

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 013 464**

51 Int. Cl.:

C12N 15/81 (2006.01)

C07K 14/39 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.12.2013 E 19186013 (9)**

97 Fecha y número de publicación de la concesión europea: **04.12.2024 EP 3584322**

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:
14.04.2025

73 Titular/es:

LONZA LIMITED (100.00%)
Lonzastrasse
3930 Visp, CH

72 Inventor/es:

MATTANOVICH, DIETHARD;
GASSER, BRIGITTE y
PRIELHOFER, ROLAND

74 Agente/Representante:

ELZABURU, S.L.P

Observaciones:

Véase nota informativa (Remarks, Remarques o
Bemerkungen) en el folleto original publicado por la
Oficina Europea de Patentes

ES 3 013 464 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).

DDESCRIPCIÓN

Promotor constitutivo

- 5 La invención se refiere a una secuencia aislada de ácido nucleico que comprende un promotor constitutivo fuerte y a un método de producción de una proteína de interés en un cultivo celular eucariota bajo el control de dicho promotor.

ANTECEDENTES DE LA INVENCION

- 10 La producción exitosa de proteínas recombinantes se ha logrado con huéspedes eucariotas. Los ejemplos más destacados son levaduras como *Saccharomyces cerevisiae*, *Pichia pastoris* o *Hansenula polymorpha*, hongos filamentosos como *Aspergillus awamori* o *Trichoderma reesei*, o células de mamífero como, por ejemplo, células CHO. Mientras que la producción de algunas proteínas se consigue fácilmente a tasas elevadas, muchas otras proteínas sólo se obtienen a niveles comparativamente bajos.

- 15 La expresión heteróloga de un gen en un organismo huésped usualmente requiere un vector que permita la transformación estable del organismo huésped. Un vector proveería al gen de un promotor funcional adyacente al extremo 5' de la secuencia codificante. De este modo, la transcripción es regulada e iniciada por esta secuencia promotora. La mayoría de los promotores utilizados hasta la fecha se han derivado de genes que codifican para proteínas que usualmente están presentes en altas concentraciones en la célula.

- 20 EP0103409A2 divulga el uso de promotores de levadura asociados a la expresión de enzimas específicas de la vía glucolítica, es decir, promotores implicados en la expresión de piruvato quinasa, triosafosfato isomerasa, fosfoglucosa isomerasa, fosfoglicerato mutasa, hexoquinasa 1 y 2, glucoquinasa, fosfofructosa quinasa, aldolasa y gen de regulación glucolítica.

- 25 WO 97/44470 describe promotores de levadura de *Yarrowia lipolytica* para la proteína del factor de elongación de la traducción 1 (TEF1 por sus siglas en inglés) y para la proteína ribosomal S7 que son adecuados para la expresión heteróloga de proteínas en levadura, y EP1951877A1 describe el uso del promotor TEF1 de *P. pastoris* para la producción de proteínas heterólogas.

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

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

- 40 El promotor AOX1 sólo se induce en respuesta al metanol y se reprime por medio de otras fuentes de carbono, como la glucosa o el etanol. El metanol tiene el inconveniente de que no es apto para usarse en la producción de determinados productos, ya que es potencialmente peligroso por su toxicidad e inflamabilidad. Por lo tanto, se buscan alternativas al promotor AOX1.

- 45 Vassileva et al. (J. Biotechnol. (2001) 88: 21-35) describe 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 con el fin de permitir el mantenimiento de las células en fase medio exponencial.

- 50 Los promotores utilizados en *Pichia pastoris* están ya sea estrechamente 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, se ha descrito que especialmente los promotores GAP y TEF son potentes y útiles para la producción de proteínas recombinantes.

- 55 Sin embargo, se demostró que la actividad de ambos promotores constitutivos no es constantemente fuerte durante un proceso de producción por lotes alimentados. Especialmente en las últimas fases del proceso, cuando las tasas de crecimiento celular son lentas, también la actividad de los promotores disminuye, limitando así los niveles de expresión del gen de interés (GOI por sus siglas en inglés) y los rendimientos de producción (Stadlmayr et al. 2010. J Biotechnol. 150: 519-529).

- 60 La selección de promotores adecuados no es intuitiva ni racional, ya que incluso enzimas glicolíticas altamente abundantes como la enolasa (ENO), la triosa fosfato isomerasa (TPI por sus siglas en inglés) o la glucosa-6-fosfato isomerasa (PGI por sus siglas en inglés) no tienen promotores que son tan potentes como pGAP y pTEF (Stadlmayr et al. 2010. J Biotechnol. 150: 519-529; Gasser et al. 2010. Metabolic Engineering 12:573-580).

- 65 Qin et al. (Applied and Environmental Microbiology (2011) 3600-3608) describe una biblioteca de promotores GAP y varios mutantes con actividades diversas.

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

5 WO2007/015178 A2 describe socios de fusión traslacional capaces de inducir la producción secretora de proteínas recombinantes, y una biblioteca de ADNc de *P. pastoris*.

CN 102180954 A describe un sistema de visualización de la superficie celular que utiliza la proteína de la pared celular GCW14 de *P. pastoris* como proteína anclada.

10

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

15

WO2014/138703A1 divulga promotores de levadura de *P. pastoris*, entre ellos un promotor denominado UPP representado en SEQ ID NO:2 de WO2014/138703 A1.

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

20

BREVE DESCRIPCIÓN DE LA INVENCION

El objeto se resuelve mediante la materia reivindicada.

25

De acuerdo con la invención se proporciona una secuencia aislada de ácido nucleico que comprende un promotor, el cual es una secuencia nativa de *Pichia pastoris* que comprende o consiste en la secuencia de ácido nucleico pCS1 de 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.

30

De acuerdo con un aspecto específico, la invención proporciona un promotor constitutivo aislado, el cual es una variante de tamaño del promotor pCS1 de *Pichia pastoris* establecido en SEQ ID NO:1, cuya variante de tamaño tiene una longitud de 500-1500 bp, al menos 90 % de identidad con la secuencia de SEQ ID NO:1, y comprende una secuencia de ácido nucleico que es al menos 90 % idéntica a cualquiera de SEQ ID NO: 2, 3, 4, o 5, donde dicha variante de tamaño exhibe la misma fuerza de expresión o transcripcional que el promotor pCS1 establecido en SEQ ID NO:1 +/- 20 %, y/o al menos un aumento de 1,1 veces en comparación con el promotor nativo pGAP de *P. pastoris* establecido en SEQ ID NO:13, donde la variante de tamaño no es SEQ ID NO:2 de WO2014/138703 A1.

35

Específicamente, la variante de tamaño comprende una secuencia de ácido nucleico que es cualquiera de SEQ ID NO: 2, 3, 4 o 5.

40

De acuerdo con un aspecto específico, el promotor exhibe una actividad promotora constitutiva determinada por el análisis de la transcripción en una célula huésped a una tasa de crecimiento elevada y a una tasa de crecimiento baja.

45

De acuerdo con un aspecto específico, el promotor está enlazado de forma operable a una secuencia de nucleótidos que codifica una proteína de interés (POI por sus siglas en inglés), cuyo ácido nucleico no está asociado de forma nativa a la secuencia de nucleótidos que codifica la POI.

50

De acuerdo con un aspecto específico, el promotor comprende además una secuencia de ácido nucleico que codifica un péptido señal que permite la secreción de la POI, preferiblemente donde la secuencia de ácido nucleico que codifica el péptido señal está situada adyacente al extremo 5' de la secuencia nucleotídica que codifica la POI.

La invención proporciona además un constructo de expresión que comprende el promotor descrito en el presente documento, preferiblemente un vector o plásmido de replicación autónoma, o uno que se integra en el ADN cromosómico de una célula huésped.

55

La invención proporciona además para una célula huésped recombinante que comprende el constructo de expresión descrito en el presente documento, preferiblemente una célula eucariota, más preferiblemente una célula de levadura o de hongo filamentoso, más preferiblemente una célula de levadura del género *Saccharomyces* o *Pichia*.

60

De acuerdo con un aspecto específico, la célula huésped recombinante comprende múltiples copias del promotor, y/o múltiples copias del constructo de expresión.

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

65

La invención proporciona además un cultivo estable de una pluralidad de la célula descrita en el presente documento.

La invención proporciona además un método de producción de una proteína de interés (POI) mediante el cultivo de una línea celular huésped recombinante, que comprende: i) el promotor descrito en el presente documento o el constructo de expresión descrito en el presente documento, donde el promotor está operablemente unido a un ácido nucleico que codifica la POI; o ii) la célula huésped recombinante descrita en el presente documento; que comprende los pasos de

- a) cultivar la línea celular en condiciones que expresen dicha POI, y
- b) recuperar la POI.

Concretamente, la POI se expresa en condiciones de crecimiento limitado.

De acuerdo con un aspecto específico, la línea celular se cultiva en condiciones de cultivo discontinuas, discontinuas alimentadas o continuas, y/o en medios que contienen un sustrato de carbono limitado.

De acuerdo con un aspecto específico, la POI es una proteína heteróloga, preferiblemente seleccionada entre proteínas terapéuticas, incluidos anticuerpos o fragmentos de los mismos, enzimas y péptidos, antibióticos proteicos, proteínas de fusión de toxinas, conjugados carbohidrato-proteína, proteínas estructurales, proteínas reguladoras, vacunas y proteínas o partículas similares a vacunas, enzimas de proceso, factores de crecimiento, hormonas y citoquinas, o un metabolito de una POI

La invención proporciona además el uso del promotor descrito en el presente documento o del constructo de expresión descrito en el presente documento en un método de producción de una proteína de interés (POI) mediante el cultivo de una célula huésped transformada con el promotor y/o el constructo de expresión, donde el promotor está unido de forma operable a una secuencia nucleotídica que codifica la POI, preferiblemente donde el cultivo se realiza en un biorreactor que comienza con una fase discontinua seguida de una fase de discontinua alimentada o una fase de cultivo continuo.

Específicamente, la presente divulgación se refiere a la secuencia de nucleótidos pCS1 la cual es idéntica a la secuencia de SEQ ID 1.

Preferiblemente, la secuencia de ácido nucleico descrita en el presente documento no es idéntica al ácido nucleico de SEQ ID 87 que se enlista en WO2007/015178 A2 (es decir, SEQ ID 24 de la presente solicitud).

En concreto, la variante funcionalmente activa es

- a) una variante de tamaño de pCS1 de SEQ ID 1, preferiblemente consistente en o que comprenda la secuencia de ácido nucleico seleccionada del grupo consistente en SEQ ID 2, 3, 4 y 5; o
- b) un mutante de la variante de tamaño de a), cuyo mutante tiene al menos un 90 % de homología con la secuencia SEQ ID 1 y al menos un 90 % de identidad con la secuencia de SEQ ID NO:1.

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

De acuerdo con un aspecto específico, el promotor se selecciona del grupo que consiste en homólogos con

- i) al menos un 90 % de identidad de secuencia nucleotídica;
- ii) homólogos obtenibles modificando la secuencia de nucleótidos de pCS1 de SEQ ID 1 o variantes de tamaño de la misma, mediante inserción, delección o sustitución de uno o más nucleótidos dentro de la secuencia o en ya sea uno o ambos de los extremos distales de la secuencia, preferiblemente con una secuencia nucleotídica de 500 bp a 1500 bp.

Específicamente, el promotor descrito en el presente documento tiene sustancialmente la misma actividad promotora que pCS1.

De acuerdo con un aspecto específico, la secuencia de ácido nucleico del promotor está operativamente unida a una secuencia de nucleótidos que codifica una proteína de interés (POI), cuyo ácido nucleico no está nativamente asociado a la secuencia de nucleótidos que codifica la POI.

De acuerdo con un aspecto específico, la secuencia de ácido nucleico comprende además un gen péptido señal que permite la secreción de la POI, preferiblemente donde el gen péptido señal está situado adyacente al extremo 5' de la secuencia nucleotídica que codifica la POI

Por consiguiente, la presente divulgación 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 de la POI, preferiblemente donde la secuencia de ácido nucleico que codifica el péptido señal está situada adyacente al extremo 5' de la secuencia de nucleótidos que codifica la POI.

La presente divulgación proporciona además un constructo de expresión que comprende una secuencia de ácido nucleico descrita en el presente documento, preferiblemente un vector o plásmido de replicación autónoma, o uno que se integra en el ADN cromosómico de una célula huésped.

5 La presente divulgación proporciona además una célula huésped recombinante que comprende la secuencia de ácido nucleico descrita en el presente documento o el constructo de expresión descrito en el presente documento, preferiblemente una célula eucariota, más preferiblemente una célula de levadura o de hongo filamentosos, más preferiblemente una célula de levadura del género *Saccharomyces* o *Pichia*.

10 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 copias del gen, GCN por sus siglas en inglés).

15 Específicamente, la célula huésped recombinante se selecciona del grupo que consiste en mamíferos, insectos, levaduras, hongos filamentosos y plantas, preferiblemente una levadura, preferiblemente cualquiera de las cepas *P. pastoris* CBS 704, CBS 2612, CBS 7435, CBS 9173- 9189, DSMZ 70877, X-33, GS115, KM71 y SMD1168. De acuerdo con un aspecto específico, la célula huésped recombinante es una cepa de *P. pastoris* distinta de X-33.

20 La presente divulgación proporciona además un cultivo estable de una pluralidad de las células descritas en el presente documento.

25 La presente divulgación proporciona además un método para producir una POI cultivando una línea celular huésped recombinante que comprende la secuencia de ácido nucleico o el promotor descritos en el presente documento, o el constructo de expresión descrito en el presente documento, y un ácido nucleico que codifica la POI bajo el control transcripcional de dicho promotor, que comprende los pasos de

- a) cultivar la línea celular en condiciones que expresen dicha POI, y
- b) recuperar la POI.

30 Específicamente, la POI se expresa en condiciones de crecimiento limitado, por ejemplo, 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 al 30 %, inferior al 20 %, inferior al 10 %, inferior al 5 %, inferior al 3 %, inferior al 2 %, inferior al 1 %, inferior al 0,5 %, inferior al 0,4 %, inferior al 0,3 %, o inferior al 0,2 % de la tasa de crecimiento máxima de las células. Normalmente, la tasa de crecimiento máxima se determina individualmente para una célula huésped específica.

35 De acuerdo con un aspecto específico, la línea celular se cultiva en condiciones de cultivo discontinuas, discontinuas alimentadas o continuas, y/o en medios que contienen un sustrato de carbono limitado.

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

45 Específicamente, las células huésped se cultivan en un medio rico en fuentes de carbono durante la fase de tasa de crecimiento elevada (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 máxima de crecimiento) y producen la POI durante una fase de baja tasa de crecimiento (por ejemplo, inferior al 90 %, preferiblemente inferior al 80 %, inferior al 70 %, inferior al 60 %, inferior al 50 %, o inferior al 40 %, inferior al 30 %, inferior al 20 %, inferior al 10 %, inferior al 5 %, inferior al 3 %, inferior al 2 %, inferior al 1 %, inferior al 0,5 %, inferior al 0,4 %, inferior al 0,3 %, o inferior al 0,2 % de la tasa de crecimiento máxima), por ejemplo, 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.

50 El POI es específicamente una proteína heteróloga, preferiblemente seleccionada de proteínas terapéuticas, incluyendo anticuerpos o fragmentos de las mismas, enzimas y péptidos, antibióticos proteicos, proteínas de fusión de toxinas, conjugados carbohidrato-proteína, proteínas estructurales, proteínas reguladoras, vacunas y proteínas o partículas similares a vacunas, enzimas de proceso, factores de crecimiento, hormonas y citoquinas, o un metabolito de una POI, incluyendo específicamente un metabolito celular del cultivo celular recombinante que expresa un gen de interés bajo el control transcripcional de un promotor descrito en el presente documento.

55 La presente divulgación proporciona además un método para identificar un promotor constitutivo de células eucariotas, que comprende los pasos de

- a) cultivar células eucariotas a una tasa de crecimiento elevada;
- b) cultivar además las células a una tasa de crecimiento baja;
- c) proporcionar muestras del cultivo celular de los pasos a) y b),
- 65 d) realizar 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 tenga una mayor fuerza de transcripción en comparación con el promotor pGAP nativo a tasas de crecimiento elevadas y bajas, preferiblemente determinando un nivel de transcripción del promotor constitutivo identificado que sea al menos 1,1 veces mayor en comparación con el 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 mayor.

Específicamente, el nivel de transcripción se determina a una tasa de crecimiento elevada y a una tasa de crecimiento baja dentro del intervalo de 0,01 a 0,2 h⁻¹, preferiblemente dentro del intervalo de 0,015 a 0,15 h⁻¹, por ejemplo examinando al menos dos muestras de un cultivo celular, una primera que represente una tasa de crecimiento elevada, tal 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 inferior a la primera, p. ej., inferior a 0,05 h⁻¹, preferiblemente inferior a 0,04 h⁻¹, inferior a 0,03 h⁻¹, o inferior a 0,02 h⁻¹, p. ej., 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 genética utilizando microarrays de ADN, por ejemplo, de acuerdo con el Ejemplo 11 a continuación.

La presente divulgación proporciona además el uso de una secuencia de ácido nucleico aislada que comprende o consiste en un promotor que cuando está unido operativamente a una secuencia de nucleótidos que codifica una POI dirige la expresión de la misma en una célula huésped a un nivel de expresión que es superior al que está bajo el control del promotor pGAP nativo a tasas de crecimiento elevadas y bajas, preferiblemente en un método de producción de una POI al mediante el cultivo de una célula huésped transformada con la secuencia de ácido nucleico, 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 continuo.

Específicamente, el nivel de expresión se determina a una tasa de crecimiento elevada y a una tasa de crecimiento baja dentro del intervalo de 0,01 a 0,2 h⁻¹, preferiblemente dentro del intervalo de 0,015 a 0,15 h⁻¹, por ejemplo, examinando al menos dos muestras de un cultivo celular, una primera que represente una tasa de crecimiento elevada, tal 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 velocidad de crecimiento de 0,15 h⁻¹, y una segunda que representa una velocidad de crecimiento baja, tal como por ejemplo inferior a la primera, por ejemplo inferior a 0,05 h⁻¹, preferiblemente inferior a 0,04 h⁻¹, inferior a 0,03 h⁻¹, o inferior a 0,02 h⁻¹, por ejemplo, a una velocidad 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, por ejemplo, en un cultivo en quimiostato con glucosa limitada, a tasas de crecimiento elevadas y bajas se proporciona 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 descrita en el presente documento o el constructo de expresión descrito en el presente documento se utiliza en un método de producción de una POI mediante el cultivo de una célula huésped transformada con la secuencia de ácido nucleico y/o el constructo de expresión, preferiblemente donde el cultivo se realiza en un biorreactor que comienza con una fase discontinua seguida de una fase de discontinua alimentada o una fase de cultivo continuo.

