

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 954 138**

51 Int. Cl.:

C07K 14/81 (2006.01)

C07K 14/78 (2006.01)

C07K 14/435 (2006.01)

C12N 15/62 (2006.01)

C12N 9/64 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.04.2017** **PCT/EP2017/059719**

87 Fecha y número de publicación internacional: **02.11.2017** **WO17186671**

96 Fecha de presentación y número de la solicitud europea: **25.04.2017** **E 17724756 (6)**

97 Fecha y número de publicación de la concesión europea: **07.06.2023** **EP 3448884**

54 Título: **Células de mamífero modificadas para la producción de proteínas recombinantes**

30 Prioridad:

25.04.2016 EP 16166789

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
20.11.2023

73 Titular/es:

DANMARKS TEKNISKE UNIVERSITET (50.0%)
Anker Engelunds Vej 101
2800 Kongens Lyngby, DK y
THE REGENTS OF THE UNIVERSITY OF
CALIFORNIA (50.0%)

72 Inventor/es:

VOLDBORG, BJØRN, GUNNAR;
PEDERSEN, LASSE, EBDROP y
PALSSON, BERNHARD

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

Observaciones:

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

ES 2 954 138 T3

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

DESCRIPCIÓN

Células de mamífero modificadas para la producción de proteínas recombinantes

5 **Campo de la invención**

La presente invención se refiere a células de mamífero modificadas para proporcionar una expresión mejorada de una proteína recombinante de interés. En concreto, la invención se refiere a células CHO y otras células huésped en las que se ha interrumpido la expresión de una o más proteínas endógenas secretadas, así como a la preparación, identificación y uso de tales células en la producción de proteínas recombinantes.

Antecedentes de la invención

Las células de ovario de hámster chino (CHO) son el principal sistema de expresión para la fabricación de productos biofarmacéuticos, que representan el 35,5 % del total acumulado (1982-2014) de aprobaciones de productos biofarmacéuticos (Walsh, 2014). Debido a su importancia en la producción comercial de proteínas recombinantes terapéuticas, se han hecho esfuerzos significativos para caracterizar y optimizar las células CHO para este fin (Kim *et al.*, 2012; Datta *et al.*, 2013; Lee *et al.*, 2015; Chaudhuri *et al.*, 2015). Estos incluyen estudios del proteoma y secretoma de células CHO, representando este último las proteínas de la célula huésped (HCP) secretadas por las células CHO (Baycin-Hizal *et al.*, 2012; Slade *et al.*, 2012; Zhang, 2010).

El documento WO 2007/006808 A1 se refiere a células CHO con genes atenuados o desactivados que codifican proteínas de células huésped que contaminan la producción recombinante de proteínas dependientes de vitamina K. Por ejemplo, se informó que la proteína S endógena contamina las preparaciones de Factor VIIa recombinante, cuando esta proteína se produjo en células CHO.

El documento WO 2012/082509 A2 se refiere a la atenuación o desactivación de los biomarcadores Acs13, Aklb8, Anapc10, Arl6ip1, Cnih2, Ctsd, Ctsl, Derl1, Dse, Ebpl, Ecm1, Elk3, Fth1, Gjb3, Hmox1, Itgb1bp1, Ldha, Lgals1, Lgals3, Lrpap1, Myl6b, Pit1, Pomp, Rps26, Sen2, Sh2d3c, S100a4, Tmed1, Tnfrsf12a, Tinagl, Tnfrsf25, Trappc6b, 0610007C21Rik y 2610209M04Rik con el fin de mejorar el crecimiento o la viabilidad celular y mejorar la producción de proteínas en las células CHO.

A pesar de estos y otros avances en la materia, todavía existe la necesidad de células huésped tales como células CHO que sean capaces de producir un alto rendimiento de proteínas recombinantes de interés.

Sumario de la invención

El (los) presente(s) inventor(es) ha(n) descubierto que las células huésped de mamífero pueden modificarse, tal como por la interrupción por desactivación de genes endógenos específicos, para una expresión mejorada o más alta de una proteína recombinante de interés producida por la célula huésped, tal como una célula CHO.

