

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 751 336**

51 Int. Cl.:

C12N 15/81 (2006.01)

C07K 14/39 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **18.12.2013 PCT/EP2013/077144**

87 Fecha y número de publicación internacional: **18.09.2014 WO14139608**

96 Fecha de presentación y número de la solicitud europea: **18.12.2013 E 13811203 (2)**

97 Fecha y número de publicación de la concesión europea: **17.07.2019 EP 2970994**

54 Título: **Promotor constitutivo**

30 Prioridad:

15.03.2013 EP 13159527
15.03.2013 US 201313835589

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
31.03.2020

73 Titular/es:

LONZA LTD (100.0%)
Lonzastrasse
3930 Visp, CH

72 Inventor/es:

MATTANOVICH, DIETHARD;
GASSER, BRIGITTE y
PRIELHOFER, ROLAND

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 751 336 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 constitutivo

Campo de la invención

5 La invención se refiere a una secuencia de ácido nucleico aislada que comprende un promotor constitutivo fuerte y un método para producir una proteína de interés en un cultivo de células eucariotas bajo el control de dicho promotor.

Antecedentes

10 La producción exitosa de proteínas recombinantes se ha logrado con huéspedes eucariotas. Los ejemplos más destacados son las levaduras como *Saccharomyces cerevisiae*, *Pichia pastoris* o *Hansenula polymorpha*, hongos filamentosos como *Aspergillus awamori* o *Trichoderma reesei*, o células de mamíferos como p. ej. células CHO. Si bien la producción de algunas proteínas se logra fácilmente con altas tasas, muchas otras proteínas solo se obtienen a niveles comparativamente bajos.

15 La expresión heteróloga de un gen en un organismo huésped generalmente requiere un vector que permita la transformación estable del organismo huésped. Un vector proporcionaría al gen un promotor funcional adyacente al extremo 5' de la secuencia de codificación. La transcripción es regulada e iniciada por esta secuencia promotora. La mayoría de los promotores utilizados hasta la fecha se derivan de genes que codifican proteínas que generalmente están presentes en altas concentraciones en la célula.

20 El documento EP0103409A2 divulga el uso de promotores de levadura asociados con la expresión de enzimas específicas en la ruta glucolítica, es decir, promotores involucrados en la expresión de piruvato quinasa, triosefosfato isomerasa, fosfoglucosa isomerasa, fosfoglucoserato mutasa, hexocinasa 1 y 2, glucoquinasa, fosfofructosa quinasa, aldolasa y el gen de regulación glucolítica.

25 El documento WO 97/44470 describe promotores de levadura de *Yarrowia lipolytica* para la proteína del factor de elongación de traducción 1 (TEF1) y para la proteína ribosómica S7 que son adecuadas para la expresión heteróloga de proteínas en levadura, y el documento EP1951877A1 describe el uso del promotor de *P. pastoris* TEF1 para la producción de proteínas heterólogas.

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

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

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

Vassileva y col. (J. Biotechnol. (2001) 88: 21-35) describen el uso del promotor GAP para expresar HBsAg en *P. pastoris*, utilizando casetes de expresión multicopia como alternativa al promotor AOX1. El sistema constitutivo se propuso para el cultivo continuo para permitir el mantenimiento de las células en la fase exponencial media.

40 Los promotores utilizados en *Pichia pastoris* están bien regulados (como pAOX o pFLD) siendo activos en sustratos específicos como el metanol, o son constitutivamente activos en muchas condiciones, medios y sustratos diferentes. Entre los constitutivos, especialmente los promotores GAP y TEF se han descrito como fuertes y útiles para la producción de proteínas recombinantes.

45 Sin embargo, se demostró que la actividad de ambos promotores constitutivos no es constantemente fuerte durante un proceso de producción de alimentación por lotes. Especialmente en las fases posteriores del proceso, cuando las tasas de crecimiento celular son lentas, también la actividad de los promotores es baja, lo que limita los niveles de expresión del gen de interés (GOI) y los rendimientos de producción (Stadlmayr et al.2010. J Biotechnol. 150: 519-529).

50 La selección de promotores adecuados no es intuitiva o racional, ya que incluso las enzimas glucolíticas muy abundantes como la enolasa (ENO), triosa fosfato isomerasa (TPI) o la glucosa-6-fosfato isomerasa (PGI) no tienen promotores que sean tan fuertes como pGAP y pTEF (Stadlmayr et al. 2010. J Biotechnol. 150: 519-529; Gasser y col. 2010 Metabolic Engineering 12: 573-580).

Qin y col. (Applied and Environmental Microbiology (2011) 3600-3608) describe una biblioteca promotora GAP y varios mutantes con diferentes actividades.

Qin et al. (Letters in Applied Microbiology 2011, 52(6):634-641) describen un enfoque de alto rendimiento para el cribado de promotores constitutivos diseñados en *P. pastoris*.

- 5 El documento WO2007/015178 A2 describe parejas de fusión traduccionales capaces de inducir la producción secretora de proteínas recombinantes, y una biblioteca de ADNc de *P. pastoris*.

El documento CN 102180954 A describe el sistema de visualización de la superficie celular usando la proteína de pared celular de *P. pastoris* GCW14 como una proteína anclada.

- 10 De Schutter et al. (Nature Biotechnology 2009 27(6):561-566) revela toda la secuencia genómica de la cepa GS115 de *P. pastoris*.

El documento WO2014/138703A1 revela promotores de levadura de *P. pastoris*, entre ellos un promotor llamado UPP representado en SEQ ID NO:2 del documento WO2014/138703 A1.

- 15 Es deseable proporcionar líneas celulares eucariotas recombinantes mejoradas para producir productos de fermentación que puedan aislarse con altos rendimientos. Por lo tanto, un objeto de la presente invención es proporcionar elementos reguladores alternativos adecuados para métodos de producción recombinantes, que sean simples y eficientes.

Sumario de la invención

El objeto es resuelto por el tema según se reivindica.

- 20 De acuerdo con la invención, se proporciona una secuencia de ácido nucleico aislada que comprende un promotor pCS1, que es una secuencia nativa de *Pichia pastoris* que comprende la secuencia de ácido nucleico pCS1 SEQ ID 1, o una variante funcionalmente activa de la misma que es una variante de tamaño, un mutante o híbrido de SEQ ID 1, o una combinación de los mismos, o una combinación de la misma, en la que dicha variante funcionalmente activa tiene una longitud de 200-1500 bp, al menos 60% de identidad de secuencia con la secuencia de SEQ ID 1, y en la que dicha variante funcionalmente activa exhibe la misma expresión o fuerza transcripcional que la secuencia de ácido nucleico pCS1 SEQ ID 1 +/- 20%, y/o al menos un aumento de 1,1 veces en comparación con el promotor nativo de *P. pastoris* pGAP establecido en SEQ ID 13, en el que la secuencia del ácido nucleico promotor pCS1 no es SEQ ID NO:2 del documento WO2014/138703 A1.

- 30 Según un aspecto específico, se proporciona una secuencia de ácido nucleico aislada que comprende un promotor, que es una secuencia nativa de *Pichia pastoris* que comprende o que consiste en la secuencia de ácido nucleico pCS1 de la SEQ ID 1.

Específicamente, la secuencia de nucleótidos pCS1 es idéntica a la secuencia de SEQ ID 1.

Preferiblemente, la secuencia de ácido nucleico de la invención no es idéntica al ácido nucleico de SEQ ID 87 como se enumera en el documento WO2007/015178 A2 (es decir, SEQ ID 24 de esta solicitud).

Específicamente, la variante funcionalmente activa es:

- 35 a) una variante de tamaño de SEQ ID 1, que consiste preferiblemente o que comprende la secuencia de ácido nucleico seleccionada del grupo que consiste en SEQ ID 2, 3, 4, 5 y 6; más preferiblemente la secuencia de ácido nucleico consiste en SEQ ID 2, 3, 4, 5 o 6;

b) un mutante de SEQ ID 1, o un mutante de la variante de tamaño de a), cuyo mutante tiene al menos un 60% de homología con la secuencia SEQ ID 1 o con la variante de tamaño;

- 40 c) un híbrido que comprende:

- una secuencia seleccionada del grupo que consiste en SEQ ID 1, una variante de tamaño de a) y un mutante de b); y

- al menos una secuencia adicional seleccionada del grupo que consiste en SEQ ID 1, una variante de tamaño de a), un mutante de b) y una secuencia heteróloga; o

- 45 d) una secuencia que se hibrida en condiciones rigurosas con cualquiera de las variantes de tamaño, o las secuencias mutantes de ácido nucleico de a), o b).

Preferiblemente, la variante de tamaño de pCS1 de SEQ ID 1 consiste en cualquiera de SEQ ID 2, SEQ ID 3, SEQ ID 4, SEQ ID 5 o SEQ ID 6

Según una realización específica, la variante funcionalmente activa se selecciona del grupo que consiste en homólogos con:

i) al menos aproximadamente 60% de identidad de secuencia de nucleótidos;

5 ii) homólogos obtenibles modificando la secuencia de nucleótidos de pCS1 de SEQ ID 1 o variantes de tamaño de la misma, mediante la inserción, delección o sustitución de uno o más nucleótidos dentro de la secuencia o en uno o ambos extremos distales de la secuencia, preferiblemente con una secuencia de nucleótidos de 200 pb a 1500 pb, y

iii) análogos derivados de especies distintas de *Pichia pastoris*.

Específicamente, la variante funcionalmente activa de la invención tiene sustancialmente la misma actividad promotora que pCS1.

10 Según una realización específica, la secuencia de ácido nucleico está operativamente unida a una secuencia de nucleótidos que codifica una proteína de interés (POI), cuyo ácido nucleico no está asociado de forma nativa con la secuencia de nucleótidos que codifica el POI.

15 De acuerdo con un aspecto específico, la secuencia de ácido nucleico comprende además un gen de péptido señal que permite la secreción del POI, preferiblemente en el que el gen del péptido señal se encuentra adyacente al extremo 5' de la secuencia de nucleótidos que codifica el POI.

Por consiguiente, la invención se refiere específicamente a la secuencia de ácido nucleico que comprende además una secuencia de ácido nucleico que codifica un péptido señal que permite la secreción del POI, preferiblemente en la que la secuencia de ácido nucleico que codifica el péptido señal se encuentra adyacente al extremo 5' de la secuencia de nucleótidos que codifica el POI.

20 La invención proporciona además una construcción de expresión que comprende una secuencia de ácido nucleico de la invención, preferiblemente un vector o plásmido que se replica autónomamente, o uno que se integra en el ADN cromosómico de una célula huésped.

25 La invención proporciona además una célula huésped recombinante que comprende la secuencia de ácido nucleico de la invención o la construcción de expresión de la invención, preferiblemente una célula eucariota, más preferiblemente una levadura o una célula fúngica filamentosa, más preferiblemente una célula de levadura del género *Saccharomyces* o *Pichia*.

De acuerdo con un aspecto específico, la célula huésped recombinante comprende múltiples copias de la secuencia de ácido nucleico y/o múltiples copias del constructo de expresión. Por ejemplo, la célula recombinante comprende 2, 3, 4 o más copias (número de copia del gen, GCN).

30 Específicamente, la célula huésped recombinante se selecciona del grupo que consiste en mamíferos, insectos, levaduras, hongos filamentosos y células vegetales, preferiblemente una levadura, preferiblemente cualquiera de las cepas de *P. pastoris* CBS 704, CBS 2612, CBS 7435, CBS 9173-9189, DSMZ 70877, X-33, GS115, KM71 y SMD1168. En algunas realizaciones, la célula huésped recombinante es una cepa de *P. pastoris* distinta de X-33.

La invención proporciona además un cultivo estable de una pluralidad de células de la invención.

35 La invención proporciona además un método para producir un POI mediante el cultivo de una línea celular recombinante del huésped que comprende el promotor de la invención, o la construcción de expresión de la invención, y un ácido nucleico que codifica el POI bajo el control transcripcional de dicho promotor, o la célula huésped recombinante de acuerdo con la invención, que comprende los pasos de

a) cultivar la línea celular en condiciones para expresar dicho POI, y

40 b) recuperar el POI.

45 Específicamente, el POI se expresa en condiciones que limitan el crecimiento, p. ej. cultivando la línea celular a una tasa de crecimiento inferior a la tasa de crecimiento máxima, típicamente inferior al 90%, preferiblemente inferior al 80%, inferior al 70%, inferior al 60%, inferior al 50%, inferior al 40%, inferior menos del 30%, menos del 20%, menos del 10%, menos del 5%, menos del 3%, menos del 2%, menos del 1%, menos del 0,5%, menos del 0,4%, menos del 0,3%, o menos del 0,2% de la tasa de crecimiento máximo de las células. Típicamente, la tasa de crecimiento máxima se determina individualmente para una célula huésped específica.

De acuerdo con una realización específica, la línea celular se cultiva en condiciones de cultivo discontinuo, de alimentación discontinua o continuo, y/o en medios que contienen sustrato de carbono limitado.

50 Específicamente, el cultivo se realiza en un biorreactor que comienza con una fase discontinua seguida de una fase discontinua alimentada o una fase de cultivo continua.

5 Específicamente, las células huésped se cultivan en un medio rico en fuentes de carbono durante la fase de alta tasa de crecimiento (por ejemplo, al menos 50%, o al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, al menos 98%, al menos 99%, o hasta la tasa de crecimiento máxima) y producir el POI durante una fase de baja tasa de crecimiento (por ejemplo, menos del 90%, preferiblemente menos del 80%, menos del 70%, menos del 60%, menos del 50% o menos del 40%, menos del 30%, menos del 20%, menos del 10%, menos del 5%, menos del 3%, menos del 2%, menos del 1%, menos del 0,5%, menos del 0,4%, menos del 0,3% o menos del 0,2% de la tasa de crecimiento máxima) mientras se limita la fuente de carbono, preferiblemente alimentando un medio mínimo definido. Específicamente, el medio mínimo definido no comprende una fuente de carbono inductora de la transcripción.

10 El POI es específicamente una proteína heteróloga, preferiblemente seleccionada de proteínas terapéuticas, que incluyen anticuerpos o fragmentos de las mismas, enzimas y péptidos, antibióticos proteicos, proteínas de fusión de toxinas, conjugados de carbohidratos y proteínas, proteínas estructurales, proteínas reguladoras, vacunas y vacunas como proteínas o partículas, enzimas de proceso, factores de crecimiento, hormonas y citocinas, o un metabolito de un POI, que incluye específicamente un metabolito celular del cultivo celular recombinante que expresa un gen de interés bajo el control transcripcional de un promotor de la invención.

15 Se describe además en este documento un método para identificar un promotor constitutivo de células eucariotas, que comprende los pasos de

- a) cultivar células eucariotas a una alta tasa de crecimiento;
- b) cultivar después las células a una baja tasa de crecimiento;
- 20 c) proporcionar muestras del cultivo celular de la etapa a) y b),

d) realizar el análisis de transcripción en dichas muestras y comparar los niveles de transcripción con los niveles de transcripción del promotor pGAP nativo de las células; y

- f) seleccionar el promotor constitutivo que tiene una fuerza de transcripción más alta en comparación con el promotor pGAP nativo a tasas de crecimiento altas y bajas, preferiblemente determinando un nivel de transcripción del promotor constitutivo identificado que sea al menos 1,1 veces mayor en comparación al promotor pGAP nativo, preferiblemente al menos 1,2 veces, preferiblemente al menos 1,3 veces, preferiblemente al menos 1,4 veces, preferiblemente al menos 1,5 veces, preferiblemente al menos 1,6 veces, preferiblemente al menos 1,7 veces, preferiblemente al menos 1,8 veces, preferiblemente al menos 1,9 veces, preferiblemente al menos 2 veces, preferiblemente al menos 3 veces, preferiblemente al menos 4 veces, preferiblemente al menos 5 veces, preferiblemente al menos 10 veces, o al menos 15 veces más.

25 Específicamente, el nivel de transcripción se determina a una tasa de crecimiento alta y una tasa de crecimiento baja dentro del rango de 0,01 a 0,2 h⁻¹, preferiblemente dentro del rango de 0,015 a 0,15 h⁻¹, p. ej. examinando al menos dos muestras de un cultivo celular, una primera que representa una alta tasa de crecimiento, como al menos 0,05 h⁻¹, preferiblemente al menos 0,06 h⁻¹, o al menos 0,07 h⁻¹, al menos 0,08 h⁻¹, al menos 0,09 h⁻¹, al menos 0,1 h⁻¹, p. ej. a una tasa de crecimiento de 0,15 h⁻¹, y una segunda que representa una tasa de crecimiento baja, como menos que la primera, por ejemplo, menos de 0,05 h⁻¹, preferiblemente menos de 0,04 h⁻¹, menos de 0,03 h⁻¹, o menos de 0,02 h⁻¹, por ejemplo, a una tasa de crecimiento de 0,015 h⁻¹. El nivel de transcripción se determina específicamente mediante el análisis de los patrones de expresión génica utilizando micromatrices de ADN, p. ej. según el ejemplo 11 a continuación.

35 Se describe además en este documento el uso de una secuencia de ácido nucleico aislada de la invención que comprende o que consiste en un promotor que cuando está operativamente unido a una secuencia de nucleótidos que codifica un POI dirige la expresión de la misma en una célula huésped a un nivel de expresión que es más alto que bajo el control del promotor pGAP nativo a tasas de crecimiento altas y bajas, preferiblemente en un método para producir un POI mediante el cultivo de una célula huésped transformada con la secuencia de ácido nucleico, en el que el cultivo se realiza en un biorreactor que comienza con una fase discontinua seguida de una fase discontinua alimentada o una fase de cultivo continuo.

40 Específicamente, el nivel de expresión se determina a una tasa de crecimiento alta y una tasa de crecimiento baja dentro del rango de 0,01 a 0,2 h⁻¹, preferiblemente dentro del rango de 0,015 a 0,15 h⁻¹, p.ej. examinando al menos dos muestras de un cultivo celular, una primera que representa una alta tasa de crecimiento, como al menos 0,05 h⁻¹, preferiblemente al menos 0,06 h⁻¹, o al menos 0,07 h⁻¹, al menos 0,08 h⁻¹, al menos 0,09 h⁻¹, al menos 0,1 h⁻¹, p. ej. a una tasa de crecimiento de 0,15 h⁻¹ y una segunda que representa una tasa de crecimiento baja, tal como menor que la primera, por ejemplo, menor que 0,05 h⁻¹, preferiblemente menor que 0,04 h⁻¹, menos de 0,03 h⁻¹, o menos de 0,02 h⁻¹, por ejemplo, a una tasa de crecimiento de 0,015 h⁻¹. El nivel de expresión se determina específicamente en condiciones de crecimiento limitado. Un ejemplo de determinación de la fuerza de expresión en condiciones de crecimiento limitado, p. ej. en un cultivo de quimiostato con glucosa limitada, proporciona tasas de crecimiento altas y bajas, como en el Ejemplo 10 a continuación.

Específicamente, el nivel de expresión es al menos 1,1 veces mayor en comparación con el promotor pGAP, preferiblemente al menos 1,2 veces, preferiblemente al menos 1,3 veces, preferiblemente al menos 1,4 veces, preferiblemente al menos 1,5 veces, preferiblemente al menos 1,6 veces, preferiblemente al menos 1,7 veces, preferiblemente al menos 1,8 veces, preferiblemente al menos 1,9 veces, preferiblemente al menos 2 veces, preferiblemente al menos 3 veces, preferiblemente al menos 4 veces, preferiblemente al menos 5 veces, preferiblemente al menos 10 veces o al menos 15 veces más.

De acuerdo con un aspecto específico, la secuencia de ácido nucleico aislada de la invención o la construcción de expresión de la invención se usa en un método para producir un POI cultivando una célula huésped transformada con la secuencia de ácido nucleico y/o la construcción de expresión, preferiblemente en donde el cultivo se realiza en un biorreactor que comienza con una fase discontinua seguida de una fase discontinua alimentada o una fase de cultivo continua.

Figuras

Figura 1: Secuencia de ácido nucleico pCS1 (985 pb, SEQ ID 1) de *P. pastoris*, y secuencias promotoras que son secuencias de ADN que incluyen pCS1 y nucleótidos adicionales en el extremo 5', o fragmentos pCS1, que promueven la expresión de CS1 en *P. pastoris*, que comprenden 1488 pb (SEQ ID 2), 767 pb (SEQ ID 3), 500 pb (SEQ ID 4), 344 pb (SEQ ID 5), 234 pb (SEQ ID 6), 138 pb (SEQ ID 7) y 85 pb (SEQ ID 8).

Figura 2: Secuencias de la secuencia de nucleótidos que codifica CS1 y la secuencia de aminoácidos

i) de las cepas GS115, CBS7435 y CBS2612 (PAS_chr1-4_0586), secuencia codificante (SEQ ID 9), secuencia traducida (XM_002490678.1, SEQ ID 10); y

ii) de la cepa DSMZ70382 (PIPA02805), secuencia codificante (SEQ ID 11), secuencia traducida (SEQ ID 12).

Figura 3: Secuencia promotora pGAP nativa de *P. pastoris* (GS115) (SEQ ID 13).

Descripción detallada del invento

Los términos específicos tal como se usan en toda la especificación tienen el siguiente significado. La expresión "fuente de carbono" o "sustrato de carbono", como se usa en el presente documento, significará un sustrato de carbono fermentable, típicamente un carbohidrato fuente, adecuado como fuente de energía para microorganismos, tales como aquellos que pueden ser metabolizados por organismos hospedadores o líneas celulares de producción, en fuentes particulares seleccionadas del grupo que consiste en monosacáridos, oligosacáridos, polisacáridos, alcoholes que incluyen glicerol, en forma purificada, en medios mínimos o provistos en materias primas, como un material nutritivo complejo. La fuente de carbono se puede usar de acuerdo con la invención como una única fuente de carbono o como una mezcla de diferentes fuentes de carbono.

La expresión "condiciones ricas en sustrato de carbono" como se usa en el presente documento se refiere específicamente al tipo y cantidad de un sustrato de carbono adecuado para el crecimiento celular, tal como un nutriente para células eucariotas. La fuente de carbono puede proporcionarse en un medio, tal como un medio basal o medio complejo, pero también en un medio definido (químicamente) que contiene una fuente de carbono purificada. La fuente de carbono para usar en una fase de crecimiento de un proceso de cultivo celular también se denomina en este documento "fuente de carbono basal" y típicamente se proporciona en una cantidad para proporcionar un crecimiento celular, por ejemplo para obtener densidades celulares 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, p. ej. exhibiendo viabilidades de más del 90% durante los pasos del subcultivo estándar, preferiblemente más del 95%.

En una fase de crecimiento, la fuente de carbono se usa típicamente en una cantidad excedente o sobrante, que se entiende como un exceso que proporciona energía para aumentar la biomasa, p. ej. durante el cultivo de una línea celular con una alta tasa de crecimiento específico.

Esta cantidad excedente es particularmente superior a la cantidad limitada de una fuente de carbono como se usa en condiciones de crecimiento limitadas, para lograr una concentración residual en el caldo de fermentación que sea medible y típicamente al menos 10 veces mayor, preferiblemente al menos 50 veces o al menos 100 veces mayor que durante la alimentación del cultivo celular con un medio que contiene sustrato de carbono limitado.

La expresión "condiciones limitadas de sustrato de carbono" o "fuente de carbono limitada", en el presente documento también denominado "fuente de carbono suplementaria", tal como se usa de acuerdo con la invención, se entiende que se refiere específicamente al tipo y cantidad de un sustrato de carbono que facilita la producción de productos de fermentación por líneas celulares de producción, en particular en un proceso de cultivo con tasas de crecimiento controladas inferiores a la tasa de crecimiento máxima. La fase de producción sigue específicamente una fase de crecimiento, p. ej. en procesos de cultivo discontinuo, alimentado por lotes y continuo.