FIGURAS

Figura 1: Secuencia de ácido nucleico pCS1 (985 bp, 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 de pCS1, que promueven la expresión de CS1 en *P. pastoris*, que comprenden 1488 bp (SEQ ID 2), 767 bp (SEQ ID 3), 500 bp (SEQ ID 4), 344 bp (SEQ ID 5), 234 bp (SEQ ID 6), 138 bp (SEQ ID 7) y 85 bp (SEQ ID 8).

Figura 2: Secuencias de secuencia de nucleótidos codificantes y secuencia de aminoácidos de CS1

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 nativa pGAP de *P. pastoris* (GS115) (SEQ ID 13).

DESCRIPCIÓN DETALLADA DE LA INVENCION

Los términos específicos utilizados en esta memoria descriptiva tienen el siguiente significado.

El término "fuente de carbono" o "sustrato de carbono", tal como se utiliza en el presente documento, se refiere a un sustrato de carbono fermentable, normalmente un carbohidrato fuente, adecuado como fuente de energía para microorganismos, tales como los que pueden ser metabolizados por organismos huéspedes o líneas celulares de producción, en particular fuentes seleccionadas del grupo que consiste en monosacáridos, oligosacáridos, polisacáridos, alcoholes, incluido el glicerol, en forma purificada, en medios mínimos o proporcionados en materias primas, tal como un material nutritivo complejo. La fuente de carbono puede utilizarse, tal como se describe en el presente documento, como una única fuente de carbono o como una mezcla de diferentes fuentes de carbono.

El término "condiciones ricas en sustrato de carbono", tal como se utiliza 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 proveerse en un medio, tal como un medio basal o un medio complejo, pero también en un medio (químicamente) definido que contenga una fuente de carbono purificada. La fuente de carbono para usarse en una fase de crecimiento de un proceso de cultivo celular también se denomina en el presente documento "fuente de carbono basal" y normalmente se proporciona en una cantidad que proporciona el crecimiento celular, por ejemplo, para obtener densidades celulares de al menos 5 g/L de masa seca celular, preferiblemente de al menos 10 g/L de masa seca celular, o de al menos 15 g/L de masa seca celular, por ejemplo, que presenten viabilidades de más del 90 % durante los pasos de subcultivo estándar, preferiblemente superiores al 95 %.

En una fase de crecimiento, la fuente de carbono se utiliza típicamente en una cantidad en exceso o excedente, lo que se entiende como un exceso que proporciona energía para aumentar la biomasa, por ejemplo, durante el cultivo de una línea celular con una tasa de crecimiento específico elevada.

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

El término "condiciones limitadas de sustrato de carbono" o "fuente limitada de carbono", también denominado en el presente documento como "fuente suplementaria de carbono", tal como se utiliza en el presente documento, se entiende que se refiere específicamente al tipo y la cantidad de un sustrato de carbono que facilita la producción de productos de fermentación por las líneas celulares de producción, en particular en un proceso de cultivo con tasas de crecimiento controladas inferiores a la tasa máxima de crecimiento. La fase de producción sigue específicamente una fase de crecimiento, por ejemplo, en procesos de cultivo discontinuos alimentados y continuos.

En general, los procesos de cultivo celular se clasifican en cultivo discontinuo, cultivo continuo y cultivo discontinuo alimentado. El cultivo discontinuo es un proceso de cultivo mediante el cual se añade una pequeña cantidad de una solución de cultivo semilla a un medio y se cultivan las células sin añadir un medio adicional ni descargar una solución de cultivo durante el cultivo. El cultivo continuo es un proceso de cultivo mediante el cual se añade y descarga continuamente un medio durante el cultivo. El cultivo continuo también incluye el cultivo en perfusión. El cultivo discontinuo alimentado, el cual es un intermedio entre el cultivo discontinuo y el cultivo continuo y también denominado cultivo semidiscontinuo, es un proceso de cultivo por el cual se añade un medio de forma continua o secuencial durante el cultivo, pero, a diferencia del cultivo continuo, no se descarga continuamente una solución de cultivo.

Específicamente se prefiere un proceso discontinuo alimentado que se basa en la alimentación de un sustrato nutritivo limitante del crecimiento a un cultivo. La estrategia discontinua alimentada, incluye la fermentación discontinua alimentada única o discontinua alimentada repetida, se utiliza normalmente en los procesos bioindustriales para alcanzar una alta densidad celular en el biorreactor. La adición controlada del sustrato de carbono afecta directamente a la tasa de crecimiento del cultivo y ayuda a evitar el metabolismo de desbordamiento o la formación de subproductos metabólicos no deseados. En condiciones de fuente de carbono limitada, la fuente de carbono puede estar contenida específicamente en la alimentación de un proceso discontinuo alimentado. De este modo, el sustrato de carbono se proporciona en una cantidad limitada.

También en el cultivo en quimiostato o continuo descrito en el presente documento, la tasa de crecimiento puede controlarse estrictamente.

Una "cantidad limitada" de una fuente de carbono se entiende en el presente documento como la cantidad de una fuente de carbono necesaria para mantener una línea celular de producción en condiciones de crecimiento limitado, por ejemplo, en una fase de producción o modo de producción. Dicha cantidad limitada puede emplearse en un proceso discontinuo alimentado, en el que la fuente de carbono está contenida en un medio de alimentación y se suministra al cultivo a bajas tasas de alimentación para un suministro sostenido de energía, por ejemplo, para producir una POI, mientras se mantiene la biomasa a bajas tasas específicas de crecimiento. Normalmente, se añade un medio de alimentación 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, el cual está por debajo de un umbral predeterminado o incluso por debajo del límite de detección medido en un ensayo estándar (carbohidratos). La cantidad residual se determinaría normalmente en el caldo de fermentación al cosechar un producto de fermentación.

5 La cantidad limitada de una fuente de carbono también puede determinarse definiendo la tasa promedio de alimentación de la fuente de carbono al fermentador, por ejemplo, determinada por la cantidad añadida durante todo el proceso de cultivo, por ejemplo, la fase discontinua alimentada, por tiempo de cultivo, para determinar una cantidad promedio calculada por tiempo. Esta tasa promedio de alimentación se mantiene baja para asegurar la utilización completa de la fuente suplementaria de carbono por el cultivo celular, por ejemplo entre $0,6 \text{ g L}^{-1} \text{ h}^{-1}$ (g de fuente de carbono por L de volumen inicial de fermentación y h de tiempo) y $25 \text{ g L}^{-1} \text{ h}^{-1}$, preferiblemente entre $1,6 \text{ g L}^{-1} \text{ h}^{-1}$ y $20 \text{ g L}^{-1} \text{ h}^{-1}$.

10 La cantidad limitada de una fuente de carbono también puede determinarse midiendo la tasa específica de crecimiento, cuya tasa específica de crecimiento se mantiene baja, por ejemplo, inferior a la tasa específica de crecimiento máxima, durante la fase de producción, por ejemplo, dentro de un intervalo predeterminado, tal como en el intervalo 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 $0,15 \text{ h}^{-1}$.

15 Puede utilizarse cualquier tipo de carbono orgánico adecuado usado para el cultivo de células eucariotas. De acuerdo con un aspecto específico, 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.

20 De acuerdo con un aspecto específicamente preferido, la fuente de carbono basal se selecciona del grupo que consiste en glucosa, glicerol, etanol, o mezclas de los mismos, y material nutritivo complejo. De acuerdo con un aspecto preferido, la fuente de carbono basal es el glicerol.

25 De acuerdo con un aspecto específico 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. De acuerdo con un aspecto preferido, una fuente suplementaria de carbono es la glucosa.

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

Específicamente, un medio de alimentación, tal como se utiliza en el presente documento, está químicamente definido y no contiene metanol.

35 El término "químicamente definido" o "definido" con respecto a un medio de cultivo celular, tal como un medio mínimo o un medio de alimentación en un proceso discontinuo alimentado, significará un medio de cultivo adecuado para el cultivo celular *in vitro* de una línea celular de producción, en la que todos los componentes químicos y (poli) péptidos son conocidos. Normalmente, un medio químicamente definido está totalmente libre de componentes de origen animal y representa un entorno de cultivo celular puro y consistente.

40 El término "línea celular", tal y como se utiliza en el presente documento, se refiere a un clon establecido de un tipo celular particular que ha adquirido la capacidad de proliferar durante un periodo de tiempo prolongado. El término "línea celular huésped" se refiere a una línea celular utilizada para expresar un gen endógeno o recombinante o productos de una vía 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 su uso para el cultivo en un biorreactor con el fin de obtener el producto de un proceso de producción, tal como una POI. El término "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. Es bien entendido que el término no incluye a los seres humanos.

50 De acuerdo con las células huésped eucariotas específicamente preferidas descritas en el presente documento, la célula o línea celular se selecciona del grupo que consiste en líneas celulares de mamíferos, insectos, levaduras, hongos filamentosos y plantas, preferiblemente una levadura.

55 Específicamente, la levadura se selecciona del grupo que consiste en *Pichia*, *Candida*, *Torulopsis*, *Arxula*, *Hansenula*, *Yarrowia*, *Kluyveromyces*, *Saccharomyces*, *Komagataella*, preferiblemente una levadura metilotrófica.

Una levadura específicamente preferida es *Pichia pastoris*, *Komagataella pastoris*, *K. phaffii* o *K. pseudopastoris*.

60 El término "cultivo celular" o "cultivo", también denominado "fermentación", con respecto a una línea celular huésped se refiere al mantenimiento de células en un entorno artificial, por ejemplo, entorno *in vitro*, en condiciones que favorecen el crecimiento, la diferenciación o la viabilidad continuada, en estado activo o quiescente, de las células, específicamente en un biorreactor controlado de acuerdo con métodos conocidos en la industria.

65 Cuando se cultiva un cultivo celular utilizando los medios de cultivo descritos en el presente 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 ciertos aspectos, un medio de cultivo como el descrito en el presente documento se utiliza para cultivar células de acuerdo con técnicas estándar de cultivo celular bien conocidas en la técnica. En varios aspectos descritos en el presente documento, se proporciona un medio de cultivo que puede utilizarse para el crecimiento de células eucariotas, específicamente levaduras u hongos filamentosos.

Los medios de cultivo celular proporcionan los nutrientes necesarios para mantener y cultivar células en un entorno controlado, artificial e *in vitro*. Las características y composiciones de los medios de cultivo celular varían en función de los requisitos celulares particulares. Parámetros importantes incluyen la osmolalidad, pH y formulaciones nutritivas. La alimentación de nutrientes puede realizarse de forma continua o discontinua de acuerdo con métodos conocidos en la técnica. Los medios de cultivo utilizados como se describe en el presente documento son particularmente útiles para la producción de proteínas recombinantes.

Mientras que un proceso discontinuo 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 más nutrientes durante la fermentación, en un proceso discontinuo alimentado, después de una fase discontinua, tiene lugar una fase de alimentación en la que se suministran uno o más nutrientes al cultivo mediante alimentación. El propósito de la alimentación nutritiva es aumentar la cantidad de biomasa con el fin de 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 divulgación que emplea el promotor descrito en el presente documento no está restringida con respecto a un determinado modo de cultivo.

En ciertos aspectos descritos en el presente documento, el método descrito en el presente documento es un proceso discontinuo alimentado. Específicamente, una célula huésped transformada con un constructo de ácido nucleico que codifica una POI recombinante deseada, se cultiva en un medio en fase de crecimiento y se transiciona a un medio en fase de producción para producir una POI recombinante deseada.

El medio de alimentación puede añadirse al medio de cultivo en forma líquida o bien en una forma alternativa, tal como un sólido, por ejemplo, en forma de tableta u otro medio de liberación sostenida, o un gas, por ejemplo, dióxido de carbono. Sin embargo, de acuerdo con un aspecto específico, la cantidad limitada de una fuente de carbono suplementaria añadida al medio de cultivo celular, puede ser incluso cero. Preferiblemente, en condiciones de un sustrato de carbono limitado, la concentración de una fuente de carbono suplementaria en el medio de cultivo es de 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, tal 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.

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.

Preferiblemente, la cantidad limitada de una fuente de carbono suplementaria es limitante del crecimiento para mantener la tasa de crecimiento específica por debajo de la tasa de crecimiento específica máxima, tal como en el intervalo de 0,001 h⁻¹ a 0,20 h⁻¹, o de 0,02 h⁻¹ a 0,20 h⁻¹, preferiblemente entre 0,02 h⁻¹ y 0,15 h⁻¹.

En otro aspecto, las células huésped descritas en el presente documento se cultivan en modo continuo, por ejemplo, en un quimiostato. Un proceso de fermentación continua se caracteriza por una tasa definida, constante y continua de alimentación de medio de cultivo fresco en el biorreactor, por lo que el caldo de cultivo se elimina al mismo tiempo del biorreactor a la misma tasa de eliminación definida, constante y continua. Al mantener el medio de cultivo, la tasa de alimentación y la tasa de eliminación en el mismo nivel constante, los parámetros y las condiciones de cultivo en el biorreactor permanecen constantes.

Se entiende específicamente que un cultivo celular estable, tal 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, por ejemplo, al menos a un nivel de pg, incluso después de aproximadamente 20 generaciones de cultivo, preferiblemente al menos 30 generaciones, más preferiblemente al menos 40 generaciones, más preferiblemente de al menos 50 generaciones. Específicamente, se proporciona una línea celular huésped recombinante estable, lo que se considera una gran ventaja cuando se utiliza para la producción a escala industrial.

El cultivo celular descrito en el presente documento es particularmente ventajoso para métodos en una escala de fabricación industrial, por ejemplo, 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 discontinuo alimentado o discontinuo, o un proceso continuo o semicontinuo (por ejemplo, quimiostato).

El término "expresión" o "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 enlace operable, de modo que los huéspedes transformados o transfectados con estas secuencias son capaces de producir las proteínas codificadas o los metabolitos

de la célula huésped. Para efectuar la transformación, el sistema de expresión puede incluirse en un vector; sin embargo, el ADN pertinente también puede integrarse en el cromosoma huésped. La expresión puede referirse a productos de expresión secretados o no secretados, incluidos polipéptidos o metabolitos.

"Constructos de expresión" o "vectores" o "plásmido" utilizados en el presente documento se definen como secuencias de ADN que son requeridas 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 de expresión o plásmidos usualmente 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 tales como la zeocina, la kanamicina, el G418 o la higromicina), varios sitios de escisión por enzimas de restricción, una secuencia promotora adecuada y un terminador de transcripción, componentes que están unidos entre sí de forma operable. Los términos "plásmido" y "vector" como se utilizan en el presente documento incluyen secuencias de nucleótidos de replicación autónoma, así como secuencias de nucleótidos de integración genómica.

El constructo de expresión descrito en el presente documento comprende específicamente un promotor descrito en el presente documento, unido de forma operable a una secuencia de nucleótidos que codifica una POI bajo el control transcripcional de dicho promotor, cuyo promotor no está asociado de forma nativa con la secuencia codificante de la POI.

El término "heterólogo", tal como se utiliza en el presente documento con respecto a una secuencia de nucleótidos o aminoácidos o a una proteína, se refiere a un compuesto que es ya sea extraño, es decir, "exógeno", tal como, que no se encuentra en la naturaleza, o una célula huésped determinada; o que se encuentra de forma natural en una célula huésped determinada, por ejemplo, es "endógeno", sin embargo, en el contexto de un constructo heterólogo, por ejemplo, empleando un ácido nucleico heterólogo. La secuencia de nucleótidos heteróloga que se encuentra endógenamente también puede producirse en una cantidad no natural, por ejemplo, mayor de la esperada o mayor de la que se encuentra naturalmente en la célula. La secuencia de nucleótidos heteróloga, o un ácido nucleico que comprende la secuencia de nucleótidos heteróloga, posiblemente difiere en secuencia de la secuencia de nucleótidos endógena 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 entiende que cualquier secuencia de nucleótidos recombinante o artificial es heteróloga. Un ejemplo de polinucleótido heterólogo es una secuencia de nucleótidos no asociada de forma nativa con el promotor descrito en el presente documento, por ejemplo, para obtener un promotor híbrido, o unir de forma operable a una secuencia codificante, tal como se describe en el presente documento. Como resultado, puede obtenerse un polinucleótido híbrido o quimérico. Otro ejemplo de compuesto heterólogo es un polinucleótido codificante de POI unido de forma operable a un elemento de control transcripcional, por ejemplo, un promotor descrito en el presente documento, al que normalmente no está unida de forma operable una secuencia codificante de POI endógena y natural.

El término "variante", como se utiliza en el presente documento en el contexto de la presente divulgación, se referirá específicamente a cualquier secuencia derivada de una secuencia madre, por ejemplo, por variación de tamaño, por ejemplo, elongación o fragmentación, mutación, hibridación (incluida la combinación de secuencias), o con un grado específico de homología, o analogía.

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

El término "nativo", tal y como se utiliza en el presente documento en el contexto de la presente divulgación, se referirá específicamente a una estructura o componente individual de un organismo, que está asociado de forma natural con su entorno. Sin embargo, es bien sabido que las estructuras o componentes nativos pueden aislarse del entorno natural asociado y proporcionarse como estructuras o componentes nativos aislados. Dichas estructuras o componentes nativos aislados pueden ser también de origen artificial o sintético, y seguir teniendo las mismas características que los de origen natural.

Aunque algunos aspectos específicos descritos en el presente documento se refieren a estructuras o componentes nativos, por ejemplo, en forma aislada, queda bien entendido que los materiales, métodos y usos descritos en el presente documento, por ejemplo, refiriéndose específicamente a secuencias aisladas de ácidos nucleicos, secuencias de aminoácidos, constructos de expresión, células huésped transformadas y proteínas recombinantes, son "hechos por el hombre" y, por lo tanto, no se consideran resultado de la "ley de la naturaleza".

