichia pastoris* o de un género distinto de *Pichia*, tal como de *K. lactis*, *Z. rouxii*, *P. stipitis*, *H. polymorpha*. Específicamente, pueden usarse las secuencias de promotores análogos nativamente asociadas con genes análogos a los correspondientes genes de *P. pastoris* como tales, o como secuencias de origen para producir sus variantes funcionalmente activos. De modo específico:

- un promotor análogo a pG1 se caracteriza porque está nativamente asociado con un gen análogo a G1 (transportador de glucosa de alta afinidad; descripción en *P. pastoris* GS115: transportador putativo, miembro de la

familia de portadores de azúcares; secuencia codificadora: SEQ ID NO:7);

- un promotor análogo a pG3 se caracteriza porque está nativamente asociado con un gen análogo a G3 (secuencia codificadora: SEQ ID NO:8);

5 - un promotor análogo a pG4 se caracteriza porque está nativamente asociado con un gen análogo a G4 (*P. pastoris* GS115: aldehído deshidrogenasa mitocondrial predicha; secuencia codificadora: SEQ ID NO:9);

- un promotor análogo a pG6 se caracteriza porque está nativamente asociado con un gen análogo a G6 (secuencia codificadora: SEQ ID NO:10);

10 - un promotor análogo a pG7 se caracteriza porque está nativamente asociado con un gen análogo a G7 (*P. pastoris* GS115: miembro de la familia de facilitadores principales de transportadores de azúcar; secuencia codificadora: SEQ ID NO:11);

- un promotor análogo a pG8 se caracteriza porque está nativamente asociado con un gen análogo a G8 (*P. pastoris* GS115: miembro de la superfamilia Gti1_Pac2; secuencia codificadora: SEQ ID NO:12).

Las propiedades de dichas secuencias de promotores análogos o de sus variantes funcionalmente activos pueden determinarse usando técnicas convencionales.

15 El variante “funcionalmente activo” de una secuencia de nucleótidos o de promotor, tal como se emplea en la presente, significa una secuencia que surge de la modificación de una secuencia de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, y dicha modificación no afecta (en particular, dificulta) la actividad de esta secuencia.

20 Específicamente, el variante funcionalmente activo de la secuencia de promotor descrita en la presente se selecciona del grupo que consiste en:

- homólogos con al menos aproximadamente 80% de identidad de secuencia de nucleótidos,

25 - homólogos que pueden obtenerse modificando la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, preferiblemente con una secuencia de nucleótidos (es decir, que la comprende o que consiste en ella) de al menos 200 pb, preferiblemente al menos 300 pb, más preferiblemente al menos 400 pb, al menos 500 pb, al menos 600 pb, al menos 700 pb, al menos 800 pb, al menos 900 pb, o al menos 1000 pb, y

- análogos derivados de especies distintas de *Pichia pastoris*.

30 Los variantes funcionalmente activos específicamente preferidos son los que se derivan de un promotor descrito en la presente y/o fragmentos de la secuencia de promotor, con una secuencia de nucleótidos (es decir, que la comprende o que consiste en ella) de al menos 200 pb, preferiblemente al menos 250 pb, preferiblemente al menos 300 pb, más preferiblemente al menos 400 pb, al menos 500 pb, al menos 600 pb, al menos 700 pb, al menos 800 pb, al menos 900 pb, o al menos 1000 pb.

35 Algunos de los variantes funcionalmente activos preferidos del promotor descrito en la presente son fragmentos de cualquiera de las secuencias de nucleótidos de los promotores pG1, pG3, pG4, pG6, pG7 o pG8, preferiblemente fragmentos que incluyen el extremo 3' de una secuencia de nucleótidos de promotor, por ejemplo, una secuencia de nucleótidos derivada de una de las secuencias de nucleótidos de promotor que tiene una longitud específica y una delección de la región 5' terminal, por ejemplo, una escisión de la secuencia de nucleótidos en el extremo 5', para obtener una longitud específica con una amplitud desde el extremo 3' hasta un extremo 5' variable, tal como con una longitud de la secuencia de nucleótidos de al menos 200 pb, preferiblemente al menos 250 pb, preferiblemente al menos 300 pb, más preferiblemente al menos 400 pb, al menos 500 pb, al menos 600 pb, al menos 700 pb, al menos 800 pb, al menos 900 pb, o al menos 1000 pb.

40 Los ejemplos de variantes que han demostrado ser funcionalmente activos comprenden o consisten en dichos fragmentos, por ejemplo, fragmentos con una longitud específica con una amplitud de 200 a 1000 pb, preferiblemente con una amplitud de 250 a 1000 pb, más preferiblemente con una amplitud de 300 a 1000 pb, por ejemplo, que incluyen la secuencia 3' terminal. Por ejemplo, un variante funcionalmente activo de pG1 se selecciona del grupo que consiste en pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46), por tanto, una secuencia de nucleótidos con una amplitud de 300-1000 pb, que incluye la secuencia 3' terminal hasta el nucleótido 1001.

45 El término “regulable” con respecto al promotor tal como se emplea en la presente, se refiere a un promotor que es reprimido en una célula eucariota en presencia de una cantidad en exceso de una fuente de carbono (sustrato nutriente) en la fase de crecimiento de un cultivo discontinuo, y es desreprimido para ejercer una potente actividad de promotor en la fase de producción de una línea celular de producción, por ejemplo, tras la reducción de la cantidad de carbono, tal como tras introducir una fuente de carbono limitante del crecimiento (sustrato nutriente) a un cultivo según la estrategia de alimentación discontinua. A este respecto, el término “regulable” se entiende como

“regulable por un límite de fuente de carbono” o “regulable por un límite de glucosa”, que se refiere a la desrepresión de un promotor por el consumo, la reducción, la falta o la merma en carbono, o por la adición limitada de la fuente de carbono de modo que pueda ser consumida con facilidad por las células.

5 El promotor funcionalmente activo descrito en la presente es un promotor regulable relativamente potente que es silenciado o reprimido bajo condiciones de crecimiento celular (fase de crecimiento) y que se activa o desreprime bajo condiciones de producción (fase de producción) y, por tanto, es adecuado para inducir la producción de PDI en una línea celular de producción mediante la limitación de la fuente de carbono. Por tanto, el variante funcionalmente activo de un promotor tiene al menos dichas propiedades regulables.

10 La potencia del promotor regulable descrito en la presente se refiere a su potencia de transcripción, representada por la eficacia de inicio de la transcripción que se produce en este promotor con alta o baja frecuencia. Cuanto más alta sea la potencia de transcripción, con más frecuencia se producirá la transcripción en ese promotor. La potencia del promotor es importante, porque determina con qué frecuencia una secuencia de ARNm concreta es transcrita, y así otorga mayor prioridad a la transcripción de algunos genes frente a otros, lo cual conduce a una mayor concentración del transcrito. Un gen que codifica una proteína necesaria en grandes cantidades, por ejemplo, generalmente tiene un promotor relativamente potente. La ARN polimerasa solo puede realizar una tarea de transcripción al mismo tiempo y, por tanto, deben priorizar su trabajo para ser eficaz. Se seleccionan diferencias en la potencia del promotor para permitir esta priorización. Específicamente, el promotor regulable descrito en la presente es relativamente potente en el estado totalmente inducido, que generalmente se entiende como el estado de actividad aproximadamente máxima. La potencia relativa normalmente se determina con respecto a un promotor patrón, tal como el respectivo promotor pGAP de la célula usada como célula hospedante. La frecuencia de la transcripción normalmente se entiende como la tasa de transcripción, por ejemplo, determinada por medio de la cantidad de un transcrito en un ensayo adecuado, por ejemplo, RT-PCR o transferencia Northern. Por ejemplo, la potencia de transcripción de un promotor descrito en la presente se determina en la célula hospedante que es *P. pastoris* y se compara con el promotor pGAP nativo de *P. pastoris*.

25 El promotor pGAP inicia la expresión del gen *gap* que codifica la gliceraldehído-3-fosfato deshidrogenasa (GAPDH), que es un promotor constitutivo presente en cualquier microorganismo capaz de crecer con glucosa. GAPDH (EC 1\2\1\12), una enzima clave de la glicolisis, desempeña un papel crucial en el metabolismo de carbohidratos catabólico y anabólico.

30 El promotor regulable descrito en la presente ejerce una potencia de transcripción relativamente alta, reflejada por una tasa de transcripción o una potencia de transcripción de al menos 20%, comparado con el promotor pGAP nativo en la célula hospedante, y a veces denominado “promotor pGAP homólogo”. Preferiblemente, la tasa o potencia de transcripción es al menos 20%, en casos específicamente preferidos al menos 30%, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90% y al menos 100% o incluso mayor, tal como al menos 150% o al menos 200%, comparado con el promotor pGAP nativo, por ejemplo, determinado en la célula eucariota seleccionada como célula hospedante para producir la PDI.

35 Se prefiere específicamente un promotor regulable que tiene, en el estado inducido, al menos una potencia de transcripción de uno de los promotores pG1, pG3, pG4, pG6, pG7 o pG8. La potencia de transcripción comparativa que emplea el promotor pGAP como referencia puede determinarse por medios convencionales, tales como midiendo la cantidad de transcritos, por ejemplo, empleando una micromatriz, o también en un cultivo celular, tal como midiendo la cantidad de los respectivos productos de expresión génica en células recombinantes. Un ejemplo de ensayo se ilustra en la sección de ejemplos.

45 Específicamente, el promotor descrito en la presente es regulable mediante una fuente de carbono, con una potencia de promotor diferencial determinada en un ensayo que compara su potencia en presencia de glucosa y con limitación de glucosa, mostrando que sigue estando reprimido a concentraciones de glucosa relativamente altas, preferiblemente a concentraciones de al menos 10 g/l, preferiblemente al menos 20 g/l. Específicamente, el promotor descrito en la presente es totalmente inducido a concentraciones limitadas de glucosa y a concentraciones umbral de glucosa que inducen totalmente al promotor, cuyo umbral es menor que 20 g/l, preferiblemente menor que 10 g/l, menor que 1 g/l, aún menor que 0,1 g/l o menor que 50 mg/l, preferiblemente con una potencia de transcripción completa, por ejemplo, de al menos 50% del promotor pGAP homólogo nativo, a concentraciones de glucosa menores que 40 mg/l.

50 Preferiblemente, la potencia del promotor diferencial se determina mediante el inicio de la producción de PDI tras cambiar a condiciones inductoras por debajo de un umbral de fuente de carbono predeterminada, y se compara con la potencia en el estado reprimido. Habitualmente, se entiende que la potencia de transcripción es la potencia en el estado totalmente inducido, es decir, que muestra unas actividades aproximadamente máximas bajo condiciones de desrepresión. La potencia del promotor diferencial se determina, por ejemplo, según la eficacia o el rendimiento de la producción de PDI en una línea de células hospedantes recombinantes bajo condiciones de desrepresión, comparado con las condiciones de represión, o también por medio de la cantidad de un transcrito. El promotor regulable descrito en la presente tiene una potencia de promotor diferencial preferida que es al menos 2 veces, más preferiblemente al menos 5 veces, aún más preferiblemente al menos 10 veces, más preferiblemente al menos 20 veces, más preferiblemente al menos 30, 40, 50, o 100 veces mayor en el estado desreprimido, comparado con el

estado reprimido, que también se entiende como inducción en número de veces. Dicha potencia del promotor diferencia puede determinarse mediante un ensayo, tal como se ilustra mediante los ejemplos adjuntos.

5 Un promotor de la técnica anterior (promotor de MLS1 o promotor de ICL1) resultó tener una potencia de promotor diferencial significativamente menor que la inducción en 2 veces. Este promotor de la técnica anterior no es útil para la producción industrial de PDI, con una potencia de promotor de aproximadamente 5% comparado con el promotor pGAP convencional. Esto se ha demostrado en una comparación directa con el promotor descrito en la presente.

10 El término "homología" indica que dos o más secuencias de nucleótidos tienen las mismas pares de bases o pares de bases conservadas en una correspondiente posición, hasta cierto grado, hasta un grado cercano al 100%. Una secuencia homóloga generalmente presenta al menos aproximadamente 50% de identidad de secuencia de nucleótidos, preferiblemente al menos aproximadamente 60% de identidad, más preferiblemente al menos aproximadamente 70% de identidad, más preferiblemente al menos aproximadamente 80% de identidad, más preferiblemente al menos aproximadamente 90% de identidad, más preferiblemente al menos aproximadamente 95% de identidad.

15 La secuencia de promotor homóloga proporcionada en la presente preferiblemente tiene cierta homología con cualquiera de las secuencias de nucleótidos de los promotores pG1, pG3, pG4, pG6, pG7 o pG8 de *P. pastoris* en al menos partes específicas de la secuencia de nucleótidos, tales como la que incluye la región 3' de la respectiva secuencia de nucleótidos de promotor, preferiblemente una parte con una longitud específica hasta el extremo 3' de la respectiva secuencia de nucleótidos de promotor, tal como una parte con una longitud de al menos 200 pb, preferiblemente al menos 250 pb, preferiblemente al menos 300 pb, más preferiblemente al menos 400 pb, al menos 500 pb, al menos 600 pb, al menos 700 pb, al menos 800 pb, al menos 900 pb, o al menos 1000 pb, y análogos derivados de especies distintas de *Pichia pastoris*. Específicamente, al menos estas partes son preferiblemente homólogas dentro de un intervalo de 300-1000 pb, incluyendo la secuencia 3' terminal de la respectiva secuencia de nucleótidos de promotor.

25 Generalmente, las secuencias análogas se derivan de otras especies o cepas. Se entiende expresamente que cualquiera de las secuencias de promotores análogas proporcionadas en la presente que se deriva de especies distintas de *Pichia pastoris* puede comprender una secuencia homóloga, es decir, una secuencia con cierta homología según se describe en la presente. Por tanto, el término "homólogo" también puede incluir secuencias análogas. Por otra parte, se entiende que las secuencias descritas en la presente también se refieren a secuencias análogas y sus homólogos que comprenden cierta homología.

30 El "porcentaje (%) de identidad" con respecto a la secuencia de nucleótidos de un gen se define como el porcentaje de nucleótidos en una secuencia de ADN candidata que son idénticos a los nucleótidos en la secuencia de ADN después de alinear las secuencias e introducir huecos, si es necesario, para lograr el máximo porcentaje de identidad de secuencia y sin considerar las sustituciones conservativas como parte de la identidad de secuencia. El alineamiento con el objetivo de determinar el porcentaje de identidad de secuencia de nucleótidos puede lograrse de varias maneras que son conocidas en la técnica, por ejemplo, usando un programa informático de acceso público. Los expertos en la técnica pueden determinar los parámetros apropiados para medir el alineamiento, que incluyen cualquier algoritmo necesario para lograr el máximo alineamiento a lo largo de la longitud completa de las secuencias que se están comparando.

40 El término "mutagénesis", tal como se emplea en la presente, se refiere a un método para proporcionar mutantes de una secuencia de nucleótidos, por ejemplo, mediante inserción, delección y/o sustitución de uno o más nucleótidos, para obtener sus variantes con al menos un cambio en la región codificadora o no codificadora. La mutagénesis puede producirse a través de mutación aleatoria, semialeatoria o dirigida a sitio. Generalmente, se producen grandes bancos de genes aleatorizados con una alta diversidad génica, que pueden seleccionarse según un genotipo o fenotipo específicamente deseado.

45 La expresión "proteína de interés (PDI)", tal como se emplea en la presente, se refiere a un polipéptido o una proteína que es producido por medio de la tecnología recombinante en una célula hospedante. Más específicamente, la proteína puede ser un polipéptido que no aparece en la naturaleza en la célula hospedante, es decir, una proteína heteróloga, o también puede ser nativo a la célula hospedante, es decir, una proteína homóloga a la célula hospedante, pero que se produce, por ejemplo, mediante transformación con un vector autorreplicante que contiene la secuencia de ácido nucleico que codifica la PDI, o tras la integración mediante técnicas recombinantes de una o más copias de la secuencia de ácido nucleico que codifica la PDI en el genoma de la célula hospedante, o mediante la modificación recombinante de una o más secuencias reguladoras que controlan la expresión del gen que codifica la PDI, por ejemplo, de la secuencia de promotor. En algunos casos, el término PDI, tal como se emplea en la presente, también se refiere a cualquier producto de metabolito de la célula hospedante, que es mediado por la proteína expresada de modo recombinante.

55 El término "recombinante", tal como se emplea en la presente, significa "preparado mediante ingeniería genética o que surge de la ingeniería genética". Por tanto, un "microorganismo recombinante" comprende al menos un "ácido nucleico recombinante". Un microorganismo recombinante comprende específicamente un vector de expresión o un vector de clonación, o se ha modificado genéticamente para que contenga una secuencia de ácido nucleico

recombinante. Una "proteína recombinante" se produce expresando un respectivo ácido nucleico recombinante en un hospedante. Un "promotor recombinante" es una secuencia de nucleótidos no codificadora modificada genéticamente adecuada para su uso como promotor funcionalmente activo según se describe en la presente.

5 De modo sorprendente, se ha determinado que las células eucariotas son capaces de inducir la producción de una PDI limitando la disponibilidad de la fuente de carbono. Se ha descubierto que las condiciones de privación de carbono activan la inducción de una potente actividad de promotor que, hasta la fecha, no era conocida en la técnica. El promotor de MLS1 de *Pichia pastoris*, que en el documento US2008299616A1 se indica que es desreprimido bajo limitaciones de azúcares, realmente es un promotor regulable comparativamente débil para la producción de una PDI. Por tanto, resultó sorprendente que dicho promotor regulable potente de *P. pastoris* pudiese identificarse y usarse en líneas celulares de producción eucariotas, en particular para la producción de una PDI recombinante.

10 Aunque se ha determinado la secuencia genómica de 9,43 Mpb de la cepa GS115 de *P. pastoris* y se ha descrito en el documento US20110021378A1, las propiedades de las secuencias individuales, tales como las secuencias de promotor, no se han investigado en detalle. Por ejemplo, la secuencia de pG4 (SEQ ID NO:4), según se describe en la presente, se identificó como una secuencia de promotor en el documento US20110021378A1, aunque no se conocían sus propiedades regulables ni su uso bajo condiciones de privación de carbono. También resultó sorprendente que dicho promotor pudiese utilizarse de modo eficaz en el método proporcionado en la presente. Los promotores regulados de la técnica anterior, tales como los que se emplean a escala industrial en la producción de PDI, se derivan principalmente de la vía metabólica del metanol y necesitan la adición de metanol para inducir la producción de PDI, lo cual a menudo no resulta deseable. El método proporcionado en la presente tiene la ventaja de que puede proporcionar un aumento de la producción mediante una expresión potenciada, y tiene un riesgo reducido de contaminación debido a la regulación específica del promotor, en particular cuando se emplea un medio químicamente definido sin metanol.

25 Se ha descubierto que el promotor regulable proporcionado en la presente ejerce su actividad regulable solo tras el uso de un medio de cultivo muy específico adecuado para establecer condiciones de represión y desrepresión del promotor. Como ejemplo, *P. pastoris* puede cultivarse con éxito bajo condiciones de un proceso de producción industrial. En primer lugar, se emplea un cultivo discontinuo sobre una fuente de carbono basal, tal como glicerol, seguido de una alimentación discontinua con una alimentación limitada de una fuente de carbono suplementaria, tal como glucosa. Se toman muestras cerca del final de la primera fase discontinua y en condiciones de crecimiento limitado, por ejemplo, empleando una cantidad limitada de la fuente de carbono suplementaria. El análisis del transcriptoma con micromatrices de ADN revela ciertos genes específicos que son muy activos sobre la fuente de carbono suplementaria y son débiles o inactivos en presencia de un exceso de carbono, concretamente, la fuente de carbono basal en una cantidad en exceso. Se identificaron al menos seis secuencias de promotores como promotores regulables según se describen en la presente, concretamente pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), y pG8 (SEQ ID NO:6). Los promotores de MLS1 o ICL1 comparables de la técnica anterior son débiles, con menos de 1/10 de la potencia del promotor pG1, y no se aprecia regulación detectable.

40 Las características de reprimir la expresión de genes recombinantes sobre la fuente de carbono basal, y la expresión potente sobre la fuente de carbono suplementaria limitada, es decir, la inducción por el cambio del sustrato, pueden verificarse en los procesos de fermentación.

45 Las secuencias de nucleótidos que pueden usarse como secuencias reguladoras descritas en la presente, que proporcionan una producción mejorada de una proteína recombinante, pueden obtenerse de una diversidad de fuentes. El origen del promotor descrito en la presente es preferiblemente una célula de levadura, lo más preferiblemente una levadura metilotrófica, tal como del género *Pichia* o de la especie *P. pastoris*, cuyo promotor puede entonces usarse como secuencia de origen para producir variantes adecuados, por ejemplo, mutantes o análogos.

Se contempla que una serie de células de levadura, en particular de cepas de *Pichia*, pueden ser adecuadas para obtener las respectivas secuencias de promotores que son responsables de la producción de proteínas bajo condiciones de privación de carbono, o los respectivos análogos en diferentes especies.

50 Los variantes del promotor de *P. pastoris* identificados, que incluyen variantes funcionalmente activos, tales como homólogos y análogos, pueden producirse empleando técnicas convencionales. El promotor puede modificarse, por ejemplo, para generar variantes de promotor con niveles de expresión alterados y propiedades reguladoras alteradas.

55 Por ejemplo, puede prepararse un banco de promotores mediante mutagénesis de las secuencias de promotores descritas en la presente, que pueden usarse como moléculas de origen, por ejemplo, para ajustar finamente la expresión génica en células eucariotas mediante el análisis de los variantes para su expresión bajo diferentes estrategias de fermentación y la selección de los variantes adecuados. Puede usarse un banco sintético de variantes, por ejemplo, para seleccionar un promotor que se ajuste a los requisitos para producir una PDI seleccionada. Estos variantes pueden tener una mayor eficacia de expresión en células hospedantes eucariotas, y

una alta expresión tras la disminución de una fuente de carbono.

Las estrategias de fermentación diferencial pueden distinguirse entre una fase de crecimiento, tal como la etapa a), según se describe en la presente, y una fase de producción, tal como la etapa b).

- 5 El crecimiento y/o la producción pueden realizarse, de forma adecuada, en un modo discontinuo, un modo de alimentación discontinua o un modo continuo. Puede usarse cualquier biorreactor adecuado, que incluye un reactor discontinuo, de alimentación discontinua, continuo, de tanque agitado o con dispositivo de elevación de aire.

Resulta ventajoso proporcionar el proceso de fermentación a escala piloto o industrial. La escala industrial del proceso preferiblemente emplea un volumen de al menos 10 l, específicamente al menos 50 l, preferiblemente al menos 1 m³, preferiblemente al menos 10 m³, lo más preferiblemente al menos 100 m³.

- 10 Se prefieren las condiciones de producción a escala industrial que se refieren, por ejemplo, a un cultivo de alimentación discontinua en volúmenes de reactor de 100 l a 10 m³ o mayores, que emplean unos tiempos de proceso típicos de varios días, o procesos continuos en volúmenes del fermentador de aproximadamente 50-1000 l o mayores, con unas tasas de dilución de aproximadamente 0,02-0,15 h⁻¹.

- 15 Las técnicas de cultivo adecuadas pueden incluir el cultivo en un biorreactor que comienza con una fase discontinua, seguida de una fase de alimentación discontinua exponencial corta a una alta tasa de crecimiento específica, seguida de una fase de alimentación discontinua a una baja tasa de crecimiento específica. Otra técnica de cultivo adecuada puede incluir una fase discontinua, seguida de una fase de cultivo continua a una baja tasa de dilución.

Una realización preferida incluye un cultivo discontinuo para proporcionar una biomasa, seguida por un cultivo de alimentación discontinua para lograr altos rendimientos de producción de PDI.

- 20 Por ejemplo, la línea celular puede cultivarse en la etapa a) según se describe en la presente sobre glicerol o glucosa para obtener una biomasa.

- 25 Según el método proporcionado en la presente, se prefiere cultivar la línea de células hospedantes en un biorreactor bajo condiciones de crecimiento para obtener una densidad celular de al menos 1 g/l de peso seco de células, más preferiblemente al menos 10 g/l de peso seco de células, preferiblemente al menos 20 g/l de peso seco de células. Resulta ventajoso proporcionar dichos rendimientos de producción de biomasa a escala piloto o industrial.

Un medio de crecimiento que permite la acumulación de biomasa, específicamente un medio de crecimiento basal, generalmente comprende una fuente de carbono, una fuente de nitrógeno, una fuente de azufre y una fuente de fosfato. Generalmente, dicho medio comprende además oligoelementos y vitaminas, y puede comprender además aminoácidos, peptona o extracto de levadura.

- 30 Las fuentes de nitrógeno preferidas incluyen NH₄H₂PO₄, o NH₃ o (NH₄)₂SO₄.

Las fuentes de azufre preferidas incluyen MgSO₄, o (NH₄)₂SO₄ o K₂SO₄.

Las fuentes de fosfato preferidas incluyen NH₄H₂PO₄, o H₃PO₄ o NaH₂PO₄, KH₂PO₄, Na₂HPO₄ o K₂HPO₄.

Otros componentes típicos del medio incluyen KCl, CaCl₂, y oligoelementos tales como: Fe, Co, Cu, Ni, Zn, Mo, Mn, I, B.

- 35 Preferiblemente, el medio se suplementa con vitamina B₇.

Un medio de crecimiento típico para *P. pastoris* comprende glicerol o glucosa, NH₄H₂PO₄, MgSO₄, KCl, CaCl₂, biotina, y oligoelementos.

En la fase de producción se emplea específicamente un medio de producción con solo una cantidad limitada de una fuente de carbono suplementaria.

- 40 Preferiblemente, la línea de células hospedantes se cultiva en un medio mineral con una fuente de carbono adecuada, simplificando con ello aún más el proceso de aislamiento de modo significativo. Un ejemplo de un medio mineral preferido es un medio que contiene una fuente de carbono utilizable (por ejemplo, glucosa, glicerol o metanol), sales que contienen macroelementos (potasio, magnesio, calcio, amonio, cloruro, sulfato, fosfato) y oligoelementos (sales de cobre, yoduro, manganeso, molibdato, cobalto, cinc y hierro, y ácido bórico), y
45 opcionalmente vitaminas o aminoácidos, por ejemplo, para complementar auxotrofías.

- 50 Las células se cultivan bajo condiciones adecuadas para llevar a cabo la expresión de la PDI deseada, que puede purificarse a partir de las células o del medio de cultivo, dependiendo de la naturaleza del sistema de expresión y la proteína expresada, por ejemplo, si la proteína está fusionada con un péptido señal y si la proteína es soluble o está unida a una membrana. Tal como entenderán los expertos en la técnica, las condiciones de cultivo varían según factores que incluyen el tipo de célula hospedante y el vector de expresión concreto empleado.

La inducción de la producción de PDI por el promotor proporcionado en la presente se controla preferiblemente mediante el cultivo de las células sobre una fuente de carbono suplementaria en una cantidad limitada como única fuente de carbono y energía. Las células crecen muy lentamente bajo condiciones de limitación de carbono, pero producen rendimientos elevados de la PDI bajo el control del promotor regulable.

- 5 La diferencia en la actividad del promotor específicamente es de al menos 2 veces, más preferiblemente al menos 5 veces, aún más preferiblemente al menos 10 veces, más preferiblemente al menos 20 veces, más preferiblemente al menos 30, 40, 50, o 100 veces mayor en el estado desreprimido, comparado con el estado reprimido.

10 Mediante la selección de la secuencia del promotor adecuada, opcionalmente en combinación con otras secuencias reguladoras preferidas, es posible proporcionar, bajo condiciones comparables, al menos la misma actividad o un aumento en al menos aproximadamente 1,5, o en al menos aproximadamente 2 veces, o en al menos aproximadamente 5 veces, 10 veces o en al menos hasta aproximadamente 15 en la actividad, según se representa por la actividad del promotor o la potencia de transcripción, o regulada por la potencia del promotor con relación a un promotor de GAP que es homólogo con la célula de producción, un pGAP nativo, o aislado a partir de *P. pastoris*.