En general, los procesos de cultivo celular se clasifican en cultivo discontinuo, cultivo continuo y cultivo alimentado por lotes. El cultivo por lotes es un proceso de cultivo mediante el cual se agrega una pequeña cantidad de una

- solución de cultivo de semillas a un medio y las células se cultivan sin agregar ningún medio adicional o descargar una solución de cultivo durante el cultivo. El cultivo continuo es un proceso de cultivo mediante el cual se agrega y descarga continuamente un medio durante el cultivo. El cultivo continuo también incluye el cultivo de perfusión. El cultivo alimentado por lotes, que es un intermedio entre el cultivo discontinuo y el cultivo continuo y que también se denomina cultivo semidiscontinuo, es un proceso de cultivo mediante el cual se agrega un medio de forma continua o secuencial durante el cultivo pero, a diferencia del cultivo continuo, la solución de cultivo no se descarga continuamente.
- Se prefiere específicamente un proceso de alimentación por lotes que se basa en la alimentación de un sustrato de nutrientes limitantes del crecimiento a un cultivo. La estrategia de alimentación por lotes, que incluye la fermentación de alimentación única o de alimentación repetida, generalmente se usa en procesos bioindustriales para alcanzar una alta densidad celular en el biorreactor. La adición controlada del sustrato de carbono afecta directamente la tasa de crecimiento del cultivo y ayuda a evitar el metabolismo de desbordamiento o la formación de subproductos metabólicos no deseados. Bajo condiciones limitadas de fuente de carbono, la fuente de carbono específicamente puede estar contenida en la alimentación de un proceso alimentado por lotes. De este modo, el sustrato de carbono se proporciona en una cantidad limitada.
- También en quimiostato o cultivo continuo como se describe en este documento, la tasa de crecimiento puede controlarse estrechamente.
- Una "cantidad limitada" de una fuente de carbono se entiende en este documento como la cantidad de una fuente de carbono necesaria para mantener una línea celular de producción en condiciones de crecimiento limitado, p. ej. en una fase de producción o modo de producción. Tal cantidad limitada puede emplearse en un proceso de alimentación por lotes, donde la fuente de carbono está contenida en un medio de alimentación y se suministra al cultivo a bajas tasas de alimentación para un suministro de energía sostenido, p. ej. para producir un POI, mientras se mantiene la biomasa a tasas de crecimiento específicas bajas. Un medio de alimentación se agrega típicamente a un caldo de fermentación durante la fase de producción de un cultivo celular.
- La cantidad limitada de una fuente de carbono puede, por ejemplo, determinarse por la cantidad residual de la fuente de carbono en el caldo de cultivo celular, que está por debajo de un umbral predeterminado o incluso por debajo del límite de detección medido en un ensayo estándar (carbohidrato). La cantidad residual típicamente se determinaría en el caldo de fermentación al cosechar un producto de fermentación.
- La cantidad limitada de una fuente de carbono también puede determinarse definiendo la velocidad de alimentación promedio de la fuente de carbono al fermentador, p. ej. según lo determinado por la cantidad agregada durante todo el proceso de cultivo, p. ej. la fase de alimentación por lotes, por tiempo de cultivo, para determinar una cantidad promedio calculada por tiempo. Esta tasa de alimentación promedio se mantiene baja para garantizar el uso completo de la fuente de carbono suplementaria por el cultivo celular, p. ej. entre $0,6 \text{ g L}^{-1} \text{ h}^{-1}$ (g fuente de carbono por L volumen de fermentación inicial y tiempo h) y $25 \text{ g L}^{-1} \text{ h}^{-1}$, preferiblemente entre $1,6 \text{ g L}^{-1} \text{ h}^{-1}$ y $20 \text{ g L}^{-1} \text{ h}^{-1}$.
- La cantidad limitada de una fuente de carbono también se puede determinar midiendo la tasa de crecimiento específica, cuya tasa de crecimiento específica se mantiene baja, p. ej. más baja que la tasa de crecimiento específica máxima, durante la fase de producción, p. ej. dentro de un rango predeterminado, como en el rango de $0,001 \text{ h}^{-1}$ a $0,20 \text{ h}^{-1}$, o de $0,02 \text{ h}^{-1}$ a $0,20 \text{ h}^{-1}$, preferiblemente entre $0,02 \text{ h}^{-1}$ y $0,15 \text{ h}^{-1}$.
- Se puede usar cualquier tipo de carbono orgánico adecuado 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 mezcla de los mismos.
- Según una realización específicamente preferida, la fuente de carbono basal se selecciona del grupo que consiste en glucosa, glicerol, etanol, o mezclas de los mismos, y material nutritivo complejo. Según una realización preferida, la fuente de carbono basal es glicerol.
- Según una realización específica adicional, una 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 mezcla de los mismos. Según una realización preferida, una fuente de carbono suplementaria es glucosa.
- Específicamente, el método puede emplear glicerol como fuente de carbono basal y glucosa como fuente de carbono suplementaria.
- Específicamente, un medio de alimentación como se usa en este documento está definido químicamente y no contiene metanol.
- La expresión "químicamente definido" o "definido" con respecto al medio de cultivo celular, tal como un medio mínimo o medio de alimentación en un proceso alimentado por lotes, significará un medio de cultivo adecuado para el cultivo celular in vitro de una línea celular de producción, en el que se conocen todos los componentes químicos y (poli) péptidos. Típicamente, un medio definido químicamente está completamente sin componentes derivados de animales y representa un ambiente de cultivo celular puro y consistente.

- La expresión "línea celular" como se usa en el presente documento se refiere a un clon establecido de un tipo celular particular que ha adquirido la capacidad de proliferar durante un período prolongado de tiempo. La expresión "línea celular huésped" se refiere a una línea celular como se usa para expresar un gen endógeno o recombinante o productos de una ruta metabólica para producir polipéptidos o metabolitos celulares mediados por dichos polipéptidos. Una "línea celular huésped de producción" o "línea celular de producción" se entiende comúnmente como una línea celular lista para usar para el cultivo en un biorreactor para obtener el producto de un proceso de producción, como un POI. La expresión "huésped eucariota" o "línea celular eucariota" significará cualquier célula u organismo eucariota, que pueda cultivarse para producir un POI o un metabolito de la célula huésped. Se entiende bien que la expresión no incluye a los seres humanos.
- 5
- 10 Según las células huésped eucariotas específicamente preferidas de la invención, la célula o línea celular se selecciona del grupo que consiste en mamíferos, insectos, levaduras, hongos filamentosos y líneas celulares vegetales, preferiblemente una levadura.
- Específicamente, la levadura se selecciona del grupo que consiste en *Pichia*, *Candida*, *Torulopsis*, *Arxula*, *Hansenula*, *Yarrowia*, *Kluyveromyces*, *Saccharomyces*, *Komagataella*, preferiblemente una levadura metilotrófica.
- 15 Una levadura específicamente preferida es *Pichia pastoris*, *Komagataella pastoris*, *K. phaffii* o *K. pseudopastoris*.
- La expresión "cultivo celular" o "cultivo", también denominado "fermentación", con respecto a una línea celular huésped significa el mantenimiento de las células en un entorno artificial, por ejemplo, *in vitro*, en condiciones que favorezcan el crecimiento, la diferenciación o la viabilidad continua, en un estado activo o inactivo, de las células, específicamente en un biorreactor controlado de acuerdo con métodos conocidos en la industria.
- 20 Cuando se cultiva un cultivo celular usando los medios de cultivo descritos en este documento, el cultivo celular se pone en contacto con los medios en un recipiente de cultivo o con un sustrato en condiciones adecuadas para soportar el cultivo del cultivo celular. En ciertas realizaciones, un medio de cultivo como se describe en este documento se usa para cultivar las células de acuerdo con técnicas de cultivo celular estándar que son bien conocidas en la técnica. En diversos aspectos de la invención, se proporciona un medio de cultivo que puede usarse para el crecimiento de células eucariotas, específicamente levaduras u hongos filamentosos.
- 25 Los medios de cultivo celular proporcionan los nutrientes necesarios para mantener y hacer crecer las células en un ambiente controlado, artificial e *in vitro*. Las características y composiciones de los medios de cultivo celular varían según los requisitos celulares particulares. Los parámetros importantes incluyen la osmolalidad, el pH y las formulaciones de nutrientes. La alimentación de nutrientes se puede realizar de forma continua o discontinua de acuerdo con los métodos conocidos en la técnica. Los medios de cultivo utilizados según la invención son particularmente útiles para producir proteínas recombinantes.
- 30 Mientras que un proceso por lotes es un modo de cultivo en el que todos los nutrientes necesarios para el cultivo de las células están contenidos en el medio de cultivo inicial, sin suministro adicional de nutrientes adicionales durante la fermentación, en un proceso de alimentación por lotes, después de una fase de lotes, tiene lugar una fase de alimentación en la que uno o más nutrientes son suministrados al cultivo mediante la alimentación. El propósito de la alimentación de nutrientes es aumentar la cantidad de biomasa para aumentar también la cantidad de proteína recombinante. Aunque en la mayoría de los procesos de cultivo el modo de alimentación es crítico e importante, la presente invención que emplea el promotor de la invención no está restringida con respecto a un cierto modo de cultivo.
- 35 En ciertas realizaciones, el método de la invención es un proceso de alimentación por lotes. Específicamente, una célula huésped transformada con una construcción de ácido nucleico que codifica un POI recombinante deseado, se cultiva en un medio en fase de crecimiento y se transita a un medio en fase de producción para producir un POI recombinante deseado.
- 40 El medio de alimentación se puede agregar al medio de cultivo en forma líquida o también en una forma alternativa, como un sólido, p. ej. como una tableta u otro medio de liberación sostenida, o un gas, p. ej. dióxido de carbono. Sin embargo, según una realización preferida, la cantidad limitada de una fuente de carbono suplementaria añadida al medio de cultivo celular, incluso puede ser cero. Preferiblemente, bajo condiciones de un sustrato de carbono limitado, la concentración de una fuente de carbono suplementaria en el medio de cultivo es 0-1 g/L, preferiblemente menos de 0,6 g/L, más preferiblemente menos de 0,3 g/L, más preferiblemente menos de 0,1 g/L, preferiblemente 1-50 mg/L, más preferiblemente 1-10 mg/L, específicamente preferido 1 mg/L o incluso por debajo, como por debajo del límite de detección medido con un ensayo estándar adecuado, por ejemplo determinado como una concentración residual en el medio de cultivo tras el consumo por el cultivo celular en crecimiento.
- 45 En un método preferido, la cantidad limitada de la fuente de carbono proporciona una cantidad residual en el cultivo celular que está por debajo del límite de detección determinado en el caldo de fermentación al final de una fase de producción o en la salida de un proceso de fermentación, preferiblemente al cosechar el producto de fermentación.
- 50
- 55

Preferiblemente, la cantidad limitada de una fuente de carbono suplementaria limita el crecimiento para mantener la tasa de crecimiento específica más baja que la tasa de crecimiento específica máxima, como en el rango de $0,001 \text{ h}^{-1}$ a $0,20 \text{ h}^{-1}$, o de $0,02 \text{ h}^{-1}$ a $0,20 \text{ h}^{-1}$, preferiblemente entre $0,02 \text{ h}^{-1}$ y $0,15 \text{ h}^{-1}$.

5 En otra realización, las células huésped de la presente invención se cultivan en modo continuo, p. ej. un quimiostato. Un proceso de fermentación continua se caracteriza por una velocidad definida, constante y continua de alimentación de medio de cultivo recién preparado en el biorreactor, por lo que el caldo de cultivo se elimina al mismo tiempo del biorreactor a la misma velocidad de eliminación definida, constante y continua. Al mantener el medio de cultivo, la tasa de alimentación y la tasa de extracción en el mismo nivel constante, los parámetros y condiciones de cultivo en el biorreactor permanecen constantes.

10 Se entiende específicamente que un cultivo celular estable como se describe en el presente documento se refiere a un cultivo celular que mantiene las propiedades genéticas, específicamente que mantiene el nivel de producción de POI alto, p. ej. 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 preferido de al menos 50 generaciones. Específicamente, se proporciona una línea celular huésped recombinante estable, que se
15 considera una gran ventaja cuando se usa para la producción a escala industrial.

El cultivo celular de la invención es particularmente ventajoso para métodos a escala de fabricación industrial, p. ej. con respecto tanto al volumen como al sistema técnico, en combinación con un modo de cultivo que se basa en la alimentación de nutrientes, en particular un proceso alimentado por lotes o por lotes, o un proceso continuo o semicontinuo (por ejemplo, quimiostato).

20 El término "expresión" o la expresión "sistema de expresión" o "casete de expresión" se refiere a moléculas de ácido nucleico que contienen una secuencia codificante deseada y secuencias de control en un enlace operable, de modo que los hospedadores transformados o transfectados con estas secuencias son capaces de producir las proteínas codificadas o metabolitos celulares huéspedes. Para efectuar la transformación, el sistema de expresión puede incluirse en un vector; sin embargo, el ADN relevante también puede integrarse en el cromosoma del huésped. La
25 expresión puede referirse a productos de expresión secretados o no secretados, incluidos polipéptidos o metabolitos.

Las "construcciones de expresión" o "vectores" o "plásmido" utilizados en el presente documento se definen como secuencias de ADN que se requieren para la transcripción de secuencias de nucleótidos recombinantes clonadas, es decir, de genes recombinantes y la traducción de su ARNm en un organismo huésped adecuado. Los vectores o
30 plásmidos de expresión generalmente comprenden un origen para la replicación autónoma en las células huésped, marcadores seleccionables (por ejemplo, un gen de síntesis de aminoácidos o un gen que confiere resistencia a antibióticos como zeocina, kanamicina, G418 o higromicina), una serie de sitios de escisión de enzimas de restricción, una secuencia promotora adecuada y un terminador de transcripción, cuyos componentes están unidos operativamente entre sí. Los términos "plásmido" y "vector" tal como se usan en el presente documento incluyen
35 secuencias de nucleótidos que se replican de forma autónoma, así como secuencias de nucleótidos que se integran en el genoma.

El constructo de expresión de la invención comprende específicamente un promotor de la invención, unido operativamente a una secuencia de nucleótidos que codifica un POI bajo el control transcripcional de dicho promotor, cuyo promotor no está asociado de forma nativa con la secuencia de codificación del POI.

40 El término "heterólogo" como se usa en el presente documento con respecto a una secuencia de nucleótidos o aminoácidos o proteína, se refiere a un compuesto que es extraño, es decir, "exógeno", tal como que no se encuentra en la naturaleza, a una célula huésped dada; o que se encuentra naturalmente en una célula huésped dada, por ejemplo, es "endógeno", sin embargo, en el contexto de una construcción heteróloga, por ejemplo empleando un ácido nucleico heterólogo. La secuencia de nucleótidos heteróloga que se encuentra endógenamente
45 también puede producirse de forma no natural, p. ej. con una cantidad mayor a la esperada o mayor a la encontrada naturalmente en la célula. La secuencia de nucleótidos heterólogos, o un ácido nucleico que comprende la secuencia de nucleótidos heterólogos, posiblemente difiere en secuencia de la secuencia de nucleótidos endógenos pero codifica la misma proteína que se encuentra endógenamente. Específicamente, las secuencias de nucleótidos heterólogas son aquellas que no se encuentran en la misma relación con una célula huésped en la naturaleza. Se
50 entiende que cualquier secuencia de nucleótidos recombinante o artificial es heteróloga. Un ejemplo de un polinucleótido heterólogo es una secuencia de nucleótidos no asociada de forma nativa con el promotor de acuerdo con la invención, p. ej. para obtener un promotor híbrido, o unido operativamente a una secuencia de codificación, como se describe en este documento. Como consecuencia, se puede obtener un polinucleótido híbrido o quimérico. Un ejemplo adicional de un compuesto heterólogo es un polinucleótido que codifica POI operativamente unido a un
55 elemento de control transcripcional, por ejemplo, un promotor de la invención, al cual una secuencia codificante de POI endógena, natural, normalmente no está operativamente unida.

El término "variante" como se usa en el presente documento en el contexto de la presente invención se referirá específicamente a cualquier secuencia derivada de una secuencia parental, p. ej. por variación de tamaño, p. ej. por

alargamiento o fragmentación, mutación, hibridación (incluida la combinación de secuencias), o con un grado específico de homología o analogía.

5 La invención proporciona específicamente un promotor que es un promotor de tipo salvaje, p. de *P. pastoris*, o una variante funcionalmente activa del mismo, p. ej. capaz de controlar la transcripción de un gen específico en una célula eucariota recombinante o de tipo salvaje.

10 El término "nativo" como se usa en el presente documento en el contexto de la presente invención se referirá específicamente a una estructura individual o componente de un organismo, que está naturalmente asociado con su entorno. Sin embargo, se entiende bien que las estructuras o componentes nativos pueden aislarse del entorno asociado naturalmente y proporcionarse como estructuras o componentes nativos aislados. Dichas estructuras o componentes nativos aislados también pueden ser de origen artificial o sintético, y aún tener las mismas características que las de origen natural.

15 Aunque algunas realizaciones de la presente invención se refieren a estructuras o componentes nativos, p. ej. en forma aislada, se entiende bien que los materiales, métodos y usos de la invención, p. ej. específicamente referidas a secuencias de ácido nucleico aisladas, secuencias de aminoácidos, construcciones de expresión, células huésped transformadas y proteínas recombinantes, están "hechas por el hombre" y, por lo tanto, no se consideran como resultado de la "ley de la naturaleza".

20 El promotor variante funcionalmente activo puede, p. ej. derivarse de la secuencia promotora pCS1 (SEQ ID 1) mediante mutagénesis, empleando así la secuencia pCS1 como secuencia "parental", para producir secuencias adecuadas para su uso como promotor en líneas celulares recombinantes. Dicho promotor variante puede obtenerse de una biblioteca (pCS1) de secuencias mutantes seleccionando aquellos miembros de la biblioteca con propiedades predeterminadas. Los promotores variantes pueden tener las mismas propiedades o incluso mejores, p. ej. mejores en la fuerza del promotor para soportar la producción de POI, todavía con sustancialmente la misma función y fuerza del promotor a tasas de crecimiento altas y bajas, específicamente entendido en este documento como "función independiente de la tasa de crecimiento".

25 El promotor variante también puede derivarse de secuencias análogas, p. ej. de especies eucariotas distintas de *Pichia pastoris* o de un género distinto de *Pichia*, como *K. lactis*, *Z. rouxii*, *P. stipitis*, *H. polymorpha*. Específicamente, las secuencias promotoras análogas asociadas de forma nativa con genes análogos a los genes de *P. pastoris* correspondientes pueden usarse como tales o como secuencias parentales para producir variantes funcionalmente activas de los mismos. Específicamente, un promotor análogo a pCS1 se caracteriza porque está asociado de forma nativa con un gen análogo a CS1 (véase la secuencia de aminoácidos de las SEQ ID 9 u 11). Las propiedades de tales secuencias promotoras análogas o variantes funcionalmente activas de las mismas pueden determinarse usando técnicas estándar.

35 La variante "funcionalmente activa" de una secuencia de nucleótidos o de promotor como se usa en el presente documento significa específicamente una secuencia mutante, p. ej. resultante de la modificación de una secuencia parental mediante la inserción, eliminación o sustitución de uno o más nucleótidos dentro de la secuencia o en uno o ambos extremos distales de la secuencia, y cuya modificación no afecta (en particular perjudica) la actividad de esta secuencia.

Específicamente, la variante funcionalmente activa de la secuencia promotora según la invención se selecciona del grupo que consiste en

40 - homólogos con al menos aproximadamente 60% de identidad de secuencia de nucleótidos, preferiblemente al menos 70%, al menos 80% o al menos 90% de grado de homología o identidad de secuencia con la secuencia original; y/o

45 - homólogos obtenibles modificando la secuencia de nucleótidos parental, como la secuencia pCS1 o la secuencia de una variante de tamaño utilizada como plantilla para proporcionar mutaciones, p. ej. por inserción, delección o sustitución de uno o más nucleótidos dentro de la secuencia o en uno o ambos extremos distales de la secuencia, con (es decir, que comprende o que consiste en) una secuencia de nucleótidos de 200 pb a 1500 pb o 234 pb a 1488 pb, preferiblemente 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*.

50 Las variantes funcionalmente activas específicamente preferidas son las derivadas de un promotor según la invención mediante modificación, extensión y/o fragmentos de la secuencia del promotor, con (es decir, que comprende o que consiste en) 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, preferiblemente hasta 1500 pb.

55 Una variante funcionalmente activa de las secuencias promotoras parentales como se describe en el presente documento puede obtenerse específicamente a través de métodos de mutagénesis. El término "mutagénesis" como

se usa en el contexto de la presente invención se referirá a un método para proporcionar mutantes de una secuencia de nucleótidos, p. ej. mediante la inserción, eliminación y/o sustitución de uno o más nucleótidos, para obtener variantes de los mismos con al menos un cambio en la región no codificante o codificante. La mutagénesis puede ser por mutación aleatoria, semialeatoria o dirigida al sitio. Típicamente, las grandes bibliotecas de genes aleatorios se producen con una alta diversidad de genes, que pueden seleccionarse de acuerdo con un genotipo o fenotipo específicamente deseado.

Algunas de las variantes funcionalmente activas preferidas del promotor según la invención son variantes de tamaño o específicamente fragmentos de pCS1, preferiblemente aquellas que incluyen el extremo 3' de una secuencia de nucleótidos del promotor, p. ej. una secuencia de nucleótidos derivada de una de las secuencias de nucleótidos del promotor que tiene una longitud e inserciones específicas o una delección de la región terminal 5', p. ej. un alargamiento o corte de la secuencia de nucleótidos en el extremo 5', para obtener una longitud específica con un rango desde el extremo 3' hasta un extremo variable de 5', 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.

La variante de tamaño alargado de la invención comprende preferiblemente uno o más nucleótidos adicionales en el extremo 5' de la secuencia pCS1, p. ej. aquellos que están asociados de forma nativa con la secuencia de pCS1 de tipo salvaje en la célula de origen.

Por ejemplo, una variante funcionalmente activa de pCS1 puede comprender una secuencia de nucleótidos o consistir en una secuencia de nucleótidos seleccionada del grupo que consiste en pCS1a (SEQ ID 2), pCS1b (SEQ ID 3), pCS1c (SEQ ID 4), pCS1d (SEQ ID 5), pCS1e (SEQ ID 6), por lo tanto, una secuencia de nucleótidos dentro del rango de 200-1500 pb. Preferiblemente, la variante funcionalmente activa de pCS1 comprende o consiste en una secuencia de nucleótidos seleccionada del grupo que consiste en pCS1a (SEQ ID 2), pCS1b (SEQ ID 3), pCS1c (SEQ ID 4), pCS1d (SEQ ID 5) y pCS1e (SEQ ID 6).

También se entiende que la variante funcionalmente activa de un promotor de la invención abarca los híbridos de pCS1 o cualquiera de sus variantes funcionalmente activas, en particular cualquiera de las variantes de tamaño parental o secuencias de fragmentos, p. ej. resultante de la combinación con una o más de cualquiera de las secuencias que califican como pCS1 o variantes funcionalmente activas de las mismas, p. ej. al menos dos de tales secuencias parentales, al menos 3, al menos 4 o al menos 5 de las secuencias, p. ej. una combinación de dos o más de las secuencias o fragmentos alargados de pCS1 seleccionados del grupo que consiste en pCS1a, pCS1b, pCS1c, pCS1d, pCS1e, pCS1f y pCS1g; preferiblemente, seleccionados del grupo que consiste en pCS1a, pCS1b, pCS1c, pCS1d, y pCS1e. En otra realización, el híbrido está compuesto por al menos una de las secuencias seleccionadas de pCS1 o cualquiera de sus variantes funcionalmente activas, en particular cualquiera de las secuencias de fragmentos o variantes de tamaño, y una secuencia heteróloga que, p. ej. no está asociada de forma nativa con la secuencia pCS1 en *P. pastoris*.

Se entiende además que la variante funcionalmente activa de un promotor de la invención abarca una secuencia de nucleótidos que se hibrida en condiciones rigurosas con el promotor pCS1 o cualquiera de las variantes o fragmentos de tamaño funcionalmente activos, mutantes o secuencias de ácido nucleico híbrido de estos.