El promotor variante funcionalmente activo puede, por ejemplo, derivarse de la secuencia promotora pCS1 (SEQ ID 1) mediante mutagénesis, empleando así la secuencia pCS1 como secuencia "madre", para producir secuencias adecuadas para usarse como un promotor en líneas celulares recombinantes. Dicho promotor variante puede obtenerse a partir 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 propiedades mejoradas, por ejemplo, una mejora en la fuerza promotora para apoyar la producción de POI, pero con una función y fuerza promotoras sustancialmente iguales a tasas de crecimiento elevadas y bajas, lo que en el presente documento se entiende específicamente como "función independiente de la tasa de crecimiento".

- El promotor variante también puede derivarse de secuencias análogas, por ejemplo, de especies eucariotas distintas de *Pichia pastoris* o de un género distinto de *Pichia*, tal como *K. lactis*, *Z. rouxii*, *P. stipitis*, *H. polymorpha*. Específicamente, las secuencias promotoras análogas asociadas nativamente con genes análogos a los correspondientes genes de *P. pastoris* pueden utilizarse como tales o como secuencias madre para producir variantes funcionalmente activas de los mismos. Específicamente, se caracteriza un promotor análogo a pCS1 que se asocia de forma nativa con un gen análogo a CS1 (véase la secuencia de aminoácidos de SEQ ID 9 u 11). Las propiedades de dichas secuencias promotoras análogas o de sus variantes funcionalmente activas pueden por lo tanto determinarse utilizando técnicas estándar.
- La variante "funcionalmente activa" de una secuencia nucleotídica o promotora, tal como se utiliza en el presente documento, se refiere específicamente a una secuencia mutante, por ejemplo, resultante de la modificación de una secuencia madre por inserción, delección o sustitución de uno o más nucleótidos dentro de la secuencia o en ya sea uno o ambos extremos distales de la secuencia, y cuya modificación no afecta (en particular, perjudica) a la actividad de esta secuencia.
- Específicamente, la variante funcionalmente activa de la secuencia promotora descrita en el presente documento se selecciona del grupo que consiste en
- homólogos con al menos aproximadamente 90 % de identidad de secuencia nucleotídica, o al menos 90 % de grado de homología o identidad de secuencia con la secuencia madre; y/o
 - homólogos obtenibles modificando la secuencia de nucleótidos parental, tal como la secuencia pCS1 o la secuencia de una variante de tamaño utilizada como plantilla para proporcionar mutaciones, por ejemplo, por inserción, delección o sustitución de uno o más nucleótidos dentro de la secuencia o en ya sea uno o ambos de los extremos distales de la secuencia, preferiblemente con (es decir, que comprenda o consista en) una secuencia de nucleótidos de 500 bp a 1500 bp, preferiblemente de al menos 600 bp, al menos 700 bp, al menos 800 bp, al menos 900 bp o al menos 1000 bp.
- Las variantes funcionalmente activas específicamente preferidas son las derivadas de un promotor descrito en el presente documento mediante modificación, extensión y/o fragmentos de la secuencia promotora, con (es decir, que comprenden o consisten en) una secuencia de nucleótidos de al menos 500 bp, preferiblemente al menos 600 bp, al menos 700 bp, al menos 800 bp, al menos 900 bp, o al menos 1000 bp, preferiblemente hasta 1500 bp.
- Una variante funcionalmente activa de una secuencia promotora madre como la descrita en el presente documento puede obtenerse específicamente mediante métodos de mutagénesis. El término "mutagénesis", tal como se utiliza en el contexto descrito en el presente documento, se referirá a un método para proporcionar mutantes de una secuencia de nucleótidos, por ejemplo, mediante la inserción, delección y/o sustitución de uno o más nucleótidos, con el fin de obtener variantes de la misma con al menos un cambio en la región no codificante o codificante. La mutagénesis puede ser a través de una mutación aleatoria, semialeatoria o dirigida. Normalmente se producen grandes bibliotecas genéticas aleatorias con una gran diversidad de genes, que pueden seleccionarse de acuerdo con un genotipo o fenotipo específicamente deseado.
- Algunas de las variantes funcionalmente activas preferidas del promotor descritas en el presente documento son variantes de tamaño o específicamente fragmentos de pCS1, preferiblemente las que incluyen el extremo 3' de una secuencia de nucleótidos del promotor, por ejemplo, una secuencia de nucleótidos derivada de una de las secuencias de nucleótidos del promotor la cual tiene de una longitud específica e inserciones o una delección de la región terminal 5', por ejemplo, un alargamiento o corte de la secuencia de nucleótidos en el extremo 5', de modo que se obtenga una longitud específica con un intervalo desde el extremo 3' hasta un extremo 5' variable, tal como con una longitud de la secuencia de nucleótidos de al menos 500 bp, preferiblemente de al menos 600 bp, al menos 700 bp, al menos 800 bp, al menos 900 bp, o al menos 1000 bp.
- La variante de tamaño alargado descrita en el presente documento comprende preferiblemente uno o más nucleótidos adicionales en el extremo 5' de la secuencia pCS1, por ejemplo, los que están asociados de forma nativa con la secuencia 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), o pCS1d (SEQ ID 5), por lo tanto, una secuencia de nucleótidos dentro del intervalo de 500-1500 bp. 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), o pCS1d (SEQ ID 5).
- También se entiende que la variante funcionalmente activa de un promotor descrita en el presente documento abarca los híbridos de la pCS1 o de cualquiera de sus variantes funcionalmente activas, en particular cualquiera de las secuencias de la variante o fragmento de tamaño madre, por ejemplo, resultantes de la combinación con una o más de cualquiera de las secuencias que se califican como pCS1 o variantes funcionalmente activas de la misma, por ejemplo, al menos dos de dichas secuencias madre, al menos 3, al menos 4 o al menos 5 de las secuencias, por ejemplo, una combinación de dos o más de las secuencias alargadas pCS1 o fragmentos 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 otro aspecto, el híbrido se compone de al menos una de las secuencias seleccionadas de pCS1 o cualquiera de las variantes funcionalmente activas del mismo, en particular cualquiera de las secuencias de variantes de tamaño o fragmentos, y una secuencia heteróloga que, por ejemplo, no está asociada de forma nativa con la secuencia pCS1 en *P. pastoris*.

Se entiende además que la variante funcionalmente activa de un promotor descrita en el presente documento abarca una secuencia de nucleótidos que hibrida en condiciones estrictas con el promotor pCS1 o cualquiera de las variantes de tamaño funcionalmente activas o fragmentos, mutantes o secuencias híbridas de ácido nucleico del mismo.

Tal como se utiliza en el presente documento, el término "hibridación" o "hibridando" se pretende que se refiera al proceso durante el cual dos secuencias de ácido nucleico se adhieren entre sí con enlaces de hidrógeno estables y específicos para formar una doble cadena en condiciones adecuadas. La hibridación entre dos secuencias complementarias o suficientemente complementarias depende de las condiciones operativas que se utilicen y, en particular, de la rigurosidad. La rigurosidad puede entenderse que denotan el grado de homología; cuanto mayor sea la rigurosidad, mayor será el porcentaje de homología entre las secuencias. La rigurosidad puede definirse en particular por la composición de bases de las dos secuencias nucleicas, y/o por el grado de discordancia entre estas dos secuencias nucleicas. Variando las condiciones, por ejemplo, la concentración de sal y la temperatura, se puede permitir que una secuencia de ácido nucleico determinada hibride sólo 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.

Tal como se utiliza en la presente divulgación, la frase "hibridación en condiciones de hibridación estrictas" se entiende preferiblemente que se refiere a la hibridación en condiciones de cierta rigurosidad. En un aspecto preferido, las "condiciones estrictas de hibridación" son condiciones en las que la homología de las dos secuencias de ácido nucleico es al menos 70 %, preferiblemente al menos 80 %, preferiblemente al menos 90 %, es decir, en condiciones en las que la hibridación sólo es posible si la doble cadena obtenida durante esta hibridación comprende preferiblemente al menos 70 %, preferiblemente al menos 80 %, preferiblemente al menos 90 % de enlaces A-T y enlaces C-G.

La rigurosidad puede depender de los parámetros de la reacción, tal 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 adecuadas, por ejemplo, como se describe en Sambrook et al. (Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1989).

La variante funcionalmente activa descrita en el presente documento se caracteriza específicamente por mostrar sustancialmente la misma actividad que pCS1.

El término "sustancialmente la misma actividad", tal como se utiliza en el presente documento, se refiere específicamente a la actividad indicada por sustancialmente la misma fuerza promotora o mejorada, específicamente la fuerza de expresión o transcripcional del promotor, y sus características sustancialmente las mismas o mejoradas con respecto a la fuerza promotora, determinadas específicamente independientemente de la tasa de crecimiento de la célula huésped, tal como una fuerza de expresión o transcripcional sustancialmente la misma que pCS1, por ejemplo +/- 20 %, y/o mayor que el promotor pGAP nativo de la célula huésped, por ejemplo, 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 en relación con la fuerza del promotor pGAP, o incluso más, por ejemplo, al menos 3 veces, al menos 4 veces, al menos 5 veces, al menos 10 veces, o al menos hasta 15 veces la actividad, según se determine en un sistema de ensayo adecuado que emplee el mismo tipo de célula huésped, las mismas condiciones de cultivo y el mismo ácido nucleico que codifica un producto de expresión tal como una POI.

El término "homología" indica que dos o más secuencias de nucleótidos tienen pares de bases iguales o conservadas en una posición correspondiente, hasta un cierto grado, cercano al 100 %. Una secuencia homóloga como la descrita en el presente documento suele tener al menos aproximadamente 60 % de identidad de secuencia nucleotídica, preferiblemente al menos aproximadamente 70 % de identidad, más preferiblemente al menos aproximadamente 80 % de identidad, más preferiblemente al menos aproximadamente 90 % de identidad, más preferiblemente al menos aproximadamente 95 % de identidad, más preferiblemente al menos aproximadamente 98 % o 99 % de identidad.

La secuencia promotora homóloga descrita en el presente documento tiene preferiblemente cierta homología con cualquiera de las secuencias nucleotídicas promotoras pCS1, pCS1a, pCS1b, pCS1c, pCS1d, pCS1e, pCS1f y pCS1g de *P. pastoris* en al menos partes específicas de la secuencia nucleotídica, tal como incluyendo la región 3' de la secuencia nucleotídica promotora respectiva, preferiblemente una parte con una longitud específica hasta el extremo 3' de la secuencia nucleotídica promotora respectiva, tal como una parte con una longitud de 500 bp a 1500 bp, más preferiblemente de 500 a 1488 bp; preferiblemente al menos 600 bp, al menos 700 bp, al menos 800 bp, al menos 900 bp, o al menos 1000 bp. Específicamente, al menos esas partes son preferiblemente homólogas dentro del intervalo de 300-1000 bp, incluyendo la secuencia terminal 3' de la secuencia nucleotídica promotora respectiva.

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 deriven de especies distintas de *Pichia pastoris* pueden comprender una secuencia homóloga, es decir, una secuencia con una cierta homología como la descrita en el presente documento. Por lo tanto, el término "homólogo" también puede incluir secuencias análogas. Por otra parte, se entiende que la divulgación también se refiere a secuencias análogas y homólogas de las mismas que presentan cierta homología.

"Porcentaje (%) de identidad" con respecto a la secuencia de nucleótidos de un gen se define como el porcentaje de nucleótidos de una secuencia de ADN candidata que es idéntica a los nucleótidos de la secuencia de ADN, tras alinear la secuencia e introducir huecos, si es necesario, para alcanzar el máximo porcentaje 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 nucleotídica puede lograrse de varias maneras que están dentro de la habilidad en la técnica, por ejemplo, utilizando software informático disponible públicamente. Los expertos en la materia pueden determinar los parámetros adecuados para medir la alineación, incluidos los algoritmos necesarios para lograr la máxima alineación en toda la longitud de las secuencias comparadas.

El término "aislado" o "aislamiento", tal como se utiliza en el presente documento con respecto a un ácido nucleico, una POI u otro compuesto, se referirá a dicho compuesto que ha sido suficientemente separado del entorno con el que se asociaría de forma natural, 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 interfieren con la actividad fundamental, y que pueden estar presentes, por ejemplo, debido a una purificación incompleta. En particular, las moléculas aisladas de ácido nucleico de la presente divulgación también significa que incluyen aquellas sintetizadas químicamente. Con referencia a los ácidos nucleicos descritos en el presente documento, a veces se utiliza el término "ácido nucleico aislado" o "secuencia de ácido nucleico aislada". Este término, cuando se aplica al ADN, se refiere a una molécula de ADN que está separada de las secuencias con las que es inmediatamente contigua en el genoma que ocurre de manera 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 plásmido o virus vector, o integrada en el ADN genómico de una célula procariota o eucariota u organismo huésped. Específicamente, el término "ácido nucleico aislado", tal como se describe en el presente documento, excluye que la secuencia pCS1 esté unida a un ácido nucleico que codifique 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.

El término "enlazado operablemente", tal como se utiliza en el presente documento, se refiere a la asociación de secuencias de nucleótidos en una única molécula de ácido nucleico, por ejemplo, un vector, de forma que la función de una o más secuencias de nucleótidos se vea afectada por al menos otra secuencia de nucleótidos presente en dicha molécula de ácido nucleico. Por ejemplo, un promotor está operativamente unido a una secuencia codificante de un gen recombinante, cuando es capaz de efectuar la expresión de dicha secuencia codificante. Como ejemplo adicional, un ácido nucleico que codifica un péptido señal está unido operativamente a una secuencia de ácido nucleico que codifica una POI, cuando es capaz de expresar una proteína en la forma secretada, tal como una preforma de una proteína madura o la proteína madura. Específicamente, tales ácidos nucleicos enlazados operablemente entre sí pueden estar inmediatamente enlazados, es decir, sin otros elementos o secuencias de ácido nucleico entre el ácido nucleico que codifica el péptido señal y la secuencia de ácido nucleico que codifica una POI.

El término "promotor" como se utiliza 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 promotora puede evaluarse por su eficacia transcripcional. Esto puede determinarse directamente midiendo la cantidad de transcripción de ARNm a partir del promotor, por ejemplo, mediante Northern Blotting, o indirectamente midiendo la cantidad de producto génico expresado a partir del promotor.

El promotor descrito en el presente documento inicia, regula o de otro modo media o controla específicamente la expresión de un ADN codificante. El ADN promotor y el ADN codificante pueden proceder del mismo gen o de genes diferentes, y pueden proceder del mismo organismo o de organismos diferentes.

El promotor descrito en el presente documento se entiende específicamente como un promotor constitutivo, es decir, un promotor que controla la expresión sin necesidad de inducción, ni posibilidad de represión. Por lo tanto, hay una expresión continua y constante a un cierto nivel. Debido a la función única de alta fuerza promotora en todas las fases de crecimiento o tasas de crecimiento de la célula huésped, el promotor constitutivo descrito en el presente documento es particularmente útil en un cultivo discontinuo alimentado de la línea celular huésped. Los promotores constitutivos de la técnica anterior tenían el inconveniente de su baja potencia en condiciones de crecimiento limitado, por ejemplo, en procesos discontinuos alimentados, o se utilizaban en realizaciones en las que las células huésped se mantenían a una tasa de crecimiento elevada, por ejemplo, en la fase exponencial media.

La fuerza del promotor descrita en el presente documento se refiere específicamente a su fuerza de transcripción, representada por la eficiencia de iniciación de la transcripción que ocurre en ese promotor con alta o baja frecuencia. Cuanto mayor sea la fuerza de transcripción, más frecuentemente se producirá la transcripción en ese promotor. La fuerza del promotor es importante porque determina la frecuencia con la que se transcribe una determinada secuencia de ARNm,

- dando efectivamente mayor prioridad a la transcripción de unos genes sobre otros, lo que conduce a una mayor concentración del transcrito. Un gen que codifica una proteína que es requerida en grandes cantidades, por ejemplo, suele tener un promotor relativamente fuerte. La ARN polimerasa sólo puede realizar una tarea de transcripción a la vez, por lo que debe priorizar su trabajo para ser eficiente. Las diferencias en la fuerza del promotor se seleccionan para permitir esta priorización. Como se describe en el presente documento, el promotor es relativamente fuerte independientemente del metabolismo o de la tasa de crecimiento de la célula huésped, por ejemplo, ambos, durante fases de crecimiento elevadas y bajas de un cultivo celular y específicamente independiente de la fuente de carbono, exhibiendo un estado de actividad aproximadamente máxima, específicamente a aproximadamente un nivel constante.
- La fuerza relativa se determina comúnmente con respecto a un promotor estándar, tal como el promotor pGAP respectivo de la célula utilizada como huésped. La frecuencia de transcripción se entiende comúnmente como la tasa de transcripción, por ejemplo, determinada por la cantidad de un transcrito en un ensayo adecuado, por ejemplo, RT-PCR o Northern blotting. 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 fuerza de expresión y/o transcripción de un promotor descrito en el presente documento se determina en la célula huésped que es *P. pastoris* y se compara con el promotor nativo pGAP de *P. pastoris*.
- La tasa de transcripción puede determinarse por la fuerza de transcripción en un microarray, o con PCR cuantitativa en tiempo real (qRT-PCR por sus siglas en inglés) donde los datos del microarray o de la qRT-PCR muestran la diferencia de nivel de expresión entre condiciones con elevada tasa de crecimiento y condiciones con baja tasa de crecimiento, o condiciones que emplean diferente composición de medios, y una alta intensidad de señal en comparación con el promotor pGAP nativo. En el ejemplo 11 a continuación, se describe específicamente un sistema de ensayo adecuado.
- El promotor descrito en el presente documento ejerce una fuerza de transcripción relativamente alta, reflejada por una tasa de transcripción o fuerza de transcripción de al menos 110 % en comparación con el promotor pGAP nativo en la célula huésped, a veces denominado "promotor pGAP homólogo". Preferiblemente, la tasa 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 superior en comparación con el promotor pGAP nativo, por ejemplo, determinado en la célula eucótica seleccionada como célula huésped para producir la POI, tal como se determina en condiciones ricas en sustrato de carbono, por ejemplo, en cultivos discontinuos, o en condiciones limitadas de sustrato de carbono, por ejemplo, cultivos en quimiostato o discontinuos alimentados.
- Preferiblemente, el análisis de la transcripción es cuantitativo o semicuantitativo, empleando preferiblemente qRT-PCR, microarrays de ADN, secuenciación de ARN y análisis del transcriptoma.
- La tasa de expresión puede, por ejemplo, determinarse por la cantidad de expresión de un gen reportero, tal como eGFP, por ejemplo, como se describe en la sección Ejemplo a continuación, un sistema de prueba se describe específicamente en el Ejemplo 10. Se pudo demostrar que el promotor pCS1 tiene una tasa de transcripción relativamente alta, de al menos el 110 % en comparación con el promotor nativo pGAP, al cultivar un clon en solución.
- El promotor descrito en el presente documento ejerce una fuerza de expresión relativamente alta, reflejada por un nivel de expresión del gen de interés, que es al menos 110 % en comparación con el promotor pGAP nativo en la célula huésped, a veces denominado "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 superior en comparación con el promotor pGAP nativo, por ejemplo, determinado en la célula eucótica seleccionada como célula huésped para producir la POI, tal como se determina en condiciones ricas en sustrato de carbono, por ejemplo, en cultivos discontinuos, o en condiciones limitadas de sustrato de carbono, por ejemplo, cultivos en quimiostato o discontinuos alimentados.
- El promotor nativo pGAP inicia la expresión del gen *gap* que codifica la gliceraldehído-3-fosfato deshidrogenasa (GAPDH por sus siglas en inglés), que es un promotor constitutivo presente en la mayoría de los organismos vivos. GAPDH (EC 1\2\1\12), enzima clave de la glucólisis y la gluconeogénesis, desempeña un papel crucial en el metabolismo catabólico y anabólico de los carbohidratos. Por lo tanto, el promotor pGAP, aunque se entiende que es un promotor constitutivo (tal como no inducible por la alimentación con fuentes de carbono específicas), es un promotor metabólico que ejerce una fuerza promotora creciente con el aumento de la tasa de crecimiento. Por lo tanto, el promotor pGAP es difícilmente 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.
- Por el contrario, el promotor descrito en el presente documento mantiene sorprendentemente su alta fuerza promotora (esencialmente) a un nivel de transcripción constantemente alto en todas las fases de crecimiento de un cultivo celular huésped.