15 Un medio de producción típico comprende una fuente de carbono suplementaria y también $\text{NH}_4\text{H}_2\text{PO}_4$, MgSO_4 , KCl, CaCl_2 , biotina, y oligoelementos.

Por ejemplo, el suministro de la fuente de carbono suplementaria añadida a la fermentación puede comprender una fuente de carbono con hasta 50% en peso de azúcares fermentables. La baja tasa de alimentación del medio suplementario limitará los efectos de la inhibición del producto sobre el crecimiento celular, y así será posible un alto rendimiento del producto basándose en el suministro del sustrato.

- 20 La fermentación preferiblemente se realiza a un pH que varía de 3 a 7,5.

Unos tiempos de fermentación típicos son de aproximadamente 24 a 120 horas con unas temperaturas en el intervalo de 20 °C a 35 °C, preferiblemente 22-30 °C.

25 En general, los organismos o ácidos nucleicos recombinantes mencionados en la presente pueden producirse mediante técnicas de recombinación muy conocidas por los expertos en la técnica. Por consiguiente, pueden emplearse técnicas de biología molecular, microbiología y de ADN recombinantes convencionales dentro de la técnica. Estas técnicas se explican a fondo en la bibliografía. Véase, por ejemplo, Maniatis, Fritsch y Sambrook, *Molecular Cloning: A Laboratory Manual* (1982).

30 Según una realización preferida, se obtiene una construcción recombinante acoplado el promotor y los genes pertinentes en un vector. Estos genes pueden integrarse de modo estable en el genoma de la célula hospedante transformando la célula hospedante usando estos vectores.

35 Los vectores de expresión pueden incluir, pero no se limitan a vectores de clonación, vectores de clonación modificados y plásmidos específicamente diseñados. El vector de expresión preferido tal como se usa en el método proporcionado en la presente puede ser cualquier vector de expresión adecuado para la expresión de un gen recombinante en una célula hospedante, y se selecciona dependiendo del organismo hospedante. El vector de expresión recombinante puede ser cualquier vector que sea capaz de replicarse o de integrarse en el genoma del organismo hospedante, también denominado vector de hospedante.

Según una realización preferida, se emplean los plásmidos derivados de pPUZZLE como el vector.

40 Los vectores de expresión apropiados comprenden generalmente otras secuencias reguladoras adecuadas para expresar un ADN que codifica una PDI en una célula hospedante eucariota. Los ejemplos de secuencias reguladoras incluyen operadores, potenciadores, sitios de unión a ribosomas y secuencias que controlan el inicio y el fin de la transcripción y la traducción. Las secuencias reguladoras pueden estar unidas operablemente a la secuencia de ADN que se va a expresar.

45 Para permitir la expresión de una secuencia de nucleótidos recombinantes en una célula hospedante, el vector de expresión puede proporcionar el promotor descrito en la presente adyacente al extremo 5' de la secuencia codificadora, por ejemplo, cadena arriba de un gen de péptido señal. Con ello, la transcripción es regulada e iniciada por esta secuencia de promotor.

50 Un péptido señal puede ser un péptido señal heterólogo o un híbrido de un péptido señal heterólogo y nativo, y específicamente puede ser heterólogo u homólogo con el organismo hospedante que produce la proteína. La función del péptido señal es permitir que la PDI sea segregada para entrar en el retículo endoplásmico. Habitualmente es una cadena peptídica corta (con una longitud de 3-60 aminoácidos) que dirige el transporte de una proteína fuera de la membrana plasmática, haciendo con ello que la proteína heteróloga sea fácil de separar y de purificar. Algunos péptidos señal se escinden de la proteína por peptidasas señal después de que las proteínas sean transportadas.

Los ejemplos de péptidos señal son secuencias señal del péptido prepro del factor de apareamiento alfa de *S. cerevisiae* y el péptido señal del gen de fosfatasa ácida (PHO1) de *P. pastoris*.

Se entiende que una secuencia de promotor está unida operablemente a una secuencia codificadora si el promotor controla la transcripción de la secuencia codificadora. Si una secuencia de promotor no está nativamente asociada con la secuencia codificadora, su transcripción no es controlada por el promotor en células nativas (de tipo salvaje) o las secuencias se recombinan con diferentes secuencias contiguas.

5 Para demostrar la función de las secuencias pertinentes, pueden construirse vectores de expresión que comprendan uno o más elementos reguladores para conducir la expresión de una PDI, y el rendimiento expresado se compara con construcciones con elementos regulatorios convencionales. En los siguientes ejemplos puede encontrarse una descripción detallada del procedimiento experimental. Los genes identificados pueden amplificarse mediante PCR a partir de *P. pastoris* empleando cebadores nucleotídicos específicos, clonarse en un vector de expresión y transformarse en una línea de células eucariotas, por ejemplo, usando un vector de levadura y una cepa de *P. pastoris* para la producción a alto nivel de diversas PDI diferentes. Para calcular el efecto del promotor proporcionado en la presente sobre la cantidad de PDI recombinante producida de esta manera, la línea de células eucariotas puede cultivarse en experimentos de matraces de agitación y fermentaciones de alimentación discontinua o en quimiostato, en comparación con cepas que comprenden un promotor convencional no regulable por una fuente de carbono, tal como, por ejemplo, el promotor pGAP convencional en la respectiva célula. En particular, la elección del promotor tiene un gran impacto sobre la producción de la proteína recombinante.

Los métodos de transformación preferidos para la captación del fragmento de ADN recombinante por el microorganismo incluyen la transformación química, la electroporación o la transformación con protoplastos. Los transformantes descritos en la presente pueden obtenerse introduciendo dicho ADN de vector, por ejemplo, ADN plasmídico, en un hospedante y seleccionando los transformantes que expresan la proteína pertinente o el metabolito de la célula hospedante con alto rendimiento.

La PDI puede producirse usando la línea de células hospedantes recombinantes mediante el cultivo de un transformante obtenido de esta manera en un medio apropiado, el aislamiento del metabolito o producto expresado del cultivo, y opcionalmente su purificación mediante un método adecuado.

25 Los transformantes descritos en la presente pueden obtenerse introduciendo dicho ADN de vector, por ejemplo, ADN plasmídico, en un hospedante y seleccionando los transformantes que expresan la PDI o el metabolito de la célula hospedante con alto rendimiento. Las células hospedantes se tratan para permitir que incorporen un ADN extraño mediante métodos que se emplean de modo convencional para la transformación de células eucariotas, tales como el método de pulsos eléctricos, el método de protoplastos, el método de acetato de litio y modificaciones de estos métodos. *P. pastoris* se transforma preferiblemente mediante electroporación.

La línea de células hospedantes preferida según la invención mantiene las propiedades genéticas empleadas en la presente, y el nivel de producción permanece elevado, por ejemplo, al menos a un nivel de µg, incluso después de aproximadamente 20 generaciones de cultivo, preferiblemente al menos 30 generaciones, más preferiblemente al menos 40 generaciones, lo más preferiblemente al menos 50 generaciones. La célula hospedante recombinante estable se considera una gran ventaja cuando se emplea para la producción a escala industrial.

Se prefieren varias estrategias diferentes para la producción de la PDI según el método proporcionado en la presente. Las sustancias pueden expresarse, procesarse y opcionalmente segregarse mediante la transformación de una célula hospedante eucariota con un vector de expresión que porta un ADN recombinante que codifica una proteína pertinente y al menos uno de los elementos reguladores según se describió anteriormente, la preparación de un cultivo de la célula transformada, el crecimiento del cultivo, la inducción de la transcripción y la producción de la PDI, y la recuperación del producto del proceso de fermentación.

La PDI se expresa preferiblemente empleando condiciones para producir rendimientos de al menos 1 mg/l, preferiblemente al menos 10 mg/l, preferiblemente al menos 100 mg/l, lo más preferiblemente al menos 1 g/l.

45 La célula hospedante descrita en la presente preferiblemente se ensaya para su rendimiento o capacidad de expresión mediante el siguiente ensayo: ELISA, ensayo de actividad, HPLC, u otros ensayos adecuados.

Se entiende que los métodos descritos en la presente pueden incluir también cultivar dichas células hospedantes recombinantes bajo condiciones que permiten la expresión de la PDI, preferiblemente en forma segregada o como un producto intracelular. Después puede aislarse una PDI producida de modo recombinante o un metabolito de la célula hospedante a partir del medio de cultivo celular y posteriormente purificarse mediante técnicas muy conocidas por los expertos en la técnica.

La PDI producida según se describe en la presente generalmente puede aislarse y purificarse usando técnicas actuales, que incluyen el aumento de la concentración de la PDI deseada y/o la disminución de la concentración de al menos una impureza.

55 Si la PDI se segrega desde las células, puede aislarse y purificarse a partir del medio de cultivo usando técnicas actuales. La secreción de los productos de expresión recombinantes a partir de las células hospedantes en general resulta ventajosa por razones que incluyen facilitar el proceso de purificación, puesto que los productos se recuperan del sobrenadante de cultivo y no de la mezcla compleja de proteínas que se produce cuando las células de levadura

se rompen para liberar las proteínas intracelulares.

Las células de transformantes cultivadas también pueden romperse de modo sónico o mecánico, enzimático o químico, para obtener un extracto celular que contiene la PDI deseada, a partir del cual se aísla y purifica la PDI.

5 Como métodos de aislamiento y purificación para obtener un producto de proteína o polipéptido recombinante, pueden usarse métodos, tales como métodos que emplean diferencias en la solubilidad, tales como desalado y precipitación con disolventes, métodos que emplean diferencias en el peso molecular, tales como ultrafiltración y electroforesis en gel, métodos que emplean diferencias en la carga eléctrica, tales como cromatografía de intercambio iónico, métodos que emplean la afinidad específica, tales como cromatografía de afinidad, métodos que emplean diferencias en la hidrofobicidad, tales como cromatografía líquida de alta resolución en fase inversa, y
10 métodos que emplean diferencias en el punto isoeléctrico, tales como enfoque isoeléctrico.

El producto muy purificado está fundamentalmente exento de proteínas contaminantes, y preferiblemente tiene una pureza de al menos 90%, más preferiblemente de al menos 95%, o incluso al menos 98%, hasta 100%. Los productos purificados pueden obtenerse mediante la purificación del sobrenadante del cultivo celular o a partir de residuos celulares.

15 Como métodos de aislamiento y purificación, se prefieren los siguientes métodos convencionales: ruptura de las células (si la PDI se obtiene de modo intracelular), separación de las células (residuos) y lavado mediante microfiltración o filtro de flujo tangencial ("Tangential Flow Filter", TFF) o centrifugación, purificación de la PDI mediante precipitación o tratamiento con calor, activación de la PDI mediante digestión enzimática, purificación de la PDI mediante cromatografía, tal como intercambio iónico ("ion exchange", IEX), cromatografía de interacción
20 hidrófoba ("hydrophobic interaction chromatography", HIC), cromatografía de afinidad, cromatografía de exclusión molecular ("size exclusion chromatography", SEC) o cromatografía HPLC, precipitación o concentración de la PDI y lavado mediante etapas de ultrafiltración.

La PDI aislada y purificada puede identificarse por métodos convencionales, tales como transferencia Western, HPLC, ensayo de actividad, o ELISA.

25 La PDI puede ser cualquier polipéptido eucariota, procariota o sintético. Puede ser una proteína segregada o una proteína intracelular. En la presente también se proporciona la producción recombinante de homólogos funcionales, variantes equivalentes funcionales, derivados y fragmentos biológicamente activos de proteínas naturales. Los homólogos funcionales son preferiblemente idénticos o correspondientes y tienen las características funcionales de una secuencia.

30 Una PDI mencionada en la presente puede ser un producto homólogo con la célula hospedante eucariota, o heterólogo, preferiblemente para un uso terapéutico, profiláctico, de diagnóstico, analítico o industrial.

La PDI es preferiblemente una proteína o polipéptido recombinante heterólogo producido en una célula eucariota, preferiblemente una célula de levadura, preferiblemente como proteínas segregadas. Los ejemplos de proteínas preferiblemente producidas son inmunoglobulinas, fragmentos de inmunoglobulina, aprotinina, inhibidor de la vía del factor tisular u otros inhibidores de proteasas, e insulina o precursores de la insulina, análogos de la insulina, hormonas del crecimiento, interleuquinas, activador del plasminógeno tisular, factor de crecimiento transformante a o b, glucagón, péptido 1 similar a glucagón ("glucagon-like peptide 1", GLP-1), péptido 2 similar a glucagón (GLP-2), GRPP, factor VII, factor VIII, factor XIII, factor 1 de crecimiento derivado de plaquetas, albúmina de suero, enzimas, tales como lipasas o proteasas, o un homólogo funcional, variante equivalente funcional, derivado y fragmento biológicamente activo con una función similar a la proteína nativa. La PDI puede ser estructuralmente similar a la proteína nativa y puede derivarse de la proteína nativa mediante la adición de uno o más aminoácidos a cualquiera o a ambos de los extremos C- y N-terminales o a la cadena lateral de la proteína nativa, la sustitución de uno o más aminoácidos en uno o en una serie de sitios diferentes en la secuencia de aminoácidos nativa, la delección de uno o más aminoácidos en cualquiera o en ambos de los extremos de la proteína nativa o en uno o varios sitios en la
40 secuencia de aminoácidos, o la inserción de uno o más aminoácidos en uno o más sitios de la secuencia de aminoácidos nativa. Estas modificaciones son muy conocidas para varias de las proteínas mencionadas anteriormente.

Un PDI también puede seleccionarse de sustratos, enzimas, inhibidores o cofactores que proporcionan las reacciones bioquímicas en la célula hospedante, con el objetivo de obtener el producto de dicha reacción bioquímica o una cascada de varias reacciones, por ejemplo, para obtener un metabolito de la célula hospedante. Los ejemplos de productos pueden ser vitaminas, tales como riboflavina, ácidos orgánicos y alcoholes, que pueden obtenerse con mayores rendimientos tras la expresión de una proteína recombinante o una PDI según se describe en la presente.

En general, la célula hospedante, que expresa un producto recombinante, puede ser cualquier célula eucariota adecuada para la expresión recombinante de una PDI.

55 Los ejemplos de células de mamífero preferidas son células BHK, CHO (CHO-DG44, CHO-DUXB11, CHO-DUKX, CHO-K1, CHOK1SV, CHO-S), HeLa, HEK293, MDCK, NIH3T3, NS0, PER.C6, SP2/0 y VERO.

Los ejemplos de células de levadura preferidas usadas como células hospedantes según se describe en la presente incluyen, pero no se limitan al género *Saccharomyces* (por ejemplo, *Saccharomyces cerevisiae*), el género *Pichia* (por ejemplo, *P. pastoris*, o *P. methanolica*), el género *Komagataella* (*K. pastoris*, *K. pseudopastoris* o *K. phaffii*), *Hansenula polymorpha* o *Kluyveromyces lactis*.

- 5 La bibliografía más reciente divide y renombra a *Pichia pastoris* como *Komagataella pastoris*, *Komagataella phaffii* y *Komagataella pseudopastoris*. En la presente, *Pichia pastoris* se emplea como sinónimo de todas, *Komagataella pastoris*, *Komagataella phaffii* y *Komagataella pseudopastoris*.

Las células hospedantes de levadura preferidas se derivan de levaduras metilotróficas, tales como *Pichia* o *Komagataella*, por ejemplo, *Pichia pastoris*, o *K. phaffii*, o *K. pseudopastoris*. Los ejemplos de hospedante incluyen levaduras, tales como *P. pastoris*. Los ejemplos de cepas de *P. pastoris* incluyen CBS 704 (=NRRL Y-1603=DSMZ 70382), CBS 2612 (=NRRL Y-7556), CBS 7435 (=NRRL Y-11430), CBS 9173-9189 (cepas CBS: CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos), y DSMZ 70877 (German Collection of Micro-organisms and Cell Cultures), pero también cepas de Invitrogen, tales como X-33, GS115, KM71 y SMD1168. Los ejemplos de cepas de *S. cerevisiae* incluyen W303, CEN.PK y la serie BY (colección EUROSCARF). Todas las cepas descritas anteriormente se han empleado con éxito para producir transformantes y expresar genes heterólogos.

Una célula hospedante de levadura preferida, tal como una célula hospedante de *P. pastoris* o *S. cerevisiae*, contiene secuencias de promotores recombinantes o heterólogos, que pueden derivarse de una cepa de *P. pastoris* o *S. cerevisiae*, diferente del hospedante de producción. En otra realización específica, la célula hospedante comprende una construcción de expresión recombinante según se describe en la presente que comprende el promotor que se origina del mismo género, especie o cepa que la célula hospedante.

El promotor puede ser un promotor según se describe en la presente o cualquier otra secuencia de ADN que muestra actividad transcripcional en la célula hospedante y que puede derivarse de genes que codifican proteínas homólogas o heterólogas con el hospedante. El promotor se deriva preferiblemente de un gen que codifica una proteína homóloga con la célula hospedante.

Por ejemplo, un promotor según se describe en la presente puede derivarse de una levadura, tal como una cepa de *S. cerevisiae*, y ser utilizado para expresar una PDI en una levadura. Una realización específicamente preferida se refiere a un promotor que se origina de *P. pastoris* para su uso en un método para producir una PDI recombinante en una línea de células hospedantes productoras de *P. pastoris*. El origen homólogo de la secuencia de nucleótidos facilita su incorporación en una célula hospedante del mismo género o especie, permitiendo así la producción estable de una PDI, con la posibilidad de aumentar el rendimiento en procesos de fabricación industriales. Además, pueden emplearse variantes funcionalmente activos del promotor procedentes de otras levaduras adecuadas o de otros hongos o de otros organismos, tales como vertebrados o plantas.

Si la PDI es una proteína homóloga con la célula hospedante, es decir, una proteína que aparece en la naturaleza en la célula hospedante, la expresión de la PDI en la célula hospedante puede modularse mediante el intercambio de su secuencia de promotor nativa por una secuencia de promotor según se describe en la presente.

Este objetivo puede alcanzarse, por ejemplo, mediante la transformación de una célula hospedante con una molécula de ADN recombinante que comprende secuencias homólogas del gen diana para permitir la recombinación específica de sitio, la secuencia de promotor y un marcador selectivo adecuado para la célula hospedante. La recombinación específica de sitio se realiza para unir operablemente la secuencia de promotor con la secuencia de nucleótidos que codifica la PDI. Esto provoca la expresión de la PDI a partir de la secuencia de promotor, según se describe en la presente, en lugar de la secuencia de promotor nativa.

En una realización específicamente preferida, la secuencia de promotor tiene una mayor actividad de promotor con relación a la secuencia de promotor nativa de la PDI.

45 Tal como se describe en la presente, se prefiere proporcionar una línea de células hospedantes de *P. pastoris* que comprende una secuencia de promotor según la invención unida operablemente a la secuencia de nucleótidos que codifica la PDI.

Según una realización específica, también es posible proporcionar una célula hospedante o vector comodín que comprende un promotor según se describe en la presente, y que está listo para incorporar un gen de interés que codifica una PDI. La línea celular comodín, por tanto, es una línea de células hospedantes preformada que se caracteriza por su capacidad de expresión. Sigue una innovadora estrategia de plataforma "comodín" para la generación de líneas de células productoras para la producción de PDI, por ejemplo, usando el intercambio de módulos mediado por recombinasa específica de sitio. Esta nueva célula hospedante facilita la clonación de un gen de interés (GDI), por ejemplo, en puntos calientes de expresión genómica predeterminados en cuestión de días para obtener líneas de células de producción reproducibles y muy eficaces.

Según una realización preferida, el método proporcionado en la presente emplea una secuencia de nucleótidos recombinante que codifica la PDI, que se proporciona sobre un plásmido adecuado para la integración en el genoma

de la célula hospedante, en una sola copia o en múltiples copias por célula. La secuencia de nucleótidos recombinante que codifica la PDI también puede proporcionarse sobre un plásmido de replicación autónoma en una sola copia o en múltiples copias por célula.

5 El método preferido según se describe en la presente emplea un plásmido, que es un vector de expresión eucariota, preferiblemente un vector de expresión de levadura. Los vectores de expresión pueden incluir, pero no se limitan a vectores de clonación, vectores de clonación modificados y plásmidos específicamente diseñados. El vector de expresión preferido tal como se usa en el método proporcionado en la presente puede ser cualquier vector de expresión adecuado para la expresión de un gen recombinante en una célula hospedante, y se selecciona dependiendo del organismo hospedante. El vector de expresión recombinante puede ser cualquier vector que sea
10 capaz de replicarse o de integrarse en el genoma de organismos hospedantes, también denominado vector de hospedante, tal como un vector de levadura, que porta una construcción de ADN. Un vector de expresión de levadura preferido es para la expresión en levaduras seleccionadas del grupo que consiste en levaduras metilotróficas representadas por los géneros *Hansenula*, *Pichia*, *Candida* y *Torulopsis*.

15 En la presente, se prefiere emplear plásmidos derivados de pPICZ, pGAPZ, pPIC9, pPICZalfa, pGAPZalfa, pPIC9K, pGAPHis o pPUZZLE como vector.

Según una realización preferida, se obtiene una construcción recombinante acoplado los genes pertinentes en un vector. Estos genes pueden integrarse de modo estable en el genoma de la célula hospedante transformando la célula hospedante usando estos vectores. Los polipéptidos codificados por los genes pueden producirse usando la línea de células hospedantes recombinantes mediante el cultivo de un transformante obtenido de esta manera en un medio apropiado, el aislamiento de la PDI expresada del cultivo, y su purificación mediante un método apropiado para el producto expresado, en particular para separar la PDI de las proteínas contaminantes.
20

Los vectores de expresión pueden comprender uno o más marcadores seleccionables fenotípicos, por ejemplo, un gen que codifica una proteína que confiere resistencia a antibióticos o que suministra un requisito autotrófico. Los vectores de levadura contienen habitualmente un origen de la replicación procedente de un plásmido de levadura, una secuencia de replicación autónoma ("autonomously replicating sequence", ARS) o, como alternativa, una secuencia usada para la integración en el genoma del hospedante, una región de promotor, secuencias para la poliadenilación, secuencias para la terminación de la transcripción, y un marcador seleccionable.
25

Los procedimientos usados para acoplar las secuencias de ADN, por ejemplo, que codifican la secuencia precursora y/o la PDI, el promotor y el terminador, respectivamente, y para insertarlas en vectores adecuados que contienen la información necesaria para la integración o la replicación en el hospedante, son muy conocidos por los expertos en la técnica, por ejemplo, según se describe en J. Sambrook *et al.*, "Molecular Cloning", 2ª edición, Cold Spring Harbor Laboratory Press (1989).
30

Se entenderá que el vector, que emplea los elementos reguladores descritos en la presente y/o la PDI como diana de integración, puede construirse preparando en primer lugar una construcción de ADN que contiene la secuencia de ADN completa que codifica los elementos reguladores y/o la PDI, y posteriormente insertando este fragmento en un vector de expresión adecuado, o mediante la inserción secuencial de fragmentos de ADN que contienen información genética para los elementos individuales, tales como la proteína señal, conductora o heteróloga, seguido del acoplamiento.
35

En la presente también pueden emplearse vectores de multiclonación, que son vectores que contienen un sitio de multiclonación, en los que puede incorporarse un gen heterólogo deseado en un sitio de multiclonación para proporcionar un vector de expresión. En los vectores de expresión, el promotor se coloca cadena arriba del gen de la PDI y regula la expresión del gen. En el caso de vectores de multiclonación, debido a que el gen de la PDI se introduce en el sitio de multiclonación, el promotor se coloca cadena arriba del sitio de multiclonación.
40

La construcción de ADN, según se proporciona para obtener una célula hospedante recombinante según se describe en la presente, puede prepararse de modo sintético por medio de métodos convencionales establecidos, por ejemplo, el método de fosforamidita. La construcción de ADN también puede tener un origen genómico o ADNc, por ejemplo, obtenido preparando un banco de ADNc o genómico y seleccionando secuencias de ADN que codifiquen todo o parte del polipéptido proporcionado en la presente mediante hibridación usando sondas oligonucleotídicas sintéticas según técnicas convencionales (Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1989). Por último, la construcción de ADN puede tener un origen mixto sintético y genómico, mixto sintético y de ADNc, o mixto genómico y de ADNc, preparándose asociando fragmentos de origen sintético, genómico o de ADNc, según sea apropiado, y los fragmentos se corresponden con diversas partes de la construcción de ADN completa, según técnicas convencionales.
45
50

En otra realización preferida, el vector de expresión de levadura es capaz de integrarse de modo estable en el genoma de levaduras, por ejemplo, mediante recombinación homóloga.
55

Una célula hospedante transformante obtenida transformando la célula con los elementos reguladores proporcionados en la presente y/o los genes de la PDI puede cultivarse preferiblemente en primer lugar a unas condiciones para que crezca con eficacia hasta alcanzar un gran número de células sin la carga de expresar una

proteína heteróloga. Cuando la línea celular se prepara para la expresión de la PDI, se eligen técnicas de cultivo para producir el producto de expresión.

Los asuntos de las siguientes definiciones se consideran realizaciones descritas en la presente:

- 5 1. Un método para producir una proteína de interés (PDI) mediante el cultivo de una línea de células eucariotas recombinantes que comprenden una construcción de expresión que comprende un promotor regulable y una molécula de ácido nucleico que codifica una PDI bajo el control transcripcional de dicho promotor, que comprende las etapas de:
 - a) cultivar la línea de células con una fuente de carbono basal que reprime al promotor,
 - 10 b) cultivar la línea de células sin una fuente de carbono suplementaria, o con una cantidad limitada de esta, que desreprime al promotor para inducir la producción de la PDI a una tasa de transcripción de al menos 15%, comparado con el promotor pGAP nativo de la célula, y
 - c) producir y recuperar la PDI.
2. El método según la definición 1, en el que la fuente de carbono basal se selecciona del grupo que consiste en glucosa, glicerol, etanol, una de sus mezclas y un material nutriente complejo.
- 15 3. El método según la definición 1 o 2, en el que la fuente de carbono suplementaria es una hexosa, tal como glucosa, fructosa, galactosa o manosa, un disacárido, tal como sacarosa, un alcohol, tal como glicerol o etanol, o una de sus mezclas.
4. El método según cualquiera de las definiciones 1 a 3, en el que la fuente de carbono basal es glicerol, y la fuente de carbono suplementaria es glucosa.
- 20 5. El método según cualquiera de las definiciones 1 a 4, en el que la etapa b) emplea un medio de alimentación que no proporciona la fuente de carbono suplementaria o la proporciona en una cantidad limitada, preferiblemente a 0-1 g/l en el medio de cultivo.
6. El método según la definición 5, en el que el medio de alimentación está químicamente definido y no contiene metanol.
- 25 7. El método según cualquiera de las definiciones 1 a 6, en el que la cantidad limitada de la fuente de carbono suplementaria limita el crecimiento para mantener la tasa de crecimiento específica dentro del intervalo de 0,02 h⁻¹ hasta 0,2 h⁻¹, preferiblemente de 0,02 h⁻¹ hasta 0,15 h⁻¹.
8. El método según la definición 7, en el que la cantidad limitada de la fuente suplementaria proporciona una cantidad residual en el cultivo de células que está por debajo del límite de detección.
- 30 9. El método según cualquiera de las definiciones 1 a 8, en el que el promotor es capaz de controlar la transcripción de un gen en una célula eucariota de tipo salvaje, y dicho gen se selecciona del grupo que consiste en G1 (SEQ ID NO:7), G3 (SEQ ID NO:8), G4 (SEQ ID NO:9), G6 (SEQ ID NO:10), G7 (SEQ ID NO:11), o G8 (SEQ ID NO:12), o uno de sus variantes funcionalmente activos.
- 35 10. El método según la definición 9, en el que dichos variantes funcionalmente activos se seleccionan del grupo que consiste en homólogos con al menos aproximadamente 60% de identidad de secuencia de nucleótidos, homólogos que pueden obtenerse por medio de la modificación de la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, que preferiblemente comprenden o consisten en una secuencia de nucleótidos de al menos 200 pb, y análogos derivados de especies distintas de *Pichia pastoris*.
- 40 11. El método según la definición 9 o 10, en el que el variante funcionalmente activo de pG1 se selecciona del grupo que consiste en pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46).
12. El método según cualquiera de las definiciones 1 a 11, en el que el promotor es un promotor de *Pichia pastoris* o uno de sus variantes funcionalmente activos.
- 45 13. El método según cualquiera de las definiciones 1 a 12, en el que la línea de células se selecciona del grupo que consiste en líneas de células de mamífero, insecto, levadura, hongo filamentoso y vegetales, preferiblemente una levadura.
- 50 14. El método según la definición 13, en el que la levadura se selecciona del grupo que consiste en *Pichia*, *Candida*, *Torulopsis*, *Arxula*, *Hansenula*, *Yarrowia*, *Kluyveromyces*, *Saccharomyces*, *Komagataella*, preferiblemente una levadura metilotrófica.