Como se usa en la presente invención, el término "hibridación" o "hibridante" pretende significar el proceso durante el cual dos secuencias de ácido nucleico se unen entre sí con enlaces de hidrógeno estables y específicos para formar una doble cadena en condiciones apropiadas. La hibridación entre dos secuencias complementarias o secuencias suficientemente complementarias depende de las condiciones operativas que se utilizan, y en particular de la rigurosidad. Se puede entender que la rigurosidad denota el grado de homología; cuanto mayor es la rigurosidad, mayor porcentaje de homología entre las secuencias. La rigurosidad puede definirse en particular por la composición base de las dos secuencias nucleicas, y/o por el grado de falta de coincidencia entre estas dos secuencias nucleicas. Variando las condiciones, p. ej. la concentración de sal y temperatura, se puede permitir que una secuencia de ácido nucleico dada se hibride solo con su complemento exacto (alta rigurosidad) o con cualquier secuencia algo relacionada (baja rigurosidad). Aumentar la temperatura o disminuir la concentración de sal puede tender a aumentar la selectividad de una reacción de hibridación.

Como se usa en la presente invención, la frase "hibridación en condiciones de hibridación rigurosas" se entiende preferiblemente que se refiere a la hibridación en condiciones de cierta rigurosidad. En una realización preferida, las "condiciones de hibridación rigurosas" son condiciones en las que la homología de las dos secuencias de ácido nucleico son al menos 70%, preferiblemente al menos 80%, preferiblemente al menos 90%, es decir, en condiciones donde la hibridación solo es posible si la cadena doble obtenida durante esta hibridación comprende preferiblemente al menos 70%, preferiblemente al menos 80%, preferiblemente al menos 90% de enlaces AT y enlaces CG.

La rigurosidad puede depender de los parámetros de reacción, tales como la concentración y el tipo de especies iónicas presentes en la solución de hibridación, la naturaleza y la concentración de los agentes desnaturizantes y/o la temperatura de hibridación. Los expertos en la materia pueden determinar las condiciones apropiadas, p. ej. como se describe en Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1989).

La variante funcionalmente activa de la invención se caracteriza específicamente por exhibir sustancialmente la misma actividad que pCS1.

La expresión "sustancialmente la misma actividad" como se usa en el presente documento se refiere específicamente a la actividad indicada por la fuerza del promotor sustancialmente igual o mejorada, específicamente la expresión o la fuerza transcripcional del promotor, y sus características sustancialmente iguales o mejoradas con respecto a la fuerza del promotor, determinada específicamente independientemente de la tasa de crecimiento de la célula huésped tal como una expresión o fuerza transcripcional sustancialmente igual a pCS1, p. ej. +/- 20% o +/- 10%, y/o mayor que el promotor pGAP nativo de la célula huésped, p. ej. un aumento de al menos 1,1 veces, o un aumento de al menos 1,2 veces, preferiblemente al menos 1,3 veces, preferiblemente al menos 1,4 veces, preferiblemente al menos 1,5 veces, preferiblemente al menos 1,6 veces, preferiblemente al menos 1,7 veces, preferiblemente al menos 1,8 veces, preferiblemente al menos 1,9 veces, y preferiblemente al menos 2 veces más en relación con la fuerza del promotor pGAP, o incluso más alto, por ejemplo al menos 3 veces, al menos 4 veces, al menos 5 veces, al menos 10 veces o al menos hasta una actividad de 15 veces, como se determina en un sistema de prueba adecuado que emplea el mismo tipo de célula huésped, las mismas condiciones de cultivo y el mismo ácido nucleico que codifica un producto de expresión como un POI.

El término "homología" indica que dos o más secuencias de nucleótidos tienen los mismos pares de bases conservadas en una posición correspondiente, hasta cierto punto, hasta un grado cercano al 100%. Una secuencia homóloga de la invención tiene típicamente al menos aproximadamente 60% de identidad de secuencia de nucleótidos, 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% identidad, más preferiblemente al menos aproximadamente 98% o 99% de identidad.

La secuencia promotora homóloga según la invención tiene preferiblemente una cierta homología con cualquiera de las secuencias de nucleótidos del promotor pCS1, pCS1a, pCS1b, pCS1c, pCS1d, pCS1e, pCS1f y pCS1g de *P. pastoris* en al menos partes específicas de la secuencia de nucleótidos, como incluir la región 3' de la secuencia de nucleótidos promotora respectiva, preferiblemente una parte con una longitud específica hasta el extremo 3' de la secuencia de nucleótidos promotora respectiva, como una parte con una longitud de 200 pb a 1500 pb, más preferiblemente de 234 a 1488 pb; preferiblemente al menos 100 pb, preferiblemente 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*. Específicamente, al menos, esas partes son preferiblemente homólogas dentro del intervalo de 300-1000 pb, incluida la secuencia terminal 3' de la secuencia de nucleótidos del promotor respectivo.

Las secuencias análogas se derivan típicamente de otras especies o cepas. Se entiende expresamente que cualquiera de las secuencias promotoras análogas de la presente invención que se derivan de especies distintas de *Pichia pastoris* puede comprender una secuencia homóloga, es decir, una secuencia con una cierta homología como se describe en el presente documento. Por lo tanto, el término "homólogo" también puede incluir secuencias análogas. Por otro lado, se entiende que la invención también se refiere a secuencias análogas y homólogas de las mismas que comprenden una cierta homología.

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 es idéntica a los nucleótidos en la secuencia de ADN, después de alinear la secuencia e introducir huecos, si es necesario, para lograr el porcentaje máximo de identidad de secuencia, y sin considerar ninguna sustitución conservadora como parte de la identidad de secuencia. La alineación con el propósito de determinar el porcentaje de identidad de secuencia de nucleótidos se puede lograr de varias maneras que están dentro de la habilidad en la técnica, por ejemplo, utilizando un software informático disponible al público. Los expertos en la materia pueden determinar los parámetros apropiados para medir la alineación, incluidos los algoritmos necesarios para lograr la máxima alineación en toda la longitud de las secuencias que se comparan.

El término "aislado" o "aislamiento", como se usa en el presente documento con respecto a un ácido nucleico, un POI u otro compuesto, se referirá a dicho compuesto que se ha separado suficientemente del entorno con el que se asociaría naturalmente, para existir en forma "sustancialmente pura". "Aislado" no significa necesariamente la exclusión de mezclas artificiales o sintéticas con otros compuestos o materiales, o la presencia de impurezas que no interfieran con la actividad fundamental, y que puedan estar presentes, por ejemplo, debido a una purificación incompleta. En particular, las moléculas de ácido nucleico aisladas de la presente invención también pretenden incluir aquellas sintetizadas químicamente. Con referencia a los ácidos nucleicos de la invención, a veces se usa la expresión "ácido nucleico aislado" o "secuencia de ácido nucleico aislado". Este término, cuando se aplica al ADN, se refiere a una molécula de ADN que se separa de las secuencias con las que es inmediatamente contigua en el genoma natural del organismo en el que se originó. Por ejemplo, un "ácido nucleico aislado" puede comprender una molécula de ADN insertada en un vector, tal como un vector de plásmido o virus, o integrada en el ADN genómico de una célula u organismo huésped procarionta o eucariota. Específicamente, la expresión "ácido nucleico aislado" según la invención excluye que la secuencia pCS1 esté unida a un ácido nucleico que codifica la proteína CS1. Un

"ácido nucleico aislado" (ya sea ADN o ARN) puede representar además una molécula producida directamente por medios biológicos o sintéticos y separada de otros componentes presentes durante su producción.

La expresión "operativamente unido" como se usa en el presente documento se refiere a la asociación de secuencias de nucleótidos en una sola molécula de ácido nucleico, p. ej. un vector, de tal manera que la función de una o más secuencias de nucleótidos se ve afectada por al menos otra secuencia de nucleótidos presente en dicha molécula de ácido nucleico. Por ejemplo, un promotor está operativamente unido con una secuencia codificante de un gen recombinante, cuando es capaz de efectuar la expresión de esa secuencia codificante. Como otro ejemplo, un ácido nucleico que codifica un péptido señal está operativamente unido a una secuencia de ácido nucleico que codifica un POI, cuando es capaz de expresar una proteína en la forma secretada, como una preforma de una proteína madura o la proteína madura. Específicamente, dichos ácidos nucleicos unidos operativamente entre sí pueden unirse inmediatamente, es decir, sin elementos adicionales o secuencias de ácido nucleico entre el ácido nucleico que codifica el péptido señal y la secuencia de ácido nucleico que codifica un POI.

El término "promotor" como se usa en el presente documento se refiere a una secuencia de ADN capaz de controlar la expresión de una secuencia codificante o ARN funcional. La actividad del promotor puede evaluarse por su eficiencia transcripcional. Esto puede determinarse directamente midiendo la cantidad de transcripción de ARNm del promotor, p. ej. mediante transferencia Northern o indirectamente mediante la medición de la cantidad de producto génico expresada desde el promotor.

El promotor de la invención inicia, regula, o media o controla la expresión de un ADN codificador. El ADN promotor y el ADN codificante pueden ser del mismo gen o de genes diferentes, y pueden ser del mismo organismo o de organismos diferentes.

El promotor de la invención se entiende específicamente como un promotor constitutivo, es decir, un promotor que controla la expresión sin la necesidad de inducción o la posibilidad de represión. Por lo tanto, hay una expresión continua y constante a cierto nivel. Debido a la función única de la alta fuerza del promotor en todas las fases de crecimiento o tasas de crecimiento de la célula huésped, el promotor constitutivo de la invención es particularmente útil en un cultivo alimentado por lotes de la línea celular huésped. Los promotores constitutivos de la técnica anterior tenían la desventaja de baja resistencia en condiciones de crecimiento limitado, p. ej. en procesos de alimentación por lotes, o se usaron en realizaciones donde las células huésped se mantenían a una alta tasa de crecimiento, p. ej. en la fase exponencial media.

La fuerza del promotor de la invención se refiere específicamente a su fuerza de transcripción, representada por la eficiencia del inicio de la transcripción que ocurre en ese promotor con alta o baja frecuencia. A mayor fuerza de transcripción, la transcripción más frecuente ocurrirá en ese promotor. La fuerza del promotor es importante, porque determina la frecuencia con la que se transcribe una secuencia de ARNm dada, dando efectivamente mayor prioridad para la transcripción a algunos genes sobre otros, lo que lleva a una mayor concentración de la transcripción. Un gen que codifica una proteína que se requiere en grandes cantidades, por ejemplo, generalmente tiene un promotor relativamente fuerte. La ARN polimerasa solo puede realizar una tarea de transcripción a la vez y, por lo tanto, debe priorizar su trabajo para ser eficiente. Las diferencias en la fuerza del promotor se seleccionan para permitir esta priorización. Según la invención, el promotor es relativamente fuerte independientemente del metabolismo o la tasa de crecimiento de la célula huésped, p. ej. ambos, durante las fases de altas y bajas tasas de crecimiento de un cultivo celular y específicamente independientes de la fuente de carbono, exhiben un estado de actividad aproximadamente máxima, específicamente a un nivel constante.

La fuerza relativa se determina comúnmente con respecto a un promotor estándar, tal como el respectivo promotor pGAP de la célula utilizada como célula huésped. La frecuencia de la transcripción se entiende comúnmente como la tasa de transcripción, p. ej. según lo determinado por la cantidad de una transcripción en un ensayo adecuado, p. ej. RT-PCR o transferencia Northern. La fuerza de un promotor para expresar un gen de interés se entiende comúnmente como la fuerza de expresión o la capacidad de soportar un alto nivel/tasa de expresión. Por ejemplo, la expresión y/o la fuerza de transcripción de un promotor de la invención se determina en la célula huésped que es *P. pastoris* y se compara con el promotor pGAP nativo de *P. pastoris*.

La velocidad de transcripción puede determinarse por la fuerza de transcripción en un micromatriz, o con PCR cuantitativa en tiempo real (qRT-PCR) donde los datos de micromatrices o q RT-PCR muestran la diferencia del nivel de expresión entre condiciones con alta tasa de crecimiento y condiciones con baja tasa de crecimiento, o condiciones que emplean diferentes composiciones de medios, y una alta intensidad de señal en comparación con el promotor pGAP nativo. Un sistema de prueba adecuado se describe específicamente en el Ejemplo 11 a continuación.

El promotor de la invención ejerce una fuerza de transcripción relativamente alta, reflejada por una velocidad de transcripción o fuerza de transcripción de al menos 110% en comparación con el promotor pGAP nativo en la célula huésped, a veces llamado "promotor pGAP homólogo". Preferiblemente, la velocidad o fuerza de transcripción es al menos 110%, preferiblemente al menos 120%, o al menos 130%, en casos específicamente preferidos al menos 140%, al menos 150%, al menos 160%, al menos 170%, al menos 180%, al menos 190%, al menos 200%, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos

aproximadamente 10 veces o al menos aproximadamente 15 veces, o incluso mayor en comparación con el promotor pGAP nativo, p. ej. determinado en la célula eucariota seleccionada como célula huésped para producir el POI, tal como se determina en condiciones ricas en sustrato de carbono, p. ej. cultivos discontinuos o en condiciones limitadas de sustrato de carbono, p. ej. quimiostato o cultivos por lotes alimentados.

- 5 Preferiblemente, el análisis de transcripción es cuantitativo o semicuantitativo, preferiblemente empleando qRT-PCR, micromatrices de ADN, secuenciación de ARN y análisis de transcriptoma.

La tasa de expresión puede, por ejemplo, estar determinada por la cantidad de expresión de un gen informador, como eGFP, p. ej. como se describe en la sección de Ejemplos a continuación, un sistema de prueba se describe específicamente en el Ejemplo 10. Podría demostrarse que el promotor pCS1 tiene una tasa de transcripción relativamente alta de al menos 110% en comparación con el promotor pGAP nativo, al cultivar un clon en solución.

El promotor de la invención ejerce una fuerza de expresión relativamente alta, reflejada por un nivel de expresión de gen de interés, que es al menos 110% en comparación con el promotor pGAP nativo en la célula huésped, a veces llamado "promotor pGAP homólogo". Preferiblemente, la fuerza de expresión es al menos 110%, preferiblemente al menos 120% o al menos 130%, en casos específicamente preferidos al menos 140%, al menos 150%, al menos 160%, al menos 170%, al menos 180%, al menos 190%, al menos 200%, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 10 veces, o al menos aproximadamente 15 veces, o incluso más en comparación con el promotor pGAP nativo, p. ej. determinado en la célula eucariota seleccionada como célula huésped para producir el POI, tal como se determina en condiciones ricas en sustrato de carbono, p. ej. cultivos discontinuos o en condiciones limitadas de sustrato de carbono, p. ej. quimiostato o cultivos por lotes alimentados.

El promotor pGAP nativo inicia la expresión del gen gap que codifica la gliceraldehído-3-fosfato deshidrogenasa (GAPDH), que es un promotor constitutivo presente en la mayoría de los organismos vivos. GAPDH (EC 1\2\1\12), una enzima clave de la glucólisis y la gluconeogénesis, desempeña un papel crucial en el metabolismo de los carbohidratos catabólicos y anabólicos. Por lo tanto, el promotor pGAP, aunque se entiende que es un promotor constitutivo (como no inducible al alimentarse con fuentes de carbono específicas), es un promotor metabólico que ejerce una fuerza promotora creciente con una tasa de crecimiento creciente. Por lo tanto, el promotor pGAP no es adecuado en un proceso de producción eficiente durante el cultivo de la línea celular huésped en fases que se caracterizan por una baja tasa de crecimiento.

En contraste, el promotor de la invención sorprendentemente mantiene su alta fuerza promotora (esencialmente) a un nivel de transcripción constantemente alto en todas las fases de crecimiento de un cultivo de células huésped.

El promotor pGAP nativo es específicamente activo en una célula eucariota recombinante de manera similar a una célula eucariota nativa de la misma especie o cepa, incluyendo la célula eucariota no modificada (no recombinante) o recombinante. Dicho promotor pGAP nativo se entiende comúnmente como un promotor endógeno, por lo tanto, homólogo a la célula eucariota, y sirve como un promotor estándar o de referencia para fines de comparación.

35 Por ejemplo, un promotor pGAP nativo de *P. pastoris* es la secuencia promotora endógena no modificada en *P. pastoris*, como se usa para controlar la expresión de GAPDH en *P. pastoris*, p. ej. que tiene la secuencia mostrada en la Figura 3: secuencia promotora pGAP nativa de *P. pastoris* (GS115) (SEQ ID 13). Si *P. pastoris* se usa como un huésped para producir un POI de acuerdo con la invención, la fuerza o tasa de transcripción del promotor de acuerdo con la invención se compara con dicho promotor pGAP nativo de *P. pastoris*.

40 Como otro ejemplo, un promotor pGAP nativo de *S. cerevisiae* es la secuencia promotora endógena no modificada en *S. cerevisiae*, como se usa para controlar la expresión de GAPDH en *S. cerevisiae*. Si se usa *S. cerevisiae* como huésped para producir un POI de acuerdo con la invención, la fuerza o tasa de transcripción del promotor de acuerdo con la invención se compara con dicho promotor pGAP nativo de *S. cerevisiae*.

45 Por lo tanto, la expresión relativa o la fuerza de transcripción de un promotor de acuerdo con la invención generalmente se compara con el promotor pGAP nativo de una célula de la misma especie o cepa que se usa como huésped para producir un POI.

Se entiende específicamente que el promotor de la invención preferiblemente no es un promotor metabólico tal como, por ejemplo, un promotor naturalmente unido a un gen que codifica una enzima glucolítica abundante o una enzima de gluconeogénesis, una proteína ribosómica o una enzima tal como una proteasa intracelular o una proteasa secretada por la célula huésped.

Se prefiere específicamente un promotor de la invención, que tiene al menos una fuerza de expresión o fuerza de transcripción de pCS1. La fuerza del promotor comparativo que emplea el promotor pGAP como referencia se puede determinar por medios estándar, como midiendo la cantidad de productos de expresión o la cantidad de transcripciones, p. ej. empleando un micromatriz, transferencia Northern, secuenciación de ARN o qRT-PCR, o bien en un cultivo celular, tal como midiendo la cantidad de productos de expresión génica respectivos en células recombinantes. Las pruebas ejemplares se ilustran en la sección de Ejemplos.

La expresión "proteína de interés (POI)", como se usa en el presente documento, se refiere a un polipéptido o una proteína que se produce mediante tecnología recombinante en una célula huésped. Más específicamente, la proteína puede ser un polipéptido que no se produzca naturalmente en la célula huésped, es decir, una proteína heteróloga, o bien puede ser nativa de la célula huésped, es decir, una proteína homóloga a la célula huésped, pero que es producida, por ejemplo, por transformación con un vector autorreplicante que contiene la secuencia de ácido nucleico que codifica el POI, o tras la integración mediante técnicas recombinantes de una o más copias de la secuencia de ácido nucleico que codifica el POI en el genoma de la célula huésped, o mediante modificación recombinante de una o más secuencias reguladoras que controlan la expresión del gen que codifica el POI, p. ej. de la secuencia promotora. En algunos casos, el término POI, como se usa en el presente documento, también se refiere a cualquier producto de metabolito por la célula huésped mediado por la proteína expresada recombinantemente.

Según una realización específica, el término POI excluiría la proteína CS1 de *Pichia pastoris*, p. ej. caracterizado por cualquiera de las secuencias de aminoácidos de la Figura 2, en particular, si el promotor de la invención está asociado de forma nativa con cualquiera de las secuencias de ácido nucleico que codifican dicha proteína CS1.

Específicamente, el POI como se describe en el presente documento es una proteína eucariota, preferiblemente una proteína de mamífero, específicamente una proteína heteróloga de la célula huésped.

Un POI producido según la invención puede ser una proteína multimérica, preferiblemente un dímero o tetrámero.

Según un aspecto de la invención, el POI es una proteína recombinante o heteróloga, preferiblemente seleccionada de proteínas terapéuticas, que incluyen anticuerpos o fragmentos de las mismas, enzimas y péptidos, antibióticos proteicos, proteínas de fusión de toxinas, conjugados de carbohidratos y proteínas, proteínas estructurales, proteínas reguladoras, vacunas y vacunas como proteínas o partículas, enzimas de proceso, factores de crecimiento, hormonas y citocinas, o un metabolito de un POI.

Un POI específico es una molécula de unión a antígeno tal como un anticuerpo, o un fragmento del mismo. Entre los POI específicos se encuentran anticuerpos tales como anticuerpos monoclonales (mAbs), inmunoglobulina (Ig) o inmunoglobulina clase G (IgG), anticuerpos de cadena pesada (HcAb), o fragmentos de los mismos, tales como las variantes de unión de antígeno de fragmento (Fab), Fd, fragmento variable de cadena sencilla (scFv), o modificadas de estos, como, por ejemplo, dímeros Fv (diacuerpos), trímeros Fv (triacuerpos), tetrámeros Fv o minicuerpos y anticuerpos de dominio único como VH o VHH o V-NAR.

Según una realización específica, un producto de fermentación se fabrica usando el POI, un metabolito o un derivado del mismo.

El POI puede recuperarse específicamente del cultivo celular en forma purificada, p. ej. sustancialmente puro.

La expresión "sustancialmente puro" o "purificado", como se usa en el presente documento, se referirá a una preparación que comprende al menos 50% (p/p), preferiblemente al menos 60%, 70%, 80%, 90% o 95% de un compuesto, como una molécula de ácido nucleico o un POI. La pureza se mide por métodos apropiados para el compuesto (por ejemplo, métodos cromatográficos, electroforesis en gel de poliacrilamida, análisis por HPLC y similares).

El término "recombinante", como se usa en el presente documento, significará "estar preparado o ser el resultado de la ingeniería genética". Por lo 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 ha sido modificado genéticamente para contener una secuencia de ácido nucleico recombinante. Una "proteína recombinante" se produce expresando un ácido nucleico recombinante respectivo en un huésped. Un "promotor recombinante" es una secuencia de nucleótidos no codificante modificada genéticamente adecuada para su uso como un promotor funcionalmente activo como se describe en el presente documento.

Por lo tanto, se identificó un nuevo promotor con propiedades funcionales únicas. Inesperadamente, PAS_chr1-4 (en este documento denominado gen CS1, SEQ ID 9) se identificó mediante análisis de la fuerza de la transcripción en condiciones de proceso de producción como el gen transcrito más fuerte en *P. pastoris*. La proteína CS1 codificada (SEQ ID 10) no es una enzima o proteína glucolítica, pero se predice que se ubicará en la superficie celular a través de un ancla GPI. Aunque la secuencia genómica de 9,43 Mbp de la cepa GS115 de *P. pastoris* se ha determinado y descrito en el documento US20110021378A1, las propiedades de las secuencias individuales, como las secuencias promotoras, no se han investigado en detalle.

Fue sorprendente que dicho promotor pudiera usarse eficazmente de acuerdo con la invención. Los promotores de *Pichia* de la técnica anterior, tales como los utilizados en la producción de POI a escala industrial, se derivaron principalmente de la ruta metabólica del metanol y necesitaban la adición de metanol para inducir la producción de POI, que a menudo no se desea. El promotor y el método de acuerdo con la invención tienen la ventaja de que pueden proporcionar una producción incrementada mediante una expresión mejorada, y tienen un menor riesgo de contaminación debido a la regulación específica del promotor, en particular cuando se usa un medio químicamente definido, sin metanol.

5 Resultó que el promotor de acuerdo con la invención ejercía su mejor actividad principalmente independientemente de si se añadía una cantidad adecuada de sustrato de carbono y medios de cultivo específicos. Como ejemplo, *P. pastoris* podía cultivarse con éxito en condiciones de un proceso de producción industrial. Primero, se emplea un cultivo discontinuo en una fuente de carbono basal, tal como glicerol, seguido de un lote alimentado con alimentación limitada de una fuente suplementaria de carbono, como glucosa. Se toman muestras cerca del final de la primera fase del lote y en condiciones de crecimiento limitadas, p. ej. utilizando una cantidad limitada de la fuente de carbono suplementaria. El análisis de transcriptoma con micromatrices de ADN revela genes específicos que son muy activos en la fuente de carbono suplementaria y en presencia de carbono excedente, es decir, la fuente basal de carbono en exceso. La secuencia del promotor pCS1 se identificó como un promotor sorprendentemente fuerte a 10 tasas de crecimiento altas y bajas. El promotor pGAP comparable de la técnica anterior fue significativamente más débil.

Las características de una fuerte expresión de genes recombinantes en la fuente de carbono basal, y una fuerte expresión en una fuente de carbono suplementaria limitada, pudieron verificarse en los procesos de fermentación.