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

Por ejemplo, un promotor pGAP nativo de *P. pastoris* es la secuencia promotora endógena no modificada en *P. pastoris*, como se utiliza para controlar la expresión de GAPDH en *P. pastoris*, por ejemplo, que tiene la secuencia mostrada en la Figura 3: secuencia promotora pGAP nativa de *P. pastoris* (GS115) (SEQ ID 13). Si se utiliza *P. pastoris* como huésped para producir una POI como el descrito en el presente documento, la fuerza o tasa de transcripción del promotor descrito en el presente documento se compara con dicho promotor pGAP nativo de *P. pastoris*.

Como otro ejemplo, un promotor pGAP nativo de *S. cerevisiae* es la secuencia promotora endógena no modificada en *S. cerevisiae*, como se utiliza para controlar la expresión de GAPDH en *S. cerevisiae*. Si se utiliza *S. cerevisiae* como huésped para producir una POI como el descrito en el presente documento, la fuerza o tasa de transcripción del promotor descrito en el presente documento se compara con dicho promotor pGAP nativo de *S. cerevisiae*.

Por lo tanto, la expresión relativa o la fuerza de transcripción de un promotor descrito en el presente documento suele compararse con el promotor pGAP nativo de una célula de la misma especie o cepa que se utiliza como huésped para producir una POI.

Se entiende específicamente que el promotor descrito en el presente documento no es preferiblemente un promotor metabólico tal como, por ejemplo, un promotor naturalmente ligado de forma operable a un gen que codifica una enzima glucolítica abundante o una enzima de la gluconeogénesis, una proteína ribosómica o una enzima tal como una proteasa intracelular o una proteasa siendo secretada por la célula huésped.

Específicamente preferido es un promotor descrito en el presente documento, el cual tiene al menos una fuerza de expresión o fuerza de transcripción de pCS1. La fuerza promotora comparativa empleando el promotor pGAP como referencia puede determinarse por medios estándar, tal como midiendo la cantidad de productos de expresión o la cantidad de transcritos, por ejemplo, empleando un microarray, Northern Blot, 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. En la sección Ejemplos se ilustran pruebas ejemplares.

El término "proteína de interés (POI)", tal y como se utiliza 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 ya sea un polipéptido que no ocurre de forma natural 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 es producida, por ejemplo, por transformación con un vector autorreplicante que contenga la secuencia de ácido nucleico que codifica la POI, o hasta la integración mediante técnicas recombinantes de una o más copias de la secuencia de ácido nucleico que codifica la POI en el genoma de la célula huésped, o por modificación recombinante de una o más secuencias reguladoras que controlan la expresión del gen que codifica la POI, por ejemplo, de la secuencia promotora. En algunos casos, el término POI, tal como se utiliza en el presente documento, también se refiere a cualquier metabolito producido por la célula huésped mediado por la proteína expresada recombinantemente.

De acuerdo con una realización específica, el término POI excluiría la proteína CS1 de *Pichia pastoris*, por ejemplo, caracterizada por cualquiera de las secuencias de aminoácidos de la Figura 2, en particular, si el promotor descrito en el presente documento está asociado de forma nativa con cualquiera de las secuencias de ácido nucleico que codifican dicha proteína CS1.

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

Una POI producida descrita en el presente documento puede ser una proteína multimérica, preferiblemente un dímero o tetramero.

De acuerdo con un aspecto descrito en el presente documento, la POI es una proteína recombinante o heteróloga, preferiblemente seleccionada entre proteínas terapéuticas, incluidos anticuerpos o fragmentos de los mismos, enzimas y péptidos, antibióticos proteicos, proteínas de fusión de toxinas, conjugados carbohidrato-proteína, proteínas estructurales, proteínas reguladoras, vacunas y proteínas o partículas similares a vacunas, enzimas de proceso, factores de crecimiento, hormonas y citoquinas, o un metabolito de una POI.

Una POI específica es una molécula de unión a antígeno, tal como un anticuerpo, o un fragmento del mismo. Entre las POI específicas se encuentran anticuerpos tales como los anticuerpos monoclonales (mAbs), inmunoglobulinas (Ig) o inmunoglobulinas de clase G (IgG), anticuerpos de cadena pesada (HcAb's), o fragmentos de los mismos tales como fragmentos de unión a antígeno (Fab), Fd, fragmento variable de cadena única (scFv), o sus variantes diseñadas tales como, por ejemplo, dímeros Fv (diabodies), trímeros Fv (triabodies), tetrameros Fv, o minicuerpos y anticuerpos de dominio único como VH o VHH o V-NAR.

De acuerdo con una realización específica, se fabrica un producto de fermentación utilizando la POI, un metabolito o un derivado del mismo.

5 La POI puede recuperarse específicamente del cultivo celular en forma purificada, por ejemplo, sustancialmente pura.

Los términos "sustancialmente puro" o "purificado" como se utilizan en el presente documento se refieren a un preparado que comprende al menos 50 % (p/p), preferiblemente al menos 60 %, 70 %, 80 %, 90 % o 95 % de un compuesto, tal como una molécula de ácido nucleico o una 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 utiliza en el presente documento, significará "siendo preparado mediante o el resultado de ingeniería genética". Así, 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 codificantes modificada genéticamente adecuada para su uso como promotor funcionalmente activo como se describe en el presente documento.

20 Por lo tanto, se identificó un nuevo promotor con propiedades funcionales únicas. Inesperadamente, PAS_chr1-4 (denominado en el presente documento gen CS1, SEQ ID 9) fue identificado mediante el análisis de la fuerza de 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 glicolítica, pero se predice que se localiza en la superficie celular a través de un anclaje GPI. Aunque la secuencia genómica de 9,43 Mbp de la cepa GS115 de *P. pastoris* se ha determinado y divulgado en el documento US20110021378A1, las propiedades de las secuencias individuales, tal como las secuencias promotoras, no se han investigado en detalle.

Resultó sorprendente que dicho promotor pudiera utilizarse eficazmente como se describe en el presente documento. Los promotores de *Pichia* de la técnica anterior, tal como los utilizados en la producción de POI a escala industrial, se derivaban principalmente de la vía metabólica del metanol y necesitaban la adición de metanol para inducir la producción de POI, lo que a menudo no se desea. El promotor y el método descritos en el presente documento tienen la ventaja de que pueden proporcionar una mayor producción mediante una expresión mejorada, y tienen el riesgo reducido de contaminación debido a la regulación específica del promotor, en particular cuando se utiliza un medio químicamente definido, libre de metanol.

Resultó que el promotor descrito en el presente documento ejercía su actividad mejorada principalmente con independencia de una cantidad adecuada de sustrato de carbono y de medios de cultivo específicos. Por ejemplo, *P. pastoris* podría cultivarse con éxito en las condiciones de un proceso de producción industrial. Primero se empleó un cultivo discontinuo con una fuente de carbono basal, tal como el glicerol, seguido de un cultivo discontinuo alimentado con una cantidad limitada de una fuente de carbono suplementaria, tal como la glucosa. Las muestras se tomaron cerca del final de la primera fase discontinua y en condiciones de crecimiento limitadas, por ejemplo, utilizando una cantidad limitada de fuente de carbono suplementaria. El análisis del transcriptoma con microarrays de ADN reveló genes específicos que son fuertemente activos en la fuente de carbono suplementario y en presencia de carbono excedente, es decir, la fuente de carbono basal en cantidad excesiva. La secuencia promotora pCS1 se identificó como un promotor sorprendentemente fuerte a tasas de crecimiento elevadas y bajas. El promotor pGAP comparable de la técnica anterior era significativamente más débil.

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

Las secuencias de nucleótidos que podrían utilizarse como secuencias promotoras constitutivas descritas en el presente documento, proporcionarían una producción mejorada de proteínas recombinantes, pueden obtenerse de diversas fuentes. El origen del promotor descrito en el presente documento es preferiblemente una célula de levadura, más preferiblemente una levadura metilotrófica tal como del género *Pichia* o de la especie *P. pastoris*, cuyo promotor puede entonces utilizarse como secuencia madre para producir variantes adecuadas, por ejemplo, mutantes o análogos.

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.

60 Las variantes del promotor de *P. pastoris* identificado, incluidas las variantes funcionalmente activas, tales como los homólogos y los análogos, pueden producirse empleando técnicas estándar. El promotor puede, por ejemplo, modificarse para generar variantes del promotor con niveles de expresión y propiedades reguladoras alterados.

65 Por ejemplo, puede prepararse una biblioteca de promotores mediante mutagénesis de las secuencias promotoras descritas en el presente documento, que pueden utilizarse como moléculas parentales, por ejemplo, para afinar la expresión génica en células eucariotas analizando las variantes por su expresión en diferentes estrategias de

- fermentación y seleccionando las variantes adecuadas. Puede utilizarse una biblioteca sintética de variantes, por ejemplo, para seleccionar un promotor que cumpla los requisitos para producir una POL seleccionada. Dichas variantes pueden tener una mayor eficiencia de expresión en células huésped eucariotas y una alta expresión en condiciones ricas y limitantes de fuentes de carbono.
- 5 Las estrategias de fermentación diferencial distinguirían entre una fase de crecimiento y una fase de producción. El crecimiento y/o la producción pueden realizarse adecuadamente en modo discontinuo, modo discontinuo alimentado o modo continuo. Puede utilizarse cualquier biorreactor adecuado, incluyendo discontinuo, discontinuo alimentado, continuo, reactor de tanque agitado o reactor aerotransportado.
- 10 Es ventajoso proporcionar el proceso de fermentación a escala piloto o industrial. La escala del proceso industrial emplearía preferiblemente un volumen de al menos 10 L, específicamente al menos 50 L, preferiblemente al menos 1 m³, preferiblemente al menos 10 m³, más preferiblemente al menos 100 m³.
- 15 Se prefieren las condiciones de producción a escala industrial, las cuales se refieren, por ejemplo, al cultivo discontinuo alimentado 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 fermentador de aproximadamente 50 - 1000 L o mayores, con tasas de dilución de aproximadamente 0,02 - 0,15 h⁻¹.
- 20 Las técnicas de cultivo adecuadas pueden incluir el cultivo en un biorreactor comenzando con una fase discontinua, seguida de una fase discontinua alimentada exponencial corta a una tasa de crecimiento específica elevada, seguida de una fase discontinua alimentada a una tasa de crecimiento específica baja.
- 25 Otra técnica de cultivo adecuada puede abarcar una fase discontinua seguida de una fase de cultivo continuo a una tasa de dilución baja.
- De acuerdo con un aspecto específico descrito en el presente documento incluye un cultivo discontinuo para proporcionar biomasa seguido de un cultivo discontinuo alimentado para la producción de POI de alto rendimiento.
- 30 Se prefiere cultivar la línea celular huésped descrita en el presente documento en un biorreactor en condiciones de crecimiento para obtener una densidad celular de al menos 1 g/L de peso seco celular, más preferiblemente de al menos 10 g/L de peso seco celular, preferiblemente de al menos 20 g/L de peso seco celular. Es ventajoso proporcionar tales rendimientos de producción de biomasa a escala piloto o industrial.
- 35 Un medio de crecimiento que permite la acumulación de biomasa, específicamente un medio de crecimiento basal, comprende típicamente 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.
- 40 Las fuentes de nitrógeno preferidas incluyen NH₄H₂PO₄, o NH₃ o (NH₄)₂SO₄;
- Las fuentes de azufre preferidas incluyen MgSO₄, o (NH₄)₂SO₄ o K₂SO₄;
- Las fuentes de fosfato preferidas incluyen NH₄H₂PO₄, o H₃PO₄ o NaH₂PO₄, KH₂PO₄, Na₂HPO₄ o K₂HPO₄;
- Otros componentes típicos del medio incluyen KCl, CaCl₂ y oligoelementos tales como: Fe, Co, Cu, Ni, Zn, Mo, Mn, I, B;
- 45 Preferiblemente, el medio se complementa con vitamina B₇;
- Un medio de crecimiento típico para *P. pastoris* comprende glicerol, sorbitol o glucosa, NH₄H₂PO₄, MgSO₄, KCl, CaCl₂, biotina y oligoelementos.
- 50 En la fase de producción se utiliza específicamente un medio de producción con una cantidad limitada de una fuente suplementaria de carbono.
- Preferiblemente, la línea celular huésped se cultiva en un medio mineral con una fuente de carbono adecuada, además simplificando así significativamente el proceso de aislamiento. Un ejemplo de medio mineral preferido es el que contiene una fuente de carbono utilizable (por ejemplo, glucosa, glicerol, sorbitol o metanol), sales que contienen macroelementos (potasio, magnesio, calcio, amonio, cloruro, sulfato, fosfato) y oligoelementos (sales de cobre, yoduro, manganeso, molibdato, cobalto, zinc y hierro, y ácido bórico), y opcionalmente vitaminas o aminoácidos, por ejemplo, para complementar las auxotrofías.
- 55 Las células se cultivan en condiciones adecuadas para llevar a cabo la expresión de la POI deseada, la cual puede purificarse a partir de las células o del medio de cultivo, dependiendo de la naturaleza del sistema de expresión y de la proteína expresada, por ejemplo, si la proteína está fusionada a un péptido señal y si la proteína es soluble o está unida a una membrana. Como comprenderá el experto, las condiciones de cultivo variarán en función de factores que incluyen el tipo de célula huésped y el vector de expresión particular empleado.
- 60 Seleccionando la secuencia promotora adecuada descrita en el presente documento, opcionalmente en combinación con otras secuencias reguladoras preferidas, es posible proporcionar, en condiciones comparables, al menos la misma, o al
- 65

menos aproximadamente 1,1 veces, o al menos aproximadamente 1,2 veces, o al menos aproximadamente 1,5 veces, o 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 aproximadamente 15 veces la actividad, por ejemplo, en condiciones de crecimiento limitado en un proceso discontinuo alimentado, representada por la actividad del promotor o la fuerza de transcripción, o regulada 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*.

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

Por ejemplo, la alimentación de la fuente de carbono suplementaria añadida a la fermentación puede comprender una fuente de carbono con hasta un 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 de producto basado en la provisión de sustrato.

La fermentación se realiza preferiblemente a un pH en el intervalo de 3 y 7,5.

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

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 un experto en la materia. De acuerdo con la presente divulgación, pueden emplearse técnicas convencionales de biología molecular, microbiología y ADN recombinante dentro de los conocimientos de la técnica. Dichas técnicas se explican detalladamente en la bibliografía. Véase, por ejemplo, Maniatis, Fritsch & Sambrook, "Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, (1982).

De acuerdo con un aspecto específico de la presente divulgación, se obtiene un constructo recombinante al ligar el promotor y los genes relevantes en un vector o constructo de expresión. Estos genes pueden integrarse de forma estable en el genoma de la célula huésped al transformar la célula huésped utilizando dichos vectores o constructos de expresión.

Los vectores de expresión pueden incluir, pero no limitarse a, vectores de clonación, vectores de clonación modificados y plásmidos diseñados específicamente. El vector de expresión preferido como se utiliza en la presente divulgación puede ser cualquier vector de expresión adecuado para la expresión de un gen recombinante en una célula huésped y se selecciona en función del organismo huésped. El vector de expresión recombinante puede ser cualquier vector capaz de replicarse en o integrarse en el genoma de los organismos huéspedes, también denominado vector huésped.

Los vectores de expresión apropiados comprenden típicamente secuencias reguladoras adicionales adecuadas para expresar ADN que codifica una POI en una célula huésped eucariota. Ejemplos de secuencias reguladoras incluyen operadores, potenciadores, sitios de unión ribosómica y secuencias que controlan la iniciación y terminación de la transcripción y la traducción. Las secuencias reguladoras pueden estar vinculadas de forma operativa a la secuencia de ADN que se va a expresar.

Para permitir la expresión de una secuencia de nucleótidos recombinante en una célula huésped, el vector de expresión puede proporcionar el promotor descrito en el presente documento adyacente al extremo 5' de la secuencia codificante, por ejemplo, corriente arriba del GOI o de un gen péptido señal que permita la secreción del POI. De este modo, la transcripción es regulada e iniciada por esta secuencia promotora.

Un péptido señal puede ser un péptido señal heterólogo o un híbrido de un péptido señal nativo y uno heterólogo, y puede ser específicamente heterólogo u homólogo al organismo huésped que produce la proteína. La función del péptido señal es permitir la secreción de la POI para que entre en el retículo endoplasmático. Suele ser una cadena peptídica corta (de 3 a 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 son escindidos de la proteína por la peptidasa señal después de que las proteínas sean transportadas.