En primer lugar, la producción por parte de las células de mamífero de estas proteínas de la célula huésped puede estar limitando la capacidad de las células para producir una proteína recombinante de interés, simplemente por el uso de la misma maquinaria de producción de proteínas de la célula, o por el suministro de nutrientes para las proteínas. Al eliminar estos genes endógenos, la célula huésped no gasta recursos en la producción de estas proteínas endógenas producto de estos genes.

En segundo lugar, algunos de estos productos de proteínas endógenas pueden ser muy difíciles de eliminar por completo de un medio de una célula huésped que produce una o más de estas proteínas de células huésped además de una proteína recombinante de interés.

Otro objetivo de las realizaciones de la presente invención es proporcionar células huésped de mamíferos, células que son capaces de crecer a una mayor densidad celular. Esto tiene la ventaja significativa de que se puede obtener una mayor producción de una proteína recombinante al tener más células en el mismo lote que produce tal proteína recombinante.

Así, en un primer aspecto, la presente invención se refiere a una célula CHO modificada para interrumpir la expresión de una o más proteínas endógenas seleccionadas entre el grupo que consiste en TIMP1, NID1-1, NID1-2 y LGALS3BP.

En un segundo aspecto, la presente invención se refiere a un método para producir una proteína recombinante de interés, que comprende

- (a) cultivar la célula CHO de algunas realizaciones de la invención en un medio de cultivo de modo que la proteína de interés se exprese a partir de la secuencia de ácido nucleico;
- (b) recolectar la proteína recombinante de interés del medio de cultivo; y

(c) opcionalmente, purificar adicionalmente la proteína recombinante recolectada de interés.

En un tercer aspecto adicional, la invención se refiere a un método para preparar la célula CHO de la invención, que comprende las etapas de

(a) modificar la célula CHO para atenuar o para inactivar el gen endógeno o los genes endógenos que codifican uno o más de TIMP1, NID1-1, NID1-2 y LGALS3BP; y

(b) opcionalmente, transfectar la célula CHO modificada con una secuencia de ácido nucleico que codifica una proteína de interés.

Breve descripción de los dibujos

Figura 1: Expresión de rituximab de un KO secretor multiplex y clones KO secretores individuales (normalizados a CHO-S WT). Títulos de rituximab de líneas de células CHO transfectadas transitoriamente cultivadas en matraces agitadores. Todos los títulos se normalizan a títulos de células CHO salvaje.

Figura 2: Densidad celular máxima de un KO secretor multiplex y clones KO secretores individuales (normalizados a CHOS WT. Máxima densidad celular alcanzada en matraces agitadores, para el 6SecKO y para las líneas celulares con los genes eliminados individualmente, en comparación con las salvajes.

Figura 3: Presencia de proteínas de células huésped (HCP) (normalizadas a CHO-S salvajes) en los medios usados medidos por cantidad (BCA y Nanodrop) e inmunogenicidad (HCP). CDCHO es medio CHO libre de HCP.

Figura 4: Análisis de la HCP presente en los medios usados a partir de células con los genes TIMP1, BGN, NID1-1 y NID1-2 eliminados individualmente. Siendo los números normalizados a las salvajes 100.

Divulgación detallada de la invención

La presente invención se refiere a células de mamíferos y, en particular, a células CHO modificadas para proporcionar una expresión mejorada de una proteína recombinante de interés. En concreto, la expresión de una o más proteínas secretadas endógenas se ha reducido o interrumpido significativamente.

Se entenderá que, en principio, se puede usar cualquier método o técnica para reducir o interrumpir la expresión de una o más proteínas endógenas secretadas. Los ejemplos de tales métodos, incluida la orientación de ARNip, inactivación génica dirigida, transfección con un factor transcripcional y la escisión específica del sitio de las hebras de ADN que codifican los contaminantes proteicos no deben interpretarse como limitantes de ningún modo. En principio, se puede usar cualquier biología molecular, biología celular o método de selección para reducir o interrumpir el nivel de expresión de proteínas secretadas endógenas particulares.