15. El método según la definición 14, en el que la levadura es *Pichia pastoris*, *Komagataella pastoris*, *K. phaffii*, o *K. pseudopastoris*.
16. El método según cualquiera de las definiciones 1 a 15, en el que el promotor no está asociado nativamente con la secuencia de nucleótidos que codifica la PDI.
- 5 17. El método según cualquiera de las definiciones 1 a 16, en el que la PDI es una proteína heteróloga, preferiblemente seleccionada de proteínas terapéuticas, que incluyen anticuerpos o sus fragmentos, enzimas y péptidos, antibióticos de proteínas, proteínas de fusión de toxinas, conjugados de carbohidrato-proteína, proteínas estructurales, proteínas reguladoras, vacunas y proteínas o partículas similares a vacunas, enzimas de procesos, factores de crecimiento, hormonas y citoquinas, o un metabolito de una PDI.
- 10 18. El método según cualquiera de las definiciones 1 a 17, en el que la PDI es una proteína eucariótica, preferiblemente una proteína de mamífero.
19. El método según cualquiera de las definiciones 1 a 18, en el que la PDI es una proteína multimérica, preferiblemente un dímero o un tetramero.
- 15 20. El método según cualquiera de las definiciones 1 a 19, en el que la PDI es una molécula de unión a un antígeno, tal como un anticuerpo, o uno de sus fragmentos.
21. El método según cualquiera de las definiciones 1 a 20, en el que se fabrica un producto de fermentación usando la PDI, un metabolito o un derivado de esta.
22. Un método para controlar la expresión de una PDI en una célula eucariota recombinante bajo el control transcripcional de un promotor regulable por una fuente de carbono que tiene una potencia de transcripción de al menos 15%, comparado con el promotor pGAP nativo de la célula, en el que la expresión se induce bajo condiciones que limitan la fuente de carbono.
- 20 23. Un método para producir una PDI en una célula eucariota recombinante bajo el control transcripcional de un promotor regulable por una fuente de carbono, en el que dicho promotor tiene una potencia de transcripción de al menos 15%, comparado con el promotor pGAP nativo de la célula.
- 25 24. El método según cualquiera de las definiciones 1 a 23, en el que el promotor regulable comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en:
- a) pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6);
- 30 b) una secuencia que tiene al menos 60% de homología con pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6);
- c) una secuencia que se hibrida bajo condiciones rigurosas con pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); y
- d) un fragmento o variante derivado de a), b) o c),
- 35 en el que dicho promotor es un promotor funcionalmente activo, que es un promotor regulable por una fuente de carbono capaz de expresar una PDI en una célula eucariota recombinante a una tasa de transcripción de al menos 15%, comparado con el promotor pGAP nativo de la célula.
25. El método según la definición 24, en el que el variante de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6) es un variante funcionalmente activo seleccionado del grupo que consiste en homólogos con al menos 60% de identidad de secuencia de nucleótidos, homólogos que pueden obtenerse por medio de la modificación de la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, que preferiblemente comprenden o consisten en una secuencia de nucleótidos de al menos 200 pb, y análogos derivados de especies distintas de *Pichia pastoris*.
- 40 26. El método según la definición 24 o 25, en el que el variante funcionalmente activo de pG1 se selecciona del grupo que consiste en pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46).
- 45 27. Un ácido nucleico aislado que comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en:
- 50 a) pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6);

b) una secuencia que tiene al menos 60% de homología con pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6);

c) una secuencia que se hibrida bajo condiciones rigurosas con pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); y

5 d) un fragmento o variante derivado de a), b) o c),

en el que dicho ácido nucleico comprende un promotor funcionalmente activo, que es un promotor regulable por una fuente de carbono capaz de expresar una PDI en una célula eucariota recombinante a una tasa de transcripción de al menos 15%, comparado con el promotor pGAP nativo de la célula.

10 28. Un ácido nucleico según la definición 27, en el que el variante de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6) es un variante funcionalmente activo seleccionado del grupo que consiste en homólogos con al menos 60% de identidad de secuencia de nucleótidos, homólogos que pueden obtenerse por medio de la modificación de la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, preferiblemente con una secuencia de nucleótidos de al menos 200 pb, y análogos derivados de especies distintas de *Pichia pastoris*.

15 29. Un ácido nucleico según la definición 27 o 28, en el que el variante funcionalmente activo de pG1 se selecciona del grupo que consiste en pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46).

20 30. Una construcción de expresión que comprende un ácido nucleico según cualquiera de las definiciones 27 a 29, unido operablemente a una secuencia de nucleótidos que codifica una PDI bajo el control transcripcional de dicho promotor, y dicho ácido nucleico no está asociado nativamente con la secuencia de nucleótidos que codifica la PDI.

31. Un vector que comprende la construcción según la definición 30.

32. Una célula eucariota recombinante que comprende la construcción de la definición 30, o el vector de la definición 31.

25 33. Una célula según la definición 31, que se selecciona del grupo que consiste en líneas de células de mamífero, insecto, levadura, hongo filamentosos y vegetales, preferiblemente una levadura.

34. Una célula según la definición 32, en la que la levadura se selecciona del grupo que consiste en *Pichia*, *Candida*, *Torulopsis*, *Arxula*, *Hansenula*, *Yarrowia*, *Kluyveromyces*, *Saccharomyces*, *Komagataella*, preferiblemente una levadura metilotrófica.

30 35. Una célula según la definición 34, en la que la levadura es *Pichia pastoris*, *Komagataella pastoris*, *K. phaffii*, o *K. pseudopastoris*.

36. Una célula según cualquiera de las definiciones 32 a 35, que tiene una tasa de crecimiento específica mayor en presencia de un exceso de fuente de carbono, con relación a unas condiciones de fuente de carbono limitada.

35 37. Un método para identificar un promotor regulable por una fuente de carbono procedente de células eucariotas, que comprende las etapas de:

a) cultivar células eucariotas en presencia de una fuente de carbono en un cultivo discontinuo bajo condiciones de crecimiento de las células,

b) cultivar después las células en un cultivo de alimentación discontinua en presencia de una cantidad limitada de una fuente de carbono suplementaria,

40 c) proporcionar muestras del cultivo de células de la etapa a) y b), y

d) realizar un análisis de la transcripción en dichas muestras para identificar un promotor regulable que muestra una potencia transcripcional mayor en las células de la etapa b) que en las células de la etapa a).

38. El método según la definición 37, en el que el análisis de la transcripción es cuantitativo o semicuantitativo, y emplea preferiblemente micromatrices de ADN, secuenciación de ARN y análisis del transcriptoma.

45 Los ejemplos específicos se refieren a una fermentación de alimentación discontinua de una línea de células de producción recombinante de *P. pastoris* que produce proteínas indicadoras, que emplea un medio discontinuo de glicerol y un medio de alimentación discontinua de glucosa. Estudios de actividad de promotor comparativos han demostrado que el promotor descrito en la presente puede activarse con éxito para inducir la producción de proteínas recombinantes.

50 Según otro ejemplo, se produjo albúmina de suero humana ("human serum albumin", HSA) como una PDI bajo el

control de los promotores inducidos por una limitación en glucosa, y se determinó el rendimiento de HSA y el número de copias del gen.

- 5 Según otro ejemplo, se realizó un cultivo de alimentación discontinua de cepas de *P. pastoris* que expresan HSA bajo el control de un promotor descrito en la presente. Se descubrió que la inducción de la actividad de promotor bajo condiciones de limitación de glucosa fue incluso mayor que 120 veces con pG1, y mayor que 20 veces con pG6, comparado con el estado reprimido.

Otros ejemplos se refieren a expresar una carboxipeptidasa B porcina como proteína modelo bajo el control transcripcional del promotor pG1 y pG6.

Otro ejemplo se refiere a la expresión de un fragmento de anticuerpo bajo el control transcripcional de pG1.

- 10 Otro ejemplo demuestra la actividad funcional de variantes de un promotor descrito en la presente, tales como fragmentos de pG1 con una longitud en el intervalo de 300 a 1000 pb. Otros experimentos han demostrado que fragmentos de pG1 incluso más cortos fueron funcionalmente activos en un escenario similar, tales como fragmentos que varían de 200 y 1000 pb, o fragmentos que varían entre 250 y 1000.

- 15 La anterior descripción se entenderá más a fondo remitiéndose a los siguientes ejemplos. Sin embargo, dichos ejemplos son simplemente representativos de métodos para practicar una o más realizaciones de la presente invención y no deben considerarse como limitantes del alcance de la invención.

Ejemplos

Los siguientes ejemplos ilustran los materiales y los métodos usados para identificar nuevos promotores regulables y para analizar sus propiedades de expresión en *Pichia pastoris*.

- 20 Ejemplo 1: Identificación de genes potentes regulados con eficacia en *P. pastoris* en condiciones de limitación de glucosa

- 25 Para identificar genes potentes regulados con eficacia y sus respectivos promotores de *P. pastoris* en condiciones de limitación de glucosa, se realizó un análisis de los patrones de expresión génica empleando micromatrices. Células de *P. pastoris* en un cultivo discontinuo de glicerol (exceso de fuente de carbono) se compararon con células que se cultivaron en condiciones en las que la glucosa limitaba el crecimiento (quimioestado), imitando con ello el desarrollo de un proceso de producción de proteínas, que habitualmente se realiza en un modo de alimentación discontinuo.

a) Cepa

- 30 Se usó una cepa de *P. pastoris* (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos), que puede crecer en medio mínimo sin suplementos.

b) Cultivo de *P. pastoris*

Se realizaron fermentaciones con reactores Minifors (Infors-HT, Suiza) con un volumen de trabajo final de 2,5 l.

Se emplearon los siguientes medios:

Disolución madre de oligosales PTM₁ contenida por litro

- 35 6,0 g de CuSO₄·5H₂O, 0,08 g de NaI, 3,36 g de MnSO₄·H₂O, 0,2 g de Na₂MoO₄·2H₂O, 0,02 g de H₃BO₃, 0,82 g de CoCl₂, 20,0 g de ZnCl₂, 65,0 g de FeSO₄·7H₂O, 0,2 g de biotina y 5,0 ml de H₂SO₄ (al 95%-98%).

Medio discontinuo de glicerol contenido por litro

- 40 2 g de ácido cítrico monohidrato (C₆H₈O₇·H₂O), 39,2 g de glicerol, 20,8 g de NH₄H₂PO₄, 0,5 g de MgSO₄·7H₂O, 1,6 g de KCl, 0,022 g de CaCl₂·2H₂O, 0,8 mg de biotina y 4,6 ml de disolución madre de oligosales PTM₁. Se añadió HCl para ajustar el pH a 5.

Medio de alimentación discontinua de glicerol contenido por litro

632 g de glicerol, 8 g de MgSO₄·7H₂O, 22 g de KCl, y 0,058 g de CaCl₂·2H₂O.

Medio de quimioestado contenido por litro

- 45 2 g de ácido cítrico monohidrato (C₆H₈O₇·H₂O), 99,42 g de glucosa monohidrato, 22 g de NH₄H₂PO₄, 1,3 g de MgSO₄·7H₂O, 3,4 g de KCl, 0,02 g de CaCl₂·2H₂O, 0,4 mg de biotina y 3,2 ml de disolución madre de oligosales PTM₁. Se añadió HCl para ajustar el pH a 5.

El oxígeno disuelto se controló a OD = 20% con la velocidad del agitador (500-1250 rpm). La tasa de aireación fue

de 60 l h⁻¹ de aire, la temperatura se controló a 25 °C, y el punto de ajuste del pH de 5 se controló con la adición de NH₄OH (al 25%).

5 Para comenzar la fermentación se esterilizaron mediante filtración 1,5 l de medio discontinuo hacia el fermentador y se inoculó *P. pastoris* (procedente de un precultivo durante la noche en YPG, 180 rpm, 28 °C) con una densidad óptica inicial (DO600) de 1. La fase discontinua de aproximadamente 25 h alcanzó una concentración de biomasa seca de aproximadamente 20 g/l, y fue seguida de una fase de alimentación discontinua exponencial de 10 h con medio de glucosa que condujo a una concentración de biomasa seca de aproximadamente 50 g/l. Después el volumen se redujo hasta 1,5 y comenzó el cultivo en un quimiostato con una tasa de alimentación/recolección de 0,15 l h⁻¹, produciendo una tasa de crecimiento constante de $\mu = 0,1$. La fermentación terminó 50 h después del inicio de la utilización del quimiostato.

Esta fermentación se realizó tres veces para obtener los duplicados biológicos necesarios para un análisis con micromatrices fiable.

Las condiciones de limitación de carbono (no existe glucosa residual detectable) durante la utilización del quimiostato se verificaron mediante un análisis de HPLC del sobrenadante del cultivo.

15 c) Toma de muestras

Las muestras se tomaron al final de la fase discontinua de glicerol y en condiciones de estado estacionario del quimiostato de glucosa. Se realizó una toma de muestra habitual como determinación de la densidad óptica o de la masa seca de levaduras, una inspección microscópica cualitativa y un análisis de la viabilidad de las células durante cada fermentación. Para el análisis de micromatrices se tomaron muestras y se trataron como sigue: para una extinción óptica, 9 ml del caldo de cultivo celular se mezcló inmediatamente con 4,5 ml de disolución de fenol al 5% enfriada en hielo (Sigma) (en etanol absoluto), y se formaron partes alícuotas. Cada 2 ml se centrifugó (13200 rpm durante 1 minuto) en tubos de recolección preenfriados (GE Healthcare, NJ), el sobrenadante se retiró completamente, y los tubos se conservaron a -80 °C hasta la purificación del ARN.

d) Purificación del ARN y preparación de las muestras para la hibridación con micromatrices

25 El ARN se aisló usando el reactivo TRI según las instrucciones de los suministradores (Ambion, EE.UU.). Los sedimentos celulares se resuspendieron en reactivo TRI y se homogeneizaron con esferas de vidrio usando un FastPrep 24 (M.P. Biomedicals, CA) a 5 m s⁻¹ durante 40 segundos. Después de la adición de cloroformo, las muestras se centrifugaron y el ARN total se precipitó de la fase acuosa añadiendo isopropanol. El sedimento se lavó con etanol al 70%, se secó y se resuspendió en agua sin ARNasa. Se determinaron las concentraciones de ARN midiendo la DO260 usando un espectrofotómetro Nanodrop 1000 (NanoDrop Products, DE). Se retiró el ADN remanente de las muestras usando el kit sin ADN DNA Free Kit (Ambion, CA). Se diluyó un volumen de muestra igual a 10 µg de ARN hasta 50 µl en agua sin ARNasa, se añadió el tampón de ADNasa I y rADNasa I y se incubó a 37 °C durante 30 minutos. Después de la adición del reactivo de inactivación de ADNasa, la muestra se centrifugó y el sobrenadante se trasladó a un tubo nuevo. Se determinaron de nuevo las concentraciones de ARN como se describió anteriormente. Además, se analizó la integridad del ARN usando nanochips de ARN (Agilent). Para controlar el flujo de trabajo de la micromatriz desde la amplificación y el marcaje hasta la hibridación de las muestras se empleó el kit Spike In Kit (Agilent, n.º de producto: 5188-5279) como control positivo. Contiene 10 transcritos poliadenilados diferentes procedentes de un adenovirus, que se amplifican, se marcan y se cohibridan junto con las propias muestras de ARN. Las muestras se marcaron con Cy 3 y Cy 5 usando el kit de marcaje Quick Amp (Agilent, n.º de producto: 5190-0444). Para ello, se diluyeron 500 ng de la muestra de ARN purificada en 8,3 µl de agua sin ARNasa, y se añadieron 2 µl de siembra A o B, y 1,2 µl de cebador del promotor de T7. La mezcla se desnaturalizó durante 10 minutos a 65 °C y se mantuvo en hielo durante 5 minutos. Después se añadieron 8,5 µl de mezcla maestra de ADNc (por muestra: 4 µl de 5× tampón de primera hebra, 2 µl de DTT 0,1 M, 1 µl de mezcla de dNTP 10 mM, 1 µl de MMLV-RT, 0,5 µl de ARNasa Out), se incubaron a 40 °C durante 2 horas, después se trasladaron a 65 °C durante 15 minutos y se colocaron sobre hielo durante 5 minutos. Se preparó la mezcla maestra de transcripción (por muestra: 15,3 µl de agua sin nucleasa, 20 µl de tampón de transcripción, 6 µl de DTT 0,1 M, 6,4 µl de PEG al 50%, 0,5 µl de inhibidor de ARNasa, 0,6 µl de fosfatasa inorgánica, 0,8 µl de ARN polimerasa de T7, 2,4 µl de cianina 3 o cianina 5) y se añadió a cada tubo y se incubó a 40 °C durante 2 horas. Para purificar el ARNc marcado obtenido se empleó el kit RNeasy Mini Kit (Qiagen, n.º de catálogo 74104). Las muestras se conservaron a -80 °C. Se realizó la cuantificación de la concentración del ARNc y la eficacia del marcaje con el espectrofotómetro Nanodrop.

e) Análisis de micromatrices

55 Para identificar genes fuertes regulados con eficacia en cultivos en quimiostato con limitación de glucosa, se compararon tres duplicados de muestras biológicas procedentes de estos con la misma referencia y en un intercambio de tintes cada una. La muestra de referencia se generó combinando las muestras de cultivo discontinuo de glicerol en cantidades iguales.

Se empleó el kit de hibridación de expresión génica (Agilent, n.º de catálogo 5188-5242) para la hibridación de los ARNc de las muestras marcados. Para la preparación de las muestras de hibridación, cada 300 ng de ARNc (Cy3 y

Cy 5) y 6 µl de agente de bloque en 10 veces se diluyeron con agua sin nucleasa hasta un volumen final de 24 µl. Después de la adición de 1 µl de tampón de fragmentación en 25 veces, la mezcla se incubó a 60 °C durante 30 minutos. Después se añadieron 25 µl de tampón de hibridación GEx Hybridisation Buffer HI-RPM para detener la reacción. Después de una centrifugación durante un minuto a 13.200 rpm, la muestra se enfrió en hielo y se usó para la hibridación inmediatamente. Se usaron matrices de oligonucleótidos específicos de *P. pastoris* diseñados en el laboratorio de los inventores (AMAD-ID: 026594, matrices básicas 8×15K, Agilent). La hibridación con micromatrices se realizó según la guía del usuario de cámaras de hibridación con micromatrices (Agilent G2534A). En primer lugar, se descubrió la platina de la junta y se colocó sobre la base de cámara con la etiqueta de Agilent boca arriba. La muestra (40 µl por matriz) se cargó en medio de cada uno de los ocho cuadrados. Después la platina de la micromatriz se colocó cuidadosamente sobre la platina de la junta (etiqueta de Agilent boca abajo) y se colocó la cubierta de la cámara y se fijó con la abrazadera. Se realizó la hibridación en el horno de hibridación durante 17 horas a 65 °C. Antes del escaneado, el chip de la micromatriz se lavó. Para ello se desmontó la cámara y las platinas intercaladas se desprendieron entre sí mientras se encontraban sumergidas en tampón de lavado 1. La micromatriz se trasladó directamente a otra placa con tampón de lavado 1, se lavó durante 1 minuto, se trasladó a tampón de lavado 2 (temperatura de al menos 30 °C) y se lavó durante otro minuto. Después de secar la platina de la micromatriz tocando el borde de la platina con una pañuelo de papel, esta se colocó en el soporte de platinas (con la etiqueta de Agilent boca arriba). El soporte de platinas se introdujo en el carrusel y se inició el escaneado.

f) Adquisición de datos y evaluación estadística de los datos de las micromatrices

Las imágenes se escanearon a una resolución de 50 nm con un escáner de micromatrices G2565AA (Agilent) y se importaron al programa informático Agilent Feature Extraction 9.5. Se empleó Agilent Feature Extraction 9.5 para la cuantificación de las intensidades de las manchas. Después los datos de intensidad de las manchas promedio brutos se importaron al programa de código abierto R para la posterior normalización y análisis de los datos.

Para el preprocesamiento y normalización de los datos, se emplearon los paquetes de R limma, vsn y marray. Los datos de intensidad no se corrigieron para el fondo y se normalizaron con VSN, y después de la normalización se transformaron a proporciones de log 2 del canal de Cy5 frente al canal de Cy3. Se calculó la expresión diferencial usando la función lmfit y eBayes del paquete limma.

Los datos de las micromatrices se exploraron para entradas con una elevada diferencia en el nivel de expresión entre el estado reprimido e inducido (cambio en número de veces), así como una elevada intensidad de señal en el estado inducido para identificar los genes fuertemente expresados y regulados de modo eficaz. En la tabla 1 se muestra una lista de los genes seleccionados, y el cambio en número de veces significa la intensidad de señal en el estado inducido dividida entre la intensidad de señal en el estado reprimido. Se añaden los datos de pGAP, pMLS1 y pICL1 como referencia.

Tabla 1. Datos de las micromatrices de los promotores seleccionados para la posterior caracterización y de pGAP, ICL1 y MLS1 como controles

Nombre	Anotación/homólogo de levadura	Identificador del gen	Cambio en n.º de veces	Intensidad*	% de intensidad/potencia de transcripción
pGAP	TDH3	PAS_chr2-1_0437	0,79	41052,5	100,0
pG1	—	PAS_chr1-3_0011	29,86	86312,9	210,2
pG3	YPR127W	PAS_chr4_0550	2,66	15644,4	38,1
pG4	—	PAS_chr4_0043	2,57	15664,8	38,2
pG6	ALD4	PAS_chr2-1_0853	2,10	26888,4	65,5
MLS1	MLS1	PAS_chr4_0191	0,81	1446,9	3,5
ICL1	ICL1	PAS_chr1-4_0338	1,71	2574,3	6,3
pG7	HXT6	PAS_chr1-4_0570	3,3	13336,5	32,5
pG8	SFL1	PAS_chr1-3_0165	2,1	9929,1	24,2

*del estado inducido en el canal verde

Ejemplo 2: Estudios de actividad de promotor comparativos de los promotores recién identificados en *P. pastoris* usando eGFP como gen indicador expresado de modo intracelular

Para analizar las propiedades de los promotores recién identificados bajo condiciones de limitación de glucosa, se realizaron selecciones en matraces de agitación como sigue: se realizó el precultivo durante 24 horas con medio rico que contenía glicerol como fuente de carbono, que simula la fase discontinua del proceso (estado reprimido de los promotores), seguido del cultivo principal con medio mínimo y esferas de alimentación de glucosa, que simula la

fase de alimentación discontinua con limitación de glucosa del proceso (estado inducido de los promotores).

a) Cepa y vector de expresión

Se usó una cepa de tipo salvaje de *P. pastoris* (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) como cepa hospedante. La transformación de la cepa se realizó con un vector del laboratorio de los inventores denominado pPUZZLE (Stadlmayr *et al.*, J. Biotechnol., diciembre de 2010, 150(4):519-529), que comprende un origen de la replicación de *E. coli* (pUC19), un módulo de resistencia a antibióticos (gene Sh ble que confiere resistencia a la zeocina) para la selección en *E. coli* y levaduras, un módulo de expresión para el gen de interés (GDI) que consiste en un sitio de clonación múltiple y el terminador de la transcripción de *S. cerevisiae* CYC1, y un locus para la integración en el genoma de *P. pastoris* (región 3'AOX1).

b) Amplificación y clonación de los promotores recién identificados pG1, pG3, pG4 y pG6 en el vector de expresión pPUZZLE que contiene eGFP como GDI

En la tabla 2 se muestra una lista de las secuencias de promotores recién identificados y sus respectivos genes (véase el ejemplo 1). Se amplificaron 1000 pb de la región no codificadora 5' de los respectivos genes hasta el codón de inicio ATG mediante PCR (Phusion Polymerase, New England Biolabs) como secuencias de promotores usando los cebadores que se muestran en la tabla 2. Estas secuencias se clonaron en el vector de expresión pPUZZLE pPM1aZ10_eGFP, produciendo pPM1aZ10_pG1_eGFP, pPM1aZ10_pG3_eGFP, pPM1aZ10_pG4_eGFP y pPM1aZ10_pG6_eGFP. Además, se usó el vector pPM1aZ10_pGAP_eGFP, que contiene el promotor utilizado habitualmente de gliceraldehído 3-fosfato deshidrogenasa (pGAP de *P. pastoris*, en la presente SEQ ID NO:25) como referencia. Los promotores se insertaron cadena arriba del codón de inicio del gen eGFP usando los sitios de restricción Apal y SbfI (véanse las tablas 2 y 3). Se verificó que las secuencias de promotores eran correctas mediante secuenciación Sanger.

Tabla 2. Cebadores para la amplificación con PCR de los promotores

Nombre	Diana	Secuencia	T _M	Sitio de restricción
pG1_dir	pG1	SEQ ID NO:14 GATAGGGCCCCAAACATTTGCT CCCCCTAGTCTC	70,8	Apal
pG1_inv	pG1	SEQ ID NO:15 GATACCTGCAGGAAGGGTGGAA TTTTAAGGATCTTTTAT	69,8	SbfI
pG3_dir	pG3	SEQ ID NO:16 GATAGGGCCCCAGCAATCCAGT AACCTTTTCTGAAT	70,4	Apal
pG3_inv	pG3	SEQ ID NO:17 GATACCTGCAGGTTGAGTTCAAT AAATTGTCCGGGA	70,2	SbfI
pG4_dir	pG4	SEQ ID NO:18 GATAGGGCCCTGGACTGTTCAAT TTGAAGTCGATG	70,4	Apal
pG4_inv	pG4	SEQ ID NO:19 GATACCTGCAGGGGATAAAGGTA AGGGAAAAAGCAA	70	SbfI
pG6_dir	pG6	SEQ ID NO:20 GATAGGGCCCCAGACCAGCAGTTT AACTACGCAAATC	70,6	Apal
pG6_inv	pG6	SEQ ID NO:21 GATACCTGCAGGCTTTTCTTTGGG CAAGGAAAAATC	70,7	SbfI

Nombre	Diana	Secuencia	T _M	Sitio de restricción
pG7_dir	pG7	SEQ ID NO:22 GATAGGGCCCAATTGATTAAGTTCAG TGAAATTTCAAAC	69,1	Apal
pG7_inv	pG7	SEQ ID NO:23 GATACCTGCAGGATTATATTATGGGG AATAATGAAGAGAAGG	70,9	Sbfl
pG8_dir	pG8	SEQ ID NO:24 GATAGGGCCCTGCACAACCATTGCC AGTAAGG	71,5	Apal
pG8_inv	pG8	SEQ ID NO:25 GATACCTGCAGGTTTTTAGAAGAGGG AGAACTTAGATTGG	70,4	Sbfl

Tabla 3. Cebadores de amplificación, enzimas de clonación y longitud de los promotores clonados

Promotor	Cebador 5'	Cebador 3'	Enzima de clonación		Longitud del fragmento
			5'	3'	
pG1	pG1_dir	pG1_inv	Apal	Sbfl	988
pG3	pG3_dir	pG3_inv	Apal	Sbfl	1011
pG4	pG4_dir	pG4_inv	Apal	Sbfl	1022
pG6	pG6_dir	pG6_inv	Apal	Sbfl	1022
pG7	pG7_dir	pG7_inv	Apal	Sbfl	1022
pG8	pG8_dir	pG8_inv	Apal	Sbfl	1022

c) Expresión de eGFP en *P. pastoris* para el análisis de la actividad de promotor

Todos los plásmidos se linealizaron con *Ascl* dentro de la región de integración en el genoma 3'AOX antes de la electroporación (2 kV, 4 ms, GenePulser, BioRad) en *P. pastoris* electrocompetente.