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

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 está controlada por el promotor en células nativas (de tipo salvaje) o las secuencias se recombinan con secuencias contiguas diferentes.

Para probar la función de las secuencias relevantes, se pueden construir vectores de expresión que comprenden uno o más de los elementos reguladores para impulsar la expresión de una POI, y el rendimiento expresado se compara con los constructos con elementos reguladores convencionales. En los ejemplos siguientes se puede encontrar una descripción detallada del procedimiento experimental. Los genes identificados pueden amplificarse mediante PCR a partir

- de *P. pastoris* utilizando cebadores nucleotídicos específicos, clonarse en un vector de expresión y transformarse en una línea celular eucariota, por ejemplo, utilizando un vector de levadura y una cepa de *P. pastoris*, para la producción a alto nivel de varias POI diferentes. Para estimar el efecto del promotor descrito en el presente documento sobre la cantidad de POI recombinante así producida, la línea celular eucariota puede cultivarse en experimentos de matraz agitado y fermentaciones discontinuas alimentadas o quimiostato en comparación con cepas que comprenden un constitutivo convencional, por ejemplo, un promotor dependiente del crecimiento, tal como, por ejemplo, el promotor estándar pGAP en la célula respectiva. En particular, la elección del promotor tiene un gran impacto en la producción de proteínas recombinantes.
- La POI puede producirse utilizando la línea celular huésped recombinante cultivando un transformante así obtenido en un medio apropiado, aislando el producto o metabolito expresado del cultivo y, opcionalmente, purificándolo mediante un método adecuado.
- Los transformantes descritos en el presente documento pueden obtenerse introduciendo dicho ADN vectorial, por ejemplo, ADN plasmídico, en un huésped y seleccionando transformantes que expresen la 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, tal como el método del pulso eléctrico, el método del protoplasto, el método del 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 la transformación química, electroporación o transformación por protoplastación. Los transformantes descritos en el presente documento pueden obtenerse introduciendo dicho ADN vectorial, por ejemplo, ADN plasmídico, en un huésped y seleccionando transformantes que expresen la proteína relevante o el metabolito de la célula huésped con altos rendimientos.
- Se prefieren varios enfoques diferentes para la producción de la POI de acuerdo con el método descrito en el presente documento. Las sustancias pueden expresarse, procesarse y, opcionalmente, secretarse transformando una célula huésped eucariota con un vector de expresión que contenga ADN recombinante que codifique una proteína relevante y al menos uno de los elementos reguladores descritos 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.
- Es preferible comprobar la capacidad de expresión o el rendimiento de la célula huésped descrita en el presente documento mediante la siguiente prueba: ELISA, ensayo de actividad, HPLC u otras pruebas adecuadas.
- La 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, más preferiblemente al menos 1 g/L.
- Se entiende que los métodos divulgados en el presente documento pueden incluir además el cultivo de dichas células huésped recombinantes en condiciones que permitan la expresión de la POI, preferiblemente en forma secretada o bien como producto intracelular. Una POI producida recombinantemente o un metabolito de la célula huésped puede aislarse entonces del medio de cultivo celular y purificarse aún más mediante técnicas bien conocidas por un experto en la materia.
- La POI producida como se describe en el presente documento puede aislarse y purificarse utilizando las técnicas del arte, incluido el aumento de la concentración de la POI deseada y/o la disminución de la concentración de al menos una impureza.
- Si la POI es secretada por las células, puede aislarse y purificarse a partir del medio de cultivo utilizando las técnicas del arte. La secreción de los productos de expresión recombinantes a partir de las células huésped es generalmente ventajosa por razones que incluyen la facilitación del proceso de purificación, ya que los productos se recuperan del sobrenadante del cultivo en lugar de la compleja mezcla de proteínas que resulta cuando las células de levadura se desorganizan para liberar proteínas intracelulares.
- Las células transformantes cultivadas también pueden romperse sónica o mecánicamente, enzimática o químicamente para obtener un extracto celular que contenga la POI deseada, a partir del cual se aísla y purifica la POI.
- Como métodos de aislamiento y purificación para obtener un polipéptido recombinante o un producto proteico, métodos, tales como métodos que utilizan la diferencia de solubilidad, tales como la salazón y la precipitación con disolventes, métodos que utilizan la diferencia de peso molecular, tales como la ultrafiltración y la electroforesis en gel, métodos que utilizan la diferencia de carga eléctrica, tales como la cromatografía de intercambio iónico, métodos que utilizan la afinidad específica, tales como la cromatografía de afinidad, métodos que utilizan la diferencia de hidrofobicidad, tales como la cromatografía líquida de alto rendimiento en fase inversa, y métodos que utilizan la diferencia de punto isoeléctrico, tales como el enfoque isoeléctrico.
- El producto altamente purificado está esencialmente libre de proteínas contaminantes, y preferiblemente tiene una pureza de al menos 90 %, más preferido al menos 95 %, o incluso al menos 98 %, hasta el 100 %. Los productos purificados pueden obtenerse por purificación del sobrenadante del cultivo celular o bien a partir de restos celulares.

- Como métodos de aislamiento y purificación se prefieren los siguientes métodos estándar: Disrupción celular (si la POI se obtiene intracelularmente), separación de células (restos) y lavado por microfiltración o filtro de flujo tangencial (TFF por sus siglas en inglés) o centrifugación, purificación de la POI por precipitación o tratamiento térmico, activación de la POI por digestión enzimática, purificación de la POI mediante cromatografía, tal como intercambio iónico (IEX por sus siglas en inglés), cromatografía de interacción hidrófoba (HIC por sus siglas en inglés), cromatografía de afinidad, exclusión por tamaño (SEC por sus siglas en inglés) o cromatografía HPLC, precipitación de la POI de concentración y lavado mediante pasos de ultrafiltración.
- La POI aislada y purificada puede identificarse mediante métodos convencionales tales como Western blot, HPLC, ensayo de actividad o ELISA.
- La POI puede ser cualquier polipéptido eucariota, procariota o sintético. Puede ser una proteína secretada o una proteína intracelular. La presente divulgación también proporciona la producción recombinante de homólogos funcionales, variantes funcionales equivalentes, derivados y fragmentos biológicamente activos de proteínas naturales. Los homólogos funcionales son preferiblemente idénticos con o corresponden a y tienen la característica funcional de una secuencia.
- Una POI mencionada en el presente documento puede ser un producto homólogo a la célula huésped eucariota o heterólogo, preferiblemente para uso terapéutico, profiláctico, diagnóstico, analítico o industrial.
- La POI es preferiblemente un polipéptido o proteína recombinante heteróloga, producida en una célula eucariota, preferiblemente una célula de levadura, preferiblemente como proteínas secretadas. Ejemplos de proteínas producidas preferiblemente son inmunoglobulinas, fragmentos de inmunoglobulina, aprotinina, inhibidor de la vía del factor tisular u otros inhibidores de la proteasa, e insulina o precursores de insulina, análogos de insulina, hormonas de crecimiento, interleucinas, activador tisular del plasminógeno, factor de crecimiento transformante α o β , glucagón, péptido similar al glucagón 1 (GLP-1), péptido similar al glucagón 2 (GLP-2), GRPP, Factor VII, Factor VIII, Factor XIII, factor de crecimiento derivado de plaquetas, albúmina sérica, enzimas, tales como lipasas o proteasas, o un homólogo funcional, variante funcional equivalente, derivado y fragmento biológicamente activo con una función similar a la proteína nativa. La 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 ya sea uno o ambos extremos C- y N-terminal o a la cadena lateral de la proteína nativa, la sustitución de uno o más aminoácidos en uno o varios sitios diferentes de la secuencia de aminoácidos nativa, la supresión de uno o más aminoácidos ya sea en uno o ambos extremos de la proteína nativa o en uno o varios sitios de la secuencia de aminoácidos, o la inserción de uno o más aminoácidos en uno o más sitios de la secuencia de aminoácidos nativa. Tales modificaciones son bien conocidas para varias de las proteínas mencionadas anteriormente.
- Una POI también puede seleccionarse entre sustratos, enzimas, inhibidores o cofactores que proporcionan reacciones bioquímicas en la célula huésped, con el fin de obtener el producto de dicha reacción bioquímica o una cascada de varias reacciones, por ejemplo, para obtener un metabolito de la célula huésped. Productos ejemplares pueden ser vitaminas, tales como la riboflavina, ácidos orgánicos y alcoholes, los cuales pueden obtenerse con mayores rendimientos tras la expresión de una proteína recombinante o una POI como se describe en el presente documento.
- En general, la célula huésped, que expresa un producto recombinante, puede ser cualquier célula eucariota adecuada para la expresión recombinante de una POI.
- Ejemplos de células de mamífero preferidas son las células BHK, CHO (CHO-DG44, CHO-DUXB11, CHO-DUKX, CHO-K1, CHOK1SV, CHO-S), HeLa, HEK293, MDCK, NIH3T3, NS0, PER.C6, SP2/0 y VERO.
- Los ejemplos de células de levadura preferidas utilizadas como células huésped descritas en el presente documento incluyen, pero no se limitan a, el género *Saccharomyces* (por ejemplo, *Saccharomyces cerevisiae*), el género *Pichia* (por ejemplo, *P. pastoris* o *P. methanolica*), el género *Komagataella* (*K. pastoris*, *K. pseudopastoris* o *K. phaffii*), *Hansenula polymorpha* o *Kluyveromyces lactis*.
- La literatura más reciente divide y renombra *Pichia pastoris* en *Komagataella pastoris*, *Komagataella phaffii* y *Komagataella pseudopastoris*. En el presente documento *Pichia pastoris* se utiliza como sinónimo de todos, *Komagataella pastoris*, *Komagataella phaffii* y *Komagataella pseudopastoris*.
- Las células huésped de levadura preferidas se derivan de levaduras metilotróficas, tales como *Pichia* o *Komagataella*, por ejemplo, *Pichia pastoris*, o *Komagataella pastoris*, o *K. phaffii*, o *K. pseudopastoris*. Ejemplos de huéspedes incluyen levaduras tales como *P. pastoris*. Ejemplos de cepas de *P. pastoris* incluyen CBS 704 (=NRRL Y-1603 = DSMZ 70382), CBS 2612 (=NRRL Y-7556), CBS 7435 (=NRRL Y-11430), CBS 9173-9189 (cepas CBS: CBS- KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos), y DSMZ 70877 (Colección Alemana de Microorganismos y Cultivos Celulares), pero también cepas de Invitrogen, tales como X-33, GS115, KM71 y SMD1168. Ejemplos de cepas de *S. cerevisiae* son 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.
- Una célula hospedadora de levadura preferida descrita en el presente documento, tal como una célula huésped de *P.*

pastoris o *S. cerevisiae*, contiene secuencias promotoras heterólogas o recombinantes, que pueden derivarse de una cepa de *P. pastoris* o *S. cerevisiae*, diferente de la huésped de producción. En otra realización específica, la célula huésped descrita en el presente documento contiene un constructo de expresión recombinante descrito en el presente documento que comprende el promotor procedente del mismo género, especie o cepa que la célula huésped.

El promotor descrito en el presente documento se deriva preferiblemente de un gen que codifica una proteína homóloga a la célula huésped.

Por ejemplo, un promotor descrito en el presente documento puede derivarse de una levadura, tal como una cepa de *S. cerevisiae*, y utilizarse para expresar una POI en una levadura. Una realización específicamente preferida se refiere a un promotor descrito en el presente documento procedente de *P. pastoris* para usarse en un método para producir una POI recombinante en una línea celular huésped productora 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 una POI, posiblemente con mayores rendimientos en los procesos de fabricación industrial. También, pueden utilizarse variantes funcionalmente activas del promotor de otras levaduras u otros hongos adecuados o de otros organismos tales como vertebrados o plantas.

Si la POI es una proteína homóloga a la célula huésped, es decir, una proteína que se encuentra de forma natural en la célula huésped, la expresión de la POI en la célula huésped puede modularse mediante el intercambio de su secuencia promotora nativa con una secuencia promotora descrita en el presente documento.

Este propósito puede lograrse, por ejemplo, mediante la transformación de una célula huésped con una molécula de ADN recombinante que comprenda 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 con el fin de unir de forma operable la secuencia promotora con la secuencia de nucleótidos que codifica la POI. Esto da lugar a la expresión de la POI a partir de la secuencia promotora de acuerdo con la invención en lugar de a partir de la secuencia promotora nativa.

De acuerdo con un aspecto específico descrito en el presente documento, la secuencia promotora tiene una actividad promotora aumentada en relación con la secuencia promotora nativa de la POI.

Como se describe en el presente documento, se prefiere proporcionar una línea celular huésped de *P. pastoris* que comprenda una secuencia promotora descrita en el presente documento enlazada de forma operable a la secuencia de nucleótidos que codifica la POI.

Como se describe en el presente documento, también es posible proporcionar un vector comodín o una célula huésped descrita en el presente documento, que comprenda un promotor descrito en el presente documento, y que esté preparada para incorporar un gen de interés que codifique una POI. La línea celular comodín es, por lo tanto, una línea celular huésped preformada, que se caracteriza por su capacidad de expresión. Esto sigue una estrategia innovadora de plataforma "comodín" para la generación de líneas celulares productoras, para la producción de POI, por ejemplo, utilizando el intercambio de casetes mediado por recombinasas sitio-específicas. Dicha nueva célula huésped facilita la clonación de un gen de interés (GOI), por ejemplo, en puntos calientes de expresión genómica predeterminados en cuestión de días, con el fin de obtener líneas celulares de producción reproducibles y altamente eficientes.

De acuerdo con una realización preferida, el método descrito en el presente documento utiliza una secuencia de nucleótidos recombinante que codifica la POI, la cual se proporciona en un plásmido adecuado para su integración en el genoma de la célula huésped, en una sola copia o en múltiples copias por célula. La secuencia de nucleótidos recombinante que codifica la POI también puede proporcionarse en un plásmido de replicación autónoma en una sola copia o en múltiples copias por célula.

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

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

De acuerdo con un aspecto específico descrito en el presente documento, se obtiene un constructo recombinante al ligar los genes relevantes en un vector. Estos genes pueden integrarse de forma estable en el genoma de la célula huésped

transformándola con dichos vectores. Los polipéptidos codificados por los genes pueden producirse utilizando la línea celular huésped recombinante al cultivar un transformante así obtenido en un medio apropiado, aislando la POI expresada del cultivo y purificándola mediante un método apropiado para el producto expresado, en particular para separar la POI de las proteínas contaminantes.

Los vectores de expresión pueden comprender uno o más marcadores seleccionables fenotípicamente, por ejemplo, un gen que codifica una proteína que confiere resistencia a los antibióticos o que suple un requerimiento autótrofo. Los vectores de levadura suelen contener un origen de replicación a partir de un plásmido de levadura, una secuencia de replicación autónoma (ARS por sus siglas en inglés) o, alternativamente, una secuencia utilizada para la integración en el genoma huésped, una región promotora, secuencias para la poliadenilación, secuencias para la terminación de la transcripción y un marcador seleccionable.

Los procedimientos utilizados para ligar las secuencias de ADN, por ejemplo, las que codifican la secuencia precursora y/o la POI, el promotor y el terminador, respectivamente, y para insertarlas en vectores adecuados que contengan la información necesaria para la integración o la replicación en el huésped, son bien conocidos por los expertos en la materia, por ejemplo, los descritos por J. Sambrook et al., (A Laboratory Manual, Cold Spring Harbor, 1989).

Se entenderá que el vector, el cual utiliza los elementos reguladores descritos en el presente documento y/o la POI como diana de integración, puede construirse ya sea preparando primero un constructo de ADN que contenga la secuencia completa de ADN que codifica los elementos reguladores y/o la POI e insertando posteriormente este fragmento en un vector de expresión adecuado, o al insertar secuencialmente fragmentos de ADN que contengan información genética para los elementos individuales, seguido de ligación.

También, se pueden utilizar vectores de multiclonación, los cuales son vectores que tienen un sitio de multiclonación, como se describe en el presente documento, donde se puede incorporar un gen heterólogo deseado en un sitio de multiclonación para proporcionar un vector de expresión. En los vectores de expresión, el promotor se sitúa corriente arriba del gen de la POI y regula la expresión del gen. En el caso de los vectores de multiclonación, dado que el gen de la POI se introduce en el sitio de multiclonación, el promotor se coloca corriente arriba del sitio de multiclonación.

El constructo de ADN proporcionado para obtener una célula huésped recombinante descrita en el presente documento puede prepararse sintéticamente mediante métodos estándar establecidos, por ejemplo, el método de la fosforamida. El constructo de ADN también puede ser de origen genómico o de ADNc, por ejemplo, obtenido preparando una biblioteca genómica o de ADNc y buscando secuencias de ADN que codifiquen la totalidad o parte del polipéptido descrito en el presente documento mediante hibridación utilizando sondas de oligonucleótidos sintéticos de acuerdo con las técnicas estándar (Sambrook et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, 1989). Por último, el constructo de ADN puede ser de origen mixto sintético y genómico, mixto sintético y ADNc o mixto genómico y ADNc, preparado por recocido de fragmentos de origen sintético, genómico o ADNc, como sea apropiado, los fragmentos correspondientes a diversas partes del constructo de ADN completa, de acuerdo con técnicas estándar.

En otra realización preferida, el vector de expresión de levadura es capaz de integrarse de forma estable en el genoma de la levadura, por ejemplo, por recombinación homóloga.

Una célula huésped transformante descrita en el presente documento obtenida mediante la transformación de la célula con los elementos reguladores descritos en el presente documento y/o los genes POI puede, preferiblemente, cultivarse primero en condiciones para crecer de manera eficiente hasta un gran número de células. Cuando se prepara la línea celular para la expresión de la POI, se eligen las técnicas de cultivo para obtener el producto de expresión.

Los ejemplos específicos descritos en el presente documento se refieren a la fermentación discontinua alimentada de una línea celular *P. pastoris* de producción recombinante que produce proteínas informadoras, empleando un medio discontinuo de glicerol y un medio discontinuo alimentado de glucosa. Los estudios comparativos de actividad promotora han demostrado que el promotor descrito en el presente documento puede utilizarse con éxito para la producción de proteínas recombinantes.

De acuerdo con otro ejemplo descrito en el presente documento, se produjo albúmina sérica humana (HSA por sus siglas en inglés) como una POI en las condiciones de glucosa-límite, y se determinó el rendimiento de thasHSA y el número de copias del gen.