15 Las secuencias de nucleótidos que pueden usarse como secuencias promotoras constitutivas de acuerdo con la invención, proporcionan una producción mejor de la proteína recombinante y pueden obtenerse de una variedad de fuentes. El origen del promotor de acuerdo con la invención es preferiblemente de una célula de levadura, más preferiblemente de levadura metilotrófica tal como del género *Pichia* o de la especie *P. pastoris*, cuyo promotor puede usarse luego como una secuencia parental para producir variantes adecuadas, p. ej. mutantes o análogos.

20 Se contempla que una serie de células de levadura, en particular de cepas de *Pichia*, puede ser adecuada para obtener secuencias promotoras respectivas o análogos respectivos en diferentes especies.

Las variantes del promotor de *P. pastoris* identificado, incluidas las variantes funcionalmente activas, tales como homólogos y análogos, pueden producirse empleando técnicas estándar. El promotor puede ser modificado p. ej. para generar variantes promotoras con niveles de expresión alterados y propiedades reguladoras.

25 Por ejemplo, se puede preparar una biblioteca promotora por mutagénesis de las secuencias promotoras de acuerdo con la invención, que se puede usar como moléculas parentales, p. ej. para ajustar la expresión génica en células eucariotas mediante el análisis de variantes para su expresión bajo diferentes estrategias de fermentación y seleccionando variantes adecuadas. Se puede usar una biblioteca sintética de variantes, p. ej. para seleccionar un promotor que coincida con los requisitos para producir un POL seleccionado. Tales variantes pueden tener una mayor eficiencia de expresión en células huésped eucariotas y una alta expresión en condiciones limitantes y ricas en fuentes de carbono. 30

Las estrategias de fermentación diferencial distinguen entre una fase de crecimiento y una fase de producción. El crecimiento y/o la producción pueden tener lugar adecuadamente en modo discontinuo, modo alimentado por lotes o modo continuo. Se puede usar cualquier biorreactor adecuado, incluidos reactores por lotes, reactor alimentado por lotes, continuo, reactor de tanque agitado o reactor aéreo.

35 Es ventajoso prever el proceso de fermentación a escala piloto o industrial. La escala de proceso industrial 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.

40 Se prefieren las condiciones de producción a escala industrial, que se refieren, por ejemplo, para un cultivo alimentado por lotes en volúmenes de reactor de 100 L a 10 m³ o mayores, empleando tiempos de proceso típicos de varios días, o procesos continuos en volúmenes de fermentadores de aproximadamente 50 - 1000 L o mayores, con velocidades de dilución de aproximadamente 0,02 - 0,15 h⁻¹.

45 Las técnicas de cultivo adecuadas pueden abarcar el cultivo en un biorreactor que comienza con una fase discontinua, seguida de una fase discontinua alimentada exponencial corta con una tasa de crecimiento específica alta, seguida además por una fase discontinua alimentada con una tasa de crecimiento específica baja. Otra técnica de cultivo adecuada puede abarcar una fase discontinua seguida de una fase de cultivo continuo a una velocidad de dilución baja.

Una realización preferida de la invención incluye un cultivo discontinuo para proporcionar biomasa seguido de un cultivo alimentado por lotes para la producción de POI de alto rendimiento.

50 Se prefiere cultivar la línea celular huésped según la invención en un biorreactor en 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. Es ventajoso proporcionar tales rendimientos de producción de biomasa a escala piloto o industrial.

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

Las fuentes de nitrógeno preferidas incluyen $\text{NH}_4\text{H}_2\text{PO}_4$ o NH_3 o $(\text{NH}_4)_2\text{SO}_4$;

Las fuentes de azufre preferidas incluyen MgSO_4 , o $(\text{NH}_4)_2\text{SO}_4$ o K_2SO_4 ;

Las fuentes de fosfato preferidas incluyen $\text{NH}_4\text{H}_2\text{PO}_4$, o H_3PO_4 o NaH_2PO_4 , KH_2PO_4 , Na_2HPO_4 o K_2HPO_4 ;

5 Otros componentes medios típicos incluyen KCl , CaCl_2 y elementos traza tales como: Fe , Co , Cu , Ni , Zn , Mo , Mn , I , B ;

Preferiblemente, el medio se complementa con vitamina B7;

Un medio de crecimiento típico para *P. pastoris* comprende glicerol, sorbitol o glucosa, $\text{NH}_4\text{H}_2\text{PO}_4$, MgSO_4 , KCl , CaCl_2 , biotina y oligoelementos.

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

15 Preferiblemente, la línea celular huésped se cultiva en un medio mineral con una fuente de carbono adecuada, lo que simplifica aún más el proceso de aislamiento significativamente. Un ejemplo de un medio mineral preferido es uno que contiene una fuente de carbono utilizable (por ejemplo, glucosa, glicerol, sorbitol o metanol), sales que contienen los macro elementos (potasio, magnesio, calcio, amonio, cloruro, sulfato, fosfato) y oligoelementos (cobre, yoduro, manganeso, molibdato, cobalto, zinc y sales de hierro y ácido bórico), y opcionalmente vitaminas o aminoácidos, por ejemplo para complementar las auxotrofías.

20 Las células se cultivan en condiciones adecuadas para efectuar la expresión del POI deseado, que puede purificarse de las células o del medio de cultivo, dependiendo de la naturaleza del sistema de expresión y la proteína expresada, p. ej. si la proteína está fusionada con un péptido señal y si la proteína es soluble o está unida a la membrana. Como entenderá el experto en la materia, las condiciones de cultivo variarán de acuerdo con factores que incluyen el tipo de célula huésped y el vector de expresión particular empleado.

25 Al seleccionar la secuencia promotora adecuada de acuerdo con la invención, opcionalmente en combinación con otras secuencias reguladoras preferidas, es posible proporcionar, en condiciones comparables, al menos la misma, o al menos aproximadamente 1,1 veces, o al menos aproximadamente 1,2 veces, al menos aproximadamente 1,5 veces, al menos aproximadamente 2 veces, al menos aproximadamente 3 veces, al menos aproximadamente 4 veces, al menos aproximadamente 5 veces, al menos aproximadamente 10 veces, o al menos hasta una actividad de aproximadamente 15 veces, p. ej. bajo condiciones de crecimiento limitado en un proceso de alimentación por lotes, como se representa por la actividad del promotor o la fuerza de transcripción, o regulado por la fuerza del promotor en relación con un promotor GAP que es homólogo a la célula de producción, un pGAP nativo o aislado de *P. pastoris*.

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

35 Por ejemplo, la alimentación de la fuente de carbono suplementaria añadida a la fermentación puede comprender una fuente de carbono con hasta 50% en peso de azúcares utilizables. La baja tasa de alimentación del medio suplementario limitará los efectos de la inhibición del producto o subproducto en el crecimiento celular, por lo que será posible un alto rendimiento del producto basado en la provisión de sustrato.

La fermentación se lleva a cabo preferiblemente a un pH que varía de 3 a 7,5.

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

40 En general, los ácidos nucleicos u organismos recombinantes a los que se hace referencia en el presente documento pueden producirse mediante técnicas de recombinación bien conocidas por cualquier experto en la materia. De acuerdo con la presente invención, se pueden emplear técnicas convencionales de biología molecular, microbiología y ADN recombinante dentro de la habilidad de la técnica. Dichas técnicas se explican completamente en la bibliografía. Véase, por ejemplo, Maniatis, Fritsch & Sambrook, "Molecular Cloning: A Laboratory Manual", Cold Spring Harbor, (1982).

45 De acuerdo con una realización preferida de la presente invención, se obtiene una construcción recombinante ligando el promotor y los genes relevantes en un vector o construcción de expresión. Estos genes pueden integrarse de manera estable en el genoma de la célula huésped transformando la célula huésped usando tales vectores o construcciones de expresión.

50 Los vectores de expresión pueden incluir, entre otros, vectores de clonación, vectores de clonación modificados y plásmidos diseñados específicamente. El vector de expresión preferido como se usa en la invención puede ser cualquier vector de expresión adecuado para la expresión de un gen recombinante en una célula huésped y se selecciona dependiendo del organismo huésped. El vector de expresión recombinante puede ser cualquier vector

que sea capaz de replicarse o integrarse en el genoma de los organismos huéspedes, también llamado vector huésped.

5 Los vectores de expresión apropiados comprenden típicamente secuencias reguladoras adicionales adecuadas para expresar ADN que codifica un POI en una célula huésped eucariota. Los ejemplos de secuencias reguladoras incluyen operadores, potenciadores, sitios de unión a ribosomas y secuencias que controlan el inicio y la terminación de la transcripción y traducción. Las secuencias reguladoras pueden estar operativamente unidas a la secuencia de ADN a expresar.

10 Para permitir la expresión de una secuencia de nucleótidos recombinante en una célula huésped, el vector de expresión puede proporcionar el promotor de acuerdo con la invención adyacente al extremo 5' de la secuencia de codificación, p. ej. arriba del GOI o un gen de péptido señal que permite la secreción del POI. La transcripción es regulada e iniciada por esta secuencia promotora.

15 Un péptido señal puede ser un péptido señal heterólogo o un híbrido de un péptido señal nativo y heterólogo, y puede ser específicamente heterólogo u homólogo con el organismo huésped que produce la proteína. La función del péptido señal es permitir que el POI sea secretado para ingresar al retículo endoplasmático. Por lo general, es una cadena peptídica corta (3-60 aminoácidos de longitud) que dirige el transporte de una proteína fuera de la membrana plasmática, lo que facilita la separación y purificación de la proteína heteróloga. Algunos péptidos señal se escinden de la proteína mediante la peptidasa señal después de que las proteínas sean transportadas.

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

20 Se entiende que una secuencia promotora está operativamente unida a una secuencia codificante, si el promotor controla la transcripción de la secuencia codificante. Si una secuencia promotora no está asociada de forma nativa con la secuencia codificante, su transcripción no estará controlada por el promotor en células nativas (de tipo salvaje) o las secuencias se recombinarán con diferentes secuencias contiguas.

25 Para probar la función de las secuencias relevantes, los vectores de expresión que comprenden uno o más de los elementos reguladores pueden construirse para conducir la expresión de un POI, y el rendimiento expresado se compara con construcciones con elementos reguladores convencionales. Se puede encontrar una descripción detallada del procedimiento experimental en los ejemplos a continuación. Los genes identificados pueden amplificarse mediante PCR de *P. pastoris* usando cebadores de nucleótidos específicos, clonados en un vector de expresión y transformados en una línea celular eucariota, p. ej. usando un vector de levadura y una cepa de *P. pastoris*, para la producción a alto nivel de varios POI diferentes. Para estimar el efecto del promotor de acuerdo con la invención sobre la cantidad de POI recombinante así producido, la línea celular eucariota se puede cultivar en experimentos con matraces con agitación y fermentaciones de alimentación o quimiostato en comparación con cepas que comprenden un componente convencional, p. ej. un promotor dependiente del crecimiento, como, por ejemplo, el promotor pGAP estándar en la célula respectiva. En particular, la elección del promotor tiene un gran impacto en la producción de proteínas recombinantes.

El POI puede producirse usando la línea celular recombinante del huésped cultivando un transformante, así obtenido en un medio apropiado, aislando el producto o metabolito expresado del cultivo, y opcionalmente purificándolo por un método adecuado.

40 Los transformantes descritos en este documento se pueden obtener mediante la introducción de dicho vector de ADN, p. ej. ADN plasmídico, en un huésped y seleccionando transformantes que expresan el POI o el metabolito de la célula huésped con altos rendimientos. Las células huésped se tratan para permitirles incorporar ADN extraño mediante métodos utilizados convencionalmente para la transformación de células eucariotas, tales como el método de pulso eléctrico, el método de protoplastos, el método de acetato de litio y métodos modificados de los mismos. *P. pastoris* se transforma preferiblemente por electroporación. Los métodos preferidos de transformación para la captación del fragmento de ADN recombinante por el microorganismo incluyen transformación química, electroporación o transformación por protoplastación. Los transformantes descritos en este documento se pueden obtener mediante la introducción de dicho vector de ADN, p. ej. ADN plasmídico, en un huésped y seleccionando transformantes que expresen la proteína relevante o el metabolito de la célula huésped con altos rendimientos.

45 Se prefieren varios enfoques diferentes para la producción del POI según el método de la invención. Las sustancias pueden expresarse, procesarse y opcionalmente secretarse transformando una célula huésped eucariota con un vector de expresión que contiene ADN recombinante que codifica una proteína relevante y al menos uno de los elementos reguladores como se describió anteriormente, preparando un cultivo de la célula transformada, haciendo crecer el cultivo, induciendo la transcripción y la producción de POI, y recuperando el producto del proceso de fermentación.

55 La célula huésped según la invención se prueba preferiblemente para determinar su capacidad de expresión o rendimiento mediante la siguiente prueba: ELISA, ensayo de actividad, HPLC u otras pruebas adecuadas.

El POI 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 preferido al menos 1 g/L.

5 Se entiende que los métodos descritos en el presente documento pueden incluir además cultivar dichas células huésped recombinantes en condiciones que permitan la expresión del POI, preferiblemente en forma secretada o también como producto intracelular. Un POI producido de forma recombinante o un metabolito de la célula huésped puede aislarse luego del medio de cultivo celular y purificarse adicionalmente mediante técnicas bien conocidas por cualquier experto en la materia.

10 El POI producido de acuerdo con la invención típicamente se puede aislar y purificar usando técnicas de vanguardia, que incluyen el aumento de la concentración del POI deseado y/o la disminución de la concentración de al menos una impureza.

15 Si el POI es secretado por las células, puede aislarse y purificarse del medio de cultivo utilizando técnicas de vanguardia. La secreción de los productos de expresión recombinantes de las células huésped es generalmente ventajosa por razones que incluyen facilitar el proceso de purificación, ya que los productos se recuperan del sobrenadante del cultivo en lugar de la compleja mezcla de proteínas que se produce cuando las células de levadura se rompen para liberar proteínas intracelulares.

Las células transformantes cultivadas también pueden romperse por vía sonora o mecánica, enzimática o química para obtener un extracto celular que contiene el POI deseado, del cual se aísla y purifica el POI.

20 Como métodos de aislamiento y purificación para obtener un polipéptido recombinante o un producto proteico, pueden usarse métodos, como los métodos que utilizan la diferencia en la solubilidad, como la salinidad y la precipitación con solventes, los métodos que utilizan la diferencia en el peso molecular, como la ultrafiltración y la electroforesis en gel, los métodos que utilizan la diferencia en carga eléctrica, como la cromatografía de intercambio iónico, los métodos que utilizan afinidad específica, como la cromatografía de afinidad, los métodos que utilizan la diferencia en hidrofobicidad, como la cromatografía líquida de alto rendimiento en fase inversa, y los métodos que utilizan la diferencia en el punto isoeléctrico, como el enfoque isoeléctrico.

25 El producto altamente purificado está esencialmente sin proteínas contaminantes, y preferiblemente tiene una pureza de al menos 90%, más preferiblemente al menos 95%, o incluso al menos 98%, hasta 100%. Los productos purificados pueden obtenerse por purificación del sobrenadante del cultivo celular o también a partir de restos celulares.

30 Como métodos de aislamiento y purificación, se prefieren los siguientes métodos estándar: disrupción celular (si el POI se obtiene intracelularmente), separación y lavado de células (restos) mediante microfiltración o filtro de flujo tangencial (TFF) o centrifugación, purificación de POI por precipitación o tratamiento térmico, activación de POI por digestión enzimática, purificación de POI por cromatografía, como intercambio iónico (IEX), cromatografía de interacción por hidrófobos (HIC), cromatografía de afinidad, cromatografía de exclusión por tamaño (SEC) o HPLC, precipitación de concentración de POI y lavado por etapas de ultrafiltración.

35 El POI aislado y purificado se puede identificar por métodos convencionales tales como transferencia Western, HPLC, ensayo de actividad o ELISA.

40 El POI puede ser cualquier polipéptido eucariota, procariota o sintético. Puede ser una proteína secretada o una proteína intracelular. La presente invención también 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 corresponden y tienen las características funcionales de una secuencia.

Un POI al que se hace referencia en este documento puede ser un producto homólogo a la célula huésped eucariota o heterólogo, preferiblemente para uso terapéutico, profiláctico, diagnóstico, analítico o industrial.

45 El POI es preferiblemente un polipéptido o proteína recombinante heterólogo, producido en una célula eucariota, preferiblemente una célula de levadura, preferiblemente como proteínas secretadas. Ejemplos de proteínas producidas preferiblemente son inmunoglobulinas, fragmentos de inmunoglobulinas, aprotinina, inhibidor de la vía del factor tisular u otros inhibidores de la proteasa, y precursores de insulina o insulina, análogos de insulina, hormonas de crecimiento, interleucinas, activador de plasminógeno tisular, factor de crecimiento transformante a o b, glucagón, péptido tipo glucagón 1 (GLP-1), péptido tipo glucagón 2 (GLP-2), GRPP, Factor VII, Factor VIII, Factor XIII, factor 1 de crecimiento derivado de plaquetas, albúmina sérica, enzimas, como lipasas o proteasas, o un homólogo funcional, una variante equivalente funcional, un derivado y un fragmento biológicamente activo con una función similar a la proteína nativa. El POI puede ser estructuralmente similar a la proteína nativa y puede derivarse de la proteína nativa mediante la adición de uno o más aminoácidos a uno o ambos extremos C y N-terminal o la cadena lateral de la proteína nativa, sustitución de uno o más aminoácidos en uno o varios sitios diferentes en la secuencia de aminoácidos nativa, delección de uno o más aminoácidos en uno o ambos extremos de la proteína nativa o en uno o varios sitios en la secuencia de aminoácidos, o inserción de uno o más aminoácidos en uno o más

sitios en la secuencia de aminoácidos nativa. Dichas modificaciones son bien conocidas para varias de las proteínas mencionadas anteriormente.

5 También se puede seleccionar un POI entre sustratos, enzimas, inhibidores o cofactores que proporcionen reacciones bioquímicas en la célula huésped, con el objetivo de obtener el producto de dicha reacción bioquímica o una cascada de varias reacciones, p. ej. para obtener un metabolito de la célula huésped. Los productos ejemplares pueden ser vitaminas, tales como riboflavina, ácidos orgánicos y alcoholes, que pueden obtenerse con mayores rendimientos después de la expresión de una proteína recombinante o un POI de acuerdo con la invención.

En general, la célula huésped, que expresa un producto recombinante, puede ser cualquier célula eucariota adecuada para la expresión recombinante de un POI.

10 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.

15 Los ejemplos de células de levadura preferidas usadas como células huésped según la invención incluyen, entre otras, el género *Saccharomyces* (p. ej. *Saccharomyces cerevisiae*), el género *Pichia* (p. ej. *P. pastoris* o *P. methanolica*), el género *Komagataella* (*K. pastoris*, *K. pseudopastoris* o *K. phaffii*), *Hansenula polymorpha* o *Kiuyveromyces lactis*.

La bibliografía más nueva divide y renombra *Pichia pastoris* en *Komagataella pastoris*, *Komagataella phaffii* y *Komagataella pseudopastoris*. En este documento, *Pichia pastoris* se usa como sinónimo de todos, *Komagataella pastoris*, *Komagataella phaffii* y *Komagataella pseudopastoris*.

20 Las células huésped de levadura preferidas se derivan de levadura metilotrófica, tal como de *Pichia* o *Komagataella*, p. ej. *Pichia pastoris*, o *Komagataella pastoris*, o *K. phaffii*, o *K. pseudopastoris*. Los ejemplos del huésped incluyen levaduras 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 Center, Centraalbureau voor Schimmel-cultures, Utrecht, Países Bajos) y DSMZ 70877 (Colección alemana de microorganismos y cultivos celulares), pero también cepas de Invitrogen, como X-33, GS115, 25 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 utilizado con éxito para producir transformantes y expresar genes heterólogos.

30 Una célula huésped de levadura preferida según la invención, tal como una célula huésped de *P. pastoris* o *S. cerevisiae*, contiene una secuencia promotora heteróloga o recombinante, que puede derivarse de una cepa de *P. pastoris* o *S. cerevisiae*, diferente del huésped de producción. En otra realización específica, la célula huésped de acuerdo con la invención comprende una construcción de expresión recombinante de acuerdo con la invención que comprende el promotor que se origina del mismo género, especie o cepa que la célula huésped.

El promotor de la invención se deriva preferiblemente de un gen que codifica una proteína homóloga a la célula huésped.

35 Por ejemplo, un promotor según la invención puede derivarse de levadura, tal como una cepa de *S. cerevisiae*, y usarse para expresar un POI en una levadura. Una realización específicamente preferida se refiere a un promotor de acuerdo con la invención que se origina en *P. pastoris* para su uso en un método para producir un POI recombinante en una línea celular productora huésped de *P. pastoris*. El origen homólogo de la secuencia de nucleótidos facilita su incorporación a la célula huésped del mismo género o especie, permitiendo así la producción estable de un POI, 40 posiblemente con mayores rendimientos en los procesos de fabricación industrial. Además, pueden usarse variantes funcionalmente activas del promotor de otras levaduras adecuadas u otros hongos o de otros organismos tales como vertebrados o plantas.

45 Si el POI es una proteína homóloga a la célula huésped, es decir, una proteína que se produce naturalmente en la célula huésped, la expresión del POI en la célula huésped se puede modular mediante el intercambio de su secuencia promotora nativa con una secuencia promotora de acuerdo con la invención.

50 Este propósito puede lograrse, p. ej. por transformación de una célula huésped con una molécula de ADN recombinante que comprende secuencias homólogas del gen diana para permitir la recombinación específica del sitio, la secuencia promotora y un marcador selectivo adecuado para la célula huésped. La recombinación específica del sitio tendrá lugar para unir operativamente la secuencia promotora con la secuencia de nucleótidos que codifica el POI. Esto da como resultado la expresión del POI a partir de la secuencia promotora de acuerdo con la invención en lugar de a partir de la secuencia promotora nativa.

En una realización específicamente preferida de la invención, la secuencia promotora tiene una actividad promotora incrementada con respecto a la secuencia promotora nativa del POI.

De acuerdo con la invención, se prefiere proporcionar una línea celular huésped de *P. pastoris* que comprenda una secuencia promotora de acuerdo con la invención operativamente unida a la secuencia de nucleótidos que codifica el POI.

5 Según la invención, también es posible proporcionar un vector comodín o una célula huésped según la invención, que comprenda un promotor según la invención y que esté listo para incorporar un gen de interés que codifique un POI. La línea celular comodín es, por tanto, una línea celular huésped preformada, que se caracteriza por su capacidad de expresión. Esto sigue una innovadora estrategia de plataforma "comodín" para la generación de líneas celulares productoras, para la producción de POI, p. ej. utilizando un intercambio de casetes mediado por recombinasa específico del sitio. Dicha nueva célula huésped facilita la clonación de un gen de interés (GOI), p. ej. en puntos calientes predeterminados de expresión genómica en cuestión de días para obtener líneas celulares de producción altamente eficientes y reproducibles.

10 Según una realización preferida, el método según la invención emplea una secuencia de nucleótidos recombinante que codifica el POI, que se proporciona en un plásmido adecuado para la integración en el genoma de la célula huésped, en una única copia o en múltiples copias por célula. La secuencia de nucleótidos recombinante que codifica el POI también puede proporcionarse en un plásmido que se replica de forma autónoma en una sola copia o en múltiples copias por célula.

15 El método preferido según la invención emplea un plásmido, que es un vector de expresión eucariota, preferiblemente un vector de expresión de levadura. Los vectores de expresión pueden incluir, entre otros, vectores de clonación, vectores de clonación modificados y plásmidos diseñados específicamente. El vector de expresión preferido como se usa en la invención puede ser cualquier vector de expresión adecuado para la expresión de un gen recombinante en una célula huésped y se selecciona dependiendo del organismo huésped. El vector de expresión recombinante puede ser cualquier vector que sea capaz de replicarse o integrarse en el genoma de los organismos huésped, también llamado vector huésped, tal como un vector de levadura, que lleva una construcción de ADN de acuerdo con la invención. Un vector de expresión de levadura preferido es para la expresión en levadura seleccionada del grupo que consiste en levaduras metilotróficas representadas por los géneros *Hansenula*, *Pichia*, *Candida* y *Torulopsis*.