- 5 Se realizó la selección de los transformantes positivos en placas YPD (por litro: 10 g de extracto de levadura, 20 g de peptona, 20 g de glucosa, 20 g de agar-agar) que contenían 25 µg/ml de zeocina (Invitrogen, CA). Se empleó una PCR de colonias para asegurar la presencia del plásmido transformado. Para ello, se adquirió el ADN genómico mediante cocido y congelación de colonias de *P. pastoris* durante 5 minutos cada uno y se aplicó directamente a la PCR con los cebadores apropiados. Para la selección de la expresión, una única colonia se inoculó en medio YPG-Zeo líquido (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de zeocina) como precultivo. Después de aproximadamente 24 h, el precultivo se usó para inocular el cultivo principal con una DO600 de 0,1 en 10 ml de medio YP (por litro: 20 g de peptona, 10 g de extracto de levadura) y 2 de esferas de alimentación de glucosa (Kuhner, CH). Se lograron las condiciones de crecimiento con limitación de glucosa debido a la lenta cinética de liberación de glucosa de estas esferas de alimentación, que se describe mediante la siguiente ecuación: (glucosa) = 1,63* $t^{0,74}$ [mg/disc]. Las muestras se tomaron al final del precultivo y 24 y 48 horas después de la inoculación del cultivo principal. Se determinó la densidad celular midiendo la DO600, se analizó la expresión de eGFP mediante citometría de flujo como se describe en Stadlmayr *et al.* (J. Biotechnology, diciembre de 2010, 150(4):519-529). Para cada muestra se analizaron 10.000 células. Se midió la autofluorescencia de *P. pastoris* usando células de tipo salvaje de *P. pastoris* no transformadas y se restó de la señal. Los niveles de expresión de eGFP relativos (intensidad de fluorescencia relacionada con el tamaño celular) como porcentaje del nivel de expresión de eGFP bajo el control del promotor pGAP constitutivo. Se realizaron estudios similares con los promotores pG7 y pG8. La clonación se realiza como se describe en el ejemplo 2b, excepto que se usó la cepa de tipo salvaje de *P. pastoris* X-33 (Invitrogen) para la transformación de pPM1aZ10_pG7_eGFP y pPM1aZ10_pG8_eGFP. Los cebadores y fragmentos de clonación usados se listan en las tablas 2 y 3. Los resultados se muestran en la tabla 4.
- 10
- 15
- 20
- 25

Tabla 4. Resultados de la selección de clones de *P. pastoris* que expresan eGFP bajo el control de los nuevos promotores; los datos mostrados (fluorescencia/tamaño de las células) se refieren a pGAP.

	Precultivo		Cultivo principal	
	Final de descont.	Desv. est.	48 h	Desv. est.
pG1	7,6	0,2	242,8	59,5
pG3	-5,1	2,4	25,4	5,5
pG4	-6,3	0,2	113,6	26,3
pG6	3,3	0,8	158,9	146,9
pG7	49,4	7,4	115,7	16,2
pG8	0,8	4,1	36,1	21,1

d) Determinación del número de copias del gen eGFP ("Gene Copy Numbers", GCN) en clones que expresan eGFP seleccionados

- 5 La potencia de expresión a menudo se correlaciona con el número de módulos de expresión integrados en el genoma de *P. pastoris*. Por tanto, se determinó el número de copias del gen de eGFP. El ADN genómico se aisló usando el kit de sangre y tejido DNeasy (Quiagen, n.º de catálogo 69504). El número de copias del gen se determinó utilizando una PCR cuantitativa. Para ello, se empleó el kit SensiMix SYBR (Bioline, QT605-05). Se mezcló Sensi Mix SYBR con los cebadores y la muestra, y se aplicó para un análisis a tiempo real en un ciclador de PCR a tiempo real (Rotor Gene, Qiagen). En la tabla 5 se muestra una lista de los cebadores. Todas las muestra se analizaron en triplicado o cuadruplicado. Se empleó el programa informático Rotor Gene para el análisis de los datos. Se empleó el gen de actina *ACT1* como calibrador. Los resultados se muestran en la tabla 6.

Tabla 5. Cebadores para la determinación del número de copias del gen mediante PCR a tiempo real

Cebador	Diana	Secuencia	T _M [°C]	Longitud del producto
PpACT1_Up	Act	SEQ ID NO:26 CCTGAGGCTTTGTTCC ACCCATCT	61,3	148 pb
PpACT1_Low	Act	SEQ ID NO:27 GGAACATAGTAGTACC ACCGGACATAACGA	61,4	148 pb
PpeGFP_Up	GFP	SEQ ID NO:28 TCGCCGACCACTACCA GCAGAA	61,4	124 pb
PpeGFP_Low	GFP	SEQ ID NO:29 ACCATGTGATCGCGCT TCTCGTT	61,6	124 pb

15 Tabla 6. Resultados de la selección (fluorescencia/tamaño de las células con relación a pGAP) y el número de copias de genes de clones de *P. pastoris* seleccionados que expresan eGFP bajo el control de pG1 y pG6

	GCN	% de fluorescencia de pGAP_eGFP/tamaño		
		Precultivo	Cultivo principal	
			24 h	48 h
pG1#8	1	7,32	33,00	184,30
pG1#9	1	7,73	33,96	303,21
pG1#12	2	7,75	33,32	240,92
pG6#48	1	3,45	2,07	56,11
pG6#50	2	4,00	23,18	327,14
pG6#53	1	2,52	9,78	93,51

e) Análisis de la potencia del promotor pG1 en una fermentación de alimentación discontinua de un clon de eGFP

Se realizaron fermentaciones con alimentación discontinua en reactores DASGIP con un volumen de trabajo final de 0,7 l.

Se emplearon los siguientes medios:

5 *Disolución madre de oligosales PTM₁ contenida por litro*

6,0 g de CuSO₄·5H₂O, 0,08 g de NaI, 3,36 g de MnSO₄·H₂O, 0,2 g de Na₂MoO₄·2H₂O, 0,02 g de H₃BO₃, 0,82 g de CoCl₂, 20,0 g de ZnCl₂, 65,0 g de FeSO₄·7H₂O, 0,2 g de biotina y 5,0 ml de H₂SO₄ (al 95%-98%).

Medio discontinuo de glicerol contenido por litro

10 2 g de ácido cítrico monohidrato (C₆H₈O₇·H₂O), 39,2 g de glicerol, 12,6 g de NH₄H₂PO₄, 0,5 g de MgSO₄·7H₂O, 0,9 g de KCl, 0,022 g de CaCl₂·2H₂O, 0,4 mg de biotina y 4,6 ml de disolución madre de oligosales PTM₁. Se añadió HCl para ajustar el pH a 5.

Medio de alimentación discontinua de glicerol contenido por litro

464 g de glucosa monohidrato, 5,2 g de MgSO₄·7H₂O, 8,4 g de KCl, 0,28 g de CaCl₂·2H₂O, 0,34 mg de biotina y 10,1 ml de disolución madre de oligosales PTM₁.

15 El oxígeno disuelto se controló a OD = 20% con la velocidad del agitador (400-1200 rpm). La tasa de aireación fue de 24 l h⁻¹ de aire, la temperatura se controló a 25 °C, y el punto de ajuste del pH de 5 se controló con la adición de NH₄OH (al 25%).

Para comenzar la fermentación se esterilizaron mediante filtración 400 ml de medio discontinuo hacia el fermentador y se inoculó *P. pastoris* clon pG1_eGFP#8 (procedente de un precultivo) con una densidad óptica inicial (DO600) de 1. La fase discontinua de aproximadamente 25 h (se alcanza una concentración de biomasa seca de aproximadamente 20 g/l) fue seguida de una fase de alimentación discontinua con limitación de glucosa que comienza con una alimentación exponencial durante 7 h y una tasa de alimentación constante de 15 g/l durante 13 h, que condujo a una concentración de biomasa seca final de aproximadamente 100 g/l. Se tomaron muestras durante la fase discontinua y la fase de alimentación discontinua, y se analizaron para la expresión de eGFP usando un lector de placas (Infinite 200, Tecan, CH). Para ello, las muestras se diluyeron hasta una densidad óptica (DO600) de 5. Los resultados se muestran en la tabla 7 como fluorescencia relativa por biorreactor (FL/r).

Tabla 7. Fluorescencia relativa por biorreactor de dos clones de *P. pastoris* diferentes que expresan eGFP bajo el control de pGAP o pG1 en una fermentación de alimentación discontinua optimizada

pGAP_eGFP#2		pG1_eGFP#8	
t [h]	FL/r	t [h]	FL/r
-1,7	176,77	-0,38	131,95
0,0	166,52	0,00	108,76
0,5	199,59	0,28	100,35
1,0	195,94	0,62	121,36
1,5	173,68	1,12	161,16
2,0	219,00	1,62	162,69
3,0	321,14	2,12	148,34
7,0	494,60	3,12	205,20
19,1	1150,96	7,12	373,08
20,0	1000,37	19,70	1745,65
		21,12	1831,52

30 Ejemplo 3: Estudios de actividad de promotor comparativos de los promotores recién identificados en *P. pastoris* usando albúmina de suero humana (HSA) como gen indicador expresado de modo extracelular

Para analizar las propiedades de los promotores recién identificados bajo condiciones de limitación de glucosa, se realizaron selecciones en matraces de agitación como sigue: se realizó el precultivo durante 24 horas con medio rico que contenía glicerol como fuente de carbono, que simula la fase discontinua del proceso (estado reprimido de los

promotores), seguido del cultivo principal con medio mínimo y esferas de alimentación de glucosa, que simula la fase de alimentación discontinua con limitación de glucosa del proceso (estado inducido de los promotores).

a) Cepa y vector de expresión

5 Se usó una cepa de tipo salvaje de *P. pastoris* (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) como cepa hospedante. La transformación de la cepa se realizó con un vector del laboratorio de los inventores denominado pPUZZLE (Stadlmayr *et al.*, J. Biotechnol., diciembre de 2010, 150(4):519-529), y la selección de los transformantes positivos se basó en la resistencia a la zeocina. Para la expresión secretora de la albúmina de suero humana (HSA) se usó su conductor de la secreción nativo.

10 b) Amplificación y clonación de los promotores recién identificados pG1, pG3, pG4 y pG6 en un vector de expresión del laboratorio de los inventores

15 Los cuatro promotores amplificados en el ejemplo 2b se clonaron en el vector de expresión pPUZZLE pPM1aZ10_HSA, produciendo pPM1aZ10_pG1_HSA, pPM1aZ10_pG3_HSA, pPM1aZ10_pG4_HSA y pPM1aZ10_pG6_HSA. Además, se usó como referencia el vector pPM1aZ10_pGAP_HSA, que contiene el promotor utilizado habitualmente de gliceraldehído 3-fosfato deshidrogenasa (pGAP). Los promotores se insertaron cadena arriba del codón de inicio del gen HSA usando los sitios de restricción Apal y SbfI (véase la tabla 3). Se verificó que las secuencias de promotores eran correctas mediante secuenciación Sanger.

c) Expresión de HSA en *P. pastoris* bajo el control de los promotores inducidos por una limitación en glucosa recién identificados

20 Todos los plásmidos se linealizaron usando la enzima de restricción Ascl antes de la electroporación (empleando un protocolo de transformación convencional para *P. pastoris*) en *P. pastoris*. Se realizó la selección de los transformantes positivos en placas YPD (por litro: 10 g de extracto de levadura, 20 g de peptona, 20 g de glucosa, 20 g de agar-agar) que contenían 25 µg/ml de zeocina. Se empleó una PCR de colonias para asegurar la presencia del plásmido transformado como se describió en el ejemplo 2c.

25 Para la selección de la expresión de HSA, una única colonia se inoculó en medio YPG-Zeo líquido (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de zeocina) como precultivo. Después de aproximadamente 24 h, el precultivo se usó para inocular el cultivo principal con una DO600 de 1 en medio YP (por litro: 20 g de peptona, 10 g de extracto de levadura) y esferas de alimentación de glucosa (Kuhner, CH). Se lograron las condiciones de crecimiento con limitación de glucosa debido a la lenta cinética de liberación de glucosa de estas esferas de alimentación, que se describe mediante la siguiente ecuación: (glucosa) = 1,63*t^{0,74} [mg/disc]. Las muestras se tomaron al final del precultivo y 24 y 48 horas después de la inoculación del cultivo principal. Se determinó la concentración de biomasa midiendo la DO600 o el peso celular en húmedo. La concentración de HSA en el sobrenadante del cultivo se cuantificó mediante el equipo de cuantificación con ELISA de albúmina humana (n.º de catálogo E80-129, Bethyl Laboratories, TX, EE.UU.) según el manual de instrucciones del suministrador. El patrón de HSA se usó con una concentración inicial de 400 ng ml⁻¹. Las muestras se diluyeron en consecuencia en diluyente de muestras (Tris-HCl 50 mM, NaCl 140 mM, BSA al 1% (en p/v), Tween20 al 0,05% (en v/v), pH 8,0). En la tabla 8 se presentan las titulaciones de HSA de la selección de varios clones de cada construcción.

Tabla 8. Resultados de la selección de clones de *P. pastoris* que expresan HSA bajo el control de pGAP, pG1 y pG6

Clon	Titulación de HSA [mg l ⁻¹]	
	Precultivo	Cultivo principal 48 h
pGAP_HSA #1	6,9	9,0
pGAP_HSA #2	9,0	8,6
pGAP_HSA #3	6,6	9,2
pGAP_HSA #4	18,9	20,4
pGAP_HSA #5	9,6	8,3
pGAP_HSA #6	10,8	8,8
pG1_HSA #19	0,6	6,9
pG1_HSA #20	0,6	6,7
pG1_HSA #21	0,1	7,0
pG1_HSA #22	—	—
pG1_HSA #23	1,3	13,5

ES 2 730 173 T3

Clon	Titulación de HSA [mg l ⁻¹]	
	Precultivo	Cultivo principal 48 h
pG1_HSA #24	1,1	13,7
pG1_HSA #25	0,5	8,9
pG1_HSA #26	0,5	9,2
pG1_HSA #27	0,6	7,3
pG1_HSA #28	0,6	6,1
pG1_HSA #29	0,6	6,4
pG1_HSA #30	0,6	7,1
pG6_HSA #31	0,3	1,8
pG6_HSA #32	0,3	1,7
pG6_HSA #33	0,3	2,0
pG6_HSA #34	0,4	2,0
pG6_HSA #35	0,2	2,2
pG6_HSA #36	0,3	2,5
pG6_HSA #37	0,3	2,3
pG6_HSA #38	0,2	1,5
pG6_HSA #39	0,7	—
pG6_HSA #40	0,2	2,4
pG6_HSA #41	0,4	—
pG6_HSA #42	—	1,9

d) Determinación del número de copias del gen HSA

El aislamiento del ADN genómico y la medición con qPCR se realizaron como en el ejemplo 2d, usando los cebadores que aparecen en la tabla 9. Los resultados se muestran en la tabla 10.

Tabla 9. Cebadores para la determinación del número de copias del gen mediante PCR a tiempo real

Cebador	Diana	Secuencia	Longitud del producto
PpACT1_Up	Act	SEQ ID NO:30 CCTGAGGCTTTGTTCCACCCATCT	148 pb
PpACT1_Low	Act	SEQ ID NO:31 GGAACATAGTAGTACCACCGGACAT AACGA	148 pb
PpHSA_Up	HSA	SEQ ID NO:32 AAACCTAGGAAAAGTGGGCAGCAA ATGT	135 pb
PpHSA_Low	HSA	SEQ ID NO:33 ACTCTGTCACTTACTGGCGTTTTTC TCATG	135 pb

Tabla 10. Resultados de la selección y el número de copias de genes de clones de *P. pastoris* seleccionados que expresan HSA bajo el control de pGAP, pG1 y pG6

Clon	GCN	HSA mg ^l ⁻¹ del cultivo principal	HSA mg ^l ⁻¹ por GCN del cultivo principal	Promedio	Desv. est.
pGAP_HSA#3	1	9,2	9,2	9,22	0,95
pGAP_HSA#4	2	20,4	10,2		
pGAP_HSA#5	1	8,3	8,3		
pG1_HSA#20	1	6,6	6,6	6,81	0,13
pG1_HSA#21	1	7,0	7,0		
pG1_HSA#23	2	13,5	6,8		
pG1_HSA#24	2	13,7	6,8		
Clon	GCN	HSA mg ^l ⁻¹ del cultivo principal	HSA mg ^l ⁻¹ por GCN del Cultivo principal	Promedio	Desv. est.
pG6_HSA#36	1	2,5	2,5	2,07	0,52
pG6_HSA#37	1	2,3	2,3		
pG6_HSA#38	1	1,5	1,5		

e) Cultivo de alimentación discontinua de cepas de *P. pastoris* que expresan HSA bajo el control del promotor pG1 y pG6

- 5 Las fermentaciones se realizaron en biorreactores DASGIP con un volumen de trabajo final de 0,7 l. La cepa pG1_HSA#23 posee dos copias del gen HSA, y la cepa pG6_HSA#36 porta solo una copia del gen HSA. Por tanto, se cultivaron dos cepas de *P. pastoris* diferentes que expresan HSA bajo el control de pGAP (pGAP_HSA#3 con una copia del gen HSA, y pGAP_HSA#4 con dos copias del gen HSA) como referencia. Todas las fermentaciones se realizaron por duplicado.
- 10 Se emplearon los siguientes medios:
- Disolución madre de oligosales PTM₁ contenida por litro*
- 6,0 g de CuSO₄.5H₂O, 0,08 g de NaI, 3,36 g de MnSO₄.H₂O, 0,2 g de Na₂MoO₄.2H₂O, 0,02 g de H₃BO₃, 0,82 g de CoCl₂, 20,0 g de ZnCl₂, 65,0 g de FeSO₄.7H₂O, 0,2 g de biotina y 5,0 ml de H₂SO₄ (al 95%-98%).
- Medio discontinuo de glicerol contenido por litro*
- 15 39,2 g de glicerol, 27,9 g de H₃PO₄ (al 85%), 7,8 g de MgSO₄.7H₂O, 2,6 g de KOH, 9,5 g de K₂SO₄, 0,6 g de CaSO₄.2H₂O, 0,4 mg de biotina y 4,6 ml de disolución madre de oligosales PTM1. El pH se ajustó a 5,85 después de una esterilización por filtración hacia el fermentador.
- Medio de alimentación discontinua de glicerol contenido por litro*
- 20 550 g de glucosa monohidrato, 6,5 g de MgSO₄.7H₂O, 10 g de KCl, 0,35 g de CaCl₂.2H₂O, 0,4 mg de biotina y 12 ml de disolución madre de oligosales PTM1.
- El oxígeno disuelto se controló a OD = 20% con la velocidad del agitador (400-1200 rpm). La tasa de aireación fue de 24 l h⁻¹ de aire, la temperatura se controló a 25 °C, y el punto de ajuste del pH de 5,85 se controló con la adición de NH₄OH (al 25%).
- 25 Para comenzar la fermentación se esterilizaron mediante filtración 400 ml de medio discontinuo hacia el fermentador y se inoculó *P. pastoris* (procedente de un precultivo) con una densidad óptica inicial (DO600) de 1. La fase discontinua de aproximadamente 25 h alcanzó una concentración de biomasa seca de aproximadamente 20 g/l y fue seguida de una fase de alimentación constante (durante 100 horas) con medio de glucosa que condujo a una concentración de biomasa seca de aproximadamente 100 g/l. El pH fue de 5,85 durante la fase discontinua y se mantuvo a 5,85 a lo largo de la fermentación. Se tomaron muestras durante la fase discontinua y la fase de alimentación discontinua. La concentración de HSA se cuantificó mediante el equipo de cuantificación con ELISA de albúmina humana (Bethyl, n.º de catálogo E80-129) como se describió en el ejemplo 3c. En la tabla 11 se muestra la concentración de biomasa y las titulaciones de HSA, y en la tabla 12 se muestra el rendimiento del producto
- 30

5 (cantidad de HSA segregada por biomasa, HSA/YDM) al final de la fase discontinua (condiciones de represión para pG1 y pG6) y al final de la alimentación discontinua (condiciones de inducción para pG1 y pG6). Con ello puede verificarse la estrategia de inducción. Los promotores pG1 y pG6 son reprimidos bajo un exceso de fuente de carbono (en la fase discontinua de glicerol), y casi no muestran HSA detectable, en contraste con los clones conducidos por pGAP. La inducción de pG1 y pG6 se produjo tras el cambio a condiciones con C limitado al inicio de la fase de alimentación discontinua. La inducción de pG1 (HSA/YDM) fue en más de 120 veces comparado con el estado reprimido, la inducción de pG6 fue en más de 20 comparado con el estado reprimido, y apenas se observó cambio para pGAP (aumento en 3 veces de HSA/YDM, comparado con la fase discontinua).

10 Tabla 11. Masa seca de levaduras y titulaciones de HSA al final de la fase discontinua y de la alimentación discontinua de 7 fermentaciones de clones de *P. pastoris* que expresan HSA bajo el control de pGAP, pG1 o pG6

CLON	n.º de fermentación	Final de la fase discontinua			Final de la alimentación discontinua		
		tiempo [h]	YDM [gl ⁻¹]	Titulación de HSA [mg l ⁻¹]	tiempo [h]	YDM [gl ⁻¹]	Titulación de HSA [mg l ⁻¹]
pG1#23	A041	-1,1	24,7	0,5	99,6	125,3	328,6
pG1#23	A048	-0,3	23,9	0,5	108,4	128,6	277,6
pG6#36	A045	-0,1	23,5	0,3	104,7	125,2	21,8
pG6#36	A049	-0,3	24,4	0,3	108,4	129,0	26,9
pGAP#4	A044	-0,1	23,6	11,2	104,7	129,1	141,4
pGAP#4	A051	-0,9	24,1	9,0	96,9	118,2	114,9
pGAP#3	A050	-0,9	24,2	5,0	96,9	117,7	57,8

Tabla 12. Titulación de HSA por masa seca de levaduras al final de la fase discontinua y al final de la alimentación discontinua de 7 fermentaciones de clones de *P. pastoris* que expresan HSA bajo el control de pGAP, pG1 o pG6

	GCN HSA	Final de la fase discontinua		Final de la alimentación discontinua		Inducción en n.º de veces
		Promedio HSA/YDM	DE	Promedio HSA/YDM	DE	
pG1#23	2	0,02	0,00	2,39	0,33	121,06
pG6#36	1	0,01	0,00	0,19	0,02	21,39
pGAP#4	2	0,42	0,07	1,03	0,09	3,16
pGAP#3	1	0,21		0,49		

Ejemplo 4: Estudios de actividad de promotor comparativos en diversas concentraciones de glucosa de los promotores recién identificados en *P. pastoris* usando eGFP como gen indicador expresado de modo intracelular

15 Para analizar las propiedades de los promotores recién identificados en diversas concentraciones de glucosa, se realizaron selecciones en matraces de agitación como sigue: se realizó el precultivo durante 24 horas con medio rico que contenía glicerol como fuente de carbono (estado reprimido de los promotores), seguido del cultivo principal con medio mínimo y glucosa como fuente de carbono (estado inducido de los promotores).

a) Cepa y vector de expresión

20 Se usó una cepa de tipo salvaje de *P. pastoris* (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) como cepa hospedante. La transformación de la cepa se realizó con un vector del laboratorio de los inventores denominado pPUZZLE (Stadlmayr *et al.*, J. Biotechnol., diciembre de 2010, 150(4):519-529), y la selección de los transformantes positivos se basó en la resistencia a la zeocina.

25 b) Amplificación y clonación de los promotores recién identificados pG1, pG3, pG4 y pG6 en el vector de expresión pPUZZLE que contiene eGFP como GDI

La amplificación y la clonación se realizan como se describió en el ejemplo 2.

c) Expresión de eGFP en *P. pastoris* para el análisis de la actividad de promotor

La transformación y la selección de clones se realizan como se describió en el ejemplo 2.

Para la selección de la expresión, una única colonia se inoculó en medio YPG-Zeo líquido (por litro: 20 g de peptona,

10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de zeocina) como precultivo. Después de aproximadamente 24 h, el precultivo se usó para inocular el cultivo principal con una DO600 de 0,01 en 10 ml de medio YP (por litro: 20 g de peptona, 10 g de extracto de levadura) y glucosa como fuente de carbono. La glucosa se emplea en diversas concentraciones de 20 a 0,001 g l⁻¹.

5 Las muestras se tomaron 1-8 después de la inoculación del cultivo principal. La expresión de eGFP se analiza mediante citometría de flujo como se describe en Stadlmayr *et al.* (J. Biotechnology, diciembre de 2010, 150(4):519-529), y la selección de los transformantes positivos se basó en la resistencia a la zeocina. Para cada muestra se analizaron 10.000 células. Se midió la autofluorescencia de *P. pastoris* usando células de tipo salvaje de *P. pastoris* no transformadas.

10 Ejemplo 5: Estudios de actividad de promotor comparativos de los promotores recién identificados en *P. pastoris* usando carboxipeptidasa B porcina (CpB) como gen indicador expresado de modo extracelular

Para analizar las propiedades de los promotores recién identificados bajo condiciones de limitación de glucosa, se realizaron selecciones en matraces de agitación como sigue: se realizó el precultivo durante 24 horas con medio rico que contenía glicerol como fuente de carbono, que simula la fase discontinua del proceso (estado reprimido de los promotores), seguido del cultivo principal con medio mínimo y esferas de alimentación de glucosa, que simula la fase de alimentación discontinua con limitación de glucosa del proceso (estado inducido de los promotores).

a) Cepa y vector de expresión

Se usó una cepa de tipo salvaje de *P. pastoris* (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) como cepa hospedante. La transformación de la cepa se realizó con un vector del laboratorio de los inventores denominado pPUZZLE (Stadlmayr *et al.*, J. Biotechnol., diciembre de 2010, 150(4):519-529), y la selección de los transformantes positivos se basó en la resistencia a la zeocina. Para la expresión secretora de la carboxipeptidasa porcina B (CpB) se usó el conductor del factor de apareamiento alfa de levaduras.

b) Amplificación y clonación de los promotores recién identificados pG1, pG3, pG4 y pG6 en un vector de expresión del laboratorio de los inventores

Dos promotores amplificados en el ejemplo 2b se clonaron en el vector de expresión pPUZZLE pPM1aZ30_aMF_CpB, produciendo pPM1aZ30_pG1_aMF_CpB y pPM1aZ30_pG6_aMF_CpB. Además, se usó como referencia el vector pPM1dZ30_pGAP_CPB, que contiene el promotor utilizado habitualmente de gliceraldehído 3-fosfato deshidrogenasa (pGAP). Los promotores se insertaron cadena arriba del codón de inicio del gen CpB usando los sitios de restricción Apal y SbfI. Se verificó que las secuencias de promotores eran correctas mediante secuenciación Sanger.

c) Expresión de CpB en *P. pastoris* bajo el control de los promotores inducidos por una limitación en glucosa recién identificados

Los plásmidos se linealizaron usando la enzima de restricción SpeI o SapI antes de la electroporación (empleando un protocolo de transformación convencional para *P. pastoris*) en *P. pastoris*. Se realizó la selección de los transformantes positivos en placas YPD (por litro: 10 g de extracto de levadura, 20 g de peptona, 20 g de glucosa, 20 g de agar-agar) que contenían 25 µg/ml de zeocina. Se empleó una PCR de colonias para asegurar la presencia del plásmido transformado como se describió en el ejemplo 2c.