De acuerdo con otro ejemplo descrito en el presente documento, se realizó el cultivo discontinuo alimentado de cepas de *P. pastoris* que expresan HSA bajo el control de un promotor descrito en el presente documento.

Otros ejemplos descritos en el presente documento se refieren a la expresión de una carboxipeptidasa B porcina como proteína modelo bajo el control transcripcional del promotor pCS1.

Sin embargo, otro ejemplo descrito en el presente documento se refiere a la expresión de un fragmento de anticuerpo bajo el control transcripcional de pCS1.

Otro ejemplo descrito en el presente documento se refiere a variantes de tamaño o longitud de un promotor descrito en el presente documento, tal como la secuencia pCS1 alargada pCS1a (SEQ ID 2) la cual comprende la secuencia pCS1 y un alargamiento en el extremo 5', o fragmentos de pCS1 con una longitud en el intervalo de aproximadamente 80 bp a 800 bp.

La descripción anterior se entenderá completamente con referencia a los siguientes ejemplos. Tales ejemplos son, sin embargo, meramente representativos de métodos de puesta en práctica de una o más realizaciones descritas en el presente documento y no deben interpretarse como limitativos del alcance de la invención.

EJEMPLOS

Los siguientes ejemplos 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*

Con el fin de identificar un gen fuerte y su respectivo promotor de *P. pastoris*, se realizó un análisis de los patrones de expresión génica utilizando microarrays de ADN. Se analizaron células de *P. pastoris* cultivadas en un lote de glicerol y en un límite de glucosa (quimioestado).

a) Cepa

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

b) Cultivo de *P. pastoris*

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

Se utilizaron los siguientes medios:

Solución madre de sales traza PTM₁ 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 biotina y 5,0 mL H₂SO₄ (95 %-98 %).

Medio Discontinuo de Glicerol contenido por litro

2 g de ácido cítrico monohidratado (C₆H₈O₇·H₂O), 39,2 g de glicerol, 20,8 g de NH₄H₂PO₄, 0,5 g de MgSO₄·7H₂O, 1,6 g de KCl, 0,022 g de CaCl₂·2H₂O, 0,8 mg de biotina y 4,6 mL de solución madre de sales traza PTM₁. Se añadió HCl para ajustar el pH a 5.

Medio discontinuo alimentado de glicerol contenido por litro

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

Medio quimioestático contenido por litro

2 g de ácido cítrico monohidratado (C₆H₈O₇·H₂O), 99,42 g de glucosa monohidratada, 22 g de NH₄H₂PO₄, 1,3 g de MgSO₄·7H₂O, 3,4 g de KCl, 0,02 g de CaCl₂·2H₂O, 0,4 mg de biotina y 3,2 mL de solución madre de sales traza PTM₁. Se añadió HCl para ajustar el pH a 5.

El oxígeno disuelto se controló a DO = 20 % con la velocidad del agitador (500 - 1250 rpm). La tasa de aireación fue de 60 L h⁻¹ de aire, la temperatura se controló a 25 °C y el valor nominal de pH de 5 se controló con adición de NH₄OH (25 %).

Para iniciar la fermentación, se filtraron estérilmente 1,5 L de medio discontinuo en el fermentador y se inoculó *P. pastoris* (de un cultivo previo de una noche en YPG, 180 rpm, 28 °C) con una densidad óptica inicial (OD₆₀₀) de 1. La fase discontinua de aproximadamente 25 h alcanzó una concentración de biomasa seca de aproximadamente 20 g/L, fue seguida por una fase discontinua alimentada exponencial de 10 h con medio de glucosa, que condujo a una concentración de biomasa seca de aproximadamente 50 g/L. A continuación, se redujo el volumen a 1,5 L y se inició el cultivo en quimioestado con una tasa de alimentación/recolección de 0,15 L h⁻¹, lo que resultó en una tasa de crecimiento constante de $\mu = 0,1$. La fermentación se interrumpió 50 h después del inicio en quimioestado.

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

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

5 c) Muestreo

Las muestras se tomaron al final de la fase discontinua de glicerol y en condiciones estables del quimiostato de glucosa. Durante cada fermentación se tomaron muestras de rutina, como la determinación de la densidad óptica o la masa seca de la levadura, la inspección microscópica cualitativa y el análisis de la viabilidad celular. Para el análisis de microarrays, se tomaron muestras y se trataron del siguiente modo: Para un enfriamiento óptimo, se mezclaron inmediatamente 9 mL de caldo de cultivo celular con 4,5 mL de solución helada de fenol al 5 % (Sigma) (en etanol abs.), y se tomaron alícuotas. Cada 2 mL se centrifugaron (13200 rpm durante 1 minuto) en tubos de recolección preenfriados (GE healthcare, NJ), se eliminó completamente el sobrenadante y los tubos se almacenaron a -80 °C hasta la purificación del ARN.

15 d) Purificación del ARN y preparación de muestras para la hibridación de microarrays

El ARN se aisló utilizando el reactivo TRI de acuerdo con las instrucciones del proveedor (Ambion, EE.UU.). Los gránulos celulares se resuspendieron en reactivo TRI y se homogeneizaron con perlas de vidrio utilizando un FastPrep 24 (M.P. Biomedicals, CA) a 5 m s^{-1} durante 40 segundos. Después de añadir cloroformo, las muestras se centrifugaron y el ARN total se precipitó de la fase acuosa añadiendo isopropanol. El gránulo se lavó con etanol al 70 %, se secó y se volvió a suspender en agua libre de ARNasa. Las concentraciones de ARN se determinaron midiendo la OD₂₆₀ utilizando un espectrofotómetro Nanodrop 1000 (NanoDrop products, DE). El ADN restante de las muestras se eliminó utilizando el kit DNA free (Ambion, CA). Se diluyó un volumen de muestra igual a 10 µg de ARN hasta 50 µL en agua libre de ARNasa, después se añadieron el amortiguador de ADNasa I y la rADNasa I y se incubaron a 37 °C durante 30 minutos. Después de añadir el reactivo de inactivación de ADNase, la muestra se centrifugó y el sobrenadante se transfirió a un tubo nuevo. Las concentraciones de ARN se determinaron de nuevo como se ha descrito anteriormente. Además, se analizó la integridad del ARN utilizando nanochips de ARN (Agilent). Para monitorizar el flujo de trabajo del microarray desde la amplificación y el etiquetado hasta la hibridación de las muestras, se utiliza el Spike In Kit (Agilent, Product Nr: 5188-5279) como control positivo. Contiene 10 transcritos poliadenilados diferentes de un adenovirus, los cuales se amplifican, etiquetan y cohibridan junto con las propias muestras de ARN. Las muestras se marcaron con Cy 3 y Cy 5 utilizando el Quick Amp Labelling Kit (Agilent, Prod. Nr.: 5190-0444). Por lo tanto, se diluyeron 500 ng de ARN de muestra purificado en 8,3 µL de agua libre de ARNasa, se añadieron 2 µL de Spike A o B y 1,2 µL de cebador del promotor T7. La mezcla se desnaturalizó durante 10 minutos a 65 °C y se mantuvo en hielo durante 5 minutos. A continuación, 8,5 µL de mastermix de ADNc (por muestra: 4 µL de amortiguador de primera cadena 5x, 2 µL de DTT 0,1 M, 1 µL de mezcla de dNTP 10 mM, 1 µL de MMLV-RT, 0,5 µL de RNase out), se incubaron a 40 °C durante 2 horas, luego se transfirieron a 65 °C durante 15 minutos y se pusieron en hielo durante 5 minutos. La mezcla maestra de transcripción (por muestra: 15,3 µL de agua libre de nucleasas, 20 µL de amortiguador de transcripción, 6 µL de DTT 0,1 M, 6,4 µL de PEG al 50 %, 0,5 µL de inhibidor de ARNasa, 0,6 µL de fosfatasa inorg., 0,8 µL de ARN polimerasa T7, 2,4 µL de cianina 3 o cianina 5), se añadió a cada tubo y se incubaron a 40 °C durante 2 horas. Con el fin de purificar el ARNc marcado obtenido, se utilizó el RNeasy Mini Kit (Qiagen, Cat.No. 74104). Las muestras se almacenaron a -80 °C. La cuantificación de la concentración de ARNc y de la eficacia del etiquetado se realizó en el espectrofotómetro Nanodrop.

e) Análisis de microarrays

Se utilizó el kit de hibridación de expresión génica (Agilent, Cat. Núm. 5188-5242) para la hibridación de los ARNc de muestra marcados. Para la preparación de las muestras de hibridación, cada 300 ng de ARNc (Cy3 y Cy 5) y 6 µL de agente de bloqueo 10 veces se diluyeron con agua libre de nucleasas hasta un volumen final de 24 µL. Tras añadir 1 µL de amortiguador de fragmentación 25 veces, la mezcla se incubó a 60 °C durante 30 minutos. A continuación, se añadieron 25 µL de amortiguador de hibridación GEx HI-RPM para detener la reacción. Después de centrifugación durante un minuto a 13.200 rpm, la muestra se enfrió en hielo y se utilizó inmediatamente para la hibridación. Se utilizaron ensayos de oligonucleótidos específicos de *P. pastoris* diseñados internamente (AMAD-ID: 026594, 8x15K custom arrays, Agilent). La hibridación de microarrays se realizó de acuerdo con Microarray Hybridisation Chamber User Guide (Agilent G2534A). En primer lugar, se descubrió el portajuntas 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 centro de cada uno de los ocho cuadrados. A continuación, se colocó con cuidado el portaobjetos de microarrays sobre el portajuntas (con la etiqueta de Agilent hacia abajo) y se colocó la tapa de la cámara 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 microarrays. Por lo tanto, se desmontó la cámara y se separaron los portaobjetos sándwich entre sí mientras se sumergían en amortiguador de lavado 1. El microarray se transfirió directamente a otro plato con amortiguador de lavado 1, se lavó durante 1 minuto, se transfirió a amortiguador de lavado 2 (temperatura de al menos 30 °C) y se lavó durante otro minuto. Después de secar el portaobjetos del microarray tocando el borde del portaobjetos con un tejido, se colocó en el soporte del portaobjetos (con la etiqueta de Agilent hacia arriba). El soporte del portaobjetos se colocó en el carrusel y se inició el escaneado.

f) Adquisición de datos y evaluación estadística de datos de microarray

Las imágenes se escanearon a una resolución de 50 nm con un escáner G2565AA Microarray (Agilent) y se importaron

en el software Agilent Feature Extraction 9.5. Para la cuantificación de las intensidades de los puntos se utilizó Agilent Feature Extraction 9.5. A continuación, los datos brutos de intensidad media de los puntos se importaron al software de código abierto R para su posterior normalización y análisis de datos.

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

10 Se examinaron los datos de microarrays en busca de entradas con alta intensidad de señal en ambos estados con el fin de 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 añaden como referencias. En promedio, pCS1 superó a pGAP en ambas condiciones (cultivo discontinuo de glicerol y en quimiostato con glucosa limitada) en a lo mucho 30 %, mientras que pTEF es aproximadamente 12 % más débil que pGAP.

15 Tabla 1: Datos de microarrays de los genes cuyos promotores se seleccionaron para su caracterización posterior y de pGAP y pTEF como controles

Promotor	identificador genético	Intensidad ¹	% de intensidad pGAP/ fuerza de transcripción	Intensidad ²	% de intensidad 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 recientemente identificado pCS1 en *P. pastoris* utilizando eGFP como gen informador de expresión intracelular

20 Con el fin de analizar las propiedades del promotor recién identificado, se realizaron cribados en matraz agitado como se indica a continuación: El precultivo durante 24 horas se realizó con medio rico que contenía glicerol como fuente de carbono -simulando la fase discontinua del proceso, al cual siguió el cultivo principal con medio mínimo y perlas de alimentación de glucosa- simulando la fase discontinua alimentada con glucosa limitada del proceso. La proteína verde fluorescente (eGFP por sus siglas en inglés) se utilizó como proteína informadora intracelular para la actividad promotora.

25 a) Cepa & vector de expresión

30 La cepa *P. pastoris* de tipo salvaje (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) se utilizó como cepa huésped. La transformación de la cepa se llevó a cabo con un vector propio denominado pPUZZLE (Stadlmayr et al. J. Biotechnol 2010;150(4):519-29), que comprende 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 consiste en un sitio de clonación múltiple y el terminador de transcripción CYC1 de *S. cerevisiae*, y un locus para la integración en el genoma de *P. pastoris* (región 3'AOX1).

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

40 El promotor pCS1 (SEQ ID) comprende 985 bp de la región 5'-no codificante del gen CS1 (véase el Ejemplo 1) hasta el codón de inicio ATG y se amplificó mediante PCR (Phusion Polymerase, New England Biolabs) a partir de ADN genómico de *P. pastoris* utilizando 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 SbfI, resultando en pPM1aZ10_pCS1_eGFP. Además, se utilizó como referencia el vector pPM1aZ10_pGAP_eGFP, que contiene el promotor comúnmente utilizado del promotor de la gliceraldehído 3-fosfato deshidrogenasa (pGAP de *P. pastoris*, aquí SEQ ID 13). Los promotores se insertaron corriente arriba del codón de inicio del gen eGFP utilizando los sitios de restricción Apal y SbfI (véanse las Tablas 2 y 3). La exactitud de las secuencias promotoras se verificó mediante secuenciación 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 GATAGGGCCCCAGGGCATCATTGAGGTTTCCAC	69,6	Apal

Nombre	Diana	Secuencia	T _M	Sitio restricción de
pCS1_back	pCS1	SEQ ID 15 GATACCTGCAGGTTTTGTTGTTGAGTGAAGCGAGTG	69,3	Sbfl

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

			Enzima de clonación		
promotor	cebador 5'	cebador 3'	5'	3'	longitud
pCS1	pCS1_fw	pCS1_back	Apal	Sbfl	985

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

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

Las transformantes positivas se seleccionaron en placas YPD (por litro: 10 g de extracto de levadura, 20 g de peptona, 20 g de glucosa, 20 g de agar-agar) que contenían 25 µg/mL de Zeocina (Invivogen, CA). Se utilizó la PCR en colonia para garantizar la presencia del plásmido transformado. Por lo tanto, se ganó ADN genómico al cocinar y congelar colonias de *P. pastoris* durante 5 minutos cada una y se aplicó directamente para PCR con los cebadores apropiados. Para el cribado de expresión, 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 Zeocina) como precultivo. Después de aproximadamente 24h, el precultivo se utilizó para inocular el cultivo principal con una OD600 de 0,1 en 2 mL de medio de cribado sintético (por litro: 22 g de glucosa monohidrato, 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 5 M; esterilizado por filtración) y 2 cuartos de perlas de alimentación de glucosa (segunda perla de alimentación añadida después de 24 horas; Kuhner, CH). Se alcanzaron condiciones de crecimiento limitantes de glucosa debido a la lenta cinética de liberación de glucosa de estas perlas de alimentación, que se describe mediante la siguiente ecuación: (Glucose)=1,63*t^{0,74} [mg/Disc]. Se tomaron muestras al final del precultivo y 24 y 48 horas después de la inoculación del cultivo principal. La densidad celular se determinó midiendo la 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 autofluorescencia de *P. pastoris* se midió utilizando células de tipo salvaje de *P. pastoris* no transformadas y se restó 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 constitutivo pGAP.

Los resultados se muestran en la Tabla 4. El clon que se expresaba bajo el control del promotor pCS1 excedía al pGAP en un 38 % al final del precultivo (discontinuo), y tenía niveles de expresión de GFP 4 veces superiores en el extremo de cultivo principal (discontinuo alimentado).

Tabla 4: Fluorescencia GFP promedio por tamaño celular 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 temporal.

	precultivo		cultivo principal	
	final discontinuo	stdev	48h	stdev
pCS1	138,2	26,6	400,7	104,87
pGAP	100,0		100,0	

d) Determinación del número de copias del gen (GCN por sus siglas en inglés) eGFP de los clones de *P. pastoris* del Ejemplo 2c

GCN representa el número de casetes de expresión de proteínas reporteras integrados en el genoma de *P. pastoris*. La determinación de GCN de clones que expresan 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 expresan eGFP bajo el control del novedoso promotor pCS1 produjeron cantidades dos veces superiores de eGFP en comparación con los clones que expresan 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 correlacionados con GCN. Por GCN, los clones pCS1 expresan el doble de la cantidad de eGFP que los clones pGAP.

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

e) Análisis de la fuerza promotora de pCS1 en la fermentación discontinua alimentada

Para evaluar la actividad promotora de pCS1 en condiciones similares a las del proceso de producción, se realizaron cultivos discontinuos alimentados de un clon de pCS1 (pCS1_eGFP#4) y un clon de pGAP (pGAP_eGFP#2), cada uno de los cuales albergaba una copia del casete de expresión de eGFP (véase el Ejemplo 2d).

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

Se utilizaron los siguientes medios:

Solución madre de sales traza PTM₁ 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 biotina y 5,0 mL H₂SO₄ (95 %-98 %).

Medio Discontinuo de Glicerol contenido por litro

2 g de ácido cítrico monohidratado (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 solución madre de sales traza PTM₁. Se añadió HCl para ajustar el pH a 5.