Definiciones

El término "interrumpir" como se usa en el presente documento se refiere a la reducción significativa hasta la eliminación completa de la expresión de una o más proteínas endógenas en una célula CHO, tal como atenuación y eliminación. Esto puede medirse como la presencia de esta o más proteínas endógenas en un medio de cultivo de la célula CHO, tal como por espectrometría de masas en la que el contenido total de una proteína endógena puede ser inferior a 500 ppm como máximo, tal como, como máximo 100 ppm, tal como, como máximo 10 ppm, por ejemplo, como máximo 1 ppm o como máximo 0,1 ppm, o no detectable.

El término "atenuación", "reducción" o "agotamiento" en el contexto de la expresión génica como se usa en el presente documento se refiere a enfoques experimentales que conducen a una expresión reducida de un gen dado en comparación con la expresión en una célula de control. La atenuación de un gen se puede lograr por varios medios experimentales, tal como la introducción de moléculas de ácido nucleico en la célula que se hibridan con partes del ARNm del gen y conducen a su degradación (por ejemplo, ARNhc, ARNi, miARN) o alteración de la secuencia del gen de una manera que conduce a una transcripción reducida, estabilidad del ARNm reducida o disminución de la traducción del ARNm.

Una inhibición completa de la expresión de un gen dado se denomina "inactivación". La inactivación de un gen significa que no se sintetizan transcritos funcionales a partir de dicho gen, lo que conduce a una pérdida de la función normalmente proporcionada por este gen. La inactivación del gen se logra alterando la secuencia de ADN que conduce a la interrupción o eliminación del gen o sus secuencias reguladoras. Las tecnologías de eliminación incluyen el uso de técnicas de recombinación homóloga para reemplazar, interrumpir o eliminar partes cruciales o toda la secuencia del gen o el uso de enzimas que modifican el ADN, como las nucleasas de dedos de zinc, para introducir roturas de doble cadena en el ADN del gen objetivo.

Un método específico para eliminar un gen específico de acuerdo con la invención son los métodos CRISPR-Cas9 como se describe en, por ejemplo, Ronda, Carlotta; Pedersen, Lasse Ebdrup; Hansen, Henning Gram; Kallehauge,

Thomas Beuchert; Betenbaugh, Michael J; Nielsen, Alex Toftgaard; Kildegaard, Helene Fastrup; Accelerating genome editing in CHO cells using CRISPR Cas9 and CRISPy, a web-based target finding tool, *Biotechnology and bioengineering*, 111,8,1604-1616,2014, o en Grav, Lise Marie; Lee, Jae Seong; Gerling, Signe; Kallehauge, Thomas Beuchert; Hansen, Anders Holmgaard; Kol, Stefan; Lee, Gyun Min; Pedersen, Lasse Ebdrup; Kildegaard, Helene Fastrup; One-step generation of triple knockout CHO cell lines using CRISPR/Cas9 and fluorescent enrichment, *Biotechnology journal*, 2015.

La expresión "línea celular huésped eucariota", como se usa en el presente documento, se refiere a cualquier línea celular adecuada de origen eucariótico que pueda usarse para la expresión de proteínas recombinantes. Los ejemplos no limitantes de líneas de células huésped adecuadas incluyen células CHO, células BKH, células HKBI I, células SP2/0 y células NSO, células L-M (ATCC CCL 1), células HEK293 (ATCC CCL 1555), HEK293 (ATCC CRL 1573; Graham *et al.*, J. Gen. Virol. 36: 59-72, 1977), células CV-1 (ATCC CCL 70), células COS-1 (ATCC CRL 1650), células COS-7 (ATCC CRL 1651), células 3T3 (ATCC CCL 92), células NIH/3T3 (ATCC CRL 1658), células HeLa (ATCC CCL 2), células C1271 (ATCC CRL 1616), células BS-C-1 (ATCC CCL 26) y células MRC-5 (ATCC CCL 171). Una célula adecuada en particular es una célula CHO (por ejemplo, ATCC CCL 61), tal como una seleccionada entre células CHO-K1, CHO-S y CHO DG44.