20 En la presente invención, se prefiere usar plásmidos derivados de pPICZ, pGAPZ, pPIC9, pPICZalfa, pGAPZalfa, pPIC9K, pGAPHis o pPUZZLE como el vector.

25 De acuerdo con una realización preferida de la presente invención, se obtiene una construcción recombinante ligando los genes relevantes en un vector. Estos genes pueden integrarse de manera estable en el genoma de la célula huésped transformando la célula huésped usando tales vectores. Los polipéptidos codificados por los genes pueden producirse usando la línea celular recombinante del huésped cultivando un transformante, así obtenido en un medio apropiado, aislando el POI expresado del cultivo y purificándolo por un método apropiado para el producto expresado, en particular para separar el POI de las proteínas contaminantes.

30 Los vectores de expresión pueden comprender uno o más marcadores seleccionables fenotípicos, p. ej. un gen que codifica una proteína que confiere resistencia a los antibióticos o que suministra un requerimiento autótrofo. Los vectores de levadura comúnmente contienen un origen de replicación a partir de un plásmido de levadura, una secuencia de replicación autónoma (ARS) o, alternativamente, una secuencia utilizada para la integración en el genoma del huésped, una región promotora, secuencias para poliadenilación, secuencias para la terminación de la transcripción y un marcador de selección.

35 Los procedimientos utilizados para ligar las secuencias de ADN, p. ej. los expertos en la técnica conocen bien la codificación de la secuencia previa y/o el POI, el promotor y el terminador, respectivamente, y cómo insertarlos en vectores adecuados que contienen la información necesaria para la integración o la replicación del huésped, p. ej., los descritos en J. Sambrook et al., (*A Laboratory Manual*, Cold Spring Harbor, 1989).

40 Se entenderá que el vector, que usa los elementos reguladores de acuerdo con la invención y/o el POI como objetivo de integración, puede construirse preparando primero una construcción de ADN que contenga la secuencia de ADN completa que codifica los elementos reguladores y/o el POI y posteriormente insertar este fragmento en un vector de expresión adecuado, o insertando secuencialmente fragmentos de ADN que contienen información genética para los elementos individuales, seguido de la ligadura.

45 También pueden usarse vectores multiclonales, que son vectores que tienen un sitio multiclonal, de acuerdo con la invención, en donde un gen heterólogo deseado puede incorporarse en un sitio multiclonal para proporcionar un vector de expresión. En los vectores de expresión, el promotor se coloca arriba del gen del POI y regula la expresión del gen. En el caso de los vectores multiclonales, debido a que el gen del POI se introduce en el sitio multiclonal, el promotor se coloca arriba del sitio multiclonal.

50 La construcción de ADN que se proporciona para obtener una célula huésped recombinante de acuerdo con la invención puede prepararse sintéticamente mediante métodos estándar establecidos, p. ej. el método de fosforamidita. La construcción de ADN también puede ser de origen genómico o de ADNc, por ejemplo, obtenida preparando una biblioteca genómica o de ADNc y seleccionando secuencias de ADN que codifiquen todo o parte del

5 polipéptido descrito en este documento mediante hibridación usando sondas de oligonucleótidos sintéticos de acuerdo con técnicas estándar (Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1989). Finalmente, el constructo de ADN puede ser de origen mixto sintético y genómico, mixto sintético y ADNc o genómico mixto y ADNc preparado por recocido de fragmentos de origen sintético, genómico o ADNc, según corresponda, los fragmentos correspondientes a varias partes de todo el constructo de ADN, de acuerdo con las técnicas estándar.

En otra realización preferida, el vector de expresión de levadura puede integrarse de manera estable en el genoma de levadura, p. ej. por recombinación homóloga.

10 Una célula huésped transformante descrita en este documento obtenida transformando la célula con los elementos reguladores descritos en este documento y/o los genes POI pueden cultivarse preferiblemente en condiciones para crecer eficientemente a un gran número de células. Cuando se prepara la línea celular para la expresión de POI, se eligen técnicas de cultivo para producir el producto de expresión.

15 Los ejemplos específicos se refieren a la fermentación alimentada por lotes de una línea celular de producción recombinante de *P. pastoris* que produce proteínas indicadoras, empleando un medio discontinuo de glicerol y un medio alimentado por lotes con glucosa. Los estudios comparativos de la actividad del promotor han demostrado que el promotor de acuerdo con la invención puede usarse con éxito para la producción de proteínas recombinantes.

Según otro ejemplo, la albúmina de suero humano (HSA) se produjo como un POI en condiciones límites de glucosa, y se determinó el rendimiento de HSA y el número de copias de genes.

20 Según otro ejemplo, se realizó un cultivo alimentado por lotes de cepas de *P. pastoris* que expresan HSA bajo el control de un promotor según la invención.

Otros ejemplos se refieren a expresar una carboxipeptidasa B porcina como proteína modelo bajo control transcripcional del promotor pCS1.

Sin embargo, un ejemplo adicional se refiere a la expresión de un fragmento de anticuerpo bajo el control transcripcional de pCS1.

25 Un ejemplo adicional se refiere a variantes de tamaño o longitud de un promotor de acuerdo con la invención, como la secuencia alargada de pCS1 pCS1a (SEQ ID 2) que comprende la secuencia de pCS1 y un alargamiento en el extremo 5', o fragmentos de pCS1 con un longitud en el rango de aproximadamente 80 pb a 800 pb.

30 La descripción anterior se entenderá más completamente con referencia a los siguientes ejemplos. Sin embargo, tales ejemplos son meramente representativos de los métodos para practicar una o más realizaciones de la presente invención y no deben leerse como limitantes del alcance de la invención.

Ejemplos

Los ejemplos a continuación ilustran los materiales y métodos utilizados para identificar un nuevo promotor y analizar sus propiedades de expresión en *Pichia pastoris*.

Ejemplo 1: Identificación de un gen fuertemente expresado en *P. pastoris*

35 Para identificar un gen fuerte y su respectivo promotor de *P. pastoris*, el análisis de los patrones de expresión génica se realizó utilizando micromatrices de ADN. Se analizaron las células de *P. pastoris* cultivadas en un lote de glicerol y en el límite de glucosa (quimiostato).

a) Cepa

40 Se utilizó una cepa de *P. pastoris* de tipo salvaje (CBS2612, Centro de Biodiversidad Fúngica CBS-KNAW, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos), que puede crecer en medios mínimos sin suplementos.

b) Cultivo de *P. pastoris*

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

Se utilizaron los siguientes medios:

45 Solución madre de sales PTM₁ traza contenida por litro

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

Medio del lote de glicerol contenido por litro

2 g de monohidrato de ácido cítrico ($C_6H_8O_7 \cdot H_2O$), 39,2 g de glicerol, 20,8 g $NH_4H_2PO_4$, 0,5 g $MgSO_4 \cdot 7H_2O$, 1,6 g KCl, 0,022 g $CaCl_2 \cdot 2H_2O$, 0,8 mg biotina y 4,6 ml de solución madre de sales traza PTM1. Se añadió HCl para establecer el pH a 5.

5 Medio alimentado con glicerol contenido por litro

632 g de glicerol, 8 g de $MgSO_4 \cdot 7H_2O$, 22 g de KCl y 0,058 g de $CaCl_2 \cdot 2H_2O$.

Medio quimiostato contenido por litro

10 2 g de monohidrato de ácido cítrico ($C_6H_8O_7 \cdot H_2O$), 99,42 g de monohidrato de glucosa, 22 g de $NH_4H_2PO_4$, 1,3 g de $MgSO_4 \cdot 7H_2O$, 3,4 g de KCl, 0,02 g de $CaCl_2 \cdot 2H_2O$, 0,4 mg de biotina y 3,2 ml de solución de reserva de sales traza de PTM1. Se añadió HCl para establecer el pH a 5.

El oxígeno disuelto se controló a $DO = 20\%$ con la velocidad del agitador (500-1250 rpm). La velocidad de aireación fue de $60 L h^{-1}$ de aire, la temperatura se controló a $25^\circ C$ y el punto de ajuste de pH de 5 se controló con la adición de NH_4OH (25%).

15 Para comenzar la fermentación, se filtraron estériles 1,5 ml de medio en el fermentador y se inoculó *P. pastoris* (de un pre-cultivo nocturno en YPG, 180 rpm, $28^\circ C$) con una densidad óptica inicial (OD_{600}) de 1. La fase del lote de aproximadamente 25 h alcanzó una concentración de biomasa seca de aproximadamente 20 g/L, fue seguida por un lote de alimentación exponencial de 10 h con medio de glucosa, lo que condujo a una concentración de biomasa seca de aproximadamente 50 g/L. Luego, el volumen se redujo a 1,5 L y el cultivo del quimiostato se inició con una tasa de alimentación/cosecha de $0,15 L h^{-1}$, lo que dio como resultado una tasa de crecimiento constante de $\mu = 0,1$.
20 La fermentación se terminó 50 h después del inicio del quimiostato.

Esta fermentación se ha realizado tres veces para obtener las réplicas biológicas necesarias para un análisis confiable de micromatrices.

Las condiciones limitadas de carbono (sin glucosa residual detectable) durante el quimiostato se verificaron mediante análisis por HPLC del sobrenadante de cultivo.

25 c) Muestreo

Se tomaron muestras al final de la fase discontinua de glicerol y en condiciones estables del quimiostato de glucosa. El muestreo de rutina como determinación de densidad óptica o masa seca de levadura, la inspección microscópica cualitativa y el análisis de viabilidad celular se realizó junto con cada fermentación. Para el análisis de micromatrices, se tomaron muestras y se trataron de la siguiente manera: para un enfriamiento óptimo, se mezclaron inmediatamente 9 ml de caldo de cultivo celular con 4,5 ml de solución de fenol al 5% (Sigma) enfriada con hielo (en etanol abs.) y se dividieron en alícuotas. Cada 2 ml se centrifugaron (13200 rpm durante 1 minuto) en tubos de recolección preenfriados (GE healthcare, NJ), el sobrenadante se eliminó por completo y los tubos se almacenaron a $-80^\circ C$ hasta la purificación de ARN.
30

d) Purificación de ARN y preparación de muestras para hibridación de micromatrices

35 El ARN se aisló utilizando reactivo TRI de acuerdo con las instrucciones del proveedor (Ambion, EE.UU.). Los sedimentos celulares se resuspendieron en reactivo TRI y se homogeneizaron con perlas 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 mediante la adición de isopropanol. El sedimento se lavó con etanol al 70%, se secó y se volvió a suspender en agua sin ARNasa. Las concentraciones de
40 ARN se determinaron midiendo OD_{260} usando un espectrofotómetro Nanodrop 1000 (productos NanoDrop, DE). El ADN restante de las muestras se eliminó usando el kit sin ADN (Ambion, CA). Se diluyó un volumen de muestra igual a 10 g de ARN a 50 μl en agua sin ARNasa, luego se añadieron el tampón de ADNsa I y la ADNrsa I y se incubaron a $37^\circ C$ durante 30 minutos. Después de la adición del reactivo de inactivación de ADNsa, la muestra se centrifugó y el sobrenadante se transfirió a un tubo nuevo. Las concentraciones de ARN se determinaron
45 nuevamente como se describió anteriormente. Además, la integridad del ARN se analizó utilizando nano chips de ARN (Agilent). Para controlar el flujo de trabajo de micromatrices desde la amplificación y el marcaje hasta la hibridación de las muestras, se usó el kit Spike In (Agilent, número de producto: 5188-5279) como control positivo. Contiene 10 transcripciones poliadeniladas diferentes de un adenovirus, que se amplifican, marcan y cohibridan
50 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, Prod. N°.: 5190-0444). Por lo tanto, 500 ng de ARN de muestra purificada se diluyeron en 8,3 μl de agua sin ARNasa, 2 μl de Spike A o B, y se añadieron 1,2 μl del cebador promotor T7. La mezcla se desnaturalizó durante 10 minutos a $65^\circ C$ y se mantuvo en hielo durante 5 minutos. Luego se agregaron 8,5 μl de la mezcla maestra de ADNc (por muestra: 4 μl 5x tampón de la primera cadena, 2 μl DTT 0,1 M, 1 μl mezcla dNTP 10 mM, 1 μl de MMLV-RT, 0,5 μl de RNase out), se incubaron a $40^\circ C$ durante 2 horas, luego se transfirieron a $65^\circ C$ durante 15
55 minutos y se pusieron en hielo durante 5 minutos. La mezcla maestra de transcripción (por muestra: 15,3 μl de agua

sin nucleasas, 20 µl tampón de transcripción, 6 µl DTT 0,1 M, 6,4 µl de PEG al 50%, 0,5 µl de inhibidor de ARNasa, 0,6 µl de fosfatasa inorg., 0,8 µl de T7 ARN polimerasa, 2,4 µl de cianina 3 o cianina 5) se preparó y se añadió a cada tubo y se incubó a 40°C durante 2 horas. Para purificar el cRNA marcado obtenido, se usó el Mini Kit RNeasy (Qiagen, Cat. No. 74104). Las muestras se almacenaron a -80°C. La cuantificación de la concentración de cRNA y la eficacia del marcado se realizó en el espectrofotómetro Nanodrop.

e) Análisis de micromatrices

El kit de hibridación de expresión génica (Agilent, cat. N°. 5188-5242) se usó para la hibridación de los ARNc de la muestra marcada. Para la preparación de las muestras de hibridación, cada 300 ng de cRNA (Cy3 y Cy5) y 6 µl 10 veces el agente de bloqueo se diluyeron con agua sin nucleasas hasta un volumen final de 24 µl. Después de la adición de 1 µl 25 veces el tampón de fragmentación, la mezcla se incubó a 60°C durante 30 minutos. Luego se añadieron 25 µl de tampón de hibridación GEx HI-RPM para detener la reacción. Después de centrifugar durante un minuto a 13.200 rpm, la muestra se enfrió en hielo y se usó para la hibridación inmediatamente. Se utilizaron conjuntos de oligonucleótidos específicos de *P. pastoris* diseñados internamente (AMAD-ID: 026594, conjuntos personalizados 8x15K, Agilent). La hibridación de micromatrices se realizó de acuerdo con la Guía del usuario de la cámara de hibridación de micromatrices (Agilent G2534A). Primero, se descubrió el portaobjetos de la junta y se colocó sobre la base de la cámara, con la etiqueta de Agilent hacia arriba. La muestra (40 µl por matriz) se cargó en el medio de cada uno de los ocho cuadrados. Luego, el portaobjetos de micromatrices se colocó cuidadosamente sobre el portaobjetos de la junta (con la etiqueta de Agilent hacia abajo) y la tapa de la cámara se colocó y se fijó con la abrazadera. La hibridación se realizó en el horno de hibridación durante 17 horas a 65°C. Antes de escanear, se lavó el chip de micromatrices. Por lo tanto, la cámara se desmontó y los portaobjetos se separaron uno del otro mientras se sumergían en el tampón de lavado 1. El micromatriz se transfirió directamente a otra placa con el tampón de lavado 1, se lavó durante 1 minuto, se transfirió al tampón de lavado 2 (temperatura de al menos 30°C) y se lavó durante otro minuto. Después de secar el portaobjetos de micromatrices tocando el borde del portaobjetos con un pañuelo, se colocó en el soporte del portaobjetos (etiqueta de Agilent hacia arriba). Se colocó el portaobjetos en el carrusel y se inició el escaneo.

f) Adquisición de datos y evaluación estadística de datos de micromatrices.

Las imágenes se escanearon a una resolución de 50 nm con un escáner de micromatrices G2565AA (Agilent) y se importaron al software Agilent Feature Extraction 9.5. Se utilizó Agilent Feature Extraction 9.5 para la cuantificación de las intensidades puntuales. Los datos brutos de intensidad de punto medio se importaron luego al software de código abierto R para una mayor normalización y análisis de datos.

Para el preprocesamiento y la normalización de los datos, se utilizaron los paquetes R limma, vsn y m. Los datos de intensidad no se corrigieron con el fondo y se normalizaron con VSN.

Los datos de micromatrices se examinaron en busca de entradas con alta intensidad de señal en ambos estados para identificar genes constitutivos fuertemente expresados. El gen más fuertemente transcrito se muestra en la Tabla 1, con la intensidad de la señal en ambos estados. Los datos de pGAP y pTEF se agregan como referencias. En promedio, pCS1 excedió el pGAP en ambas condiciones (lote de glicerol y cultivo de quimiostato con glucosa limitada) en aproximadamente un 30%, mientras que el pTEF es aproximadamente un 12% más débil que el pGAP.

Tabla 1: Datos de micromatrices de los genes, cuyos promotores se seleccionaron para una caracterización adicional y de pGAP y pTEF como controles

Promotor	Identificador del gen	Intensidad ¹	% de intensidad de pGAP/fuerza de transcripción	Intensidad ²	% de intensidad de pGAP/fuerza de transcripción
pGAP	PAS_chr2-1_0437	56235,2	100,0	44411,4	100,0
pTEF	PAS_FragB_0052	47046,3	83,4	39956,2	89,9
pCS1	PAS_chr1-4_0586	83570,4	148,6	54723,2	122,7
¹ en fase discontinua de glicerol (promedio de ambos canales)					
² en quimiostato con glucosa limitada (promedio de ambos canales)					

Ejemplo 2: Estudios comparativos de la actividad promotora del promotor pCS1 recientemente identificado en *P. pastoris* usando eGFP como gen indicador expresado intracelularmente

Con el fin de analizar las propiedades del promotor recientemente identificado, se realizaron pruebas del matraz de agitación de la siguiente manera: se realizó un cultivo previo durante 24 horas con medio rico que contenía glicerol como fuente de carbono, simulando la fase discontinua del proceso, lo que fue seguido por el cultivo principal con medio mínimo y perlas de alimentación de glucosa, que simulaban la fase discontinua en alimentación limitada en

glucosa del proceso. La proteína de fluorescencia verde (eGFP) se usó como proteína indicadora intracelular para la actividad promotora.

a) Cepa y vector de expresión

5 La cepa *P. pastoris* de tipo salvaje (CBS2612, CBS-KNAW Fungal Biodiversity Center, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) se utilizó como cepa huésped. La transformación de la cepa se realizó con un vector propio denominado pPUZZLE (Stadlmayr et al. J. Biotechnol 2010; 150 (4): 519-29), que comprendía un origen de replicación para *E. coli* (pUC19), un casete de resistencia a antibióticos (gen *Sh ble* que confiere resistencia a Zeocin) para la selección en *E. coli* y levadura, un casete de expresión para el gen de interés (GOI) que consistía en un sitio de clonación múltiple y el terminador de la transcripción CYC1 de *S. cerevisiae*, y un locus para la integración en el genoma de *P. pastoris* (región 3AOX1).

b) Amplificación y clonación del promotor pCS1 recientemente identificado en el vector de expresión pPUZZLE que contenía eGFP como GOI

15 El promotor pCS1 (SEQ ID) comprende 985 pb de la región 5' no codificante del gen CS1 (ver Ejemplo 1) hasta el codón de inicio ATG y se amplificó por PCR (Phusion Polymerase, New England Biolabs) del ADN genómico de *P. pastoris* usando los cebadores que se muestran en la Tabla 2. La secuencia se clonó en el vector de expresión pPUZZLE pPM1aZ10_eGFP, se cortó con *Apal* y *Sbfl*, dando como resultado pPM1aZ10_pCS1_eGFP. Además, el vector pPM1aZ10_pGAP_eGFP, que contenía el promotor comúnmente utilizado del promotor de la gliceraldehído 3-fosfato deshidrogenasa (pGAP de *P. pastoris*, en este documento SEQ ID 13) se usó como referencia. Los promotores se insertaron arriba del codón de inicio del gen eGFP usando los sitios de restricción *Apal* y *Sbfl* (ver 20 Tablas 2 y 3). La corrección de las secuencias promotoras se verificó mediante secuenciación de Sanger.

Tabla 2: Cebadores para la amplificación por PCR de los promotores.

Nombre	Diana	Secuencia	T _M	Sitio de restricción
pCS1_fw	pCS1	SEQ ID 14 GATAGGGCCCCAGGGCATCATTGAGGTTT CCAC	69,6	<i>Apal</i>
pCS1_back	pCS1	SEQ ID 15 GATACCTGCAGGTTTTGTTGTTGAGTGAAG CGAGTG	69,3	<i>Sbfl</i>

Tabla 3: Cebadores de amplificación, enzimas de clonación y la longitud del promotor clonado

Promotor	cebador 5'	cebador 3'	Clonación de enzima		longitud
			5'	3'	
pCS1	pCS1_fw	pCS1_back	<i>Apal</i>	<i>Sbfl</i>	985

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

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

30 Se seleccionaron 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), placas que contenían 25 µg/ml de Zeocin (Invivogen, CA). Se usó PCR de colonia para asegurar la presencia del plásmido transformado. Por lo tanto, el ADN genómico se obtuvo cocinando y congelando las colonias de *P. pastoris* durante 5 minutos cada una y se aplicó directamente para la PCR con los cebadores apropiados. Para el cribado de la expresión, se inoculó una única colonia en medio líquido YPG-Zeo (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de Zeocin) como cultivo previo. Después de aproximadamente 24 h, el precultivo se usó para inocular el cultivo principal con una OD600 de 0,1 en 2 ml de medio de cribado sintético (por litro: 22 g de monohidrato de glucosa, 22 g de ácido cítrico, 3,15 g de (NH₄)₂HPO₄, 0,027 g de CaCl₂·2H₂O, 0,9 g de KCl, 0,5 g de MgSO₄·7H₂O, 2 ml de 500 x biotina y 1,47 ml de solución madre de sales traza [por litro: 6 g CuSO₄·5H₂O, 0,08 g NaI, 3 g MnSO₄·H₂O, 0,2 g Na₂MoO₄·2H₂O, 0,02 g H₃BO₃, 0,5 g CoCl₂, 20 g ZnCl₂, 5 g FeSO₄·7H₂O y 5 ml H₂SO₄]; pH ajustado a 5 con KOH 5M; esterilizado por filtración) y 2 cuartos de perlas de alimentación de glucosa (segunda perla de alimentación añadida después de 24 horas; Kuhner, CH). Las condiciones de crecimiento limitantes de glucosa se lograron debido a la cinética de liberación lenta de glucosa de estas perlas de alimentación, que se describe mediante la siguiente ecuación: 40 (Glucosa) = 1,63*t^{0,74} [mg/Disco]. Se tomaron muestras al final del pre-cultivo, y 24 y 48 horas después de la inoculación del cultivo principal. La densidad celular se determinó midiendo OD600, la expresión de eGFP se analizó por citometría de flujo como se describe en Stadlmayr et al. (J. Biotechnology 2010 dic; 150 (4): 519-29). Para cada

muestra se analizaron 10.000 células. La auto-fluorescencia de *P. pastoris* se midió usando células de tipo salvaje de *P. pastoris* no transformadas y se sustrajo de la señal. Los niveles relativos de expresión de eGFP (intensidad de fluorescencia relacionada con el tamaño celular) se muestran como porcentaje del nivel de expresión de eGFP de un clon que expresa eGFP bajo el control del promotor pGAP constitutivo.

- 5 Los resultados se muestran en la Tabla 4. El clon que se expresó bajo el control del promotor pCS1 excedió el pGAP en un 38% en el final previo al cultivo (lote) y tenía niveles de expresión de GFP 4 veces más altos en el final del cultivo principal (lote alimentado).

Tabla 4: Promedio de fluorescencia de GFP por tamaño de célula de clones de *P. pastoris* que expresan eGFP bajo el control del nuevo promotor pCS1. Los datos se muestran en relación con pGAP en el mismo punto de tiempo.

	Pre-cultivo		Cultivo principal	
	Final del lote	desv. st.	48 h	desv. st.
pCS1	138,2	26,6	400,7	104,87
pGAP	100,0		100,0	

- 10 d) Determinación del número de copias del gen eGFP (GCN) de los clones de *P. pastoris* del Ejemplo 2c

GCN representa el número de casetes de expresión de proteína informadora integrados en el genoma de *P. pastoris*. La determinación de GCN de los clones que expresaban eGFP bajo el control de pCS1 o pGAP se realizó como se describe en el Ejemplo 5 más adelante. Los niveles de expresión de eGFP se analizaron como en el Ejemplo 2c. A modo de ejemplo, los resultados de un clon de cada promotor se muestran en la Tabla 5. Los clones que expresaban eGFP bajo el control del nuevo promotor pCS1 produjeron dos cantidades de eGFP en comparación con los clones que se expresaban bajo el promotor pGAP con el mismo GCN en cultivos de cribado.