Medio discontinuo alimentado de glucosa contenido por litro

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

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

Para iniciar la fermentación, se filtraron estérilmente 400 mL de medio discontinuo en el fermentador y se inoculó el clon de *P. pastoris* pCS1_eGFP#1 (del precultivo) con una densidad óptica inicial (OD₆₀₀) de 1. La fase discontinua de aproximadamente 25 h (que alcanzó una concentración de biomasa seca de aproximadamente 20 g/L) fue seguida por una fase discontinua alimentada con glucosa limitada que comenzó con una alimentación exponencial durante 7 h y una tasa de alimentación constante de 15 g/L durante 13 h, lo que condujo a una concentración final de biomasa seca de aproximadamente 110 g/L. Se tomaron muestras durante las fases discontinua y discontinua alimentada, y se analizó la expresión de eGFP utilizando un lector de placas (Infinite 200, Tecan, CH). Por lo tanto, las muestras se diluyeron hasta 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 expresó bajo el control del promotor pCS1 tuvo una expresión promedio de eGFP 4,2 veces superior 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 discontinua alimentada optimizada, t indica el tiempo de alimentación.

t[h]	pGAP_eGFP#2		pCS1_eGFP#1		comparación FL/r pCS1/pGAP
	FL/r	STDEV	FL/r	STDEV	
-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

Con el fin de obtener más información sobre la actividad promotora de pCS1 en diferentes composiciones de medio y condiciones de crecimiento, la cepa pCS1_eGFP#4 se cultivó en medio YP que contenía 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 sola colonia 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 precultivo. Después de aproximadamente 24 horas, el precultivo se utilizó 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 + glucosa al 2 %, YPG: YP + glicerol al 2 %; YPM: YP + metanol al 1 %; YPE: YP + etanol al 1 %; perla YPfeed: YP + 1 perla de alimentación de glucosa (Kuhner, CH, diámetro 6 mm); YPD pH4,5: YPD ajustado a pH 4,5 con HCl; SCD: medio de cribado sintético (por litro: 22 g de glucosa monohidrato, 22 g de ácido cítrico, 3,15 g de $(\text{NH}_4)_2\text{HPO}_4$, 0,027 g de $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0,9 g de KCl, 0,5 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 2 mL de 500 x biotina y 1,47 mL de solución madre de sales traza [por litro: 6 g $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0,08 g NaI, 3 g $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0,2 g $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$, 0,02 g H_3BO_3 , 0,5 g CoCl_2 , 20 g ZnCl_2 , 5 g $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ y 5 mL H_2SO_4]; pH ajustado a 5 con KOH 5M; esterilizado por filtración). Todos los cultivos, excepto el que tenía la perla 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 eGFP por tamaño celular 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 superiores después de 24 h y 24 h. Los niveles de expresión bajo el control del promotor pCS1 novedoso fueron incluso mayores cuando se utilizó metanol o etanol como fuente de carbono, o cuando se cultivó a pH 4,5 (Tabla 7).

Tabla 7: Fluorescencia relativa de eGFP por tamaño celular 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 cribado. Se indican los valores promedio y la SD de 2 cultivos.

clon	Medio de cultivo principal	24 h	48 h
pGAP_eGFP#2	YPD	291,0	392,9
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+perla de alimentación de glucosa	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 la actividad promotora del promotor recientemente identificado pCS1 en *P. pastoris* utilizando albúmina sérica humana (HSA) como gen informador de expresión extracelular

Con el fin de analizar las propiedades del promotor recién identificado para la expresión de la proteína informadora secretada HSA, se realizaron cribados en matraz agitado de la siguiente manera: El precultivo durante 24 horas se realizó en medio rico que contenía glicerol como fuente de carbono -simulando la fase discontinua del proceso, al cual siguió el cultivo principal en medio rico amortiguado (glucosa al 2 %). El cultivo principal se alimentó con glucosa al 0,5 % cada 12 horas.

a) Cepa & vector de expresión

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

b) Amplificación y clonación del promotor pCS1 recién 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, se utilizó como referencia el vector pPM1aZ10_pGAP_HSA, que contiene el promotor comúnmente utilizado de la gliceraldehído 3-fosfato deshidrogenasa (pGAP). Los promotores se insertaron corriente arriba del codón de inicio del gen HSA utilizando los sitios de restricción Apal y SbfI (véase la Tabla 3). La

exactitud de las secuencias promotoras se verificó mediante secuenciación Sanger.

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

Todos los plásmidos se linealizaron utilizando la enzima de restricción AscI antes de la electroporación (utilizando 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) que contenían 25 µg/mL de Zeocina. Se utilizó la PCR en colonia 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 precultivo. Después de aproximadamente 24 h, el precultivo se utilizó 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 nitrogenada de levadura con sulfato de amonio, 0,4 mg de biotina y amortiguador de fosfato de potasio 100 mM a pH 6,0). El cultivo principal se alimentó con glucosa al 0,5 % cada 12 horas. Las muestras se tomaron al final del cultivo principal. La concentración de biomasa se determinó midiendo la OD₆₀₀ o el peso celular húmedo. La concentración de HSA en el sobrenadante del cultivo se cuantificó mediante Human Albumin ELISA Quantitation Set (Cat.No. E80-129, Bethyl Laboratories, TX, USA) siguiendo el manual de instrucciones del proveedor. El estándar de HSA se utilizó con una concentración inicial de 400 ng mL⁻¹. Las muestras se diluyeron convenientemente en diluyente de muestras (50 mM Tris-HCl, 140 mM NaCl, BSA 1 % (p/v), Tween20 0,05 % (v/v), pH 8,0). En la Tabla 8 se presentan los títulos de HSA procedentes 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. Todos los clones controlados por pCS1 secretan el doble de cantidad de HSA que el clon controlado por pGAP con una copia del gen. GCN de HSA se determinaron como se describe en el Ejemplo 5. Todos los clones analizados bajo control de pCS1 albergaban una copia del casete de expresión (1 GCN). Así, todos los clones bajo el control del promotor pCS1 novedoso tenían un rendimiento de secreción de HSA dos veces mayor (mg de HSA secretada/g de biomasa) que los clones pGAP con el mismo GCN en cultivos de cribado.

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

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

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

Para analizar la capacidad de pCS1 para impulsar la expresión de HSA en condiciones similares a las del proceso de producción, se realizaron cultivos discontinuos alimentados de un clon de pCS1 (pCS1_HSA#1) que albergaba una copia del casete de expresión de HSA (véase el Ejemplo 3).

Las fermentaciones se realizaron en biorreactores DASGIP con un volumen de trabajo final de 1,0 L. Dos cepas diferentes de *P. pastoris* que expresan HSA bajo control de pGAP (pGAP_HSA#3 que tiene una copia del gen HSA, y pGAP_HSA#4 que tiene dos copias del gen HSA, descritas 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

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 biotina y 5,0 mL H₂SO₄ (95 %-98 %).

Medio Discontinuo de Glicerol contenido por litro

39,2 g de 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 PTM₁. El pH se ajustó a 5,85 después de filtrarlo estérilmente en el fermentador.

Medio discontinuo alimentado de glucosa contenido por litro

550 g de glucosa monohidrato, 6,5 g de MgSO₄·7H₂O, 10 g de KCl, 0,35 g de CaCl₂·2H₂O, 0,4 mg de biotina y 12 mL de solución madre de sales traza PTM₁.

El oxígeno disuelto se controló a DO = 20 % con la velocidad del agitador (400 - 1200 rpm). La tasa de aireación es de 24 L h⁻¹ de aire, la temperatura se controló a 25 °C y el valor nominal de pH de 5,85 se controla con la adición de NH₄OH

(25 %).

Para iniciar la fermentación, se filtraron estérilmente 400 mL de medio discontinuo en el fermentador y se inoculó *P. pastoris* (procedente de precultivo) con una densidad óptica inicial (OD₆₀₀) de 1. La fase discontinua de aproximadamente 25 h alcanzó una concentración de biomasa seca de aproximadamente 20 g/L y fue seguida por discontinua alimentada con una alimentación constante de medio de glucosa (2 g L⁻¹ h⁻¹) durante 100 horas, lo que condujo a una concentración de biomasa seca de aproximadamente 100 g/L. El pH fue de 5,85 durante la fase discontinua, y se mantuvo en 5,85 durante toda la fermentación. Se tomaron muestras durante la fase discontinua y discontinua alimentada. La concentración de HSA se cuantificó utilizando Human Albumin ELISA Quantitation Set (Bethyl, Cat.No. E80-129) como se describe en el Ejemplo 3c.

Como se demostró anteriormente, el clon pGAP con 2 copias de HSA secretaba 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 superiores al final de la fase discontinua y discontinua alimentada en comparación con un clon pGAP de copia única (resultados mostrados en la Tabla 9). En relación con la biomasa y el GCN, los clones pCS1 produjeron un 390 % de HSA secretada que el clon pGAP con el mismo número de copias del gen.

Tabla 9: Concentraciones de masa seca de levadura y títulos de HSA en el sobrenadante al final de la fase discontinua y al final de la fase discontinua alimentada, y título de HSA por masa seca de levadura al final de la fase discontinua alimentada de cultivos en biorreactor de clones de *P. pastoris* que expresan HSA bajo el control de pCS1 o pGAP.

Promotor	GCN	final discontinuo		final discontinuo alimentado		
		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 del número de copias de genes (GCN) de clones seleccionados

La fuerza de expresión suele estar correlacionada 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 utilizando DNeasy Blood&Tissue Kit (Quiagen, Cat.No. 69504). El número de copias de genes se determinó mediante PCR cuantitativa. Por lo tanto, se utilizó SensiMix SYBR kit (Bioline, QT605-05). Sensi Mix SYBR se mezcló con los cebadores respectivos (indicados en Prielhofer *et al.* 2013. Microb. Cell. Fact. 12:5) y la muestra, y se aplicó para el análisis en tiempo real en un ciclador de PCR en tiempo real (Rotor Gene, Qiagen). Todas las muestras se analizaron por triplicado o cuadruplicado. Para el análisis de los datos se utilizó el programa informático Rotor Gene.

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

Con el fin de analizar las propiedades del promotor identificado recientemente, se realizan ensayos en matraz agitado como se indica a continuación: El precultivo durante 24 horas se realiza con medio rico que contiene glicerol como fuente de carbono, al que sigue el cultivo principal con medio rico.

a) Cepa & vector de expresión

Como cepa huésped se utiliza la cepa de tipo salvaje de *P. pastoris* (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos). La transformación de la cepa se lleva a cabo con un vector propio denominado pPUZZLE (Stadlmayr *et al.* J. Biotechnol 2010 dic; 150(4):519-29), la selección de transformantes positivos se basa en la resistencia a Zeocina. Para la expresión secretora de la carboxipeptidasa porcina B (CpB) se utiliza la levadura alfa líder del factor de acoplamiento.

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

El promotor amplificado en el Ejemplo 2b se clona en el vector de expresión pPUZZLE pPM1aZ30_aMF_CpB, dando como resultado pPM1aZ30_pCS1_aMF_CpB. Además, se utiliza como referencia el vector pPM1dZ30_pGAP_CpB, que contiene el promotor comúnmente utilizado de la gliceraldehído 3-fosfato deshidrogenasa (pGAP). Los promotores se insertan corriente arriba del codón de inicio del gen CpB utilizando los sitios de restricción Apal y SbfI. La corrección de las secuencias promotoras se verifica mediante secuenciación Sanger.

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

identificados

Los plásmidos se linealizan utilizando la enzima de restricción AscI antes de la electroporación (utilizando 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) que contienen 25 µg/mL de Zeocina. Se utiliza la PCR en colonia para asegurar la presencia del plásmido transformado como se describe en el Ejemplo 2c.

Para el cribado de la expresión de CpB, se inocula 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 precultivo. Después de aproximadamente 24h, el precultivo se utiliza 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 glucosa al 0,5 % cada 12 horas. Las muestras se toman al final del precultivo y 24 y 48 horas después de la inoculación del cultivo principal. La concentración de biomasa se determina midiendo la 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 la CpB. La cinética de la 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 amortiguaron con un amortiguador de ensayo (25 mM Tris, 100 mM HCl, pH 7,65) y se activan utilizando amortiguador de activación (0,01 mgL⁻¹ Tripsina, 300 mM Tris, 1µM ZnCl₂, pH 7,65). Se utiliza amortiguador de activación sin tripsina en lugar de la muestra como control negativo. La reacción se inicia añadiendo la solución de sustrato (1mM de hippuril-L-arginina en amortiguador de ensayo).

d) Cultivo discontinuo alimentado de cepas de *P. pastoris* que expresan CpB bajo el control del promotor pCS1 La fermentación discontinua alimentada se realiza como se describe en el ejemplo 4 utilizando los medios descritos en el ejemplo 2d.

Ejemplo 7: Estudios comparativos de la actividad promotora del promotor identificado recientemente pCS1 en clones multicopia de *P. pastoris* utilizando albúmina sérica humana (HSA) como gen informador de expresión extracelular

Para investigar si la producción de HSA podría mejorarse aún más con un mayor GCN, se amplificaron clones de Pcs1 que expresaban HSA mediante el método de amplificación de vectores post-transformación, descrito por Marx et al., (2009). Se generaron vectores que se integran en el locus ADN_r por recombinación homóloga. La amplificación se realizó por selección en concentraciones crecientes paso a paso de antibióticos.

a) Cepa & vector de expresión

La cepa *P. pastoris* de tipo salvaje (CBS2612, CBS-KNAW Fungal Biodiversity Centre, Centraalbureau voor Schimmelcultures, Utrecht, Países Bajos) se utilizó como cepa huésped. La transformación de la cepa se llevó a cabo con una variante del vector pPUZZLE que contenía 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 la Zeocina. Para la expresión secretora de la albúmina sérica humana (HSA) se utilizó 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.

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

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

Los plásmidos se linealizaron utilizando la enzima de restricción SpeI antes de la electroporación (utilizando 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) que contenían 25 µg/mL de Zeocina. Se utilizó la PCR en colonia para asegurar la presencia del plásmido transformado como se describe en el Ejemplo 2c. La amplificación del número de copias de genes se realizó como se describe en Marx et al. (FEMS Yeast Res. 2009 dic;9(8): 1260-70) repitiendo la siembra de clones en placas de agar YPD con concentraciones más altas de Zeocina (50, 100 y 500 µg/mL de Zeocina).

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

Al aumentar el GCN de la HSA, también se pudieron aumentar los niveles de secreción de HSA (mostrados 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 superiores por WCW en comparación con el clon de copia única. También cabría esperar un

aumento similar de la productividad de los clones pCS1 multicopia en cultivos en biorreactores discontinuos alimentados, tal como se presenta en el Ejemplo 4.

Tabla 10: Resultados del cribado de la expresión de HSA con clones simples y multicopia bajo el control de pGAP o pCS1. Se muestran los GCN y los títulos por GCN.

Clon	$\mu\text{g HSA/g WCW}$	GCN	$\mu\text{g HSA/g WCW/GCN}$
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 identificado recientemente pCS1 en *P. pastoris* utilizando un fragmento de anticuerpo (Fab) como gen informador de expresión extracelular

Con el fin de analizar las propiedades del promotor recién identificado, se realizaron cribados en matraz agitado como se indica a continuación: Se realizó un precultivo durante 24 horas con medio rico que contenía glicerol como fuente de carbono, al que siguió el cultivo principal con medio rico.

a) Cepa & vector de expresión

La cepa *P. pastoris* de tipo salvaje (CBS2612, CBS-KNAW Fungal Biodiversity Centre, 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 corriente arriba del codón de inicio de los genes Fab utilizando los sitios de restricción Apal y SbfI. Después de verificar la secuencia, los casetes de expresión de ambas cadenas se combinaron en un vector utilizando las enzimas de restricción compatibles MreI y AgeI.

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

Los plásmidos se linealizaron utilizando la enzima de restricción AscI antes de la electroporación (utilizando 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) que contenían 25 $\mu\text{g/mL}$ de Zeocina. Se utilizó la PCR en colonia para asegurar la presencia del plásmido transformado como se describe en el Ejemplo 2c.

El cribado de expresión de Fab se realizó de forma similar al cribado de expresión de HSA descrito en el Ejemplo 3c. La cuantificación de Fab intacto se realizó mediante ELISA utilizando un anticuerpo IgG antihumano (Abcam ab7497) como anticuerpo de recubrimiento (1:1.000), y un anticuerpo de cabra anti-cadenas ligeras Kappa humanas (unidas y libres) - conjugado con fosfatasa alcalina (Sigma A3813) como anticuerpo de detección (1:1.000). Se utilizó como estándar el fragmento humano Fab/Kappa, IgG (Bethyl P80-115) con una concentración inicial de 50 ng/mL. Las muestras sobrenadantes se diluyen en consecuencia. La detección se realizó con el sustrato pNPP (Sigma S0942). Los amortiguadores de recubrimiento, dilución y lavado se basaron en PBS (2 mM KH_2PO_4 , 10 mM $\text{Na}_2\text{HPO}_4 \cdot 2 \text{H}_2\text{O}$, 2,7 mM g KCl, 8 mM NaCl, pH 7,4) y se completaron 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 por sus siglas en inglés) y de la cadena ligera (LC por sus siglas en inglés) como se describió en el Ejemplo 5. Los GCN se relacionaron con el clon PCS1#5, para el que GCN HC y LC se fijó en uno. Los rendimientos de expresión Fab ($\mu\text{g Fab/g WCW}$), GCN relativos y rendimiento Fab por GCN se muestran en la Tabla 11. En cuanto a las otras proteínas modelo, se obtuvo un rendimiento Fab por GCN aproximadamente dos veces superior (un promedio de 35,4 $\mu\text{g Fab/g WCW}$) para los clones que se expresaban bajo el control de pCS1 en comparación con los clones que se expresaban bajo el control de pGAP.

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

Clon	$[\mu\text{g Fab g}^{-1} \text{WCW}]$	GCN relativo relacionado con pCS1_Fab#5		Rendimiento Fab $[\mu\text{g Fab g}^{-1} \text{WCW}]$ por GCN	
		Cadena pesada	Cadena ligera	Cadena pesada	Cadena ligera
pGAP_Fab#34	73,3	4	3	18	24

Clon	[µg Fab g ⁻¹ WCW]	GCN relativo relacionado con pCS1_Fab#5		Rendimiento Fab [µg Fab g ⁻¹ WCW] por GCN	
		Cadena pesada	Cadena ligera	Cadena pesada	Cadena ligera
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 discontinuo alimentado de cepas de *P. pastoris* que expresan Fab bajo el control del promotor pCS1.

Se realizaron fermentaciones discontinuas alimentadas similares a las descritas en el ejemplo 4 utilizando medios como los descritos en el ejemplo 2d.

En el cultivo en biorreactor, la expresión de Fab bajo el control de pCS1 dio lugar a productividades específicas (qP) por GCN más de 3,0 veces superiores en comparación con pGAP (véase la Tabla 12).

10 Tabla 12: Resultados del cultivo en biorreactor de la expresión de Fab bajo el control de pGAP o pCS1. Se muestran los rendimientos de Fab al final del cultivo discontinuo alimentado, la productividad específica media qP (tasa promedio de producción de Fab por biomasa (peso celular seco DOW) durante toda la fase discontinua alimentada), el GCN relativo de los clones respectivos y el qP medio por GCN.

clon	Rendimiento Fab [mg Fab g ⁻¹ DCW]	qP [µg Fab g ⁻¹ DCW h ⁻¹]	GCN relativo relacionado con 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

15 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 someten a un cribado similar al descrito en el ejemplo 2c. Los clones que expresaban bajo el control de pCS1 (longitud estándar) y pGAP se utilizaron como controles. Los cebadores sentido y las longitudes de pCS1 y sus variantes se listan en la Tabla 13.