Para la selección de la expresión de CpB, una única colonia se inoculó en medio YPG-Zeo líquido (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de zeocina) como precultivo. Después de aproximadamente 24 h, el precultivo se usó para inocular el cultivo principal con una DO600 de 1 en medio YP (por litro: 20 g de peptona, 10 g de extracto de levadura) y esferas de alimentación de glucosa (Kuhner, CH). Se lograron las condiciones de crecimiento con limitación de glucosa debido a la lenta cinética de liberación de glucosa de estas esferas de alimentación, que se describe mediante la siguiente ecuación: (glucosa) = 1,63*t^{0,74} [mg/disc]. Las muestras se tomaron al final del precultivo y 24 y 48 horas después de la inoculación del cultivo principal. Se determinó la concentración de biomasa midiendo la DO600 o el peso celular en húmedo. La concentración de CpB en el sobrenadante del cultivo se cuantificó mediante un ensayo enzimático, basado en la conversión de hipuril-L-arginina en ácido hipúrico por la CpB. La cinética de reacción se midió controlando la absorción a 254 nm a 25 °C usando un espectrofotómetro Hitachi U-2910 cuando se inicia la reacción. Las muestras y los patrones se tamponan con tampón de ensayo (Tris 25 mM, HCl 100 mM, pH 7,65) y se activaron usando tampón de activación (tripsina 0,01 mg/l, Tris 300 mM, ZnCl₂ 1 µM, pH 7,65). Se empleó tampón de activación sin tripsina en lugar de muestra como control negativo. La reacción comienza añadiendo la disolución de sustrato (hipuril-L-arginina 1 mM en tampón de ensayo).

d) Cultivo de alimentación discontinua de cepas de *P. pastoris* que expresan CpB bajo el control del promotor pG6

La fermentación de alimentación discontinua se realiza como se describió en el ejemplo 3e. El clon pPM1aZ10_pG6_CpB#4 no produjo CpB detectable en la fase discontinua y produjo más de 210 mg/l de CpB al final

de la alimentación discontinua.

Ejemplo 6: Estudios de actividad de promotor comparativos de los promotores recién identificados pG1 y pG6 en clones de múltiples copias de *P. pastoris* usando albúmina de suero humana (HSA) como gen indicador expresado de modo extracelular

5 Para analizar las propiedades de los promotores recién identificados bajo condiciones de limitación de glucosa, se realizaron selecciones en matraces de agitación como sigue: se realizó el precultivo durante 24 horas con medio rico que contenía glicerol como fuente de carbono, que simula la fase discontinua del proceso (estado reprimido de los promotores), seguido del cultivo principal con medio mínimo y esferas de alimentación de glucosa, que simula la fase de alimentación discontinua con limitación de glucosa del proceso (estado inducido de los promotores).

10 a) Cepa y vector de expresión

Se usó una cepa de tipo salvaje de *P. pastoris* (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) como cepa hospedante. La transformación de la cepa se realizó con un vector del laboratorio de los inventores denominado pPUZZLE (Stadlmayr *et al.*, J. Biotechnol., diciembre de 2010, 150(4):519-529), y la selección de los transformantes positivos se basó en la resistencia a la zeocina. Para la expresión secretora de la albúmina de suero humana (HSA) se usó su conductor de la secreción nativo.

b) Amplificación y clonación de los promotores recién identificados pG1 y pG6 en un vector de expresión del laboratorio de los inventores

Dos promotores amplificados en el ejemplo 2b se clonaron en el vector de expresión pPUZZLE pPM1nZ30_HSA, produciendo pPM1nZ30_pG1_HSA y pPM1 nZ30_pG6_HSA.

20 Los promotores se insertaron cadena arriba del codón de inicio del gen HSA usando los sitios de restricción Apal y SbfI. Se verificó que las secuencias de promotores eran correctas mediante secuenciación Sanger.

c) Expresión de HSA en *P. pastoris* bajo el control de los promotores inducidos por una limitación en glucosa recién identificados

25 Todos los plásmidos se linealizaron usando la enzima de restricción AscI antes de la electroporación (empleando un protocolo de transformación convencional para *P. pastoris*) en *P. pastoris*. Se realizó la selección de los transformantes positivos en placas YPD (por litro: 10 g de extracto de levadura, 20 g de peptona, 20 g de glucosa, 20 g de agar-agar) que contenían 25 µg/ml de zeocina. La amplificación del número de copias del gen se realizó como se describe en Marx *et al.* (FEMS Yeast Res., diciembre de 2009, 9(8):1260-1270). Se empleó una PCR de colonias para asegurar la presencia del plásmido transformado como se describió en el ejemplo 2c.

30 Para la selección de la expresión de HSA, una única colonia se inoculó en medio YPG-Zeo líquido (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de zeocina) como precultivo. Después de aproximadamente 24 h, el precultivo se usó para inocular el cultivo principal con una DO600 de 1 en medio YP (por litro: 20 g de peptona, 10 g de extracto de levadura) y esferas de alimentación de glucosa (Kuhner, CH). Se lograron las condiciones de crecimiento con limitación de glucosa debido a la lenta cinética de liberación de glucosa de estas esferas de alimentación, que se describe mediante la siguiente ecuación: (glucosa) = 1,63*t^{0,74} [mg/disc]. Las muestras se tomaron al final del precultivo y 24 y 48 horas después de la inoculación del cultivo principal. Se determinó la concentración de biomasa midiendo la DO600 o el peso celular en húmedo. La concentración de HSA en el sobrenadante del cultivo se cuantificó mediante el equipo de cuantificación con ELISA de albúmina humana (n.º de catálogo E80-129, Bethyl Laboratories, TX, EE.UU.) según el manual de instrucciones del suministrador. El patrón de HSA se usó con una concentración inicial de 400 ng ml⁻¹. Las muestras se diluyeron en consecuencia en diluyente de muestras (Tris-HCl 50 mM, NaCl 140 mM, BSA al 1% (en p/v), Tween20 al 0,05% (en v/v), pH 8,0). En la tabla 13 se presentan las titulaciones de HSA de la selección de varios clones de múltiples copias y de clones de una sola copia del ejemplo 3c.

45 Tabla 13. Resultados de la selección de clones de múltiples copias de *P. pastoris* que expresan HSA bajo el control de pGAP, pG1 y pG6

Clon	Titulación de HSA (mg/l)
pPM1aZ10_pG1_HSA#23	8,20
pPM1nZ30_pG1_HSA#C2	19,55
pPM1nZ30_pG1_HSA#4*1000	21,59
pPM1nZ30_pG1_HSA#5*1000	21,33
pPM1nZ30_pG1_HSA#X4	27,22
pPM1nZ30_pG1_HSA#X5	6,90

Clon	Titulación de HSA (mg/l)
pPM1nZ10_pG6_HSA#36	1,55
pPM1nZ30_pG6_HSA#C6	14,12
pPM1nZ30_pG6_HSA#2*1000	15,85
pPM1nZ30_pG6_HSA#X5	11,52
pPM1nZ30_pG6_HSA#X8	7,87

d) Determinación del número de copias del gen HSA

El aislamiento del ADN genómico y la medición con qPCR se realizaron como en el ejemplo 2d, usando los cebadores que aparecen en la tabla 9. Los resultados se muestran en la tabla 14.

5

Tabla 14. Resultados de la selección y el número de copias de genes de clones de múltiples copias de *P. pastoris* que expresan HSA bajo el control de pGAP, pG1 y pG6

Clon	GCN	HSA mg l ⁻¹ del cultivo principal	HSA mg l ⁻¹ por GCN del cultivo principal
pPM1nZ30_pG1_HSA#4*1000	11	21,59	1,89
pPM1nZ30_pG1_HSA#X4	17	27,22	1,64
pPM1nZ30_pG6_HSA#C6	12	14,12	1,23
pPM1nZ30_pG6_HSA#2*1000	6	15,85	2,50

e) Cultivo de alimentación discontinua de cepas de múltiples copias de *P. pastoris* que expresan HSA bajo el control del promotor pG1 y pG6

10

Las fermentaciones de alimentación discontinua se realizan como se describió en el ejemplo 3e. Los clones pPM1nZ30_pG1_HSA#4*1000 y pPM1nZ30_pG6_HSA#C6 alcanzaron 1060 y 728 mg/l de HSA al final de la alimentación discontinua, respectivamente.

Ejemplo 7: Estudios de actividad de promotor comparativos del promotor recién identificado en *P. pastoris* PG1 usando un fragmento de anticuerpo (Fab) como gen indicador expresado de modo extracelular

15

Para analizar las propiedades de los promotores recién identificados bajo condiciones de limitación de glucosa, se realizaron selecciones en matraces de agitación como sigue: se realizó el precultivo durante 24 horas con medio rico que contenía glicerol como fuente de carbono, que simula la fase discontinua del proceso (estado reprimido de los promotores), seguido del cultivo principal con medio mínimo y esferas de alimentación de glucosa, que simula la fase de alimentación discontinua con limitación de glucosa del proceso (estado inducido de los promotores).

a) Cepa y vector de expresión

20

Se usó una cepa de tipo salvaje de *P. pastoris* (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) como cepa hospedante. La transformación de la cepa se realizó con un vector del laboratorio de los inventores denominado pPUZZLE (Stadlmayr *et al.*, J. Biotechnol., diciembre de 2010, 150(4):519-529), y la selección de los transformantes positivos se basó en la resistencia a la zeocina. Para la expresión secretora de un fragmento Fab de anticuerpo, se usó el conductor del factor de apareamiento alfa de levaduras.

25

b) Amplificación y clonación del promotor recién identificado pG1 en un vector de expresión del laboratorio de los inventores

30

El promotor pG1 amplificado en el ejemplo 2b se clonó en el vector de expresión pPUZZLE que contiene Fab como GDI, según se describe en el ejemplo 5b. El promotor se inserta cadena arriba del codón de inicio del gen Fab usando los sitios de restricción ApaI y SbfI. Se verificó que la secuencia de promotor era correcta mediante secuenciación Sanger.

c) Expresión de Fab en *P. pastoris* bajo el control del promotor inducido por una limitación en glucosa recién identificado pG1

35

Los plásmidos se linealizaron usando la enzima de restricción SpeI o SapI antes de la electroporación (empleando un protocolo de transformación convencional para *P. pastoris*) en *P. pastoris*. Se realizó la selección de los transformantes positivos en placas YPD (por litro: 10 g de extracto de levadura, 20 g de peptona, 20 g de glucosa,

20 g de agar-agar) que contenían 25 µg/ml de zeocina. Se empleó una PCR de colonias para asegurar la presencia del plásmido transformado como se describió en el ejemplo 2c.

5 Para la selección de la expresión de Fab, una única colonia se inoculó en medio YPG-Zeo líquido (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de zeocina) como precultivo. Después de aproximadamente 24 h, el precultivo se usó para inocular el cultivo principal con una DO600 de 1 en medio YP (por litro: 20 g de peptona, 10 g de extracto de levadura) y esferas de alimentación de glucosa (Kuhner, CH). Se lograron las condiciones de crecimiento con limitación de glucosa debido a la lenta cinética de liberación de glucosa de estas esferas de alimentación, que se describe mediante la siguiente ecuación: (glucosa) = $1,63 \cdot t^{0,74}$ [mg/disc]. Las muestras se tomaron al final del precultivo y 24 y 48 horas después de la inoculación del cultivo principal. Se determinó la concentración de biomasa midiendo la DO600 o el peso celular en húmedo. Los niveles de expresión de Fab se cuantifican mediante ELISA usando anticuerpos anticadenas ligeras kappa humanas (unidos y libres)-fosfatasa alcalina producidos en cabras. En la tabla 15 se presentan las titulaciones de Fab de la selección de varios clones que expresan Fab bajo el control de pGAP y pG1.

Tabla 15. Resultados de la selección de clones de *P. pastoris* que expresan Fab bajo el control de pGAP y pG1

	Fab (mg/l)
pPM1dZ30_pGAP_Fab#2	0,00
pPM1dZ30_pGAP_Fab#5	0,70
pPM1dZ30_pGAP_Fab#7	0,68
pPM1aZ30_pG1_Fab#2	2,02
pPM1aZ30_pG1_Fab#3	0,70
pPM1aZ30_pG1_Fab#4	1,10
pPM1aZ30_pG1_Fab#5	0,00
pPM1aZ30_pG1_Fab#6	0,56
pPM1aZ30_pG1_Fab#9	0,66
pPM1aZ30_pG1_Fab#10	1,80
pPM1aZ30_pG1_Fab#11	1,64
pPM1aZ30_pG1_Fab#12	2,31
pPM1aZ30_pG1_Fab#13	2,35
pPM1aZ30_pG1_Fab#14	2,27
pPM1aZ30_pG1_Fab#15	1,60
pPM1aZ30_pG1_Fab#16	1,45
pPM1aZ30_pG1_Fab#B9	2,89
pPM1aZ30_pG1_Fab#B10	2,32
pPM1aZ30_pG1_Fab#B11	6,45
pPM1aZ30_pG1_Fab#B12	3,24
pPM1aZ30_pG1_Fab#B13	2,57
pPM1aZ30_pG1_Fab#B14	3,14
pPM1aZ30_pG1_Fab#B15	3,23
pPM1aZ30_pG1_Fab#B16	2,61
pPM1aZ30_pG1_Fab#C1	10,58
pPM1aZ30_pG1_Fab#C2	1,46
pPM1aZ30_pG1_Fab#C3	12,38
pPM1aZ30_pG1_Fab#C4	9,91
pPM1aZ30_pG1_Fab#C5	1,96

	Fab (mg/l)
pPM1aZ30_pG1_Fab#C6	2,87
pPM1aZ30_pG1_Fab#C7	7,03
pPM1aZ30_pG1_Fab#C8	6,37

d) Cultivo de alimentación discontinua de cepas de *P. pastoris* que expresan Fab bajo el control del promotor pG1

Las fermentaciones de alimentación discontinua se realizan como se describió en el ejemplo 3e, pero se emplea la alimentación discontinua de glucosa como se describe en el ejemplo 2e. Los clones pPM1aZ30_pG1_Fab#C4 y pPM1aZ30_pG1_Fab#C7 alcanzaron 165 y 131 mg/l de Fab al final de la alimentación discontinua, respectivamente.

5 Ejemplo 8: Fermentación de alimentación discontinua exponencial para controlar la tasa de crecimiento específica a la máxima productividad volumétrica de los promotores recién identificados

10 Se emplean cultivos en quimiostato de clones de *P. pastoris* que expresan un gen indicador bajo el control de los promotores recién identificados para determinar la productividad específica y volumétrica a diferentes tasas de crecimiento. Según se describe en Maurer *et al.* (Microb. Cell Fact., diciembre de 2006, 11, 5:37), pueden usarse fermentaciones de alimentación discontinua exponencial para cultivar un clon de *P. pastoris* a cierta tasa de crecimiento para mejorar la producción durante la fase de alimentación completa. Con ello puede optimizarse el rendimiento espacio-tiempo. Se aplicó una alimentación optimizada, y el rendimiento espacio-tiempo de la fase de alimentación discontinua mejoró en más del 35%.

15 Ejemplo 9: Determinación de la potencia de transcripción/promotor: estudio de actividad de promotor comparativo para identificar la regulación del promotor con diferentes concentraciones de glucosa usando eGFP como gen indicador expresado de modo intracelular

20 Las propiedades de regulación de un promotor se analizan mediante la selección de clones que expresan eGFP bajo el control de dicho promotor. Para ello, una única colonia se inoculó en medio YPG-Zeo líquido (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de zeocina) como precultivo. Después de aproximadamente 24 h, el precultivo se usó para inocular el cultivo principal con una DO600 de 0,01 en 10 ml de medio YP (por litro: 20 g de peptona, 10 g de extracto de levadura) y glucosa en diferentes concentraciones (20, 10, 5, 2,5, 1,25, 0,625, 0,313, 0,156, 0,078, 0,039, 0,020, 0,010, 0,005 y 0,002 g/l). Se toma una muestra después de 6 horas y se analiza mediante citometría de flujo como se describe en Stadlmayr *et al.* (J. Biotechnol., diciembre de 2010, 150(4):519-529). Se calcula la fluorescencia relacionada con el tamaño celular (dispersión directa hasta la potencia de 1,5) para cada célula/punto de datos, y se emplea su promedio geométrico para comparar los niveles de expresión de eGFP producidos en diferentes concentraciones de glucosa. Se emplea un clon que expresa eGFP bajo el control de pGAP como referencia (pGAP de *P. pastoris*, en la presente SEQ ID NO:25). Se midió la autofluorescencia de *P. pastoris* usando células de tipo salvaje de *P. pastoris* no transformadas y se restó de la señal. La tabla 16 muestra la inducción completa del promotor pG1 a aproximadamente 40 mg/l de glucosa o menos y su potencia de transcripción, comparado con el promotor pGAP nativo.

Tabla 16. Expresión de eGFP relativa (con relación a pGAP) de un clon de *P. pastoris* que expresa eGFP bajo el control del promotor pG1 en diferentes concentraciones de glucosa (20-0,002 g/l)

% de pGAP	Glucosa (g/l)
14,7	20
17,4	10
23,7	5
25,4	2,5
28,2	1,25
30,6	0,625
36,9	0,3125
44,5	0,15625
50,9	0,078125
56,2	0,0390625
55,0	0,0195313
57,5	0,0097656

% de pGAP	Glucosa (g/l)
59,2	0,0048828
59,6	0,0024414

Se realizaron estudios similares para comparar la potencia de transcripción relativa de los promotores desreprimidos pG1, pG3, pG4, pG6 y pG7. Un clon que expresa eGFP bajo el control de uno de los promotores se cultivó en YPG (glicerol 20 g/l, estado reprimido) y después se inoculó en medio YP que contenía diferentes cantidades de glucosa (de 20 a 0,002 g/l (D20, D10,...D0,002), estado inducido) y se cultivó durante 5-6 horas. Las células se analizaron mediante citometría de flujo y los resultados se evaluaron como sigue: la fluorescencia se relacionó con el tamaño celular (dispersión directa hasta la potencia de 1,5) para cada célula, y se usó su promedio geométrico para comparar las diferentes concentraciones de glucosa. El resultado concluyente de estas selecciones se muestra en la figura 14, un diagrama que muestra las concentraciones de glucosa logarítmicas frente la fluorescencia relativa, que ofrece una buena ilustración del comportamiento de inducción de promotores regulables por la limitación de glucosa. La figura 14 muestra la inducción completa del promotor pG1 a aproximadamente 40 mg/l de glucosa o menos, y de los promotores pG3, pG4 y pG6 a aproximadamente 4 g/l o menos, y su potencia de transcripción, comparado con el promotor pGAP nativo. El comportamiento de inducción de pG7 es similar a pG1 (los datos no se muestran). Basándose en los resultados previos con pG8, se supuso que su comportamiento de inducción está en el intervalo de los otros promotores.

Ejemplo 10: Comparación de los promotores de la técnica anterior pICL1 y pMLS1 con pG1 en el ensayo de selección de concentración de glucosa

El estudio de actividad de promotor comparativo se realiza según el ejemplo 9, empleando los promotores pICL1 y pMLS1 como referencia para la comparación con el promotor pG1 descrito en la presente.

Se descubrió que la actividad de ambos promotores pICL1 y pMLS1 era muy débil, sin diferencias significativas a una concentración de glucosa alta (D20: 20 g/l, represión) o baja (D0,04: 0,04 g/l, inducción = desrepresión). En cualquier caso, la actividad es mucho menor que la actividad del promotor pG1 reprimido en el mismo escenario. Los resultados se muestran en la tabla 17 como actividad de promotor en porcentaje con relación al promotor pGAP.

Tabla 17. Fluorescencia relativa de cepas que expresan eGFP bajo el control de pG1, pICL1 y pMLS1, respectivamente, cultivadas en medio que contiene glucosa 20 g/l (D20) o 0,04 g/l (D0,04)

	D20/represión		D0,04/inducción	
pG1#8	9,95	2,60	48,41	+/- 2,76
pICL1	2,68	+/- 1,78	5,07	+/- 0,90
pMLS1	-1,26	+/- 0,54	0,58	+/- 0,22

Ejemplo 11: Comparación de variantes de pG1

Se clonaron variantes más cortos del promotor pG1 como se describió en el ejemplo 2a y se seleccionaron de modo similar al descrito en el ejemplo 2c, pero en una configuración transformada a escala menor usando placas de 24 pocillos (Whatman, Reino Unido, artículo n.º 7701-5110) y cuartos de esferas de alimentación (12 mm, Kuhner, CH) en lugar de esferas completas. Se emplearon clones que se expresan bajo el control de pG1 y pGAP como controles. En la tabla 18 se listan los cebadores directos y las longitudes de pG1 y sus variantes. No se produjo una diferencia significativa en la fluorescencia relativa de las células que expresan eGFP bajo el control de pG1 y los variantes de pG1 a-f.

Tabla 18. pG1 y sus variantes: cebadores directos y posiciones de inicio 5' y de fin 3' en la secuencia de pG1 (SEQ ID NO:1). Secuencias de pG1a-f, véase la figura 15 (SEQ ID NO:41-46).

Promotor	Cebador	5'	3'	Longitud (pb)
pG1	GATAGGGCCCCAAACATTTGCTCCCCCTAGTCTC SEQ ID NO:34	36	1001	988
pG1 a	GATAGGGCCCCGGAATCTGTATTGTTAGAAAGAACGAGAG SEQ ID NO:35	143	1001	881
pG1 b	GATAGGGCCCCCATATTCAGTAGGTGTTTCTTGAC SEQ ID NO:36	338	1001	686

ES 2 730 173 T3

Promotor	Cebador	5'	3'	Longitud (pb)
pG1 c	GATAGGGCCCCTGCAGATAGACTTCAAGATCTCAGG SEQ ID NO:37	509	1001	515
pG1 d	GATAGGGCCCGACCCCGTTTTTCGTGACAAATT SEQ ID NO:38	632	1001	392
pG1 e	GATAGGGCCCCCGGATAAGAGAATTTTGTGGATTAT SEQ ID NO:39	674	1001	350
pG1 f	GATAGGGCCCGCCTGCTCCATATTTTCCGG SEQ ID NO:40	719	1001	305

Listado de secuencias

- <110> Lonza Ltd.
- <120> Promotor regulable
- <130> LO001P
- 5 <160> 46
- <170> PatentIn version 3.3
- <210> 1
- <211> 1001
- <212> ADN
- 10 <213> Pichia pastoris
- <400> 1

ES 2 730 173 T3

atttccaccc ccatcccagt agaatgtagg gtccccaac atttgctccc cctagtctcc 60
 agggaaatgt aaaatatact gctaatagaa aacagtaaga cgctcagttg tcaggataat 120
 tacgttcgac tgtagtaaaa caggaatctg tattgtaga aagaacgaga gttttttacg 180
 gcgcccgcatt attgggcccgt gtgaaaacag cttgaaaccc cactactttc aaagggtctg 240
 ttgctataca cgaacccatgt ttaaccaacc tcgcttttga cttgactgaa gtcacgcggtt 300
 aacaatcaag taccctagtc tgtctgaatg ctcttttcca tattcagtag gtgtttcttg 360
 cactttttgca tgcactgcgg aagaattagc caatagcgcg tttcatatgc gcttttacc 420
 cctcttttgt caagcgcaaa atgcctgtaa gatttggtgg ggggtgtgagc cgtagctga 480
 agtacaacag gctaattccc tgaaaaaact gcagatagac ttcaagatct cagggattcc 540
 cactatttgg tattctgata tgtttttcct gatatgcatc aaaactctaa tctaaaacct 600
 gaatctccgc tatttttttt ttttttttga tgaccccgtt ttcgtgacaa attaatttcc 660
 aacgggggtct tgtccggata agagaatttt gtttgattat ccggttcggat aaatggacgc 720
 ctgctccata tttttccggt tattaccca cctggaagtg cccagaattt tccggggatt 780
 acggataata cggtggtctg gattaattaa tacgccaagt cttacatttt gttgcagtct 840
 cgtgcgagta tgtgcaataa taaacaagat gagccaattt attggattag ttgcagcttg 900
 accccgccat agctaggcat agccaagtgc tatgggtggt agatgatgca cttggatgca 960
 gtgagttttg gagtataaaa gatccttaaa attccaccct t 1001

<210> 2
 <211> 1000
 <212> ADN
 <213> Pichia pastoris

5

<400> 2

gtaaatagcg gcagcaatcc agtaaccttt tctgaatagc agagccttaa ctaaaataat 60
 ggccagggtgta aaaaattcga aatttgacac caaaaataaa gacttgtcgt tataagtctt 120

ES 2 730 173 T3

aacaaagtcc gcaatthttgg agctaacggt ggcggttgct gggatattca ataatggtag 180
aatgttgctg cgggtatatg acagagcgtg aaacacactg aacaaggtaa atggaacaac 240
agcaattgca atatggggga ggatagtcaa gaacaaagca gcaatggcaa agtactgaat 300
attctccaaa gccaaaaggt ccagtggttt caacgacaaa gtcttggttg tatagctttg 360
gaacaaaagg acaccgaaag actcgacagc gccacaaaat acagcgttgt agaagaacga 420
attgattgct ccagagcttc taatagtcag aagatacccc aaacctccga gcaacgtag 480
cacatgacct aagaaccagg cgaagtgaag agtctggaat aacgacacct agtcagtttt 540
tcctgagctc ctggtgggat tggtagaagc atttgatttg cttggagtgg ttttatttga 600
agatggtggt gaagccattg ttgctaaaga gtcggagttt tgcttttagg gtttgtaag 660
caaaggagga aaaactgctc cgtttgaagt cccaggtagt ttcgctgtg aggccagcca 720
gggaaagctt ccttcggtac tttttttct tttgcaggtt ccggacggat taagcttcgg 780
gttatgaggg gggcggtagc caattccgga cacaatattg cgtcgcagct agtcacccccg 840
ccataaatat acgcaggatt gaggtataa catcgatagt cttagtaatt aatacaattc 900
agtggcgaat ttggcaacat gacgtaaggc cactgttgt ctataaaagg ggatgaattt 960
tcatgttttt gaggcctccc ggacaattta ttgaactcaa 1000

<210> 3
<211> 1001
<212> ADN
5 <213> Pichia pastoris

<400> 3
agaccagcag ttttaactacg caaatccaca ggaatttcta catcacaata ccaatggtaa 60
taccacgacg tcaaggaatg gaaacgacga cttggaggaa gacttcgtca acctcttgcg 120
gagtaccoga ggctaagaca ataagaagaa aaaaaaaga aaagcgggtg gggagggatt 180
attaataaag gattatgtaa ccccagggta ccgttctata catatttaag gattatttag 240
gacaatcgat gaaatcggca tcaaactgga tgggagtata gtgtccggat aatcggataa 300
atcatcttgc gaggagccgc ttggttggtt ggtgagagga gtgaaatatg tgtctcctca 360
cccaagaatc gcgatatcag caccctgtgg gggacactat tggcctccct cccaaacctt 420
cgatgtggtg gtgctttatt atattgatta cattgattac atagctaac cctgcctggt 480
tgcaagttga gctccgaatt ccaatattag taaaatgcct gcaagataac ctcggtatgg 540
cgtccgaccc cgcttaatta ttttaactcc tttccaacga ggacttcgta atttttgatt 600
agggagttga gaaacggggg gtcttgatac ctctcgtatt tcagatccca ccccctctca 660
gtcccaagtg ggacccccct cggccgtgaa atgcgcgcac tttagttttt ttcgcatgta 720
aacgcccgtg tccgtcaatt aaaagtcgca gactaggggtg aactttacca tttttgtcgc 780