Tabla 5: Expresión de eGFP en cultivos de cribado bajo el control de pGAP o pCS1 correlacionado con GCN. Por GCN, los clones pCS1 expresan el doble de la cantidad de eGFP que los clones pGAP.

Clon	FL/Tamaño de célula	GCN	Fluorescencia/Tamaño de célula/GCN
pGAP_eGFP#2	6,0	1	6,0
pCS1_eGFP#4	13,0	1	13,0

- e) Análisis de la fuerza del promotor pCS1 en la fermentación alimentada por lotes

- 20 Para evaluar la actividad del promotor pCS1 en condiciones similares al proceso de producción, se realizaron cultivos por lotes alimentados de un clon pCS1 (pCS1_eGFP N° 4) y un clon pGAP (pGAP_eGFP N° 2), cada uno de los cuales albergaba una copia del casete de expresión eGFP (ver Ejemplo 2d).

Las fermentaciones alimentadas por lotes se realizaron en reactores DASGIP con un volumen de trabajo final de 1,0 L.

- 25 Se utilizaron los siguientes medios:

Solución de reserva de sales traza PTM₁ contenida por litro

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

Medio del lote de glicerol contenido por litro

- 30 2 g de monohidrato de ácido cítrico (C₆H₈O₇·H₂O), 39,2 g de glicerol, 12,6 g de NH₄H₂PO₄, 0,5 g MgSO₄·7H₂O, 0,9 g KCl, 0,022 g CaCl₂·2H₂O, 0,4 mg de biotina y 4,6 ml de solución madre de sales traza de PTM₁. Se añadió HCl para ajustar el pH a 5.

Medio alimentado por lotes con glucosa contenido por litro

- 35 464 g de monohidrato de glucosa, 5,2 g de MgSO₄·7H₂O, 8,4 g KCl, 0,28 g CaCl₂·2H₂O, 0,34 mg de biotina y 10,1 ml de solución madre de sales traza de PTM₁.

El oxígeno disuelto se controló a DO = 20% con la velocidad del agitador (400 - 1200 rpm). La velocidad 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 (25%).

Para comenzar la fermentación, se filtraron estériles 400 ml de medio del lote en el fermentador y se inoculó el clon pCS1_eGFP N° 1 de *P. pastoris* (del pre-cultivo) con una densidad óptica inicial (OD600) de 1. La fase del lote de aproximadamente 25 h (que alcanza una concentración de biomasa seca de aproximadamente 20 g/L) fue seguida por un lote de alimentación limitada en glucosa que comenzó con una alimentación exponencial durante 7 h y una velocidad de alimentación constante de 15 g/L durante 13 h, lo que lleva a una concentración final de biomasa seca de aproximadamente 110 g/L. Se tomaron muestras durante la fase discontinua y alimentada por lotes, y se analizaron para determinar la expresión de eGFP usando un lector de placas (Infinite 200, Tecan, CH). Por lo tanto, las muestras se diluyeron a una densidad óptica (OD600) de 5. Las fermentaciones se realizaron por duplicado. Los resultados se muestran en la Tabla 6 como fluorescencia relativa por biorreactor (FL/r). El clon que se expresaba bajo control del promotor pCS1 tenía en promedio una expresión de eGFP 4,2 veces mayor en comparación con pGAP durante todo el proceso de fermentación.

Tabla 6: Fluorescencia relativa por biorreactor de dos clones diferentes de *P. pastoris* que expresan eGFP bajo el control de pGAP o pCS1 en una fermentación de alimentación por lotes optimizada. t indica el tiempo de alimentación.

	pGAP_eGFP N° 2	pCS1_eGFP N° 1		comparación FL/r	
t [h]	FL/r	DESV. ST.	FL/r	DESV. ST.	pCS1/pGAP
-4,5	2,1	0,1	9,3	0,6	4,4
0,8	4,7	0,1	16,8	0,9	3,6
2,3	5,9	1,0	20,2	1,5	3,4
4,3	9,2	0,1	35,3	1,4	3,8
6,8	14,2	0,6	57,6	4,0	4,1
17,9	57,6	4,2	288,2	15,6	5,0
19,9	86,0	16,6	337,3	26,5	3,9
21,6	73,9	9,7	379,9	41,7	5,1

f) Actividad promotora de pCS1 en diferentes condiciones de crecimiento y sustratos

Para obtener más información sobre la actividad promotora de pCS1 en diferentes composiciones de medios y condiciones de crecimiento, la cepa pCS1_eGFP # 4 se cultivó en medio YP con diferentes fuentes de carbono, a diferentes valores de pH y en medio mínimo sintético. Se tomaron muestras 24 y 48 horas después de la inoculación del cultivo principal y se analizaron por citometría de flujo como se describe en el Ejemplo 2c.

Se inoculó una colonia única de pCS1-eGFP # 4 o pGAP-eGFP # 2 como referencia en medio líquido YPG-Zeo (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de Zeocin) como pre-cultivo. Después de aproximadamente 24 h, el pre-cultivo se usó para inocular el cultivo principal con una OD600 de 0,1 en 2 ml de medio de cultivo principal. Los principales medios de cultivo fueron los siguientes: YP por litro: 20 g de peptona, 10 g de extracto de levadura, pH 7,0-7,5; YPD: YP + 2% de glucosa, YPG: YP + 2% de glicerol; YPM: YP + 1% metanol; YPE: YP + 1% de etanol; Perlas de alimentación YP: perlas de alimentación de glucosa YP + 1 (Kuhner, CH, diámetro 6 mm); YPD pH 4,5: YPD ajustado a pH 4,5 con HCl; SCD: medio de cribado sintético (por litro: 22 g de monohidrato de glucosa, 22 g de ácido cítrico, 3,15 g (NH₄)₂HPO₄, 0,027 g CaCl₂·2H₂O, 0,9 g KCl, 0,5 g MgSO₄·7H₂O, 2 mL 500 x biotina y 1,47 mL de solución madre de sales traza [por litro: 6 g CuSO₄·5H₂O, 0,08 g NaI, 3 g MnSO₄·H₂O, 0,2 g Na₂MoO₄·2H₂O, 0,02 g H₃BO₃, 0,5 g CoCl₂, 20 g ZnCl₂, 5 g FeSO₄·7H₂O y 5 mL H₂SO₄]; pH ajustado a 5 con KOH 5M; esterilizado por filtración). Todos los cultivos, excepto el que tiene las perlas de alimentación, se alimentaron con 0,5% de la fuente de carbono respectiva después de 19 h y 43 h. Los resultados de la fluorescencia de eGFP por tamaño de célula se muestran en la Tabla 7.

En YPD, los niveles de expresión de pCS1 superan los niveles de expresión de pGAP en 2 veces después de 24 y 48 h, mientras que en YPG los niveles de expresión de pCS1 fueron 3,9 veces más altos después de 24 h y 48 h. Los niveles de expresión bajo el control del nuevo promotor pCS1 fueron incluso mayores cuando se usaba metanol o etanol como fuente de carbono, o cuando se cultivaba a pH 4,5 (Tabla 7).

Tabla 7: Fluorescencia relativa de eGFP por tamaño de célula de clones de *P. pastoris* pGAP_eGFP # 2 o pCS1-eGFP # 4 después de 24 h y 48 h de cultivo en diferentes medios de detección. Se dan valores medios y SD de 2 cultivos.

Clon	Medio de cultivo principal	24 h	48 h
pGAP_eGFP#2	YPD	291,0	392,9

Clon	Medio de cultivo principal	24 h	48 h
pGAP_eGFP#2	YPG	194,8	393,0
pCS1_eGFP#4	YPD	598,7 ± 6,4	890,2 ± 52,2
pCS1_eGFP#4	YPG	770,8 ± 17,5	1521,4 ± 77,2
pCS1_eGFP#4	YP+glucose feed bead	724,3 ± 4,6	1559,3 ± 28,3
pCS1_eGFP#4	YPM	844,0 ± 25,4	1526,4 ± 65,3
pCS1_eGFP#4	YPE	873,1 ± 16,0	1658,0 ± 344,4
pCS1_eGFP#4	YPD, pH 4,5	742,1 ± 62,8	1938,8 ± 62,9
pCS1_eGFP#4	SCD, pH 5,0	382,4 ± 11,5	1072,2 ± 167,1

Ejemplo 3: Estudios comparativos de actividad promotora del promotor pCS1 recientemente identificado en *P. pastoris* usando albúmina de suero humano (HSA) como gen indicador expresado extracelular

Con el fin de analizar las propiedades del promotor recientemente identificado para la expresión de la proteína informadora secretada HSA, se realizaron pruebas de matraz de agitación de la siguiente manera: El cultivo previo durante 24 horas se realizó en medio rico que contenía glicerol como fuente de carbono, simulando la fase discontinua del proceso, que fue seguido por el cultivo principal en medio rico tamponado (2% de glucosa). El cultivo principal se alimentó con 0,5% de glucosa cada 12 horas,

a) Cepa y vector de expresión

La cepa *P. pastoris* de tipo salvaje (CBS2612, CBS-KNAW Fungal Biodiversity Center, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) se utilizó como cepa huésped. La transformación de la cepa se realizó con un vector propio llamado pPUZZLE (Stadlmayr et al. J. Biotechnol 2010 dic; 150 (4): 519-29), la selección de los transformantes positivos se basó en la resistencia a Zeocin. Para la expresión secretora de albúmina de suero humano (HSA) se usó su líder de secreción nativo.

b) Amplificación y clonación del promotor pCS1 recientemente identificado en un vector de expresión propio

El promotor amplificado en el Ejemplo 2b se clonó en el vector de expresión pPUZZLE pPM1aZ10_HSA, dando como resultado pPM1aZ10_pCS1_HSA. Además, el vector pPM1aZ10_pGAP_HSA, que contenía el promotor comúnmente utilizado del promotor de la gliceraldehído 3-fosfato deshidrogenasa (pGAP) se usó como referencia. Los promotores se insertaron arriba del codón de inicio del gen HSA usando los sitios de restricción Apal y SbfI (ver Tabla 3). La corrección de las secuencias promotoras se verificó mediante secuenciación de Sanger.

c) Expresión de HSA en *P. pastoris* bajo el control del nuevo promotor identificado pCS1

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

Para el cribado de la expresión de HSA, se inoculó una sola colonia en medio líquido YPG-Zeo (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de Zeocin) como cultivo previo. Después de aproximadamente 24 h, el precultivo se usó para inocular el cultivo principal con una OD600 de 0,1 en medio BM (por litro: 20 g de peptona, 10 g de extracto de levadura, 20 g de glucosa, 13,4 g de base de nitrógeno de levadura con sulfato de amonio, 0,4 mg de biotina y 100 mM de tampón fosfato de potasio pH 6,0). El cultivo principal se alimentó con 0,5% de glucosa cada 12 horas. Se tomaron muestras al final del cultivo principal. La concentración de biomasa se determinó midiendo OD600 o el peso celular húmedo. La concentración de HSA en el sobrenadante de cultivo se cuantificó mediante el conjunto de cuantificación ELISA de albúmina humana (Cat. No. E80-129, Bethyl Laboratories, TX, EE.UU.) siguiendo el manual de instrucciones del proveedor. El estándar HSA se utilizó con una concentración inicial de 400 ng mL⁻¹. Las muestras se diluyeron en consecuencia en diluyente de muestra (Tris-HCl 50 mM, NaCl 140 mM, BSA al 1% (p/v), Tween20 al 0,05% (v/v), pH 8,0). Los títulos de HSA del cribado de clones que expresan HSA bajo el control de pGAP (1 copia del gen HSA) y de 9 clones que expresan eGFP bajo el control de pCS1 se presentan en la Tabla 8. Todos los clones controlados por pCS1 secretan el doble de la cantidad de HSA que el clon controlado por pGAP con una copia génica. Los GCN de HSA se determinaron como se describe en el Ejemplo 5. Todos los clones analizados bajo el control de pCS1 albergaban una copia del casete de expresión (1GCN). Por lo tanto, todos los clones bajo el control del nuevo promotor pCS1 tenían un rendimiento de secreción de HSA dos veces mayor (mg de HSA secretada/g de biomasa) que los clones de pGAP con el mismo GCN en cultivos de cribado.

Tabla 8: Cuantificación de los niveles de HSA secretados en sobrenadantes de clones de *P. pastoris* que expresaban HSA bajo el control de pGAP y pCS1 después de 48 h del cultivo de cribado.

Clon	HSA [mg/L] 48 h cultivo principal
pGAP_HSA #3 (1GCN)	32,8
pCS1_HSA	76,5 +/- 5,0

Ejemplo 4: Cultivo de alimentación por lotes de cepas de *P. pastoris* que expresan HSA bajo el control del promotor pCS1

- 5 Para analizar la capacidad de pCS1 para impulsar la expresión de HSA en condiciones similares a las de un proceso de producción, se realizaron cultivos de alimentación por lotes de un clon pCS1 (pCS1_HSA # 1) que albergaba una copia del casete de expresión de HSA (ver Ejemplo 3).

10 Las fermentaciones se realizaron en biorreactores DASGIP con un volumen de trabajo final de 1,0 L. Dos cepas diferentes de *P. pastoris* que expresaban 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, descrito en Prielhofer et al. 2013. *Microb. Cell. Fact.* 12: 5) se cultivaron como referencia.

Se utilizaron los siguientes medios:

Solución madre de sales traza PTM₁ contenida por litro

- 15 6,0 g CuSO₄·5H₂O, 0,08 g NaI, 3,36 g MnSO₄·H₂O, 0,2 g Na₂MoO₄·2H₂O, 0,02 g H₃BO₃, 0,82 g CoCl₂, 20,0 g ZnCl₂, 65,0 g FeSO₄·7H₂O, 0,2 g de biotina y 5,0 ml de H₂SO₄ (95%-98%).

Medio del lote de glicerol contenido por litro

39,2 g glicerol, 27,9 g de H₃PO₄ (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 solución madre de sales traza de PTM₁. El pH se ajustó a 5,85 después de la filtración estéril en el fermentador.

- 20 Medio alimentado por lotes con glucosa contenido por litro

550 g de monohidrato de glucosa, 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 solución madre de sales traza de PTM₁.

- 25 El oxígeno disuelto se controló a DO = 20% con la velocidad del agitador (400 - 1200 rpm). La velocidad de aireación es 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 (25%).

- 30 Para comenzar la fermentación, se filtraron 400 ml de medio discontinuo en el fermentador y se inoculó *P. pastoris* (del pre-cultivo) con una densidad óptica inicial (OD₆₀₀) de 1. La fase del lote de aproximadamente 25 h alcanzó una concentración de biomasa seca de aproximadamente 20 g/L y fue seguida por un lote alimentado con alimentación constante de medio de glucosa (2 g L⁻¹h⁻¹) durante 100 horas, lo que llevó a una concentración de biomasa seca de aproximadamente 100 g/L. El pH fue de 5,85 durante el lote y se mantuvo a 5,85 durante toda la fermentación. Se tomaron muestras durante el lote y se alimentó la fase del lote. La concentración de HSA se cuantificó usando el kit de cuantificación ELISA de albúmina humana (Bethyl, Cat. E80-129) como se describe en el Ejemplo 3c.

- 35 Como se mostró anteriormente, el clon pGAP con 2 copias de HSA secretó el doble que el clon pGAP con una copia (Prielhofer et al. 2013. *Microb. Cell. Fact.* 12:5). Dos clones que secretaban HSA bajo el control de pCS1 alcanzaron títulos de HSA más de 4 veces más altos al final del lote y se alimentaron por lotes en comparación con un clon de pGAP de copia única (los resultados se muestran en la Tabla 9). En correlación con la biomasa y la GCN, los clones pCS1 produjeron 390% de HSA secretada del clon pGAP con el mismo número de copia génica.

- 40 Tabla 9: Concentraciones de masa seca de levadura y títulos de HSA en el sobrenadante al final del lote y al final del lote alimentado, y título de HSA por masa seca de levadura al final del lote alimentado de cultivos de biorreactor de clones de *P. pastoris* que expresaban HSA bajo el control de pCS1 o pGAP.

Promotor	GCN	final del lote		final de lote de alimentación		
		HSA [mg L ⁻¹]	YDM [g L ⁻¹]	HSA [mg L ⁻¹]	YDM [g L ⁻¹]	[% HSA/YDM de P _{GAP}]
P _{CS1} #1	1	38,1	22,2	377,7	119,9	390,7
P _{GAP} #3	1	8,4	22,1	78,6	97,7	100,0

P _{GAP} # 4	2	17,4	24,4	167,7	121,0	172,0
----------------------	---	------	------	-------	-------	-------

Ejemplo 5: Determinación de números de copia génica (GCN) de clones seleccionados

La intensidad de la expresión a menudo se correlaciona con el número de casetes de expresión integrados en el genoma de *P. pastoris*. Por lo tanto, se determinó el número de copias del gen de los clones seleccionados. El ADN genómico se aisló usando el kit DNeasy Blood & Tissue (Quiagen, Cat. No. 69504). Los números de copias de genes se determinaron usando PCR cuantitativa. Por lo tanto, se utilizó el kit SensiMix SYBR (Bioline, QT605-05). Sensi Mix SYBR se mezcló con los cebadores respectivos (dados en Prielhofer et al. 2013. Microb. Cell. Fact. 12: 5) y la muestra, y se aplicó para el análisis a tiempo real en un ciclador PCR a tiempo real (Rotor Gene, Qiagen). Todas las muestras fueron analizadas por triplicado o cuadruplicado. Se usó el software Rotor Gene para el análisis de datos.

5 Ejemplo 6: Estudios comparativos de la actividad promotora del promotor pCS1 recientemente identificado en *P. pastoris* usando carboxipeptidasa B porcina (CpB) como gen informador expresado extracelular

Para analizar las propiedades del nuevo promotor identificado, las pruebas del matraz de agitación se realizaron de la siguiente manera: el cultivo previo durante 24 horas se realizó con un medio rico que contenía glicerol como fuente de carbono, seguido del cultivo principal con medios ricos.

15 a) Cepa y vector de expresión

La cepa de tipo salvaje *P. pastoris* (CBS2612, CBS-KNAW Fungal Biodiversity Center, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) se usó como cepa huésped. La transformación de la cepa se llevó a cabo con un vector propio denominado pPUZZLE (Stadlmayr et al. J. Biotechnol 2010 Dec; 150 (4): 519-29), la selección de los transformantes positivos se basa en la resistencia a Zeocin. Para la expresión secretora de la carboxipeptidasa B porcina (CpB) se utiliza el factor de apareamiento principal alfa de levadura.

20 b) Amplificación y clonación del promotor pCS1 recientemente identificado en un vector de expresión propio

Los promotores amplificados en el Ejemplo 2b se clonan en el vector de expresión pPUZZLE p PM1aZ30_aMF_CpB, dando como resultado pPM1aZ30_pCS1_aMF_CpB. Además, el vector pPM1dZ30_pGAP_CpB, que contenía el promotor comúnmente usado del promotor de la gliceraldehído 3-fosfato deshidrogenasa (pGAP) se usó como referencia. Los promotores se insertaron arriba del codón de inicio del gen CpB usando los sitios de restricción Apal y SbfI. La secuencia de Sanger verifica la corrección de las secuencias del promotor.

25 c) Expresión de CpB en *P. pastoris* bajo el control de los promotores inducidos por límite de glucosa recientemente identificados

30 Los plásmidos se linealizan usando la enzima de restricción AclI antes de la electroporación (usando un protocolo de transformación estándar para *P. pastoris* en *P. pastoris*. La selección de transformantes positivos se realiza en placas YPD (por litro: 10 g de extracto de levadura, 20 g de peptona, 20 g de glucosa, 20 g de agar-agar), placas que contienen 25 pg/ml de Zeocin. La colonia PCR se usa para asegurar la presencia del plásmido transformado como se describe en el ejemplo 2c.

35 Para el cribado de la expresión de CpB, se inocula una única colonia en medio líquido YPG-Zeo (por litro: 20 g de peptona, 10 g de extracto de levadura, 12,6 g de glicerol y 25 mg de Zeocin) como cultivo previo. Después de aproximadamente 24 h, el precultivo se usa para inocular el cultivo principal con una OD600 de 1 en medio YPD (por litro: 20 g de peptona, 10 g de extracto de levadura, 20 g de glucosa). El cultivo principal se alimenta con 0,5% de glucosa cada 12 horas. Se toman muestras al final del pre-cultivo, y 24 y 48 horas después de la inoculación del cultivo principal. La concentración de biomasa se determina midiendo OD600 o el peso celular húmedo. La concentración de CpB en el sobrenadante del cultivo se cuantifica mediante un ensayo enzimático, basado en la conversión de hippuril-L-arginina en ácido hipúrico por el CpB. La cinética de reacción se mide controlando la absorción a 254 nm a 25°C utilizando un espectrofotómetro Hitachi U-2910 cuando se inicia la reacción. Las muestras y los estándares se tamponan con tampón de ensayo (Tris 25 mM, HCl 100 mM, pH 7,65) y se activan usando tampón de activación (tripsina 0,01 mgL⁻¹, Tris 300 mM, ZnCl₂ 1 μM, pH 7,65). El tampón de activación sin tripsina se usa en lugar de la muestra como control negativo. La reacción se inicia mediante la adición de la solución de sustrato (hipuril-L-arginina 1 mM en tampón de ensayo).

45 d) Cultivo alimentado por lotes de cepas de *P. pastoris* que expresan CpB bajo el control del promotor pCS1. La fermentación alimentada por lotes se realiza como se describe en el ejemplo 4 usando medios como se describe en el ejemplo 2d.

Ejemplo 7: Estudios comparativos de actividad promotora del promotor pCS1 recientemente identificado en clones multicopia de *P. pastoris* usando albúmina sérica humana (HSA) como gen indicador expresado extracelularmente

Para investigar si la producción de HSA podía mejorarse aún más mediante un GCN más alto, los clones P_{CS1} que expresaban HSA se amplificaron mediante el método de amplificación de vectores postransformationales, como lo describen Marx et al. (2009). Se generaron vectores que se integraban en el locus de ADN_r por recombinación homóloga. La amplificación se realizó mediante selección a una concentración gradual creciente de antibióticos.

5 a) Cepa y vector de expresión

La cepa *P. pastoris* de tipo salvaje (CBS2612, CBS-KNAW Fungal Biodiversity Center, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) se utilizó como cepa huésped. La transformación de la cepa se realizó con una variante del vector pPUZZLE que contiene la región NTS del locus de ADN ribosómico como sitio de integración (Marx et al. 2009. FEMS Yeast Res. 9(8): 1260-70). La selección de transformantes positivos se basó en la resistencia a Zeocin. Para la expresión secretora de albúmina de suero humano (HSA) se usó su líder de secreción nativo.

El promotor pCS1 amplificado en el Ejemplo 2b se clonó en el vector de expresión pPUZZLE pPM1nZ30_HSA, dando como resultado pPM1nZ30_pCS1_HSA.

15 El promotor se insertó arriba del codón de inicio del gen HSA usando los sitios de restricción Apal y SbfI. La corrección de las secuencias promotoras se verificó mediante secuenciación de Sanger.

b) Amplificación del vector postransformacional y expresión de HSA en *P. pastoris* bajo el control del promotor recién identificado pCS1

20 Los plásmidos se linealizaron usando la enzima de restricción SpeI antes de la electroporación (usando un protocolo de transformación estándar para *P. pastoris*) en *P. pastoris*. La selección inicial de transformantes positivos se realizó en placas YPD (por litro: 10 g de extracto de levadura, 20 g de peptona, 20 g de glucosa, 20 g de agar-agar), placas que contienen 25 g/ml de Zeocin. Se usó la colonia PCR para asegurar la presencia del plásmido transformado como se describe en el ejemplo 2c. La amplificación del número de copia génica se realizó como se describe en Marx et al. (FEMS Yeast Res. 2009 Dec; 9 (8): 1260-70) mediante rayas repetidas de clones en placas de agar YPD que contienen mayores concentraciones de Zeocin (50, 100 y 500 g/ml de Zeocin).