20

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

variantes	cebador sentido	Longitud (bp)
pCS1 (estándar) SEQ ID 1	GATAGGGCCCCAGGGCATCATTGAGGTTTCCAC SEQ ID 16	985
pCS1-1488 SEQ ID 2	GATAGGGCCCCGATAGTTCTAGAAGACCTGGCGTCG SEQ ID 17	1488
pCS1-767F SEQ ID 3	GATAGGGCCCAGCCAACCATCTTTTGTTCG SEQ ID 18	767
pCS1-500F SEQ ID 4	GATAGGGCCCCGTGGTTTCCAGGACAACACCC SEQ ID 19	500
pCS1-344F SEQ ID 5	GATAGGGCCCCGACCGCAATTACCATGATGC SEQ ID 20	344
pCS1-234F SEQ ID 6	GATAGGGCCCCAGCCTGCTTCATTCCTGCC SEQ ID 21	234
pCS1-138F SEQ ID 7	GATAGGGCCCCCGCAAAAAGGTTTGTATAG SEQ ID 22	138
pCS1-85F SEQ ID 8	GATAGGGCCCCATACTCTCCTCCCCCCTG SEQ ID 23	85

Los cribados de clones que expresan eGFP bajo el control de pGAP, pCS1 o pCS1 reveló que pCS1 requiere una longitud mínima de aproximadamente 500 bp para una actividad óptima. Las variantes pCS1 más cortas pierden actividad a medida que disminuye su longitud (véase la Tabla 14).

25

Clon	FL/Tamaño celular, 24 h	FL/Tamaño celular, 48 h
pGAP_eGFP#2	15,39	17,59
pCS1_eGFP#4	25,30	37,77
pCS1_85pool	-0,15	0,04

Clon	FL/Tamaño celular, 24 h	FL/Tamaño celular, 48 h
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

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

5 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 a tasas de crecimiento elevadas y bajas

a) Cepa

10 Para comparar los niveles de expresión se utilizan una cepa huésped (por ejemplo, *Pichia pastoris*) que expresa eGFP bajo el control de un "promotor de interés" y una cepa que expresa eGFP bajo el control de pGAP.

b) Cultivo de cepas que expresan eGFP para comparar promotores

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

Se utilizan los siguientes medios:

20

La solución madre de sales traza PTM1 contiene por litro

6,0 g de $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$, 0,08 g de NaI, 3,36 g de $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0,2 g de $\text{Na}_2\text{MoO}_4 \cdot 2\text{H}_2\text{O}$,
 25 0,02 g de H_3BO_3 , 0,82 g de CoCl_2 , 20,0 g de ZnCl_2 , 65,0 g de $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$, 0,2 g de biotina y 5,0 mL de H_2SO_4 (95 %-98 %).

Medio Discontinuo de Glicerol que contiene por litro

30 2 g de ácido cítrico monohidratado ($\text{C}_6\text{H}_8\text{O}_7 \cdot \text{H}_2\text{O}$), 39,2 g de glicerol, 12,6 g de $\text{NH}_4\text{H}_2\text{PO}_4$, 0,5 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,9 g de KCl, 0,022 g de $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0,4 mg de biotina y 4,6 mL de PTM1
 solución madre de sales traza. Se añade HCl para ajustar el pH a 5.

Medio quimiostático que contiene por litro

35 2,5 g de ácido cítrico monohidratado ($\text{C}_6\text{H}_8\text{O}_7 \cdot \text{H}_2\text{O}$), 55,0 g de glucosa monohidratada, 21,75 g de $(\text{NH}_4)_2\text{HPO}_4$, 1,0 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 2,5 g de KCl, 0,04 g de $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0,4 mg de biotina y 2,43 mL de solución madre de sales traza PTM1.
 Se añade HCl para ajustar el pH a 5.

40 El oxígeno disuelto se controla a $\text{DO} = 20\%$ con la velocidad del agitador (400 - 1200 rpm). La tasa de aireación es de 24 L h^{-1} de aire, la temperatura se controla a 25°C y el valor nominal de pH de 5 se controla con adición de NH_4OH (25 %). Para iniciar la fermentación, se filtran estérilmente 400 mL de medio discontinuo en el fermentador y se inocula un clon de *P. pastoris* (procedente de precultivo) con una densidad óptica inicial (OD_{600}) de 1. A la fase discontinua de aproximadamente 25 h (que alcanza una concentración de biomasa seca de aproximadamente 20 g L^{-1}) le sigue el cultivo en quimiostato limitado por glucosa. La tasa de alimentación del medio quimiostático y la tasa de cosecha se utilizan para mantener constante la tasa de crecimiento específica deseada. Durante este cultivo, el volumen del caldo de cultivo se
 45 mantiene constante y se determina el peso seco de las células para garantizar una tasa de crecimiento constante. Las células se ultivan a una tasa de crecimiento elevada y baja de $0,15$ y $0,015 \text{ h}^{-1}$, respectivamente. Por lo tanto, la tasa de alimentación/recolección se controla a $150 \text{ mL h}^{-1} \text{ L}^{-1}$ (mL de medio quimiostático por litro de caldo de cultivo y hora) y $15 \text{ mL h}^{-1} \text{ L}^{-1}$, respectivamente.

50 c) Muestreo

Se toman muestras en condiciones estables (después de al menos 5 intercambios de volumen) y se analiza la expresión de eGFP utilizando un lector de placas (Infinite 200, Tecan, CH). Por lo tanto, las muestras se diluyen hasta una densidad óptica (OD_{600}) de 5. Las fermentaciones se realizan por duplicado. Los datos de expresión se comparan calculando la
 55 fluorescencia relativa por biorreactor como se describe en el ejemplo 2d).

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

- 5 Con el fin de identificar un promotor que permitiera una transcripción elevada a tasas de crecimiento elevadas y bajas, se analizaron los patrones de expresión génica mediante microarrays de ADN. A partir de los datos transcriptómicos se seleccionaron los genes que mostraban una elevada fuerza de transcripción a tasas de crecimiento elevadas 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 elevadas y bajas de 0,15 y 0,015 h⁻¹, respectivamente. El muestreo, la purificación del ARN, la
- 10 preparación de la muestra para la hibridación de microarrays, el análisis de microarrays, la adquisición de datos y la evaluación estadística se realizan como se describe en los ejemplos 1c), 1 d), 1e) y 1f). Los genes y sus respectivos promotores con una elevada fuerza de transcripción en condiciones de elevada y baja tasa de crecimiento se identificaron examinando los datos de microarrays en busca de genes con intensidades de señal elevadas tanto en condiciones de elevada como de baja tasa de crecimiento. Como segundo criterio, las intensidades de señal deben ser superiores a las
- 15 del gen de la gliceraldehído-3-fosfato deshidrogenasa (GAP, sinónimos GAPDH y TDH3) en ambas condiciones. Para aislar el promotor, se amplificó un fragmento de ácido nucleico de aproximadamente 1000 bp corriente arriba del codón de inicio ATG del gen respectivo.

REIVINDICACIONES

1. Un promotor constitutivo aislado, que es una variante de tamaño del promotor pCS1 de *Pichia pastoris* establecido en SEQ ID NO:1, cuya variante de tamaño tiene una longitud de 500-1500 bp, al menos 90 % de identidad con la secuencia de SEQ ID NO:1, y comprende una secuencia de ácido nucleico que es al menos 90 % idéntica a cualquiera de SEQ ID NO: 2, 3, 4 o 5,
 donde dicha variante de tamaño exhibe la misma expresión o fuerza transcripcional que el promotor pCS1 establecido en SEQ ID NO:1 +/- 20 %, y/o al menos un aumento de 1,1 veces en comparación con el promotor nativo pGAP de *P. pastoris* establecido en SEQ ID NO: 13, donde la variante de tamaño no es SEQ ID NO:2 de WO2014/138703 A1.
2. El promotor de conformidad con la reivindicación 1, donde la variante de tamaño comprende una secuencia de ácido nucleico que es cualquiera de SEQ ID NO: 2, 3, 4 o 5.
3. El promotor de conformidad con la reivindicación 1 o 2, que exhibe actividad promotora constitutiva como se determina por análisis de transcripción en una célula huésped a elevada tasa de crecimiento y a baja tasa de crecimiento.
4. El promotor de conformidad con cualquiera de las reivindicaciones 1 a 3, que está unido de forma operable 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 la POI.
5. El promotor de conformidad con la reivindicación 4, que comprende además una secuencia de ácido nucleico que codifica un péptido señal que permite la secreción de la POI, preferiblemente donde la secuencia de ácido nucleico que codifica el péptido señal está situada adyacente al extremo 5' de la secuencia de nucleótidos que codifica la POI.
6. Un constructo de expresión que comprenda el promotor de conformidad con cualquiera de las reivindicaciones 1 a 5, preferiblemente un vector o plásmido de replicación autónoma, o uno que se integre en el ADN cromosómico de una célula huésped.
7. Una célula huésped recombinante que comprende un constructo de expresión de conformidad con la reivindicación 6, preferiblemente una célula eucariota, más preferiblemente una célula de levadura o de hongo filamentoso, más preferiblemente una célula de levadura del género *Saccharomyces* o *Pichia*.
8. La célula huésped recombinante de conformidad con la reivindicación 7, que comprende múltiples copias del promotor, y/o múltiples copias del constructo de expresión.
9. La célula huésped recombinante de conformidad con la reivindicación 7 u 8, que se selecciona del grupo que consiste en células de mamíferos, insectos, levaduras, hongos filamentosos y plantas, preferiblemente una levadura, preferiblemente cualquiera de las cepas *P. pastoris* CBS 704, CBS 2612, CBS 7435, CBS 9173-9189, DSMZ 70877, X-33, GS115, KM71 y SMD1168.
10. Un cultivo estable de una pluralidad de la célula de conformidad con cualquiera de las reivindicaciones 7 a 9.
11. Un método de producción de una proteína de interés (POI) mediante el cultivo de una línea celular huésped recombinante, que comprende:
 - i) el promotor de conformidad con cualquiera de las reivindicaciones 1 a 5 o el constructo de expresión de conformidad con la reivindicación 6, donde el promotor está operativamente unido a un ácido nucleico que codifica la POI; o
 - ii) la célula huésped recombinante de conformidad con cualquiera de las reivindicaciones 7 a 9;
 que comprende los pasos de
 - a) cultivar la línea celular en condiciones que expresen dicha POI, y
 - b) recuperar la POI.
12. El método de conformidad con la reivindicación 11, donde la POI se expresa en condiciones limitantes del crecimiento.
13. El método de conformidad con la reivindicación 11 o 12, donde la línea celular se cultiva en condiciones de cultivo discontinuo, discontinuo alimentado o continuo, y/o en medios que contienen sustrato de carbono limitado.
14. El método de conformidad con cualquiera de las reivindicaciones 11 a 13, donde la 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 carbohidrato-proteína, proteínas estructurales, proteínas reguladoras, vacunas y proteínas o partículas similares a vacunas, enzimas de proceso, factores de crecimiento, hormonas y citoquinas, o un metabolito de una POI.

- 5 15. Uso del promotor de conformidad con cualquiera de las reivindicaciones 1 a 5 o del constructo de expresión de conformidad con la reivindicación 6 en un método de producción de una proteína de interés (POI) mediante el cultivo de una célula huésped transformada con el promotor y/o el constructo de expresión, donde el promotor está unido de forma operable a una secuencia nucleotídica que codifica la POI, preferiblemente donde el cultivo se realiza en un biorreactor que comienza con una fase de cultivo discontinua seguida de una fase discontinua alimentada o una fase continua.

SEQ ID 4

GTGGTTTCAGGACACACCCCAAAAAAAGGTATCAATGGCCACTAGGCCAGTCGGTTTTATTTTTGGTCAACCCACGCA
AAGAAGCACCACCCACCTCTTTTAGGTTTTAAGTTGTGGGACACAGTAACACCGCCTAGAGCTTCAGGAAAAACCACTAC
CTGTGACCCGCAATTACCATGATGCAGAAATGTTAATTTAAACGAGTGGCAAAATCAAGATTTCAACAGACAAATCAA
TCGATCCATAGTTACCCATTCCAGCCCTTTTGGTCTGTCGAGCCTGCTTCATTCCCTGCTCAGGTGCATAACTTTGCA
TGAAAACTCCAGATTAGGGCAGATTTTGAGTTTAAAAATAGGAAATATAAACAAATATACCGCGAAAAAAGCTTTGTT
TATAGCTTTTCCCTTGGTGGCGTACCGTATAAAATACATACTCTCCCTCCCCCCCCCTGGTCTCTCTTTTCTTTTGTTA
CTTACATTTTACCGTTCCGTCACCTCGCTTCACTCAACAACAAAA

SEQ ID 5

GACCCGCAATTCACCATGATGCAGAAATGTTAATTTAAACGAGTGGCAAAATCAAGATTTCAACAGACAAATCAATCGA
TCCATAGTTAGCCATTCACCGCTTTTGGTCTGTCGAGCCTGCTTCATTCCCTGCTCAGGTGCATAACTTTGCATCAA
AAGTCCAGATTAGGGCAGATTTTGAGTTTAAAAATAGGAAATATAAACAAATATACCGCGAAAAAAGCTTTGTTTATA
GCTTTTCCCTTGGTGGCGTACCGTATAAAATACATACTCTCCCTCCCCCCCCCTGGTCTCTCTTTTCTTTTGTACTTA
CATTTTACCGTTCCGTCACCTCGCTTCACTCAACAACAAAA

SEQ ID 6

AGCCTGCTTCATTCTGCTCAGGTGCATAACTTTGCATGAAAACTCCAGATTAGGGCAGATTTTGAGTTTAAAAAT
AGCAAAATATAAACAAATATACCGCGAAAAAAGCTTTCTTTATAGCTTTTCCCTTGGTGGCGTACCGTATAAAATACA
ACTCTCCCTCCCCCTGGTCTCTCTTTTCTTTTGTACTTACATTTTACCGTTCCGTCACCTCGCTTCACTCAACA
ACAAAA

SEQ ID 7

CCGCGAAAAAGGTTTGTATAGCTTTTCCCTTGGTGGCGTACCGTATAAAATACATACTCTCTCCCTCCCCCCCCCTGGT
TCTCTTTTCTTTTGTACTTACATTTTACCGTTCCGTCACCTCGCTTCACTCAACAACAAAA

SEQ ID 8

CATACTCTCTCCCTCCCCCTGGTCTCTCTTTTCTTTTGTACTTACATTTTACCGTTCCGTCACCTCGCTTCACTCA
ACAACAAAA

Fig. 2

SEQ ID 9

ATGCAATTCTCTATCGTCGCTACTTTGGCTCTTGGCTGGTTCCGCTCTGGCTGCTTACTCTAACGTAACCTTACACTT
ACGAGACTACCATCACCGATGTTGTCACCGAGCTCACCACCTTACTGCCCGAGAGCCCAACCACCTTCGTTCCACAAGAA
CAAGACCATCACTSTGACCGCCCAACCACTTTGACCATCACTGACTGTCTTGCACCATCTCCAAGACCACCAAG
ATCACCACCTGATGTTCCACCAACCAACCACTCCACCCACACACCAACCACTCAGGTGCCATCTACCTCTACCC
CAGCTCCAAACCACTCTGTTTCTACCATCTCTCAGGCTGCTGCTGCTAAGGCTGGTGTGCTGGTTGGCGGGTGT
TGCTGCTGCCGCTGCTTACTTCTTGTA

SEQ ID 10

MQFSIVATLALAGSALAAYSNVITYTYETITDVVTELTTCFEPPTFVNKNTITVTAFTTLITDCPCTISRTTK
ITTDVFPPTHSTPHTTTTTHVPSTSTPAPTHSVSTISHGGAAGACVAGLAGVAAAAAYFL

SEQ ID 11

ATGCAATTCTCTATCGTCGCTACTTTGGCTCTTGGCTGGTTCCGCTCTGGCTGCTTACTCTAACG
TAACTTACACTTACGAGACTACCATCACCGATGTTGTCACTGAGTTGACCACTTACTGCCCGAGA
GCCAACCACCTTTGTTTACAAGAACAAGACCATCACCGTGACTGAGCCAACCACCTTTGACCATC
ACTGACTGCCCATGCACCATCTCAAAGACCACCAAGATCACCACTGATGTTCCACCAACCACCC
ACGTCACCCCATCCACCACTCAGGTGCCATCTACCTCTACCCAGCTCCAACCCACTCTGTTTC
TACCATCTCTCAGGTTGGTGCTGCTAAGGCTGGTGTGCTGGTTTGGCCGGTGTGCTGCTGCC
GCTGCTTACTTCTTGTA

SEQ ID 12

MQFSIVATLALAGSALAAYSNVITYTYETITDVVTELTTCFEPPTFVYKNTITVTEFTTLITDCPCTISRTTK
ITTDVFPPTHVTPSTTHVDSSTPAPTHSVSTISHGGAAGACVAGLAGVAAAAAYFL

Fig. 3

SEQ ID 13

CTTTTTGTAGAAATGTCTTGGTGTCTCTCGTCCAATCAGGTAGCCATCTCTGAAATATCTGGCT
 CCGTTGCAACTCCGAACGACCTGCTGGCAACGTAAAATTCTCCGGGGTAAAACCTTAAATGTGGA
 GTAATGGAACCAGAAACGTCTCTTCCCTTCTCTCTCCTTCCACCGCCCGTTACCGTCCCTAGGA
 AATTTTACTCTGCTGGAGAGCTTCTTCTACGGCCCCCTTGCAGCAATGCTCTTCCCAGCATTAC
 GTTGCGGGTAAAACGGAGGTCTGTGTACCCGACCTAGCAGCCCAGGGATGGAAAAGTCCCGGCCG
 TCGCTGGCAATAATAGCGGGCGGACGCATGTCATGAGATTATTGGAAACCACCAGAATCGAATA
 TAAAAGGCGAACACCTTTCCTCAATTTTGTTTCTCCTGACCCAAAGACTTTAAATTTAATTTAT
 TTGTCCCTATTTCAATCAATTGAACAACCTATCACCTGCAGGCC