ES 2 730 173 T3

actccgtctc ctcggaatag ggggtgtagta attctgcagt agtgcaatth ttaccccgcc 840
 aaggggggggc gaaaagagac gacctcatca cgcattctcc agtcgctctc tacgcctaca 900
 gcaccgacgt agttaactth ctcccatata taaagcaatt gccattcccc tgaaaactth 960
 aacctctgct ttttcttgat ttttccttgc ccaaagaaaa g 1001

<210> 4
 <211> 1000
 <212> ADN
 <213> Pichia pastoris

5

<400> 4
 tggactgttc aatttgaagt c gatgctgac gatgtcaaga gagatgctca attatatttg 60
 tcatttgctg gttacactgg aaacgctact tttgttgccg gaaactctac cagtttgcc 120
 gtccatgtaa acgatgtcgt tctgggccgt gaccgttca acacgaacat aaccaatgac 180
 aaatccactt acaggtctag ttcatatgga ggcaattggc accttacttc tttggatgtc 240
 ccaagtgggg ctttaacgtc tggtaactaac aatgtctcgt ttgtcactac aaactccgag 300
 gtaaataaag gattcttgct ggattctctc aagtttgtht ggaagttgta acaggttht 360
 aagcatatcg tgcgcttgct cacaattgaa tcatttattg ttgogagata catgaacaaa 420
 gtgtgaactg ggaccatta ctacaattcc cacgcaaccg ttgtttcaaa gcccatattt 480
 tttgacaatt gtttcgttac acccccagtt tgatgtacat cgcttgcaat gatgtgtgct 540
 ccggagtatt ttccatattc agcttgaatt cgtatactca accaatatct ggggggtatac 600
 ttttatgtaa cctatacaaa tcaactatac tatttcacct ttcgaccaat catctcccat 660
 cttgttaagt tttgcttctt atatccctga ccctgacatc acccatgatt ccgctcaacg 720
 gttctcctct acatcgtccc tcttttgag aggggttca gtttgacatt caaattacc 780
 cccgccatca cgcgcaaccg agaccgcacc cccgaattht cacaaattac cccacaccct 840
 atactccacc actatgaggg ttattagaac tgatcacgta taaataccac cgcaagttcc 900
 caagggatcg tgttcttctt ctccaattgc aatcatattt ctgactctth ctagttcaga 960
 ttaattcctt tacacttgct tttttccctt acctttatcc 1000

<210> 5
 <211> 1000
 <212> ADN
 <213> Pichia pastoris

10

<400> 5
 aattgattaa gttcagtgaa atttcaaacc gctatacaca acaggacaac tttgagttha 60
 gaaaaatccg atgtagtgta acggctagca cggctccgctt tcaccgggca gaccgggtt 120
 cgactcccgg catcggagtt tatttttcca tttcgttctt tagagtattc tcctcagcat 180
 gccccctga atttttcctt ttttccatgt gtcccattth tccactthtt ttacagttth 240

ES 2 730 173 T3

cctcgtgatg ttaattggct acacaaaagc tgccacacga aaccttaatc acgaaaaact 300
 atacagcctt cactaatccg tagccccata atatgttgtc cacgtgctgt tgggtactac 360
 ctgtagactc tcatacccca ctccgtcttt ctccaacaat taacgcagta ccgagattta 420
 tcagcagact caaattgggc aaactctgta tttttccttg cccgcataat ttatgggtct 480
 caggcctcca cgtttcctgt ttacttgaag aatattggct gcgaaaaag tggtaaggac 540
 aacccccttt taattggatc cagtttttcc gaaatgttcc gatccgtacg tcatctccga 600
 agccgtacat tttcactcaa tctacgtagc tttggactca gcgctcctgg aattgccagg 660
 acagtttact tgagttgata ttcccttgta gattgtgtgc ttctttttcc aaaatttgag 720
 gcttcgtttg aaaagtggaa tctggtcgct agatcacttc atgcctatth ttcacggaaa 780
 aataagtggc actatgcacc ccttaaacct aaagaaaaac ggaaaaatta ccccaaaacc 840
 tggatgatgtt tttcgcccct ttctttttat ccgagttttt cttttttctt gtctgccaaa 900
 ttctctcct gaccttagcg tccccgaaa aaattaacta cttaaggacc gaatgagccc 960
 cagcttttcc ctttctcttc attattcccc ataataaat 1000

<210> 6
 <211> 1000
 <212> ADN
 <213> Pichia pastoris

5

<400> 6
 ctgcacaacc attgccagta aggacgaaga gaaggcccca ctacccaaaa ttcaggataa 60
 cgtcttcata ccatgcagcg acgcctacaa gacgctgtca agacatgcca acttcaacga 120
 agtgaacttt aacacattga tcgggaaatt gaccaccaag ggaatgctgg ttgaggctgg 180
 aagcgttgcc agtgtcctga gggaaactgga ccgaaagttt agtaatgcat aagaggatat 240
 atataggaat gcagtaataa tattagtacc cattaagtgg gctaagccat tggaggccg 300
 tctgactgat ggtggtgttc ttctcattta gatagtgcac ttgcaactac cgtctgagat 360
 tgagtttgat gtgaagctcc agcgccaaaa cagtataaga accttatctc cgcattattg 420
 ttcttgcgta aaagtttgtg tgaagaaaca ggggtagttg cgcagattag ttgtaatatg 480
 cgcataaggat gggtcattga cttctttcct cgaaagagcc acaccgtag ctaaaaaagg 540
 acgcgcatct accccaaaat agaatgtggg gaaataggac gcgcaacttc ctctcaatca 600
 ctggacgtca gaaaaacaaa tgcgcaatcg agtcaccctc cgtgatacc tccgtgatac 660
 cccctctccg tctattctga cagcgtctcc ccatgacggt tcaatctact tagaaaagat 720
 ttcgtttttt tttccttcaa ttacacgatc tcatcttctg caagggctctg gaggacatca 780
 ccaatctgcg actccataac ttagtcctga gtttatattt acgcttcatc tgatgagtag 840
 gaagaaaaag tttcacgaaa ttcccccgcc aacttgcctt tcggaataag cagccactct 900

ES 2 730 173 T3

ccttctgccc atagtaagct tgcgcgaggc cccaacttgg ccagaaactt taaatatgcc 960
aaacaatctc cccaatcta agttctccct cttctaaaaa 1000
<210> 7
<211> 1662
<212> ADN
5 <213> Pichia pastoris
<400> 7
atgtcctcgt tttttctaaa caaccaaaaca gtaaagatga tgacgccttt gggaagagct 60
agagctctaa attttcaggg caaagtttat gataaatctc caaaaactta caatatctac 120
gctattgcaa taacagccac cgtttctgga ctgatgttcg gttttgatat ttcttctgtg 180
tcctcgttcg taagtcagga tcattacaga aactacttca accgtcccga cagtttgacg 240
caagggggta tcaccgcaag tatggctgga ggttctttct tggggttcgtt attttcttct 300
gacttcagg atattcttgg aagaagagtt gctctgcata tgtgcagtgt cctctggatt 360
atcggggcca ttcttcaatg cgctgcacaa aaccaaggta tgctgatcgc agggagattg 420
atctccggta tcgggtgctgg gtttggttca gcttcagctc cagtctattg ttctgaagtt 480
gctccagcaa agattagagg aatgattgga ggattatctc aattttctgt cactgtgggt 540
atcatgataa tgttttatat cggatatgga tgtcactaca ttgacggcgt tgcattcattt 600
agactggcct gggggttgca aatggttcca ggtcttattc ttttggtcgg tgtattcttc 660
cttctgagt ctccaagatg gctggctaac cacaaccgct gggaagacgc agttgaggtt 720
attgctaattg ttggtgcaaa aggtgacaga gaaaacgccg atgtgcgtct gcaattggat 780
gaagttcagg agcaactatt gattgacaaa gatgcttctg attttggtta ccttgatttg 840
tttaagaaag attgtatcaa acgtaccttc attggagtgt cagctcaagt gtggcaacaa 900
ctttgtggta ttaatggtgc aatgtactac gttgtgtatc tcttccaaat ggctggtttt 960
actggaaatg tggcgttggg atcgtcctca attcaatatg ttttgaatgt tgttatgact 1020
gttccagctt tgtttctaata ggaccgtata ggcagacgac ccctactaat tgggtggtgg 1080
attttcatgt gtatttggct gtttggagtg gcaggattat taggcactta ctctgaacca 1140
attgaaaatt tcagcgggta tgatactgtc agaattacta ttcctgacca gcacaaggct 1200
gcagcaaggg gtgttattgc ctgttcctat ctattcgtgt gctcctttgc tccaacctgg 1260
ggtatctgca tttgggttta tgcctctgaa attttcaaca acagacaaag agcaaagggg 1320
gcagcatttg ctgcctccgc taactggatt ttcaactttg ccttggctat gttcgtgcca 1380
tcagccttta gaaacattac atggaagact tacatcattt ttggagtatt ttcgttctgc 1440
ttaacaatcc atgttttctt acaattccca gaaaccagag gtaagacttt ggaagaaatt 1500
gatcaaatgt ttaaggacaa tattccagct tggagaagtg cttcgtacgt tccagatatg 1560

ES 2 730 173 T3

ccaattttca acaaagagaa ggtagtatct actgagcatg cagaaaatgc ttccagctcg 1620

tccgaaaaag ccttgatggt tcaggaagag gaatctgtat aa 1662

<210> 8
 <211> 987
 <212> ADN
 5 <213> Pichia pastoris

<400> 8

atggtagttg caatcgaagg tggtagcaggc ttaggcctta tgaatcttac ttggaaacca 60

actccaaccc caattgatga tgcaattgag acaattagat atgctggtga ggaagctggt 120

gtcagatact tgaacggagg agagttctac aactttcctc ttgattcaaa cctgaatttg 180

cagtacattc aggaatttgc aaaaaggtac cccgagctat ataaaaaggt gagtctgtcg 240

gtaaaaggtg ctgtcagttt ggtcgatgtg agccccgatt cttccccgga gaaccttgaa 300

aaatcgattt caaacataac caaacatttg ccgaacaact tcctgccaat ttttgagcct 360

gctagaatcg ataaacgta ctccattgag gagacaataa agaatctctc taagtctgctc 420

gaagatggca gaattggagg tatttcactt agtgaagttg gtgctgacac tatcagaaga 480

gctgcgaaag tggctcccat cgcctgtgtg gaagtggagt tttctctatt gactagagat 540

attcttcata atggagttct tgctgcttgt gaggatttga acattcctat tattgcctac 600

agtcccttgg gaagaggatt tttgactgga acgataaaca gcaaagctga cattcctgaa 660

ggtgatataca ggttaagttt ggaaagattc aatgacgatg aagttattga acacaatttg 720

aaacttgttc acggtttgaa aaagatagcc gacaaaaaag gagtacatt ggctcaattg 780

tctcttgctg ggttacgaaa gtttgagat aaacacgtca aggtgcttcc tattccaagc 840

tgctcatctc ctcgtagagt tgcagaaaac acaaaagaga tttccttgac tgatagcagag 900

ttccaggaga ttactgactt tgcagagtcg gttccaatca aaggtggtcg ttacaacaaa 960

gcaagtgagg ctgttcttaa cggttag 987

<210> 9
 <211> 1506
 <212> ADN
 10 <213> Pichia pastoris

<400> 9

atgacatttg ctctccctt agaattcgag attgaccttc ctaacggatt gaagtacact 60

caaccattgg gactcttcat caacaatgag tttggtgaag gtgtagaggg aaagctctta 120

ccagtgatca atccttgatga tgagactaaa ataaccaag tttgggaagc ttctgcagcg 180

gatggtgacc gtgctggtga tgccgctgaa gatgctttca acaactccgt atgggctact 240

caggacccat tagagagggg aaagctgatg aacaaattgg cagaccttat cgatcgtgac 300

ES 2 730 173 T3

ttcaacatct tggctggtat cgaatccatc gacaatggta aggcctatac ctctgccag 360
 ggtgatgta ctcttgctgt caactacatc agatcctgtg ctggatgggc cgacaagatt 420
 ttgggaaacg ttggtgattc cggaaacacc caccttaact tggttaaaag agagccattg 480
 ggtggttggg gacaaattat cccatggaac tttcctctcc tgatgttggc ttggaagttg 540
 ggacctgccc tggccacagg taacactggt gttttgaaga ctgccgagtc taccocctctg 600
 tggggtttat acgttgccaa attgatcaag gaggccggtt tcccacctgg tgtggttaac 660
 attctcagtg gtttcggtaa cccagctgga gctgccatcg ctgctcatcc cagaatcaag 720
 aagattgctt tcaccggatc cactgcaaca ggccgtaaga tcatggaagc agccgctaaa 780
 tctaacctga aaaaagtcac tttggaacta ggtggtaaat ctccaaacat tgtgtttgaa 840
 gatgctgata tccagaagac tatccataac attatthttgg gaatcttctt caattctggt 900
 gaagtctggt gtgcaggttc cagagtctac attcaagaca ctgtgtatga agaagtgctt 960
 gaagccttca agaaggagac tgataacggt aaggttggtg gaccattcga agaaggtgtc 1020
 ttccaagggc ctccagacctc tgagttgcaa cttaacagaa tccttagtta catcaaacac 1080
 ggtaaggatg aaggtgctcg tgtaattacc ggtggttcaa gataccgtaa ccgaggttac 1140
 tacattaagc ccacaatttt tgctgacggt actgaagaca tgaagattgt caaggaggag 1200
 atthtttggtc ctgtggttac taccactaag ttctctaccg tggatgaggt tgttgatata 1260
 gccacaaca ccaactatgg tctagctgct ggtattcaca caaacaactt gaacaaagcc 1320
 attgatggtt ccagtagaat caaggcgggt gtcgthttgga ttaacaccta caacgatttc 1380
 caccacatgg ttcctthtcgg aggttatgga gaatctggtt ttggcagaga gcttggtgct 1440
 gaggctthtg ataactacac tcaagccaag gctatcagaa ttgcttacac tcctgaacat 1500
 aagtag 1506

<210> 10
 <211> 1578
 <212> ADN
 <213> Pichia pastoris

5

<400> 10
 atgcttagaa cttctccagc tactaagaaa gctctcaagt cgcagattaa cgccttcaac 60
 gttgctgcct tgagattcta ctctcattg cctthtcagc ttccaattac cttgccaaac 120
 ggtaagacct acaatcagcc aacaggttht tttatcaaca atgagthtcgt tcctthctaag 180
 caaggtaaga cctthtcgct tthaaacct tccactgagg aggagattac tcactgctac 240
 gagtccagag aggacgacgt tgagthtagc gthtcagccg ctcaaaaggc tthtcgactca 300
 acctggtcca cccaggacct tgctgagaga ggtaaggtct tgaacaagtt ggctgacctg 360
 atcgaggagc actctgagac cctthtcgcc atcgagthct tggacaacgg taaggccatt 420

ES 2 730 173 T3

tcctccgcta gaggtgatgt tggctctggtt gtcgcctact tgaagtcctg tgccggttgg 480
 gccgacaagg ttttcggtag agttggtgaa accggaagct cccacttcaa ctacgtaga 540
 agagagccat tgggtgtttg tggctcagatt atcccatgga actttcctct tctgatgtgg 600
 tcctggaaag ttggtccagc tttggccact ggtaaactg ttgtcctgaa gacagccgag 660
 tctactcctc tgtccgcctt gtacgtttcc caattggtca aggaggccgg tatcccagct 720
 ggtgtccaca acattgtgtc cggtttcggt aagattactg gtgaagctat tgctactcat 780
 cctaagatca agaaggttgc cttcactggt tctaccgcca ctggtcgtca catcatgaag 840
 gctgctgccg aatccaactt gaagaaggtt actttggagt tgggtggtaa atctcctaac 900
 atcgtgttca acgatgctaa cattaagcaa gctgtcgcca acatcatcct cggattttac 960
 tacaactctg gagaagtttg ttgtgctggt tccagagttt atgttcaatc cggattttac 1020
 gacgagcttt tggccgaatt caagactgct gctgagaatg tcaaggttgg taaccattc 1080
 gacgaggaca ccttccaagg tgctcaaacc tctcagcaac aattggagaa gattttgggt 1140
 ttcggtgagc gtggaagaa ggacggtgct actttgatta ctggtggtgg cagattaggt 1200
 gacaaggggtt acttcgtcca gccaaactatc ttcggtgatg ttacaccaga gatggagatt 1260
 gtcaaggaag agatctttgg tcctgttgtc actatcagca agtttgacac cattgatgag 1320
 gttgtcgacc ttgctaacga ctctcaatac ggtcttgctg ctggtatcca ctctgacgat 1380
 atcaacaagg tcattgacgt tgctgctaga atcaagtccg gtaccgtgtg ggtcaacacc 1440
 tacaacgatt tccaccaa at ggttccattc ggtggatttg gccaatccgg tattggtcgt 1500
 gagatgggtg ttgaagcttt ggaaaactac acccaataca aggctatccg tgtcaagatc 1560
 aaccacaaga acgagtaa 1578

<210> 11
 <211> 1614
 <212> ADN
 <213> Pichia pastoris

5

<400> 11
 atgagttcaa cagatatcca aggtgatcaa ggtgacaatg aaaagatata cgccattgag 60
 agcagtcctt ccaatgagca aataaaagat attcatgagg ctccggccga caacaaaagt 120
 gaactagaca tcccagtcaa acccaaggggt tcctatatct tgggtgtctgt gttatgtctt 180
 ctagtcgcat tcggtggttt cgtgttcggt tgggataccg gtaccatctc aggtttcgtt 240
 aacatgtctg actttacgag acgtttcggga cagtttaacg gtgaaacgta ttacctttct 300
 aaagtgagag ttggtttaat tgtttctatt ttcaacattg gttgtgctat cggaggtgtc 360
 actctaggta aacttggtga catttgggggt agaaagaagg ctttgatggt cgtcatggtc 420
 atctatatgg tcggtatfff gattcaaatt gcttccattg acaaatggta ccagtatctc 480

ES 2 730 173 T3

attggaagaa ttattgcagg tctggccgtc ggtgcagttt ccgttttatc ccccatgttc 540
atcagtgaga cttctcctaa acacatcaga ggttccttag tctcctgcta ccaattaatg 600
attacagccg gtatthttctt gggttactgt accacttacg gaaccaagac ttacaccgac 660
tccaccaat ggagagttcc tttgggattg tgtttcgctt gggccattct gatgattgtt 720
ggtatgacct tcatgccaga gtccccacgt ttcttggttg aggttaacag agtcgacgag 780
gctatgaagt ccattgccag agttaacaag gtctctatcg acgatccatc tgtctacaat 840
gagatgagac ttatthctga cggattgag aaggagaagg aggctggtag cgtthcttgg 900
ggtgaactgt tcaactgtaa gccaaagatt ttctaccgtc tattgattgg taththcatg 960
caatctthgc aacaattgac cggtaacaac taththctct actacggaac taccaththc 1020
aaggctgtcg gattggacga ttctthccaa actthctatca ttcttggtgt tgtcaathth 1080
gctthccacat tcttaggtat ctacaccatg gataaathth gtagaagaag aacactthta 1140
ggaggtthctg gagccatggt tgtthgtthg gtcaththca gthccgthtg tgtcaagtct 1200
ctthtatgaga acggtaagga tgatccatcc aaaccagcag gtaacgccat gattgtcttc 1260
acctgtctgt tcaaththctt ctthgcatgt acctgggctc cagggtthth cgtcgthtg 1320
tctgaaacct accacttag aattagatcc aagggtatgg ccatcgctca aggtthccat 1380
tggtthggg gththctcat tgctthctc actccathth tctcaggtgc cattgaththc 1440
gcctacggtt acgtctthtat gggatgtact ctgttcgctt tctthththgt gtactthctc 1500
gthctgaaa ccaagggtct gtcgctggaa gacgttgatg aagtctatga gaaccttacc 1560
thcggaagag catatgcata cagccacacg athaaagaca agggcgccct ataa 1614

<210> 12
<211> 1032
<212> ADN
5 <213> Pichia pastoris

<400> 12
atggccctat ctctaccta tcagggctac atatctacca ctggcgacgc gttgatcgtg 60
atccaggcag ctctaaataa ccatttgaat cttcttccc gaagaccaag agaaagagag 120
cgagatgggc taatacgatc aggtaacgta tttgtththg tcgagcaacg gtctcatatc 180
aaacgatgga ccgatggtat cccctggtct ccatctagag tccttgaaa gthctthgtg 240
tatcgggaac tggacaagga taccctaaa aactcgcaa gtgacgaaga tactgaggag 300
gggagaaaaga ggcgaaagac thctgtggat gtaaccgatc caataccag gcagthgtg 360
ggatcattgg tgactthctta tgactthcaa gaggatggac thattaahhaa aacactctcc 420
thgactthcc agaccggtgc taatgaagaa agggaaacag tgcactthgat tagthattat 480
actccggaag atgtaacgaa ccatcgththg aacaggccgt ctgacaatcc atactggcc 540

ES 2 730 173 T3

aatatcactg tttcagagtc attattgact gccttgagag agagtaccct tggaggaaga 600
 gcaacgtctg atgacgagct ttcttttagtc agaagtaact cgttagagta ccaagaggta 660
 ccaatgaaca tatctatgtc tttaccttta tcaactccac tttccttgaa cacaggagta 720
 aactcaacta cccagctgca acagcaacaa ctacaacaac aacaacagca acagcaacag 780
 cagcagcaac aacagcagca acaacagcaa ccggtagcat cccttccaaa atttgatgga 840
 tcctttctat tacaacaggg tgtaattcca gttcctcatt tcatggacca aaaaatggga 900
 agtagcaatt cgtggattaa caattggttt cgtccaaatt cgtcagaatc aatggggcta 960
 tcggttatcg gacctcaciaa gggatatgac gaacaaagtc cagcaacgag ttatactttg 1020
 aatgaacgtt ga 1032
 <210> 13
 <211> 491
 <212> ADN
 5 <213> Pichia pastoris
 <400> 13
 cttttttgta gaaatgtctt ggtgtcctcg tccaatcagg tagccatctc tgaatatct 60
 ggctccggtg caactccgaa cgacctgctg gcaacgtaaa attctccggg gtaaaactta 120
 aatgtggagt aatggaacca gaaacgtctc ttcccttctc tctccttcca ccgcccgtta 180
 ccgtccctag gaaattttac tctgctggag agcttcttct acggccccct tgcagcaatg 240
 ctcttcccag cattacggtg cgggtaaaac ggaggtcgtg taccogacct agcagcccag 300
 ggatggaaaa gtcccggccg tcgctggcaa taatagcggg cggacgcatg tcatgagatt 360
 attggaaacc accagaatcg aatataaaag gcgaacacct ttcccaattt tggtttctcc 420
 tgacccaaag actttaaatt taatttattt gtccctattt caatcaattg aacaactatc 480
 acctgcaggc c 491
 <210> 14
 <211> 34
 10 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 <400> 14
 15 gataggccc caaacattg ctccccctag tctc 34
 <210> 15
 <211> 39
 <212> ADN
 <213> artificial
 20 <220>
 <223> Cebador
 <400> 15
 gatactgca ggaagggtg aatttaagg atctttat 39

ES 2 730 173 T3

<210> 16
 <211> 36
 <212> ADN
 <213> artificial
 5 <220>
 <223> Cebador
 <400> 16
 gatagggccc cagcaatcca gtaaccttt ctgaat 36
 10 <210> 17
 <211> 36
 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 15 <400> 17
 gatactgca gggtgagtc aataaattgt cggga 36
 <210> 18
 <211> 35
 <212> ADN
 20 <213> artificial
 <220>
 <223> Cebador
 <400> 18
 gatagggccc tggactgttc aattgaagt cgatg 35
 25 <210> 19
 <211> 37
 <212> ADN
 <213> artificial
 <220>
 30 <223> Cebador
 <400> 19
 gatactgca ggggataaag gtaaggaaa aaagcaa 37
 35 <210> 20
 <211> 36
 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 <400> 20
 40 gatagggccc agaccagcag ttaactacg caaatc 36
 <210> 21
 <211> 36
 <212> ADN
 <213> artificial
 45 <220>
 <223> Cebador
 <400> 21

gatacctgca ggcttttctt tgggcaagga aaaatc 36
 <210> 22
 <211> 39
 <212> ADN
 5 <213> artificial
 <220>
 <223> Cebador
 <400> 22

gatagggccc aattgattaa gttcagttaa attcaaac 39
 10 <210> 23
 <211> 42
 <212> ADN
 <213> artificial
 <220>
 15 <223> Cebador
 <400> 23

gatacctgca ggattatatt atggggaata atgaagagaa gg 42
 20 <210> 24
 <211> 33
 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 <400> 24

gatagggccc ctgcacaacc attgccagta agg 33
 25 <210> 25
 <211> 40
 <212> ADN
 <213> artificial
 30 <220>
 <223> Cebador
 <400> 25

gatacctgca ggtttttaga agaggagaa ctagattgg 40
 35 <210> 26
 <211> 24
 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 40 <400> 26

cctgaggctt tgtccaccc atct 24
 45 <210> 27
 <211> 30
 <212> ADN
 <213> artificial
 <220>
 <223> Cebador

ES 2 730 173 T3

<400> 27
 ggaacatagt agtaccaccg gacataacga 30
 <210> 28
 <211> 22
 5 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 <400> 28
 10 tcgccgacca ctaccagcag aa 22
 <210> 29
 <211> 23
 <212> ADN
 <213> artificial
 15 <220>
 <223> Cebador
 <400> 29
 accatgtgat cgcgcttctc gtt 23
 <210> 30
 20 <211> 24
 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 25 <400> 30
 cctgaggctt tgtccaccc atct 24
 <210> 31
 <211> 30
 <212> ADN
 30 <213> artificial
 <220>
 <223> Cebador
 <400> 31
 ggaacatagt agtaccaccg gacataacga 30
 35 <210> 32
 <211> 28
 <212> ADN
 <213> artificial
 <220>
 40 <223> Cebador
 <400> 32
 aaacctagga aaagtgggca gcaaatgt 28
 <210> 33
 <211> 29
 45 <212> ADN
 <213> artificial
 <220>

<223> Cebador
 <400> 33
 actctgtcac ttactggcgt ttctcatg 29
 <210> 34
 5 <211> 34
 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 10 <400> 34
 gatagggccc caaacatttg ctcccctag tctc 34
 <210> 35
 <211> 39
 <212> ADN
 15 <213> artificial
 <220>
 <223> Cebador
 <400> 35
 gatagggccc ggaatctgta ttgtagaaa gaacgagag 39
 20 <210> 36
 <211> 36
 <212> ADN
 <213> artificial
 <220>
 25 <223> Cebador
 <400> 36
 gatagggccc ccatattcag tagtgtttc tgcac 36
 <210> 37
 <211> 36
 30 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 <400> 37
 35 gatagggccc ctgcagatag actcaagat ctcagg 36
 <210> 38
 <211> 32
 <212> ADN
 <213> artificial
 40 <220>
 <223> Cebador
 <400> 38
 gatagggccc gaccccgttt tcgtgacaaa tt 32
 45 <210> 39
 <211> 37
 <212> ADN
 <213> artificial