25 El cribado de expresión de HSA y la cuantificación del producto mediante ELISA de HSA se realizaron como se describe en el Ejemplo 3c. Se determinó GCN de algunos de los clones con los niveles de secreción de HSA más altos como se describe en el Ejemplo 5. Los clones P_{CS1} amplificados con GCN y los clones P_{GAP} y P_{CS1} con GCN conocidos (uno o dos) como control se cultivaron en medio BM durante 48 h.

30 Al aumentar la GCN de HSA, los niveles de secreción de HSA también podían aumentarse (como se muestra en la Tabla 10). Los clones multicopia que expresan HSA bajo el control de pCS1 con 3 copias del gen HSA produjeron cantidades de HSA de dos a tres veces más altas por WCW en comparación con el clon de copia única. También podía esperarse un aumento similar en la productividad de los clones pCS1 multicopia en cultivos de biorreactor alimentado por lotes como se presenta en el Ejemplo 4.

35 Tabla 10: Resultados de la detección de la expresión de HSA con clones únicos y multicopia bajo el control de pGAP o pCS1. Se muestran GCN y títulos por GCN.

Clon	µg HAS/g WCW	GCN	µg HAS/g WCW/GCN
pGAP_HSA#3	291,0	1	291,0
pGAP_HSA#4	476,2	2	238,1
pCS1_HSA#1	540,7	1	540,7
pCS1_HSA#1_25-100#7	1464,1	3	488,0
pCS1_HSA#1_50-100#5	1178,3	4	294,6
pCS1_HSA#1_50-100#9	1580,7	3	526,9

Ejemplo 8: Estudios comparativos de la actividad promotora del promotor recién identificado pCS1 en *P. pastoris* usando el fragmento de anticuerpo (Fab) como gen indicador expresado extracelularmente

40 Con el fin de analizar las propiedades del promotor recientemente identificado, las pruebas de matraz de agitación se realizaron de la siguiente manera: el cultivo previo durante 24 horas se realizó con medio rico que contenía glicerol como fuente de carbono, seguido por el cultivo principal con medio rico.

a) Cepa y vector de expresión

La cepa de *P. pastoris* de tipo salvaje (CBS2612, CBS-KNAW Fungal Biodiversity Center, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) se utilizó como cepa huésped. El promotor pCS1 amplificado en el

Ejemplo 2b se clonó en el vector de expresión pPUZZLE que contenía LC o Fab-HC del anticuerpo HyHEL como GOI. El promotor se insertó arriba del codón de inicio de los genes Fab usando los sitios de restricción Apal y SbfI. Después de la verificación de la secuencia, los casetes de expresión para ambas cadenas se combinaron en un vector usando las enzimas de restricción compatibles Mrel y AgeI.

5 b) Expresión de Fab en *P. pastoris* bajo el control del nuevo promotor identificado pCS1

Los plásmidos se linealizaron usando la enzima de restricción Ascl antes de la electroporación (usando un protocolo de transformación estándar para *P. pastoris*) en *P. pastoris*. La selección de transformantes positivos se realizó en placas YPD (por litro: 10 g de extracto de levadura, 20 g de peptona, 20 g de glucosa, 20 g de agar-agar), placas que contienen 25 µg/ml de Zeocin. Se usó PCR de la colonia para asegurar la presencia del plásmido transformado como se describe en el ejemplo 2c.

El cribado de la expresión de Fab se realizó de forma similar al cribado de la expresión de HSA descrito en el Ejemplo 3c. La cuantificación de Fab intacto se realizó mediante ELISA usando anticuerpo IgG antihumano (Abeam ab7497) como anticuerpo de recubrimiento (1:1.000), y un anticuerpo conjugado de cadenas ligeras kappa anti-humano de cabra (unidas y libres) - fosfatasa alcalina (Sigma) A3813) como anticuerpo de detección (1:1.000). Se usó Fab/Kappa humano, fragmento de IgG (Bethyl P80-115) como estándar con una concentración inicial de 50 ng/mL. Las muestras de sobrenadante se diluyeron en consecuencia. La detección se realizó con sustrato pNPP (Sigma S0942). El tampón de revestimiento, dilución y lavado se basó en PBS (KH₂PO₄ 2 mM, Na₂HPO₄·2H₂O 10 mM, KCl 2,7 mM, NaCl 8 mM, pH 7,4) y se completó con BSA (1% (p/v)) y/o Tween20 (0,1% (v/v)) en consecuencia.

Se aisló el ADN genómico de los clones seleccionados y se determinaron los GCN de la cadena pesada (HC) y la cadena ligera (LC) como se describe en el Ejemplo 5. Los GCN se relacionaron con el clon PCS1 # 5, para el cual HC y LC GCN se estableció en uno. Los rendimientos de expresión de Fab (µg Fab/g WCW), los GCN relativos y el rendimiento de Fab por GCN se muestran en la Tabla 11. En cuanto a las otras proteínas modelo, se obtuvo un rendimiento de Fab aproximadamente dos veces mayor por GCN (en promedio 35,4 µg Fab/gWCW) para los clones que se expresan bajo el control de pCS1 en comparación con los clones que se expresaban bajo el control de pGAP.

Tabla 11: Resultados de la detección de la expresión Fab bajo el control de pGAP o pCS1. Los rendimientos de GCN y Fab por GCN se muestran a la derecha.

Clon	[µg Fab g ⁻¹ WCW]	GCN relativo con respecto a pCS1_Fab#5		Rendimiento de Fab [µg Fab g ⁻¹ WCW] por GCN	
		Cadena pesada	Cadena ligera	Cadena pesada	Cadena ligera
pGAP_Fab#34	73,3	4	3	18	24
pCS1_Fab#2	35,9	1	1	36	36
pCS1_Fab#4	66,8	2	2	33	33
pCS1_Fab#5	34,7	1	1	34,7	34,7
pCS1_Fab#7	37,5	1	1	38	37,5

c) Cultivo alimentado por lotes de cepas de *P. pastoris* que expresan Fab bajo el control del promotor pCS1.

Las fermentaciones alimentadas por lotes se realizaron de forma similar a la descrita en el ejemplo 4 usando medios como se describe en el ejemplo 2d.

En el cultivo de biorreactores, la expresión de Fab bajo el control de pCS1 dio como resultado productividades específicas (qP) más de 3,0 veces mayores por GCN en comparación con pGAP (ver Tabla 12).

Tabla 12: Resultados del cultivo de biorreactor de la expresión de Fab bajo el control de pGAP o pCS1. Se muestran los rendimientos de Fab al final del cultivo alimentado por lotes, la productividad específica media qP (tasa de producción promedio de Fab por biomasa (peso de células secas DCW) durante la fase del lote alimentado completo), GCN relativo de los clones respectivos y qP media por GCN.

Clon	Rendimiento de Fab [mg Fab g ⁻¹ DCW]	qP [µg Fab g ⁻¹ DCW h ⁻¹]	GCN relativo con respecto a pCS1_Fab#5		qP por GCN	
			Cadena pesada	Cadena ligera	Cadena pesada	Cadena ligera
pGAP_Fab#34	0,220	2,18	4	3	0,55	0,73
pCS1_Fab#4	0,262 ± 0,000	3,51 ± 0,14	2	2	1,76	1,76

pCS1_Fab#5	0,178	2,54	1	1	2,54	2,54
------------	-------	------	---	---	------	------

Ejemplo 9: Comparación de variantes de pCS1

Las variantes de longitud del promotor pCS1 se clonan como se describe en el ejemplo 2a y se seleccionan de forma similar a la descrita en el ejemplo 2c. Los clones que se expresaron bajo el control de pCS1 (longitud estándar) y pGAP se usaron como controles. Los cebadores directos y las longitudes de pCS1 y sus variantes se enumeran en la Tabla 13.

5

Tabla 13: pCS1 y sus variantes: cebadores directos y longitudes de las variantes de tamaño de pCS1 (SEQ ID 1)

Variante	cebador directo	Longitud (bp)
pCS1 (estándar) SEQ ID 1	GATAGGGCCCCAGGGCATCATTGAGGTTTCCAC SEQ ID 16	985
pCS1-1488 SEQ ID 2	GATAGGGCCCGATAGTTCTAGAAGACCTGGCGTCG SEQ ID 17	1488
pCS1-767F SEQ ID 3	GATAGGGCCCAGCCAACCATCTTTTGTTCG SEQ ID 18	767
pCS1-500F SEQ ID 4	GATAGGGCCCGTGGTTTCCAGGACAACACCC SEQ ID 19	500
pCS1-344F SEQ ID 5	GATAGGGCCCGACCGCAATTCACCATGATGC SEQ ID 20	344
pCS1-234F SEQ ID 6	GATAGGGCCCAGCCTGCTTCATTCTGCC SEQ ID 21	234
pCS1-138F SEQ ID 7	GATAGGGCCCCCGGAAAAGGTTTGTATAG SEQ ID 22	138
pCS1-85F SEQ ID 8	GATAGGGCCCATACTCTCCTCCCCCCTG SEQ ID 23	85

Los exámenes de los clones que expresaban eGFP bajo el control de pGAP, pCS1 o pCS1 revelaron que pCS1 requiere una longitud mínima de aproximadamente 500 pb para una actividad óptima. Las variantes pCS1 más cortas pierden actividad con una longitud decreciente (ver Tabla 14).

10 Tabla 14: Resultados de detección (fluorescencia relativa) de la expresión de eGFP bajo el control de las variantes pGAP, pCS1 o pCS1.

Clon	FL/tamaño célula, 24 h	FL/tamaño célula, 48 h
pGAP_eGFP#2	15,39	17,59
pCS1_eGFP#4	25,30	37,77
pCS1_85pool	-0,15	0,04
pCS1_138pool	7,70	6,30
pCS1_234pool	18,57	19,71
pCS1_344pool	25,19	27,63
pCS1_500pool	28,50	38,82
pCS1_767pool	26,88	36,47
pCS1_1488pool	30,59	44,40

Ejemplo 10: Verificación de la fuerza de expresión de un promotor en un clon que expresa eGFP bajo el control de dicho promotor en condiciones de crecimiento limitado con tasas de crecimiento altas y bajas

a) Cepa

15 Una cepa huésped (p. ej. *Pichia pastoris*) que expresa eGFP bajo el control de un "promotor de interés" y una cepa que expresa eGFP bajo el control de pGAP se usan para comparar los niveles de expresión.

b) Cultivo de cepas que expresan eGFP para la comparación de promotores

20 Estas cepas se cultivan en quimiostato a dos tasas de crecimiento específicas fijas (una tasa de crecimiento específica baja y otra alta al establecer la tasa de dilución), utilizando biorreactores DASGIP con un volumen de trabajo final de 1,0 L.

Se utilizan los siguientes medios:

La solución de reserva de sales traza PTM1 contiene por litro

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

El medio del lote de glicerol contiene por litro

5 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 PTM1

solución madre de trazas de sales. Se agrega HCl para establecer el pH a 5.

El medio de quimiostato contiene por litro

10 2,5 g de monohidrato de ácido cítrico (C₆H₈O₇·H₂O), 55,0 g de monohidrato de glucosa, 21,75 g (NH₄)₂HPO₄, 1,0 g MgSO₄·7H₂O, 2,5 g KCl, 0,04 g CaCl₂·2H₂O, 0,4 mg de biotina y 2,43 mL de solución madre de sales traza de PTM1. Se agrega HCl para ajustar el pH a 5.

15 El oxígeno disuelto se controla a DO = 20% con la velocidad del agitador (400 - 1200 rpm). La velocidad de aireación es de 24 L h⁻¹ de aire, la temperatura se controla a 25°C y el punto de ajuste del pH de 5 se controla con la adición de NH₄OH (25%). Para comenzar la fermentación, se filtran estéril 400 ml de medio discontinuo en el fermentador y se inocula un clon de *P. pastoris* (del pre-cultivo) con una densidad óptica inicial (OD₆₀₀) de 1. La fase discontinua de aproximadamente 25 h (que alcanza una concentración de biomasa seca de aproximadamente 20 g L⁻¹) es seguida por un cultivo de quimiostato limitado en glucosa. La velocidad de alimentación del medio quimiostato y la velocidad de cosecha se usan para mantener una velocidad de crecimiento específica constante según se desee. Durante este cultivo, el volumen del caldo de cultivo se mantiene constante y se determina el peso seco de la célula para asegurar una tasa de crecimiento constante. Las células se cultivan a una tasa de crecimiento alta y baja de 20 0,15 y 0,015 h⁻¹, respectivamente. Por lo tanto, la velocidad de alimentación/cosecha se controla a 150 ml h⁻¹ L⁻¹ (ml de medio quimiostato por litro de caldo de cultivo y hora) y 15 ml h⁻¹ 5 L⁻¹, respectivamente,

c) Muestreo

25 Las muestras se toman en condiciones de estado estacionario (después de al menos 5 intercambios de volumen) y se analizan para determinar la expresión de eGFP usando un lector de placas (Infinite 200, Tecan, CH). Por lo tanto, las muestras se diluyen a una densidad óptica (OD₆₀₀) de 5. Las fermentaciones se realizan por duplicado. Los datos de expresión se comparan calculando la fluorescencia relativa por biorreactor como se describe en el ejemplo 2d).

Ejemplo 11: Identificación de un promotor de *P. pastoris* que permite una alta transcripción a tasas de crecimiento específicas altas y bajas

30 Para identificar un promotor que permita una alta transcripción a tasas de crecimiento altas y bajas, el análisis de los patrones de expresión génica se realizó utilizando micromatrices de ADN. A partir de los datos de transcriptómica se seleccionaron genes que mostraban una alta fuerza de transcripción a tasas de crecimiento altas y bajas. Por lo tanto, las células de *P. pastoris* se cultivaron en cultivo de quimiostato como se describe en el ejemplo 10b) a tasas de crecimiento específicas altas y bajas de 0,15 y 0,015 h⁻¹, respectivamente. El muestreo, la purificación de ARN, la preparación de muestras para la hibridación de micromatrices, el análisis de micromatrices, la adquisición de datos y 35 la evaluación estadística se realizan como se describe en el ejemplo 1c), 1d), 1e) y 1f). Los genes y los promotores respectivos con una alta fuerza de transcripción a una tasa de crecimiento alta y baja se identificaron explorando los datos de micromatrices en busca de genes con intensidades de señal altas en condiciones de tasa de crecimiento alta y baja. Como segundo criterio, las intensidades de señal deberían ser mayores que las del gen de la gliceraldehído-3-fosfato deshidrogenasa (GAP, sinónimos GAPDH y TDH3) en ambas condiciones. Para aislar el 40 promotor, se amplificó un fragmento de ácido nucleico de aproximadamente 1000 bp arriba del codón de inicio ATG del gen respectivo.

Listado de secuencias

<110> Lonza Ltd.
45 <120> Promotor constitutivo
<130> L0003P
<160> 24
<170> PatentIn versión 3.5
<210> 1
50 <211> 985
<212> ADN
<213> *Pichia pastoris*

ES 2 751 336 T3

<400> 1

```

agggcatcat tgaggtttcc acaaaaggaa gaaacatgga tccagagaca tcaacagaga      60
ggaaagcggg tagtgaagcc gaagccacaa cacagcccga tttggaaggg agttcacaat     120
caaggtgagt ccagccattt tttttctttt tttttttttt attcaggtga acccacctaa     180
ctatttttta ctgggatcca gtgagctcgc tgggtgaaag ccaaccatct tttgtttcgg     240
ggaaccgtgc tcgccccgta aagtaattt ttttttcccg cgcagcttta atctttcggc     300
agagaaggcg ttttcatcgt agcgtgggaa cagaataatc agttcatgtg ctatacaggc     360
acatggcagc agtcactatt ttgcttttta accttaaagt cgttcatcaa tcattaactg     420
accaatcaga ttttttgcat ttgccactta tctaaaaata cttttgtatc tcgcagatac     480
gttcagtggg ttccaggaca acacccaaaa aaaggtatca atgccactag gcagtcgggt     540
ttatttttgg tcaccacgc aaagaagcac ccacctcttt taggttttaa gttgtgggaa     600
cagtaacacc gcctagagct tcaggaaaaa ccagtacctg tgaccgcaat tcaccatgat     660
gcagaatggt aatttaaacy agtgccaaat caagatttca acagacaaat caatcgatcc     720
atagttaccg attccagcct tttcgtcgtc gagcctgott cattcctgcc tcaggtgcat     780
aactttgcat gaaaagtcca gattagggca gatthttgagt ttaaaatagg aaatataaac     840
aaatataccg cgaaaaaggt ttgtttatag cttttcgcct ggtgccgtac ggtataaata     900
catactctcc tccccccctt ggttctcttt ttcttttggt acttacattt taccgttccg     960
tcaactcgctt cactcaacaa caaaa                                     985

```

<210> 2

<211> 1488

5 <212> ADN

<213> Pichia pastoris

<400> 2

```

gatagttcta gaagacctgg cgtcgtggt caactactcg tttcacaagt tgacaaagac      60
ttttaccaag agaagagcag ttgtcaccga ccacaataac aacctggaag ccgagaaaaa     120

```

ES 2 751 336 T3

acttggaaac tccaaggtga gactgcctag aaatccgtac tctgcggccg atcgaagact 180
gcatttcctc caagaattga tgctcatggt ctcatcttac aatcgcacac acaacagtgt 240
cagtcttctt tgcggtccct tgaacacaac caaccgaaag gtggggaagt ctaatgtcac 300
gcaaacgata ttgcaaccaa tgttgggctc tactggcgtc tggctgcatc aaatagctga 360
tcggttcgta atcttcaaag attggtgtag gacgaacgag tctgctgggc tacaagtttt 420
gccccatata gctggtcaag ccaaccccgcg gaatccaaa acaccccatc cgacaaaagt 480
tgttgttttc agcagatcta gggagggcat cattgagggt tccacaaaag gaagaaacat 540
ggatccagag acatcaacag agaggaaagc gggtagtgaa gccgaagcca caacacagcc 600
cgatttggaa gggagttcac aatcaaggtg agtccagcca ttttttttct tttttttttt 660
tttattcagg tgaaccacc taactatttt taactgggat ccagtgagct cgctgggtga 720
aagccaacca tcttttgttt cggggaaccg tgctcgcccc gtaaagttaa tttttttttc 780
ccgcgagct ttaatctttc ggagagaag gcgttttcat cgtagcgtgg gaacagaata 840
atcagttcat gtgctataca ggacatggc agcagtcact attttgcttt ttaaccttaa 900
agtgttcat caatcattaa ctgaccaatc agattttttg catttgccac ttatctaaaa 960
atacttttgt atctcgaga tacgttcagt ggtttccagg acaacacca aaaaaggtta 1020
tcaatgccac taggcagtcg gttttatfff ttggtcaccca cgcaaagaag caccacctc 1080
ttttaggttt taagttgtgg gaacagtaac accgcctaga gcttcaggaa aaaccagtac 1140
ctgtgaccgc aattcacat gatgcagaat gtttaattaa acgagtgcc aatcaagatt 1200
tcaacagaca aatcaatcga tccatagtta ccattccag ctttttcgtc gtcgagcctg 1260
cttcattcct gcctcaggtg cataactttg catgaaaagt ccagattagg gcagattttg 1320
agtttaaaat aggaaatata acaaatata ccgcgaaaaa ggtttgttta tagcttttcg 1380
cctggtgccg tacggtataa atacatactc tcctcccccc cctggttctc tttttctttt 1440
gttacttaca ttttaccgtt ccgtcactcg cttcactcaa caacaaaa 1488

<210> 3
<211> 767
<212> ADN
5 <213> Pichia pastoris

<400> 3
agccaacat cttttgtttc ggggaaccgt gctcgccccg taaagttaat ttttttttcc 60
cgcgagctt taatctttcg gcagagaagg cgttttcatc gtagcgtggg aacagaataa 120
tcagttcatg tgctatacag gcacatggca gcagtcacta ttttgctttt taaccttaaa 180
gtcgttcatc aatcattaac tgaccaatca gattttttgc atttgccact tatctaaaaa 240
tacttttgta tctcgagat acgttcagtg gtttccagga caacacccaa aaaaaggtat 300

ES 2 751 336 T3

caatgccact aggcagtcgg ttttattttt ggtcaccac gcaaagaagc acccacctct 360
 tttaggTTTT aagttgtggg aacagtaaca cgcctagag cttcaggaaa aaccagtacc 420
 tgtgaccgca attcaccatg atgcagaatg ttaatttaaa cgagtGCCAA atcaagattt 480
 caacagacaa atcaatcgat ccatagttac ccattccagc cttttcgtcg tcgagcctgc 540
 ttcattcctg cctcaggtgc ataactttgc atgaaaagtc cagattaggg cagattttga 600
 gtttaaaata ggaaatataa acaaataac cgcgaaaag gtttgttat agcttttcgc 660
 ctggtgccgt acggtataaa tacatactct cctccccccc ctggttctct ttttctttg 720
 ttacttacat tttaccgttc cgtcactcgc ttcactcaac aacaaaa 767

<210> 4
 <211> 500
 <212> ADN
 5 <213> Pichia pastoris
 <400> 4

gtggtttcca ggacaacacc caaaaaaagg tatcaatgcc actaggcagt cggttttatt 60
 tttggtcacc cacgcaaaga agcaccacc tcttttaggt ttttaagttgt gggAACAGTA 120
 acaccgccta gagcttcagg aaaaaccagt acctgtgacc gcaattcacc atgatgcaga 180
 atgTTAATTT aaacgagtgc caaatcaaga tttcaacaga caaatcaatc gatccatagt 240
 taccattcc agccttttcg tcgctgagcc tgcttcattc ctgcctcagg tgcataactt 300
 tgcatgaaaa gtccagatta gggcagattt tgagtttaaa ataggaaata taaacaaata 360
 taccgcgaaa aaggtttggt tatagctttt cgcctgggtgc cgtacggtat aaatacatac 420
 tctcctccc cccctggttc tcttttctt ttgttactta cttttaccg ttccgctact 480
 cgcttcactc aacaacaaaa 500

<210> 5
 <211> 344
 10 <212> ADN
 <213> Pichia pastoris
 <400> 5

gaccgcaatt caccatgatg cagaatgta atttaaacga gtGCCAAATC aagatttcaa 60
 cagacaaatc aatcgatcca tagttacca ttccagcctt ttcgctcgtcg agcctgcttc 120
 attcctgcct caggtgcata actttgcatg aaaagtccag attagggcag attttgagtt 180
 taaaatagga aatataaaca aatataccgc gaaaaagggt tgtttatagc ttttcgcctg 240
 gtGCCGTACG gtataaatac atactctcct cccccccctg gttctctttt tcttttgta 300
 cttacatttt accgttccgt cactcgttc actcaacaac aaaa 344

<210> 6
 <211> 234
 15 <212> ADN
 <213> Pichia pastoris
 <400> 6

ES 2 751 336 T3

agcctgcttc attcctgcct caggtgcata actttgcatg aaaagtccag attagggcag 60
 attttgagtt taaaatagga aatataaaca aatataccgc gaaaaagggtt tgtttatagc 120
 ttttcgcctg gtgccgtacg gtataaatac atactctcct cccccccctg gttctctttt 180
 tcttttgтта cttacatttt accgttccgt cactcgcttc actcaacaac aaaa 234
 <210> 7
 <211> 138
 <212> ADN
 5 <213> Pichia pastoris
 <400> 7
 ccgcgaaaaa ggtttgтта tagcttttcg cctggtgccg tacggтатаa atacatactc 60
 tcctcccccc cctggttctc tttttctttt gttacttaca ttttaccggt ccgtcactcg 120
 cttcactcaa caacaaaa 138
 <210> 8
 <211> 85
 10 <212> ADN
 <213> Pichia pastoris
 <400> 8
 catactctcc tccccccct ggttctcttt ttcttttgtt acttacattt taccgttccg 60
 tcactcgctt cactcaacaa caaaa 85
 <210> 9
 <211> 408
 15 <212> ADN
 <213> Pichia pastoris
 <400> 9
 atgcaattct ctatcgtcgc tactttggct cttgctggtt ccgctctggc tgcttactct 60
 aacgtaactt aacttacga gactaccatc accgatgttg tcaccgagct caccacttac 120
 tgcccagagc caaccacett cgttcacaag aacaagacca tcaactgtgac cgcccccaacc 180
 actttgacca tcaactgactg tccttgacc atctccaaga ccaccaagat caccactgat 240
 gttccaccaa ccaccactc caccaccac accaccacca ctcaactgac atctacctct 300
 accccagctc caaccactc tgtttctacc atctctcagc gtggctgctg taaggctggc 360
 gttgctggtt tggccgggtg tgctgctgcc gctgcttact tcttgtaa 408
 <210> 10
 <211> 135
 <212> PRT
 20 <213> Pichia pastoris
 <400> 10