ES 2 730 173 T3

<220>
 <223> Cebador
 <400> 39
 gatagggcc cccgataaga gaatttggt tgattat 37

5 <210> 40
 <211> 31
 <212> ADN
 <213> artificial

10 <220>
 <223> Cebador
 <400> 40
 gatagggcc gcctgctcca tattttccg g 31

15 <210> 41
 <211> 859
 <212> ADN
 <213> artificial

<220>
 <223> variante de promotor
 <400> 41

20 ggaatctgta ttgtagaaa gaacgagagt tttttacggc gccgccatat tgggccgtgt 60
 gaaaacagct tgaacccca ctactttcaa aggttctggt gctatacacg aaccatgttt 120
 aaccaacctc gcttttgact tgactgaagt catcggttaa caatcaagta ccctagtctg 180
 tctgaatgct cctttccata ttcagtaggt gtttcttgca cttttgcatg cactgcggaa 240
 gaattagcca atagcgcgtt tcatatgcmc ttttaccccc tcttttgtca agcgcaaaat 300
 gcctgtaaga tttggtgggg gtgtgagccg ttagctgaag tacaacaggc taattccctg 360
 aaaaaactgc agatagactt caagatctca gggattccca ctatttggtt ttctgatatg 420
 tttttcctga tatgcatcaa aactctaate taaaacctga atctccgcta tttttttttt 480
 ttttttgatg accccgtttt cgtgacaaat taatttccaa cgggggtcttg tccggataag 540
 agaattttgt ttgattatcc gttcggataa atggacgcct gctccatatt tttccggtta 600
 ttaccaccacc tggaagtgcc cagaattttc cggggattac ggataatacg gtggtctgga 660
 ttaattaata cgccaagtct tacattttgt tgcagtctcg tgcgagtatg tgcaataata 720
 aacaagatga gccaatat tggattagtt gcagcttgac cccgccatag ctaggcatag 780
 ccaagtgcta tgggtgtag atgatgcact tggatgcagt gagttttgga gtataaaaaga 840
 tccttaaaat tccaccctt 859

<210> 42
 <211> 664
 <212> ADN
 <213> artificial

25 <220>
 <223> variante de promotor
 <400> 42

ES 2 730 173 T3

ccatattcag taggtgtttc ttgcactttt gcatgactg cggaagaatt agccaatagc 60
 gcgtttcata tgcgctttta ccccctcttt tgtcaagcgc aaaatgcctg taagatttgg 120
 tgggggtgtg agccgtagc tgaagtacaa caggctaatt ccctgaaaaa actgcagata 180
 gacttcaaga tctcagggat tcccactatt tggatttctg atatgttttt cctgatatgc 240
 atcaaaaactc taatctaaaa cctgaatctc cgctattttt tttttttttt tgatgacccc 300
 gttttcgtga caaattaatt tccaacgggg tcttgtccgg ataagagaat tttgtttgat 360
 tatccgttcg gataaatgga cgcctgctcc atatttttcc ggttattacc ccacctggaa 420
 gtgccagaa ttttccgggg attacggata atacggtggt ctggattaat taatacgcca 480
 agtcttacat tttgttgag tctcgtgca gtatgtgcaa taataaaca gatgagccaa 540
 tttattggat tagttgcagc ttgacccgc catagctagg catagccaag tgctatgggt 600
 gttagatgat gcacttggat gcagtgagtt ttggagtata aaagatcctt aaaattccac 660
 cctt 664

- <210> 43
- <211> 493
- <212> ADN
- 5 <213> artificial

- <220>
- <223> variante de promotor
- <400> 43

ctgcagatag acttcaagat ctcagggatt cccactatth ggtattctga tatgtttttc 60
 ctgatatgca tcaaaaactct aatctaaaac ctgaatctcc gctatthttt tttttttttt 120
 gatgaccccg ttttcgtgac aaattaattt ccaacggggt cttgtccgga taagagaatt 180
 ttgtttgatt atccgttcgg ataaatggac gcctgctcca tathttttccg gttattacc 240
 cacctggaag tgcccagaat tttccgggga ttacggataa tacggtggtc tggattaatt 300
 aatacgccaa gtcttacatt ttgttgagc ctcgtgagc tatgtgcaat aataaacaag 360
 atgagccaat ttattggatt agttgcagc tgaccccgcc atagctaggc atagccaagt 420
 gctatgggtg ttagatgatg cacttggatg cagtgagttt tggagtataa aagatcctta 480
 aaattccacc ctt 493

- 10 <210> 44
- <211> 370
- <212> ADN
- <213> artificial

- <220>
- 15 <223> variante de promotor
- <400> 44

ES 2 730 173 T3

	gacccccgttt tcgtgacaaa ttaatttcca acgggggtctt gtccggataa gagaatthttg	60
	tttgattatc cgttcggata aatggacgcc tgctccatat ttttccggtt attacccac	120
	ctggaagtgc ccagaatthtt ccggggatta cggataatac ggtgggtctgg attaattaat	180
	acgccaagtc ttacatthttg ttgcagtctc gtgcgagtat gtgcaataat aaacaagatg	240
	agccaatthta ttggattagt tgcagcttga ccccgccata gctaggcata gccaaagtgct	300
	atgggtgtta gatgatgcac ttggatgcag tgagthttgg agtataaaaag atccttaaaa	360
	ttccaccctt	370
	<210> 45	
	<211> 328	
	<212> ADN	
5	<213> artificial	
	<220>	
	<223> variante de promotor	
	<400> 45	
	ccggataaga gaatthttggt tgattatccg ttcggataaa tggacgcctg ctccatattt	60
	ttccggttat taccacacct ggaagtgcc agaatthttcc ggggattacg gataatacgg	120
	tggtctggat taattaatac gccaaagtctt acatthttggt gcagtctcgt gcgagtatgt	180
	gcaataataa acaagatgag ccaatthtatt ggattagttg cagcttgacc ccgccatagc	240
	taggcatagc caagtgcctat ggggtgttaga tgatgcactt ggatgcagtg agthtttggag	300
10	tataaaaagat ccttaaaaatt ccaccctt	328
	<210> 46	
	<211> 283	
	<212> ADN	
	<213> artificial	
15	<220>	
	<223> variante de promotor	
	<400> 46	
	gcctgctcca tathtttccg gttattacc cacctggaag tgcccagaat tttccgggga	60
	ttacggataa tacggtggtc tggattaatt aatacgccaa gtcttacatt ttgttgcagt	120
	ctcgtgcgag tatgtgcaat aataacaag atgagccaat ttattggatt agttgcagct	180
	tgaccccgcc atagctagcc atagccaagt gctatgggtg ttagatgatg cacttggatg	240
	cagtgagthtt tggagtataa aagatcctta aaattccacc ctt	283

REIVINDICACIONES

- 5 1.- Un método para producir una proteína de interés (PDI) mediante el cultivo de una línea de células eucariotas recombinantes que comprenden una construcción de expresión que comprende un promotor regulable y una molécula de ácido nucleico que codifica una PDI bajo el control transcripcional de dicho promotor, que comprende las etapas de:
- a) cultivar la línea de células con una fuente de carbono basal que reprime al promotor, en el que la fuente de carbono basal es una fuente de carbono adecuada para el crecimiento de las células,
- b) cultivar la línea de células sin una fuente de carbono suplementaria, o con una cantidad limitada de esta, que desreprime al promotor para inducir la producción de la PDI a una tasa de transcripción de al menos 20%, comparado con el promotor pGAP nativo de la célula, y
- 10 c) producir y recuperar la PDI;
- en el que el promotor se selecciona del grupo que consiste en:
- (i) pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); y
- 15 (ii) una secuencia que comprende una secuencia de nucleótidos de al menos 200 pb y que tiene al menos 80% de identidad de secuencia con una cualquiera de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); o con una parte de cualquiera de las anteriores con una longitud de al menos 200 pb.
- 20 2.- El método según la reivindicación 1, en el que la fuente de carbono basal se selecciona del grupo que consiste en glucosa, glicerol, etanol, una de sus mezclas y un material nutriente complejo.
- 3.- El método según la reivindicación 1 o 2, en el que la fuente de carbono suplementaria es una hexosa, tal como glucosa, fructosa, galactosa o manosa, un disacárido, tal como sacarosa, un alcohol, tal como glicerol o etanol, o una de sus mezclas.
- 25 4.- El método según cualquiera de las reivindicaciones 1 a 3, en el que la etapa b) emplea un medio de alimentación que no proporciona la fuente de carbono suplementaria o la proporciona en una cantidad limitada, preferiblemente a 0-1 g/l en el medio de cultivo.
- 5.- El método según cualquiera de las reivindicaciones 1 a 4, en el que la cantidad limitada de la fuente de carbono suplementaria limita el crecimiento para mantener la tasa de crecimiento específica dentro del intervalo de 0,02 h⁻¹ hasta 0,2 h⁻¹, preferiblemente de 0,02 h⁻¹ hasta 0,15 h⁻¹.
- 30 6.- El método según las reivindicaciones 1 a 5, en el que el promotor es capaz de controlar la transcripción de un gen de una célula eucariota de tipo salvaje, y dicho gen se selecciona del grupo que consiste en G1 (SEQ ID NO:7), G3 (SEQ ID NO:8), G4 (SEQ ID NO:9), G6 (SEQ ID NO:10), G7 (SEQ ID NO:11) o G8 (SEQ ID NO:12), o uno de sus variantes funcionalmente activos que comprende una secuencia de nucleótidos de al menos 200 pb y una identidad de secuencia de al menos 80% con cualquiera de las secuencias de promotores o con una de sus partes
- 35 con una longitud de al menos 200 pb.
- 7.- El método según la reivindicación 6, en el que el promotor es pG1 (SEQ ID NO:1) o uno de sus variantes funcionalmente activos.
- 8.- El método según las reivindicaciones 6 o 7, en el que dichos variantes funcionalmente activos se seleccionan del grupo que consiste en homólogos que pueden obtenerse por medio de la modificación de la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, y análogos derivados de especies distintas de *Pichia pastoris*.
- 40 9.- El método según cualquiera de las reivindicaciones 1 a 5, en el que el promotor se selecciona del grupo que consiste en pG1 (SEQ ID NO:1), pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46).
- 45 10.- El método según cualquiera de las reivindicaciones 1 a 9, en el que la línea de células se selecciona del grupo que consiste en líneas de células de mamífero, insecto, levadura, hongo filamentoso y vegetales, preferiblemente una levadura.
- 50 11.- El método según cualquiera de las reivindicaciones 1 a 10, en el que la PDI es una proteína heteróloga, preferiblemente seleccionada de proteínas terapéuticas, que incluyen anticuerpos o sus fragmentos, enzimas y péptidos, antibióticos de proteínas, proteínas de fusión de toxinas, conjugados de carbohidrato-proteína, proteínas estructurales, proteínas reguladoras, vacunas y proteínas o partículas similares a vacunas, enzimas de procesos,

factores de crecimiento, hormonas y citoquinas, o un metabolito de una PDI.

5 12.- El método según la reivindicación 1, en el que el promotor es un variante funcionalmente activo de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6), que se selecciona del grupo que consiste en homólogos que pueden obtenerse por medio de la modificación de la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, y análogos derivados de especies distintas de *Pichia pastoris*.

10 13.- El método según la reivindicación 12, en el que el promotor se selecciona del grupo que consiste en pG1 (SEQ ID NO:1), pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46).

14.- Un ácido nucleico de promotor aislado que está unido operablemente a una secuencia de nucleótidos que codifica una PDI, y dicho ácido nucleico de promotor no está asociado nativamente con la secuencia de nucleótidos que codifica la PDI, en el que el ácido nucleico de promotor comprende una secuencia de ácido nucleico seleccionada del grupo que consiste en:

15 a) pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); y

20 b) una secuencia que comprende una secuencia de nucleótidos de al menos 200 pb y que tiene al menos 80% de identidad de secuencia con una cualquiera de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6); o con una parte de cualquiera de las anteriores con una longitud de al menos 200 pb;

en el que dicho ácido nucleico comprende un promotor funcionalmente activo, que es un promotor regulable por una fuente de carbono capaz de expresar una PDI en una célula eucariota recombinante a una tasa de transcripción de al menos 20%, comparado con el promotor pGAP nativo de la célula.

25 15.- Un ácido nucleico de promotor según la reivindicación 14, en el que el promotor es un variante funcionalmente activo de pG1 (SEQ ID NO:1), pG3 (SEQ ID NO:2), pG4 (SEQ ID NO:4), pG6 (SEQ ID NO:3), pG7 (SEQ ID NO:5), o pG8 (SEQ ID NO:6), que se selecciona del grupo que consiste en homólogos que pueden obtenerse por medio de la modificación de la secuencia de nucleótidos de origen mediante la inserción, la delección o la sustitución de uno o más nucleótidos dentro de la secuencia o en cualquiera o ambos de los extremos distales de la secuencia, y análogos derivados de especies distintas de *Pichia pastoris*.

30 16.- Un ácido nucleico de promotor según la reivindicación 14 o 15, en el que el promotor se selecciona del grupo que consiste en pG1 (SEQ ID NO:1), pG1a (SEQ ID NO:41), pG1b (SEQ ID NO:42), pG1c (SEQ ID NO:43), pG1d (SEQ ID NO:44), pG1e (SEQ ID NO:45) y pG1f (SEQ ID NO:46).

35 17.- Un vector que comprende el ácido nucleico de promotor según las reivindicaciones 14 a 16, en el que la secuencia de nucleótidos que codifica un PDI está bajo el control transcripcional de dicho promotor, y dicho ácido nucleico de promotor no está asociado nativamente con la secuencia de nucleótidos que codifica la PDI.

18.- Una célula eucariota recombinante que comprende el vector de la reivindicación 17.

19.- Un método para producir la proteína de interés (PDI) usando la célula eucariota recombinante de la reivindicación 18.

Fig. 1: pG1 (SEQ ID 1)

ATTTCCACCCCATCCCAGTAGAATGTAGGGTCCCCAACATTTGCTCCCCCTAG
TCTCCAGGGAAATGTAAAATATACTGCTAATAGAAAACAGTAAGACGCTCAGTTGT
CAGGATAATTACGTTTCGACTGTAGTAAACAGGAATCTGTATTGTTAGAAAGAACG
AGAGTTTTTTACGGCGCCGCCATATTGGGCCGTGTGAAAACAGCTTGAAACCCCA
CTACTTTCAAAGGTTCTGTTGCTATACACGAACCATGTTTAACCAACCTCGCTTTT
GACTTGACTGAAGTCATCGGTTAACAATCAAGTACCCTAGTCTGTCTGAATGCTCC
TTTCCATATTCAGTAGGTGTTTCTTGCACTTTTGCATGCACTGCGGAAGAATTAGC
CAATAGCGCGTTTCATATGCGCTTTTACCCCTCTTTTGTCAAGCGCAAATGCCT
GTAAGATTTGGTGGGGGTGTGAGCCGTTAGCTGAAGTACAACAGGCTAATTCCT
GAAAAAAGTGCAGATAGACTTCAAGATCTCAGGGATTCCCACTATTTGGTATTCTG
ATATGTTTTTCTGATATGCATCAAACTCTAATCTAAAACCTGAATCTCCGCTATT
TTTTTTTTTTTTTTGATGACCCCGTTTTCGTGACAAATTAATTTCCAACGGGGTCTT
GTCCGGATAAGAGAATTTTGTGTTGATTATCCGTTCCGATAAATGGACGCCTGCTCC
ATATTTTTCCGGTTATTACCCACCTGGAAGTGCCAGAAATTTCCGGGGATTACG
GATAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACATTTTGTGTCAGTC
TCGTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGATTAGTTGCA
GCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCA
CTTGGATGCAGTGAGTTTTGGAGTATAAAAGATCCTTAAAATTCCACCCTT

Fig. 2: pG3 (SEQ ID 2)

GTAATAGCGGCAGCAATCCAGTAACCTTTTCTGAATAGCAGAGCCTTAACTAAAA
TAATGGCCAGGGTAAAAAATTCGAAATTTGACACCAAAAATAAAGACTTGTCGTTA
TAAGTCTTAACAAAGTCCGCAATTTTGGAGCTAACGGTGGCGGTTGCTGGGATAT
TCAATAATGGTAGAATGTTGCTGCGGGTATATGACAGAGCGTGAAACACACTGAA
CAAGGTAAATGGAACAACAGCAATTGCAATATGGGGGAGGATAGTCAAGAACAAA
GCAGCAATGGCAAAGTACTGAATATTCTCCAAAGCCAAAAGGTCCAGTGGTTTCA
ACGACAAAGTCTTGTTGGTATAGCTTTGGAACAAAAGGACACCGAAAGACTCGAC
AGCGCCCACAAATACAGCGTTGTAGAAGAACGAATTGATTGCTCCAGAGCTTCTA
ATAGTCAGAAGATACCCCAAACCTCCGAGCAACGTTAGCACATGACCTAAGAACC
AGGCGAAGTGAAGAGTCTGGAATAACGACACCCAGTCAGTTTTTTCCTGAGCTCCT
GGTGGGATTGGTAGAAGCATTGATTTGCTTGGAGTGGTTTTATTTGAAGATGGT
GTTGAAGCCATTGTTGCTAAAGAGTCGGAGTTTTGCTTTTAGGGTTTGTAAAGCAA
AGGAGGAAAAACTGCGCCGTTTGAAGTCCCAGGTAGTTTTCGCGTGTGAGGCCAG
CCAGGGAAAGCTTCCTTCGGTACTTTTTTTTTCTTTTGCAGGTTCCGGACGGATTAA
GCTTCGGGTTATGAGGGGGGCGGTAGCCAATTCCGGACACAATATTGCGTCGCA
GCTAGTCACCCCGCCATAAATATACGCAGGATTGAGGTAATAACATCGATAGTCTT
AGTAATTAATACAATTCAGTGGCGAATTTGGCAACATGACGTAAGGCCCACTGTTG
TCTATAAAAGGGGATGAATTTTCATGTTTTTGAGGCCTCCCGGACAATTTATTGAA
CTCAA

Fig. 3: pG4 (SEQ ID 4)

TGGACTGTTCAATTTGAAGTCGATGCTGACGATGTCAAGAGAGATGCTCAATTATA
TTTGTCAATTTGCTGGTTACACTGGAAACGCTACTTTTGTGGCGGAAACTCTACCA
GTTTGGCCGTCCATGTAAACGATGTCGTTCTGGGCCGTGACCGTTTCAACACGAA
CATAACCAATGACAAATCCACTTACAGGTCTAGTTCATATGGAGGCAATTGGTACC
TACTTCTTTGGATGTCCCAAGTGGGGCTTTAACGTCTGGTACTAACAATGTCTCG
TTTGTCACTACAAACTCCGAGGTAAATAAAGGATTCTTGTGGGATTCTCTCAAGTT
TGTTTGGAAGTTGTAACAGGTTTATAAGCATATCGTGCGCTTGTCCACAATTGAAT
CATTTATTGTTGCGAGATACATGAACAAAGTGTGAACTGGGACCCATTACTACAAT
TCCCACGCAACCGTTGTTTCAAAGCCCATATTTTTTGACAATTGTTTCGTTACACC
CCCAGTTTGATGTACATCGCTTGCAATGATGTGTGTCCCGGAGTATTTTCCATATT
CAGCTTGAATTCGTATACTCAACCAATATCTGGGGGTATACTTTTATGTAACCTATA
CAAATCAACTATACTATTTACCTTTGACCAATCATCTCCCATCTTGTTAAGTTTT
GCTTCCTATATCCCTGACCCTGACATCACCCATGATTCCGCTCAACGGTTCTCCTC
TACATCGTCCCTCTTTTGGAGAGGGTGTTGAGTTTGACATTCAAATTACCCCCCGC
CATCACGCGCAACCGAGACCGCACCCCCGAATTTTCACAAATTACCCACACCCT
ATACTCCACCACTATGAGGGTTATTAGAACTGATCACGTATAAATACCACCGCAAG
TTCCCAAGGGATCGTGTTCTTCTTCTCCAATTGCAATCATATTTCTGACTCTTTCTA
GTTTCAAGATTAATTCCTTTACACTTGCTTTTTTCCCTTACCTTTATCC

Fig. 4: pG6 (SEQ ID 3)

AGACCAGCAGTTTAACTACGCAAATCCACAGGAATTTCTACATCACAATACCAATG
GTAATACCACGACGTCAAGGAATGGAACGACGACTTGGAGGAAGACTTCGTCAA
CCTCTTGCGGAGTACCCGAGGCTAAGACAATAAGAAGAAAAAAAAAAGAAAAGCG
GTGGGGGAGGGATTATTAATAAGGATTATGTAACCCCGAGGGTACCGTTCTATAC
ATATTTAAGGATTATTTAGGACAATCGATGAAATCGGCATCAAACCTGGATGGGAGT
ATAGTGTCCGGATAATCGGATAAATCATCTTGCGAGGAGCCGCTTGGTTGGTTGG
TGAGAGGAGTGAAATATGTGTCTCCTCACCCAAGAATCGCGATATCAGCACCCCTG
TGGGGGACACTATTGGCCTCCCTCCCAAACCTTCGATGTGGTAGTGCTTTATTAT
ATTGATTACATTGATTACATAGCTAAACCCTGCCTGGTTGCAAGTTGAGCTCCGAA
TTCCAATATTAGTAAAATGCCTGCAAGATAACCTCGGTATGGCGTCCGACCCCGC
TTAATTATTTTAACTCCTTTCCAACGAGGACTTCGTAATTTTTGATTAGGGAGTTGA
GAAACGGGGGGTCTTGATACCTCCTCGATTTTCAGATCCCACCCCTCTCAGTCCC
AAGTGGGACCCCTCGGCCGTGAAATGCGCGCACTTTAGTTTTTTTCGCATGTA
AACGCCGGTGTCCGTCAATTAAGTTCGCAGACTAGGGTGAACCTTACCATTTTT
GTCGCACTCCGTCTCCTCGGAATAGGGGTGTAGTAATTCTGCAGTAGTGCAATTT
TTACCCCGCCAAGGGGGGGCGAAAAGAGACGACCTCATCACGCATTCTCCAGTC
GCTCTCTACGCCTACAGCACCGACGTAGTTAACTTTCTCCCATATATAAAGCAATT
GCCATTCCCCTGAAAACCTTAACCTCTGCTTTTTCTTGATTTTTCTTGCCCAAAGA
AAAG

Fig. 5: pG7 (SEQ ID 5)

AATTGATTAAGTTCAGTGAAATTTCAAACCGCTATACACAACAGGACAACCTTTGAG
TTTAGAAAAATCCGATGTAGTGTAACGGCTAGCACGGTCCGCTTTCACCGGGCAG
ACCCGGGTTCGACTCCCGGCATCGGAGTTTATTTTTCCATTTTCGTTCTTTAGAGTA
TTCTCCTCAGCATGCCCCCTGAATTTTTCTTTTTTCCATGTGTCCCATTTTTCCA
CTTTTTTTACAGTTTTCTCGTGATGTTAATTGGCTACACAAAAGCTGCCACACGA
AACCTTAATCACGAAAACTATACAGCCTTCACTAATCCGTAGCCCCATAATATGT
TGTCCACGTGCTGTTGGGTAACCTGTAGACTCTCATACCCCACTCCGTCTTTCT
CCAACAATTAACGCAGTACCGAGATTTATCAGCAGACTCAAATTGGGCAAACCTCT
GTATTTTTCTTGCCCGCATAATTTATGGGTCTCAGGCCTCCACGTTTCTGTTTA
CTTGAAGAATATTGGCTGCGGAAAAAGTGGTAAGGACAACCCCTTTTAATTGGA
TCCAGTTTTTCCGAAATGTTCCGATCCGTACGTCATCTCCGAAGCCGTACATTTTC
ACTCAATCTACGTAGCTTTGGACTCAGCGCTCCTGGAATTGCCAGGACAGTTTAC
TTGAGTTGATATCCCTTGTAGATTGTGTGCTTCTTTTTCCAAAATTTGAGGCTTCG
TTTGAAAAGTGGAATCTGGTCGCTAGATCACTTCATGCCTATTTTTCACGGAAAAA
TAAGTGGTACTATGCACCCCTTAAACCTAAAGAAAAACGGAAAAATTACCCCAAAA
CCTGGTGATGTTTTTCGCCCTTTCTTTTTATCCGAGTTTTTCTTTTTTCTTGTCTG
CCAAATTCCTCTCCTGACCTTAGCGTCCCGGAAAAAATTACTACTTAAGGACCG
AATGAGCCCCAGCTTTTCCCCTTCTCTTCATTATTCCCATAATATAAT

Fig. 6: pG8 (SEQ ID 6)

CTGCACAACCATTGCCAGTAAGGACGAAGAGAAGGCCCCACTACCCAAAATTCAG
GATAACGTCTTCATACCATGCAGCGACGCCTACAAGACGCTGTCAAGACATGCCA
ACTTCAACGAAGTGAACCTTTAACACATTGATCGGGAAATTGACCACCAAGGGAAT
GCTGGTTGAGGCTGGAAGCGTTGCCAGTGTCTGAGGGAACTGGACCGAAAGTT
TAGTAATGCATAAGAGGATATATATAGGAATGCAGTAATAATATTAGTACCCATTAA
GTGGGCTAAGCCATTGGAAGGCCGTCTGACTGATGGTGGTGTTCCTTCATTTAG
ATAGTGCATTTGCAACTACCGTCTGAGATTGAGTTTGATGTGAAGCTCCAGCGCC
AAAACAGTATAAGAACCTTATCTCCGCATTATTGTTCTTGCGTAAAAGTTTGTGTGA
AGAAACAGGGGTAGTTGCGCAGATTAGTTGTAATATGCGCATAGGATGGGTCATT
GACTTCTTTCCTCGAAAGAGCCACACCGTTAGCTAAAAAAGGACGCGCATCTACC
CCAAAATAGAATGTGGGGAAATAGGACGCGCAACTTCCTCTCAATCACTGGACGT
CAGAAAAACAAATGCGCAATCGAGTCACCCTCCGTGATACCCTCCGTGATACCCC
CTCTCCGTCTATTCTGACAGCGTCTCCCCATGACGTTTCAATCTACTTAGAAAAGA
TTTCGTTTTTTTTTTCCTTCAATTACACGATCTCATCTTCTGCAAGGGTCTGGAGGAC
ATCACCAATCTGCGACTCCATAACTTAGTCCTGAGTTTATATTTACGCTTCATCTGA
TGAGTAGGAAGAAAAAGTTTCACGAAATTCCCCCGCCAACCTTGCCCTTCGGAATA
AGCAGCCACTCTCCTTCTGCCCATAGTAAGCTTGCGCGAGGCCCAACTTGGCC
AGAACTTTAAATATGCCAAACAATCTCCCCCAATCTAAGTTCTCCCTCTTCTAAAA
A

Fig. 7: G1 (SEQ ID 7)

ATGTCCTCGTTTTTTCTAAACAACCAAACAGTAAAGATGATGACGCCTTTGGGAAG
AGCTAGAGCTCTAAATTTTCAGGGCAAAGTTTATGATAAATTTCCAAAACTTACAA
TATCTACGCTATTGCAATAACAGCCACCGTTTCTGGACTGATGTTCCGGTTTTGATA
TTTTCTTCTGTGTCCTCGTTTCGTAAGTCAGGATCATTACAGAACTACTTCAACCGT
CCCGACAGTTTGACGCAAGGGGGTATCACCGCAAGTATGGCTGGAGGTTCTTTCT
TGGGTTTCGTTATTTTCTTCTGACTTCCAGGATATCTTTGGAAGAAGAGTTGCTCTG
CATATGTGCAGTGTCTCTGGATTATCGGGGCCATTCTTCAATGCGCTGCACAAA
ACCAAGGTATGCTGATCGCAGGGAGATTGATTTCCGGTATCGGTGTCCGGTTTTGG
TTCAGCTTCAGCTCCAGTCTATTGTTCTGAAGTTGCTCCAGCAAAGATTAGAGGAA
TGATTGGAGGATTATTTCAATTTTCTGTCACTGTGGGTATCATGATAATGTTTTATA
TCGGATATGGATGTCACTACATTGACGGCGTTGCATCATTTAGACTGGCCTGGGG
TTTGCAAATGGTTCCAGGTCTTATTCTTTTGGTCGGTGTATTCTTCCTTCCTGAGT
CTCCAAGATGGCTGGCTAACCACAACCGCTGGGAAGACGCAGTTGAGGTTATTG
CTAATGTTGTTGCAAAAGGTGACAGAGAAAACGCCGATGTGCGTCTGCAATTGGA
TGAAGTTCAGGAGCAACTATTGATTGACAAAGATGCTTCTGATTTTGGTTACCTTG
ATTTGTTTAAGAAAGATTGTATCAAACGTACCTTCATTGGAGTGTCAGCTCAAGTG
TGGCAACAACCTTTGTGGTATTAATGTTGCAATGTACTACGTTGTGTATCTCTTCCA
AATGGCTGGTTTTACTGGAAATGTGGCGTTGGTATCGTCCTCAATTCAATATGTTT
TGAATGTTGTTATGACTGTTCCAGCTTTGTTTCTAATGGACCGTATAGGCAGACGA
CCCCTACTAATTGGTGGTGGTATTTTCATGTGTATTTGGCTGTTTGGAGTGGCAG
GATTATTAGGCACTTACTCTGAACCAATTGAAAATTTAGCGGTGATGATACTGTC
AGAATTACTATTCCTGACCAGCACAAGGCTGCAGCAAGGGGTGTTATTGCCTGTT
CCTATCTATTTCGTGTGCTCCTTTGCTCCAACCTGGGGTATCTGCATTTGGGTTTAT
GCCTCTGAAATTTTCAACAACAGACAAAGAGCAAAGGGAGCAGCATTGCTGCCT
CCGCTAACTGGATTTTCAACTTTGCCTTGGCTATGTTTCGTGCCATCAGCCTTTAGA
AACATTACATGGAAGACTTACATCATTTTTGGAGTATTTTCGTTCTGCTTAACAATC
CATGTTTTCTTACAATTCCCAGAAACCAGAGGTAAGACTTTGGAAGAAATTGATCA
AATGTTTAAGGACAATATTCCAGCTTGGAGAAGTGCTTCGTACGTTCCAGATATGC
CAATTTTCAACAAAGAGAAGGTAGTATCTACTGAGCATGCAGAAAATGCTTCCAGC
TCGTCCGAAAAAGCCTTGATGGTTCAGGAAGAGGAATCTGTATAA

Fig. 8: G3 (SEQ ID 8)

ATGGTAGTTGCAATCGAAGGTGGTACAGGCTTAGGCCTTATGAATCTTACTTGGA
AACCAACTCCAACCCCAATTGATGATGCAATTGAGACAATTAGATATGCTGTTGAG
GAAGCTGGTGTGAGATACTTGAACGGAGGAGAGTTCTACAACCTTTCCTCTTGATT
CAAACCTGAATTTGCAGTACATTCAGGAATTTGCAAAAAGGTACCCCGAGCTATAT
AAAAAGGTGAGTCTGTGCGTAAAAGGTGCTGTGAGTTTGGTTCGATGTGAGCCCCG
ATTCTTCCCCGGAGAACCTTGAAAAATCGATTTCAAACATAACCAAACATTTGCCG
AACAACTTCCTGCCAATTTTTGAGCCTGCTAGAATCGATAAACGTTACTCCATTGA
GGAGACAATAAAGAATCTCTCTAAGTTCGTGCGAAGATGGCAGAATTGGAGGTATT
TCACTTAGTGAAGTTGGTGTGACTACTATCAGAAGAGCTGCGAAAGTGGCTCCCA
TCGCCTGTGTGGAAGTGGAGTTTTCTCTATTGACTAGAGATATTCTTCATAATGGA
GTTCTTGCTGCTTGTGAGGATTTGAACATTCCTATTATTGCCTACAGTCCCTTGGG
AAGAGGATTTTTGACTGGAACGATAAACAGCAAAGCTGACATTCCTGAAGGTGAT
ATCAGGTTAAGTTTGGAAAGATTCAATGACGATGAAGTTATTGAACACAATTTGAA
ACTTGTTACGGTTTAAAAAGATAGCCGACAAAAAAGGAGTCACATTGGCTCAAT
TGTCTCTTGCGTGGTTACGAAAGTTTGGAGATAAACACGTCAAGGTGCTTCCTATT
CCAAGCTGCTCATCTCCTCGTAGAGTTGCAGAAAACACAAAAGAGATTTCTTGA
CTGATAGCGAGTTCCAGGAGATTACTGACTTTGCAGAGTCGGTTCCAATCAAAGG
TGGTCGTTACAACAAAGCAAGTGAGGCTGTTCTTAACGGTTAG

Fig. 9: G4 (SEQ ID 9)

ATGACATTTGCTCCTCCCTTAGAATTCGAGATTGACCTTCCTAACGGATTGAAGTA
CACTCAACCATTGGGACTCTTCATCAACAATGAGTTTGTGGAAGGTGTAGAGGGA
AAGCTCTTACCAGTGATCAATCCTTGTGATGAGACTAAAATAACCCAAGTTTGGGA
AGCTTCTGCAGCGGATGTTGACCGTGCTGTTGATGCCGCTGAAGATGCTTTC AAC
AACTCCGTATGGGCTACTCAGGACCCATTAGAGAGGGGAAAGCTGATGAACAAAT
TGGCAGACCTTATCGATCGTGACTTCAACATCTTGGCTGGTATCGAATCCATCGA
CAATGGTAAGGCCTATACCTCTGCCAGGGTGATGTTACTCTTGCTGTCAACTAC
ATCAGATCCTGTGCTGGATGGGCCGACAAGATTTTGGGAAACGTTGTTGATTCCG
GAAACACCCACCTTAACTTGGTTAAAAGAGAGCCATTGGGTGTTGTGGGACAAAT
TATCCCATGGAACCTTTCCTCTCCTGATGTTGGCTTGGAAGTTGGGACCTGCGCTG
GCCACAGGTAACACTGTTGTTTTGAAGACTGCCGAGTCTACCCCTCTGTCTGGGTT
TATACGTTGCCAAATTGATCAAGGAGGCCGTTTCCCACCTGGTGTGGTTAACAT
TCTCAGTGGTTTTCGGTAACCCAGCTGGAGCTGCCATCGCTGCTCATCCCAGAATC
AAGAAGATTGCTTTCACCGGATCCACTGCAACAGGCCGTAAGATCATGGAAGCAG
CCGCTAAATCTAACCTGAAAAAGTCACTTTGGAAGTGGTAAATCTCCAAAC
ATTGTGTTTGAAGATGCTGATATCCAGAAGACTATCCATAACATTATTTTGGGAAT
CTTCTTCAATTCTGGTGAAGTCTGTTGTGCAGGTTCCAGAGTCTACATTCAAGACA
CTGTGTATGAAGAAGTGCTTGAAGCCTTCAAGAAGGAGACTGATAACGTTAAGGT
TGGTGGACCATTGAAGAAGGTGTCTTCCAAGGGCCTCAGACCTCTGAGTTGCAA
CTTAACAGAATCCTTAGTTACATCAAACACGGTAAGGATGAAGGTGCTCGTGTAAAT
TACCGGTGGTTCAAGATACCGTAACCGAGGTTACTACATTAAGCCACAATTTTGTG
CTGACGTTACTGAAGACATGAAGATTGTCAAGGAGGAGATTTTGGTCCTGTGGT
TACTATCACTAAGTTCTCTACCGTGGATGAGGTTGTTGGATATGCCAACAACACCA
ACTATGGTCTAGCTGCTGGTATTCACACAAACAACCTTGAACAAAGCCATTGATGTT
GCCAGTAGAATCAAGGCGGGTGTCTGTTTGGATTAACACCTACAACGATTTCCACC
ACATGGTTCCTTTCGGAGGTTATGGAGAATCTGGTATTGGCAGAGAGCTTGGTGC
TGAGGCTTTGGATAACTCACTCAAGCCAAGGCTATCAGAATTGCTTACACTCCTG
AACATAAGTAG

Fig. 10: G6 (SEQ ID 10)

ATGCTTAGAACTTCTCCAGCTACTAAGAAAGCTCTCAAGTCGCAGATTAACGCCTT
CAACGTTGCTGCCTTGAGATTCTACTCCTCATTGCCTTTGCAGGTTCCAATTACCT
TGCCAAACGGTAAGACCTACAATCAGCCAACAGGTTTGTATTATCAACAATGAGTTC
GTTCCCTTCTAAGCAAGGTAAGACCTTTGCTGTTTTAAACCCTTCCACTGAGGAGGA
GATTACTCACGTCTACGAGTCCAGAGAGGACGACGTTGAGTTAGCCGTTGCAGC
CGCTCAAAGGCTTTGACTCAACCTGGTCCACCCAGGACCCTGCTGAGAGAGG
TAAGGTCTTGAACAAGTTGGCTGACCTGATCGAGGAGCACTCTGAGACCCTTGCC
GCCATCGAGTCCTTGGACAACGGTAAGGCCATTTCTCCGCTAGAGGTGATGTTG
GTCTGGTTGTGCCTACTTGAAGTCCTGTGCCGGTTGGGCCGACAAGGTTTTTCG
GTAGAGTTGTTGAAACCGGAAGCTCCCACTTCAACTACGTTAGAAGAGAGCCATT
GGGTGTTTGTGGTCAGATTATCCCATGGAACCTTTCTCTTCTGATGTGGTCCTGG
AAAGTTGGTCCAGCTTTGGCCACTGGTAACACTGTTGTCCTGAAGACAGCCGAGT
CTACTCCTCTGTCCGCCCTGTACGTTTCCCAATTGGTCAAGGAGGCCGGTATCCC
AGCTGGTGTCCACAACATTGTGTCCGGTTTTCGGTAAGATTACTGGTGAAGCTATT
GCTACTCATCCTAAGATCAAGAAGGTTGCCTTCACTGGTTCTACCGCCACTGGTC
GTCACATCATGAAGGCTGCTGCCGAATCCAACCTTGAAGAAGGTTACTTTGGAGTT
GGGTGGTAAATCTCCTAACATCGTGTTCAACGATGCTAACATTAAGCAAGCTGTC
GCCAACATCATCCTCGGTATTTACTACAACCTCTGGAGAAGTTTTGTTGTGCTGGTTC
CAGAGTTTATGTTCAATCCGGTATTTACGACGAGCTTTTGGCCGAATTCAAGACTG
CTGCTGAGAATGTCAAGGTTGGTAACCCATTGACGAGGACACCTTCCAAGGTGC
TCAAACCTCTCAGCAACAATTGGAGAAGATTTTGGGTTTTCGTTGAGCGTGGTAAG
AAGGACGGTGCTACTTTGATTACTGGTGGTGGCAGATTAGGTGACAAGGGTACT
TCGTCCAGCCAACCTATCTTCGGTGATGTTACACCAGAGATGGAGATTGTCAAGGA
AGAGATCTTTGGTCCTGTTGTCACTATCAGCAAGTTTGACACCATTGATGAGGTTG
TCGACCTTGCTAACGACTCTCAATACGGTCTTGCTGCTGGTATCCACTCTGACGA
TATCAACAAGGTCATTGACGTTGCTGCTAGAATCAAGTCCGGTACCGTGTGGGTC
AACACCTACAACGATTTCCACCAAATGGTTCCATTCCGGTGGATTTGGCCAATCCG
GTATTGGTCGTGAGATGGGTGTTGAAGCTTTGGAAAACCTACACCCAATACAAGGC
TATCCGTGTCAAGATCAACCACAAGAACGAGTAA

Fig. 11: G7 (SEQ ID 11)

ATGAGTTCAACAGATATCCAAGGTGATCAAGGTGACAATGAAAAGATATACGCCA
TTGAGAGCAGTCCCTCCAATGAGCAAATAAAAGATATTCATGAGGCTCCGGCCGA
CAACAAAAGTGAAGTAGACATCCCAGTCAAACCCAAGGGTTCCTATATCTTGGTGT
CTGTGTTATGTCTTCTAGTCGCATTCGGTGGTTTCGTGTTCCGGTTGGGATACCGG
TACCATCTCAGGTTTCGTTAACATGTCTGACTTTACGAGACGTTTCGGACAGTTTA
ACGGTGAAACGTATTACCTTTCTAAAGTGAGAGTTGGTTTAATTGTTTCTATTTTCA
ACATTGGTTGTGCTATCGGAGGTGTCACTCTAGGTAAACTTGGTGACATTTGGGG
TAGAAAGAAGGCTTTGATGTTTCGTCATGGTCATCTATATGGTCCGGTATTTTGATTC
AAATTGCTTCCATTGACAAATGGTACCAGTATTTCAATTGGAAGAATTATTGCAGGT
CTGGCCGTCGGTGCAGTTTCCGTTTTATCCCCATGTTTCATCAGTGAGACTTCTC
CTAAACACATCAGAGGTTCCCTTAGTCTCCTGCTACCAATTAATGATTACAGCCGGT
ATTTTCTTGGGTTACTGTACCACTTACGGAACCAAGACTTACACCGACTCCACCCA
ATGGAGAGTTCCTTTGGGATTGTGTTTCGCTTGGGCCATTCTGATGATTGTTGGTA
TGACCTTCATGCCAGAGTCCCCACGTTTCTTGGTTGAGGTTAACAGAGTCGACGA
GGCTATGAAGTCCATTGCCAGAGTTAACAAGGTCTCTATCGACGATCCATCTGTC
TACAATGAGATGAGACTTATTTCTGACGGTATTGAGAAGGAGAAGGAGGCTGGTA
GCGTTTCTTGGGGTGAAGTGTTCCTGTTCAAGCCAAAGATTTTCTACCGTCTATTG
ATTGGTATTTTCATGCAATCTTTGCAACAATTGACCGGTAACAACACTATTTCTTCTAC
TACGGAACCTACATTTTCAAGGCTGTCGGATTGGACGATTCTTTCCAAACTTCTAT
CATTCTTGGTGTGTCATTTTGTCTTCCACATTCCTAGGTATCTACACCATGGATAA
ATTTGGTAGAAGAAGAACAACACTTTTAGGAGGTTCTGGAGCCATGGTTGTTTGGTGG
TCATTTTCAGTTCGGTTGGTGTCAAGTCTCTTTATGAGAACGGTAAGGATGATCCA
TCCAAACCAGCAGGTAACGCCATGATTGTCTTACCTGTCTGTTCATTTTCTTCTT
TGCATGTACCTGGGCTCCAGGTGTTTTCGTCGTTGTGTCTGAAACCTACCCACTT
AGAATTAGATCCAAGGGTATGGCCATCGCTCAAGGTTCCAATTGGCTTTGGGGTT
TCCTCATTGCCTTCTTCACTCCATTTATCTCAGGTGCCATTGATTTGCCTACGGT
TACGTCTTTATGGGATGTACTCTGTTTCGCCTTCTTCTTTGTGTACTTCTTCGTTCCCT
GAAACCAAGGGTCTGTGCTGGAAGACGTTGATGAAGTCTATGAGAACCTTACCT
TCGGAAGAGCATATGCATACAGCCACACGATTAAGACAAGGGCGCCCTATAA

Fig. 12: G8 (SEQ ID 12)

ATGGCCCTATCTCCTACCTATCAGGGCTACATATCTACCACTGGCGACGCGTTGA
TCGTGATCCAGGCAGCTCTAAATAACCATTTGAATCTTCTTCCCCGAAGACCAAGA
GAAAGAGAGCGAGATGGGCTAATACGATCAGGTAACGTATTTGTTTTTGTTCGAGC
AACGGTCTCATATCAAACGATGGACCGATGGTATCCCCTGGTCTCCATCTAGAGT
CCTTGAAAGTTCTTGTTGTATCGGGAAGTGGACAAGGATACCCCCAAAACTCG
CAAAGTGACGAAGATACTGAGGAGGGGAGAAAGAGGCGAAAGACTTCTGTGGAT
GTAACCGATCCAAATACCAGGCAGTTGGTGGGATCATTGGTGACTTCCTATGACT
TCAAAGAGGATGGACTTATTAACAAAAACACTCTCCTTGACTTTCCAGACCGGTGCT
AATGAAGAAAGGGAAACAGTGCACTTGATTAGTTATTATACTCCGGAAGATGTAAC
GAACCATCGTTTGAACAGGCCGTCTGACAATCCATATCTGGCCAATATCACTGTTT
CAGAGTCATTATTGACTGCCTTGAGAGAGAGTACCCTTGGAGGAAGAGCAACGTC
TGATGACGAGCTTTCTTTAGTCAGAAGTAACTCGTTAGAGTACCAAGAGGTACCAA
TGAACATATCTATGTCTTTACCTTTATCAACTCCACTTTCCCTTGAACACAGGAGTAA
ACTCAACTACCCAGCTGCAACAGCAACAACACTACAACAACAACAACAGCAACAGCA
ACAGCAGCAGCAACAACAGCAGCAACAACAGCAACCCGGTAGCATCCCTTCCAAA
ATTTGATGGATCCTTTCTATTACAACAGGGTGTAAATCCAGTTCCTCATTTCATGGA
CCAAAAAATGGGAAGTAGCAATTCGTGGATTAACAATTGGTTTTCGTCCAAATTCGT
CAGAATCAAATGGGCTATCGGTTATCGGACCTCACAAGGGATATGACGAACAAAG
TCCAGCAACGAGTTATACTTTGAATGAACGTTGA

Fig. 13: pGAP (SEQ ID 13)

CTTTTTGTAGAAATGTCTTGGTGTCTCGTCCAATCAGGTAGCCATCTCTGAAAT
ATCTGGCTCCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGG
TAAACTTAAATGTGGAGTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCTT
CCACCGCCCGTTACCGTCCCTAGGAAATTTTACTCTGCTGGAGAGCTTCTTCTAC
GGCCCCCTTGCAGCAATGCTCTTCCCAGCATTACGTTGCGGGTAAAACGGAGGT
CGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCGTCGCTGGCAA
TAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAGAATCGAATAT
AAAAGGCGAACACCTTTCCCAATTTTGGTTTCTCCTGACCCAAAGACTTTAAATTT
AATTTATTTGTCCCTATTTCAATCAATTGAACAACCTATCACCTGCAGGCC

Fig. 14

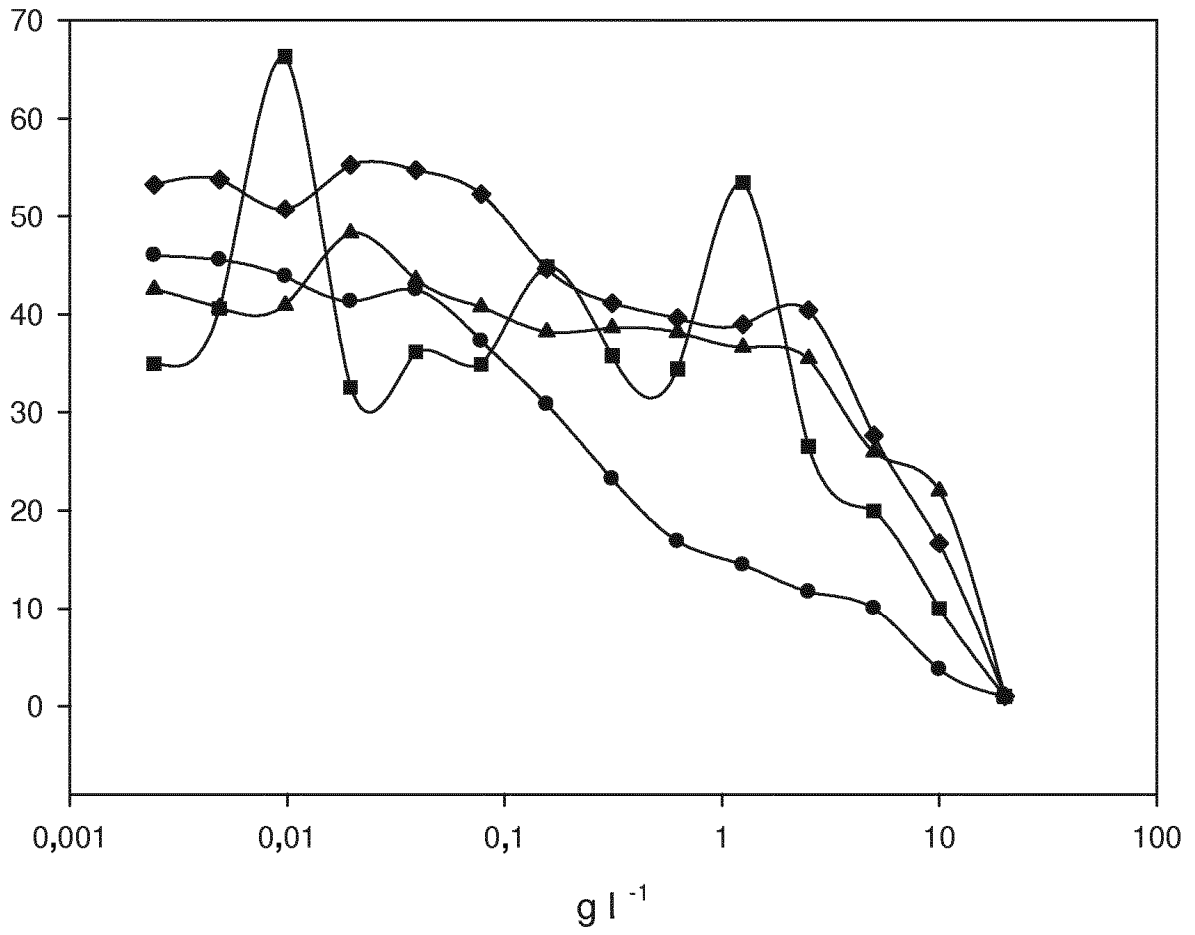


Fig. 15

pG1a (SEQ ID 41)

GGAATCTGTATTGTTAGAAAGAACGAGAGTTTTTTACGGCGCCGCCATATTGGGC
CGTGTGAAAACAGCTTGAAACCCCACTACTTTCAAAGGTTCTGTTGCTATACACGA
ACCATGTTTAAACCAACCTCGCTTTTGACTTGACTGAAGTCATCGGTAAACAATCAA
GTACCCTAGTCTGTCTGAATGCTCCTTTCCATATTCAGTAGGTGTTTCTTGCACTT
TTGCATGCACTGCGGAAGAATTAGCCAATAGCGCGTTTTCATATGCGCTTTTACCC
CCTCTTTTGTCAAGCGCAAATGCCTGTAAGATTTGGTGGGGGTGTGAGCCGTTA
GCTGAAGTACAACAGGCTAATTCCCTGAAAAAAGTGCAGATAGACTTCAAGATCTC
AGGGATTCCCCTATTTGGTATTCTGATATGTTTTTCTGATATGCATCAAACTCT
AATCTAAAACCTGAATCTCCGCTATTTTTTTTTTTTTTTTTGATGACCCCGTTTTCGT
GACAAATTAATTTCCAACGGGGTCTTGTCCGGATAAGAGAATTTGTTTGATTATC
CGTTCGGATAAATGGACGCCTGCTCCATATTTTTCCGGTTATTACCCACCTGGAA
GTGCCCAGAATTTCCGGGGATTACGGATAATACGGTGGTCTGGATTAATTAATA
CGCCAAGTCTTACATTTTGTGTCAGTCTCGTGCGAGTATGTGCAATAATAACAAG
ATGAGCCAATTTATTGGATTAGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGC
CAAGTGCTATGGGTGTTAGATGATGCACTTGGATGCAGTGAGTTTTGGAGTATAA
AAGATCCTTAAAATTCCACCCTT

Fig. 15 (continuación)

pG1b (SEQ ID 42)

CCATATTCAGTAGGTGTTTCTTGCACTTTTGCATGCACTGCGGAAGAATTAGCCAA
TAGCGCGTTTCATATGCGCTTTTACCCCTCTTTTGTCAAGCGCAAATGCCTGTA
AGATTTGGTGGGGGTGTGAGCCGTTAGCTGAAGTACAACAGGCTAATTCCTGAA
AAACTGCAGATAGACTTCAAGATCTCAGGGATTCCCACTATTTGGTATTCTGATA
TGTTTTTCTGATATGCATCAAACTCTAATCTAAAACCTGAATCTCCGCTATTTTT
TTTTTTTTTTGATGACCCCGTTTTCGTGACAAATTAATTTCCAACGGGGTCTTGTC
CGGATAAGAGAATTTTGTGATTATCCGTTCCGATAAATGGACGCCTGCTCCATA
TTTTTCCGGTTATTACCCACCTGGAAGTGCCCAGAATTTTCCGGGGATTACGGA
TAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACATTTTGTGTCAGTCTC
GTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGATTAGTTGCAGC
TTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACT
TGGATGCAGTGAGTTTTGGAGTATAAAAGATCCTTAAAATTCCACCCTT

pG1c (SEQ ID 43)

CTGCAGATAGACTTCAAGATCTCAGGGATTCCCACTATTTGGTATTCTGATATGTT
TTTCCTGATATGCATCAAACTCTAATCTAAAACCTGAATCTCCGCTATTTTTTTTTT
TTTTTTGATGACCCCGTTTTCGTGACAAATTAATTTCCAACGGGGTCTTGTCGGGA
TAAGAGAATTTTGTGATTATCCGTTCCGATAAATGGACGCCTGCTCCATATTTTT
CCGGTTATTACCCACCTGGAAGTGCCCAGAATTTTCCGGGGATTACGGATAATA
CGGTGGTCTGGATTAATTAATACGCCAAGTCTTACATTTTGTGTCAGTCTCGTGCG
AGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGATTAGTTGCAGCTTGAC
CCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACTTGGAT
GCAGTGAGTTTTGGAGTATAAAAGATCCTTAAAATTCCACCCTT

Fig. 15 (continuación)

pG1d (SEQ ID 44)

GACCCCGTTTTTCGTGACAAATTAATTTCCAACGGGGTCTTGTCCGGATAAGAGAA
TTTTGTTTGATTATCCGTTCCGATAAATGGACGCCTGCTCCATATTTTTCCGGTTAT
TACCCACCTGGAAGTGCCCAGAATTTCCGGGGATTACGGATAATACGGTGGTC
TGGATTAATTAATACGCCAAGTCTTACATTTTGTTCAGTCTCGTGCGAGTATGTG
CAATAATAACAAGATGAGCCAATTTATTGGATTAGTTGCAGCTTGACCCCGCCAT
AGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCACTTGGATGCAGTGAG
TTTTGGAGTATAAAAGATCCTTAAAATTCCACCCTT

pG1e (SEQ ID 45)

CCGGATAAGAGAATTTTGTTCGTTCCGATAAATGGACGCCTGCTCCAT
ATTTTTCCGGTTATTACCCACCTGGAAGTGCCCAGAATTTCCGGGGATTACGG
ATAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACATTTTGTTCAGTCT
CGTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGATTAGTTGCAG
CTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTAGATGATGCAC
TTGGATGCAGTGAGTTTTGGAGTATAAAAGATCCTTAAAATTCCACCCTT

pG1f (SEQ ID 46)

GCCTGCTCCATATTTTTCCGGTTATTACCCACCTGGAAGTGCCCAGAATTTCCG
GGGATTACGGATAATACGGTGGTCTGGATTAATTAATACGCCAAGTCTTACATTTT
GTTGCAGTCTCGTGCGAGTATGTGCAATAATAACAAGATGAGCCAATTTATTGGA
TTAGTTGCAGCTTGACCCCGCCATAGCTAGGCATAGCCAAGTGCTATGGGTGTTA
GATGATGCACTTGGATGCAGTGAGTTTTGGAGTATAAAAGATCCTTAAAATTCCAC
CCTT