ES 2 751 336 T3

Met Gln Phe Ser Ile Val Ala Thr Leu Ala Leu Ala Gly Ser Ala Leu
1 5 10 15

Ala Ala Tyr Ser Asn Val Thr Tyr Thr Tyr Glu Thr Thr Ile Thr Asp
20 25 30

Val Val Thr Glu Leu Thr Thr Tyr Cys Pro Glu Pro Thr Thr Phe Val
35 40 45

His Lys Asn Lys Thr Ile Thr Val Thr Ala Pro Thr Thr Leu Thr Ile
50 55 60

Thr Asp Cys Pro Cys Thr Ile Ser Lys Thr Thr Lys Ile Thr Thr Asp
65 70 75 80

Val Pro Pro Thr Thr His Ser Thr Pro His Thr Thr Thr Thr His Val
85 90 95

Pro Ser Thr Ser Thr Pro Ala Pro Thr His Ser Val Ser Thr Ile Ser
100 105 110

His Gly Gly Ala Ala Lys Ala Gly Val Ala Gly Leu Ala Gly Val Ala
115 120 125

Ala Ala Ala Ala Tyr Phe Leu
130 135

<210> 11
<211> 402
<212> **ADN**
5 <213> Pichia pastoris
<400> 11

atgcaattct ctatcgtcgc tactttggct cttgctgggt cgcctctggc tgcttactct 60
aacgtaactt aacttacga gactaccatc accgatggtg tcaactgagtt gaccacttac 120
tgcccagagc caaccacctt tgtttacaag aacaagacca tcaccgtgac tgagccaacc 180
actttgacca tcaactgactg cccatgcacc atctcaaaga ccaccaagat caccactgat 240
gttccaccaa ccaccacgt caccatcc accactcaag tgccatctac ctctaccca 300
gctccaacc actctgtttc taccatctct cacggtggtg ctgctaaggc tgggtgtgct 360
ggtttgccg gtgttgctgc tgccgctgct tacttcttgt aa 402

<210> 12
<211> 133
10 <212> PRT
<213> Pichia pastoris
<400> 12

ES 2 751 336 T3

Met Gln Phe Ser Ile Val Ala Thr Leu Ala Leu Ala Gly Ser Ala Leu
1 5 10 15

Ala Ala Tyr Ser Asn Val Thr Tyr Thr Tyr Glu Thr Thr Ile Thr Asp
20 25 30

Val Val Thr Glu Leu Thr Thr Tyr Cys Pro Glu Pro Thr Thr Phe Val
35 40 45

Tyr Lys Asn Lys Thr Ile Thr Val Thr Glu Pro Thr Thr Leu Thr Ile
50 55 60

Thr Asp Cys Pro Cys Thr Ile Ser Lys Thr Thr Lys Ile Thr Thr Asp
65 70 75 80

Val Pro Pro Thr Thr His Val Thr Pro Ser Thr Thr His Val Pro Ser
85 90 95

Thr Ser Thr Pro Ala Pro Thr His Ser Val Ser Thr Ile Ser His Gly
100 105 110

Gly Ala Ala Lys Ala Gly Val Ala Gly Leu Ala Gly Val Ala Ala Ala
115 120 125

Ala Ala Tyr Phe Leu
130

<210> 13

<211> 491

<212> ADN

5 <213> Pichia pastoris

<400> 13

cttttttgta gaaatgtctt ggtgtcctcg tccaatcagg tagccatctc tgaaatatct 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 taccgacct agcagcccag 300

ggatggaaaa gtcccggccg tcgctggcaa taatagcggg cggacgcatg tcatgagatt 360

attggaaacc accagaatcg aatataaaag gcgaacacct ttccaattt tggtttctcc 420

tgacccaaag actttaaatt taatttattt gtccctattt caatcaattg aacaactatc 480

acctgcaggc c 491

<210> 14

<211> 33

10 <212> ADN

<213> artificial

<220>

ES 2 751 336 T3

<223> Cebador
 <400> 14
 gatagggcc cagggcatca ttgaggttc cac 33
 5 <210> 15
 <211> 36
 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 10 <400> 15
 gatactgca ggtttgtg ttgagtgaag cgagtg 36
 <210> 16
 <211> 33
 <212> ADN
 15 <213> artificial
 <220>
 <223> Cebador
 <400> 16
 gatagggcc cagggcatca ttgaggttc cac 33
 20 <210> 17
 <211> 35
 <212> ADN
 <213> artificial
 <220>
 25 <223> Cebador
 <400> 17
 gatagggcc gatagtcta gaagacctgg cgtcg 35
 <210> 18
 <211> 31
 30 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 <400> 18
 35 gatagggcc agccaacat ctttgttc g 31
 <210> 19
 <211> 31
 <212> ADN
 <213> artificial
 40 <220>
 <223> Cebador
 <400> 19
 gatagggcc gtggttcca ggacaacacc c 31
 45 <210> 20
 <211> 31
 <212> ADN
 <213> artificial

ES 2 751 336 T3

<220>
 <223> Cebador
 <400> 20
 gatagggccc gaccgcaatt caccatgatg c 31

5 <210> 21
 <211> 29
 <212> ADN
 <213> artificial
 <220>
 10 <223> Cebador
 <400> 21
 gatagggccc agcctgcttc attcctgcc 29
 <210> 22
 <211> 33
 15 <212> ADN
 <213> artificial
 <220>
 <223> Cebador
 <400> 22
 20 gatagggccc cgcgaaaaa ggttgttta tag 33
 <210> 23
 <211> 31
 <212> ADN
 <213> artificial
 25 <220>
 <223> Cebador
 <400> 23
 gatagggccc catactctcc tccccccct g 31
 <210> 24
 30 <211> 345
 <212> ADN
 <213> artificial
 <220>
 <223> pareja de fusión traduccional
 35 <400> 24
 ggccattacg gccgggggac ttacatttta ccgttccgtc actcgcttca ctcaacaaca 60
 aaaatgcaat tctctatcgt cgctactttg gctcttgctg gttccgctct ggctgcttac 120
 tctaacgtaa cttacactta cgagactacc atcaccgatg ttgtcaccga gctcaccact 180
 tactgcccag agccaaccac cttcgttcac aagaacaaga ccatcactgt gaccgcccc 240
 accactttga ccatcactga ctgtccttgc accatctcca agaccaccaa gatcaccact 300
 gatgttccac caaccaccca ctccacccca ctggccgcct cggcc 345

REIVINDICACIONES

1. Una secuencia aislada de un ácido nucleico promotor pCS1, que es una secuencia nativa de *Pichia pastoris* que comprende la secuencia de ácido nucleico pCS1 SEQ ID 1, o una variante funcionalmente activa de la misma que es una variante de tamaño, un mutante o híbrido de SEQ ID 1, o una combinación de los mismos, donde dicha variante funcionalmente activa tiene una longitud de 200-1500 bp, al menos 60% de identidad de secuencia con respecto a la secuencia de SEQ ID 1, y en la que dicha variante funcionalmente activa exhibe la misma expresión o fuerza transcripcional que la secuencia de ácido nucleico de pCS1 SEQ ID 1 +/- 20%, y/o al menos un aumento de 1,1 veces en comparación con el promotor nativo de *P. pastoris* pGAP establecido en SEQ ID 13, donde la secuencia de ácido nucleico promotora pCS1 no es la SEQ ID NO:2 del documento WO2014/138703 A1.
2. La secuencia de ácido nucleico según la reivindicación 1, que consiste en la secuencia de ácido nucleico de pCS1 SEQ ID 1.
3. La secuencia de ácido nucleico según la reivindicación 1, en la que la variante funcionalmente activa es:
- a) una variante de tamaño de SEQ ID 1, preferiblemente compuesta o consistente en la secuencia de ácido nucleico seleccionada del grupo que consiste en SEQ ID 2, 3, 4, 5 y 6, más preferiblemente la secuencia de ácido nucleico consiste en SEQ ID 2, 3, 4, 5 o 6;
- b) un mutante de SEQ ID 1, o un mutante de la variante de tamaño de a), cuyo mutante tiene al menos un 60% de homología con respecto a la secuencia SEQ ID 1 o a la variante de tamaño;
- c) un híbrido que comprende:
- una secuencia seleccionada del grupo que consiste en SEQ ID 1, una variante de tamaño de a), y un mutante de b); y
- al menos una secuencia más seleccionada del grupo que consiste en SEQ ID 1, una variante de tamaño de a), un mutante de b) y una secuencia heteróloga; o
- d) una secuencia que se hibrida bajo condiciones estrictas a cualquiera de la variante de tamaño, o las secuencias de ácido nucleico mutante de a), o b).
4. La secuencia de ácido nucleico según cualquiera de las reivindicaciones 1 a 3, en la que la variante funcionalmente activa se selecciona del grupo que consiste en:
- i) los homólogos que se pueden obtener modificando la secuencia de nucleótidos de SEQ ID 1 o las variantes de tamaño de los mismos, mediante la inserción, eliminación o sustitución de uno o más nucleótidos dentro de la secuencia o en uno o ambos extremos distales de la secuencia; y
- ii) análogos derivados de especies distintas de *Pichia pastoris*.
5. La secuencia de ácido nucleico según cualquiera de las reivindicaciones 1 a 4, que está unida operablemente a una secuencia de nucleótidos que codifica una proteína de interés (POI), cuyo ácido nucleico no se asocia de forma nativa con la secuencia de nucleótidos que codifica el POI.
6. La secuencia de ácido nucleico según la reivindicación 5, que comprende además una secuencia de ácido nucleico que codifica un péptido de señal que permite la secreción del POI, preferiblemente en el que la secuencia de ácido nucleico que codifica el péptido de señal se encuentra adyacente al extremo de 5' de la secuencia de nucleótidos que codifica el PDI.
7. Una construcción de expresión que comprende la secuencia de ácido nucleico según cualquiera de las reivindicaciones 1 a 6, preferiblemente un vector o plásmido replicante autónomamente, o uno que se integra en el ADN cromosómico de una célula huésped.
8. Una célula huésped recombinante que comprende una construcción de expresión de acuerdo con la reivindicación 7, preferiblemente una célula eucariota, más preferiblemente una levadura o célula fúngica filamentosa, más preferiblemente una célula de levadura del género *Saccharomyces* o *Pichia*.
9. La célula huésped recombinante según la reivindicación 8, que comprende múltiples copias de la secuencia de ácido nucleico, y/o múltiples copias de la construcción de la expresión.
10. La célula huésped recombinante según la reivindicación 8 o 9, que se selecciona del grupo formado por mamíferos, insectos, levaduras, hongos filamentosos y células vegetales, preferiblemente una levadura, preferiblemente cualquiera de las cepas de *P. pastoris* CBS 704, CBS 2612, CBS 7435, CBS 9173-9189, DSMZ 70877, X-33, GS115, KM71 y SMD1168.
11. Un cultivo estable de una pluralidad de células según cualquiera de las reivindicaciones 8 a 10.

12. Un método de producción de un POI mediante el cultivo de una línea de células huésped recombinantes que comprende al promotor de acuerdo con cualquiera de las reivindicaciones 1 a 6 o la construcción de la expresión según la reivindicación 7, y un ácido nucleico que codifica el POI bajo el control transcripcional de dicho promotor, o la célula huésped recombinante según cualquiera de las reivindicaciones 8 a 10, que comprende las etapas de:
- 5 a) cultivar la línea celular en condiciones que expresen dicho POI, y
b) recuperar el POI.
13. El método según la reivindicación 12, en el que el POI se expresa en condiciones de limitación del crecimiento.
14. El método según la reivindicación 12 o 13, en el que la línea celular se cultiva bajo condiciones de cultivo por lotes, alimentación por lotes o continua, y/o en medios que contengan sustrato de carbono limitado.
- 10 15. El método según la reivindicación 14, en el que el cultivo se realiza en un biorreactor a partir de una fase discontinua seguida de una fase alimentada por lotes o una fase de cultivo continuo.
16. El método según cualquiera de las reivindicaciones 12 a 15, en el que el POI es una proteína heteróloga, preferiblemente seleccionada de proteínas terapéuticas, incluidos anticuerpos o fragmentos de los mismos, enzimas y péptidos, antibióticos proteicos, proteínas de fusión de toxinas, conjugados carbohidratos proteicos, proteínas
- 15 estructurales, proteínas reguladoras, vacunas y vacunas como proteínas o partículas, enzimas de proceso, factores de crecimiento, hormonas y citoquinas, o un metabolito de un POI.
17. El uso de la secuencia aislada de ácido nucleico de cualquiera de las reivindicaciones 1 a 6 o la construcción de la expresión de la reivindicación 7 en un método de producción de un POI mediante el cultivo de una célula anfitriona transformada con la secuencia de ácido nucleico y/o la construcción de la expresión, preferiblemente en la que el
- 20 cultivo se realiza en un biorreactor a partir de una fase discontinua seguida de una fase de alimentación por lotes o una fase de cultivo continua.

Fig. 1

SEQ ID 1

AGGGCATCATTTGAGGTTTTCCACAAAAGGAAGAAACATGGATCCAGAGACATCAACAGAGAGGAAAGCGGGTAGTGA
 AGCCGAAGCCACAACACAGCCCGATTTGGAAGGGAGTTCAACAATCAAGGTGAGTCCAGCCATTTTTTTTTCTTTTT
 TTTTTTTTATTCAGGTGAACCCACCTAACTATTTTTAACTGGGATCCAGTGAGCTCGCTGGGTGAAAGCCAACCAT
 CTTTGTTCGGGGAAACCGTGCTCGCCCCGTAAAGTTAATTTTTTTTTCCCGCGCAGCTTTAATCTTTCGGCAGAG
 AAGGCGTTTTTCATCGTAGCGTGGGAACAGAATAATCAGTTCATGTGCTATACAGGCACATGGCAGCAGTCACTATT
 TTGCTTTTTAACCTTAAAGTCGTTTCATCAATCATTAACCTGACCAATCAGATTTTTTGCAATTTGCCACTTATCTAAA
 AATACTTTTGTATCTCGCAGATACGTTTCAGTGGTTTTCCAGGACAACACCCAAAAAAGGTATCAATGCCACTAGGC
 AGTCGGTTTTTATTTTTGGTCAACCCACGCAAGAAGCACCACCTCTTTTAGGTTTTAAGTTGTGGGAACAGTAACA
 CCGCTAGAGCTTCAGGAAAAACCAGTACCTGTGACCGCAATTCACCATGATGCAGAATGTTAATTTAAACGAGTG
 CCAATCAAGATTTCAACAGACAAATCAATCGATCCATAGTTACCCATTCCAGCCTTTTCGTCGTCGAGCCTGCTT
 CATTCCTGCCTCAGGTGCATAACTTTGCATGAAAAGTCCAGATTAGGGCAGATTTTGAGTTTTAAAATAGGAAATAT
 AAACAAATATACCGCGAAAAAGGTTTGTATAGCTTTTCGCTGGTGCCTGACGGTATAAATACATACTCTCCTC
 CCCCCCTGGTTCTCTTTTTCTTTTGTACTTACATTTTACCGTTCGGTCACTCGCTTCACTCAACAACAAAA

SEQ ID 2

GATAGTCTTAGAAGACCTGGCGTCGCTGGTCAACTACTCGTTTCAACAAGTTGACAAAGACTTTTTACCAAGAGAAGA
 GCAGTTGTCACCGACCACAATAACAACCTGGAAGCCGAGAAAAAACTTGAATCTCCAAGGTGAGACTGCCTAGAA
 ATCCGTACTCTGGGGCCGATCGAAGACTGCATTTTCCCTCCAAGAATTGATGCTCATGGTCTCATCTTACAATCGCAC
 ACACAACAGTGTGAGTCTTCTTTGCGGTCCCTTGAACACAACCAACCGAAAGGTGGGGAAGTCTAATGTCACGCAA
 ACGATATTGCAACCAATGTTGGGCTCTACTGGCGTCTGGCTGCATCAAATAGCTGATCGGTTGTAATCTTCAAAG
 ATTGGTGTAGGACGAACGAGTCTGCTGGGCTACAAGTTTTGCCCATATCGCTGTTCAAGCCAACCCGCGGAATCC
 CAAAACACCCCATCCGACAAAAGTTGTTGTTTTCAGCAGATCTAGGGAGGGCATCATTGAGGTTTTCCACAAAAGGA
 AGAAACATGGATCCAGAGACATCAACAGAGAGGAAAGCGGGTAGTGAAGCCGAAGCCACAACACAGCCCGATTTGG
 AAGGGAGTTCACAATCAAGGTGAGTCCAGCCATTTTTTTTTCTTTTTTTTTTTTTTATTCAGGTGAACCCACCTAAC
 TATTTTTAACTGGGATCCAGTGAGCTCGCTGGGTGAAAGCCAACCATCTTTGTTTCGGGGAAACCGTGCTCGCCCC
 GTAAAGTTAATTTTTTTTTTCCCGCGCAGCTTTAATCTTTTCGGCAGAGAAGGCGTTTTTCATCGTAGCGTGGGAACAG
 AATAATCAGTTCATGTGCTATACAGGCACATGGCAGCAGTCACTATTTTGCCTTTTAACTTAAAGTCGTTTCATCA
 ATCATTAACCTGACCAATCAGATTTTTTGCATTTGCCACTTATCTAAAAATACTTTTGTATCTCGCAGATACGTTCA
 GTGGTTTTCCAGGACAACACCCAAAAAAGGTATCAATGCCACTAGGCAGTCCGTTTTATTTTTGGTCAACCCACGCA
 AAGAAGCACCCACCTCTTTTAGGTTTTAAGTTGTGGGAACAGTAACACCCGCTAGAGCTTCAGGAAAAACCAGTAC
 CTGTGACCGCAATTCACCATGATGCAGAATGTTAATTTAAACGAGTGCCAAATCAAGATTTCAACAGACAAATCAA
 TCGATCCATAGTTACCCATTCCAGCCTTTTCGTCGTCGAGCCTGCTTCATTCCCTGCTCAGGTGCATAACTTTGCA
 TGAAAAGTCCAGATTAGGGCAGATTTTGAGTTTTAAAATAGGAAATATAAACAAATATACCGCGAAAAAGGTTTGT
 TATAGCTTTTCGCTGGTGCCTGACGGTATAAATACATACTCTCTCCCCCCCCCTGGTTCTCTTTTTCTTTTGTTA
 CTTACATTTTACCGTTCGGTCACTCGCTTCACTCAACAACAAAA

SEQ ID 3

AGCCAACCATCTTTTGTTCGGGGAAACCGTGCTCGCCCCGTAAAGTTAATTTTTTTTTTCCCGCGCAGCTTTAATCT
 TTCGGCAGAGAAGGCGTTTTTCATCGTAGCGTGGGAACAGAATAATCAGTTCATGTGCTATACAGGCACATGGCAGC
 AGTCACTATTTTGTCTTTTAACTTAAAGTCGTTTCATCAATCATTAACCTGACCAATCAGATTTTTTGCAATTTGCCA
 CTTATCTAAAAATACTTTTGTATCTCGCAGATACGTTTCAGTGGTTTTCCAGGACAACACCCAAAAAAGGTATCAAT
 GCCACTAGGCAGTCCGTTTTATTTTTGGTCAACCCACGCAAGAAGCACCACCTCTTTTAGGTTTTAAGTTGTGGG
 AACAGTAACACCGCTAGAGCTTCAGGAAAAACCAGTACCTGTGACCGCAATTCACCATGATGCAGAATGTTAAT
 TAAACGAGTGCCAAATCAAGATTTCAACAGACAAATCAATCGATCCATAGTTACCCATTCCAGCCTTTTCGTCGTC
 GAGCCTGCTTCATTCTGCCTCAGGTGCATAACTTTGCATGAAAAGTCCAGATTAGGGCAGATTTTGAGTTTTAAA
 TAGGAAATATAAACAAATATACCGCGAAAAAGGTTTGTATAGCTTTTCGCTGGTGCCTGACGGTATAAATACA
 TACTCTCTCCCCCCCCCTGGTTCTCTTTTTCTTTTGTACTTACATTTTACCGTTCGGTCACTCGCTTCACTCAAC
 AACAAAA

SEQ ID 4

GTGGTTTCCAGGACAACACCCAAAAAAGGTATCAATGCCACTAGGCAGTCGGTTTTATTTTTGGTCACCCACGCA
AAGAAGCACCCACCTCTTTTAGGTTTTAAGTTGTGGGAACAGTAACACCGCTAGAGCTTCAGGAAAAACCAGTAC
CTGTGACCGCAATTCACCATGATGCAGAATGTTAATTTAAACGAGTGCCAAATCAAGATTTCAACAGACAAATCAA
TCGATCCATAGTTACCCATTCCAGCCTTTTCGTCGTCGAGCCTGCTTCATTCCTGCCTCAGGTGCATAACTTTGCA
TGAAAAGTCCAGATTAGGGCAGATTTTGAGTTTAAAATAGGAAATATAAACAAATATACCGGAAAAAGGTTTGT
TATAGCTTTTCGCCTGGTGCCGTACGGTATAAATACATACTCTCCTCCCCCCCCTGGTTCTCTTTTTCTTTTGTTA
CTTACATTTTACCGTTCCGTCACTCGCTTCACTCAACAACAAAA

SEQ ID 5

GACCGCAATTCACCATGATGCAGAATGTTAATTTAAACGAGTGCCAAATCAAGATTTCAACAGACAAATCAATCGA
TCCATAGTTACCCATTCCAGCCTTTTCGTCGTCGAGCCTGCTTCATTCCTGCCTCAGGTGCATAACTTTGCATGAA
AAGTCCAGATTAGGGCAGATTTTGAGTTTAAAATAGGAAATATAAACAAATATACCGGAAAAAGGTTTGTGTTATA
GCTTTTCGCCTGGTGCCGTACGGTATAAATACATACTCTCCTCCCCCCCCTGGTTCTCTTTTTCTTTTGTTACTTA
CATTTTACCGTTCCGTCACTCGCTTCACTCAACAACAAAA

SEQ ID 6

AGCCTGCTTCATTCCTGCCTCAGGTGCATAACTTTGCATGAAAAGTCCAGATTAGGGCAGATTTTGAGTTTAAAAT
AGGAAATATAAACAAATATACCGGAAAAAGGTTTGTGTTATAGCTTTTCGCCTGGTGCCGTACGGTATAAATACAT
ACTCTCCTCCCCCCCCTGGTTCTCTTTTTCTTTTGTTACTTACATTTTACCGTTCCGTCACTCGCTTCACTCAACA
ACAAAA

SEQ ID 7

CCGCGAAAAAGGTTTGTGTTATAGCTTTTCGCCTGGTGCCGTACGGTATAAATACATACTCTCCTCCCCCCCCTGGT
TCTCTTTTTCTTTTGTTACTTACATTTTACCGTTCCGTCACTCGCTTCACTCAACAACAAAA

SEQ ID 8

CATACTCTCCTCCCCCCCCTGGTTCTCTTTTTCTTTTGTTACTTACATTTTACCGTTCCGTCACTCGCTTCACTCA
ACAACAAAA

Fig. 3

SEQ ID 13

CTTTTTGTAGAAATGTCTTGGTGTCCCTCGTCCAATCAGGTAGCCATCTCTGAAATATCTGGCT
CCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAAAACCTTAAATGTGGA
GTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGCCCGTTACCGTCCCTAGGA
AATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCCCCTTGCAGCAATGCTCTTCCCAGCATTAC
GTTGCGGGTAAAACGGAGGTCGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCG
TCGCTGGCAATAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAGAATCGAATA
TAAAAGGCGAACACCTTTCCCAATTTGGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTAT
TTGTCCCTATTTCAATCAATTGAACAACCTATCACCTGCAGGCC