La expresión "proteína de interés", como se usa en el presente documento, se refiere a cualquier proteína que pueda producirse por medios recombinantes, normalmente una proteína que requiere expresión en una célula huésped eucariota debido a modificaciones posteriores a la traducción, incluida la glicosilación. Esto incluye proteínas de origen humano y animal, pero también proteínas de otras fuentes, tales como plantas, insectos, etc., y proteínas mutadas, artificiales, sintéticas, de fusión o quiméricas. En particular, "proteína de interés" incluye proteínas plasmáticas, hormonas peptídicas, factores de crecimiento, citocinas y anticuerpos. Con más detalle, las proteínas plasmáticas incluyen factores de coagulación de la sangre humanos y animales, tales como fibrinógeno, protrombina, trombina, FX, FXa, FIX, FIXa, FVII, FVIIa, FVIII, FVIIa, FXI, FXIa, FXII, FXIIa, FXIII, FXIIIa, factor de von Willebrand, etc., proteínas transportadoras, tales como albúmina, transferrina, ceruloplasmina, haptoglobina, hemoglobina, hemopexina, etc., inhibidores de la proteasa, tal como β -antitrombina, α -antitrombina, α 2-macroglobulina, inhibidor de CI, inhibidor de la vía del factor tisular (TFPI), cofactor II de la heparina, inhibidor de la proteína C (PAI-3), Proteína C, proteína S, etc., proteínas antiangiogénicas, tal como antitrombina latente, etc., proteínas altamente glicosiladas, incluida la glicoproteína ácida alfa-1, antitripsina, inhibidor de la inter- α -tripsina, glucoproteína α -2-HS, Proteína C-reactiva, etc. y otras proteínas, tales como glicoproteína rica en histidina, lectina de unión a manano, proteína de unión a C4, fibronectina, GC-globulina, plasminógeno, factores sanguíneos, tales como la eritropoyetina, interferón, factores tumorales, tPA, gCSF y derivados y muteínas de las mismas. El anticuerpo puede ser cualquier anticuerpo terapéutico seleccionado entre el grupo que consiste en abagovomab, adalimumab, alemtuzumab, aurograb, bapineuzumab, basiliximab, belimumab, bevacizumab, briakinumab, canakinumab, catumaxomab, certolizumab pegol, cetuximab, daclizumab, denosumab, efalizumab, galiximab, gemtuzumab, ozogamicina, golimumab, ibritumomab tiuxetán, infliximab, ipilimumab, lumiliximab, mepolizumab, motavizumab, muromonab, mycograb, natalizumab, nimotuzumab, ocrelizumab, ofatumumab, omalizumab, palivizumab, panitumumab, pertuzumab, ranibizumab, reslizumab, rituximab, teplizumab, tocilizumab/atlizumab, tositumomab, trastuzumab, Proxinium™, Rencarex™, ustekinumab y zalutumumab.

"TIMP1" o TIMP-1 se refiere a los inhibidores tisulares de metaloproteasas de tipo 1 derivados endógenamente de la célula huésped, tal como TIMP-1 de hámster chino de una célula CHO.

"LGALS3BP" se refiere a lectina, proteína soluble 3 de unión a galactósido derivada endógenamente de la célula huésped, tal como LGALS3BP de hámster chino de una célula CHO.

NID1 o NID1-1 como se usa en el presente documento se refiere a Nidogen 1. En el genoma de referencia de CHO K1, este gen existe en 2 isoformas, de ahí NID1-1 y NID1-2.

"BGN" se refiere al gen de biglicano.

"CTSD" se refiere a una proteína de cathepsina D endógenamente derivada de la célula huésped, tal como CTSD de hámster chino de una célula CHO.

Identificadores de Timp1 de NCBI:

ID del gen: 100770193; ARNm: XM_007630118.1, XM_007630119.1, JP061669.1, JI873443.1
Proteína: XP_003512290.1, XP_007628308.1, XP_007628309.1, EGW14768.1, ERE65804.1

Identificadores de LGALS3BP de NCBI:

ID del gen: 100767541; ARNm: XM_003498817.2, XM_007622088.1
Proteína: XP_007620278.1, XP_003498865.1, EGV97878.1

Identificadores de NID1 de NCBI:

ID del gen: 100766031; ARNm: XM_003509885.2, XM_003507587.2, XM_007632645.1, KE673075.1
Proteínas: ERE78436.1, XP_007630835.1, XP_003507635.2, XP_003509933.2

5 Identificadores BGN del NCBI:

ID del gen: 100771022; ARNm: XM_003506540.2, XM_007628144.1, KE685970.1:158610-162415 (+),
JH000676.1:731964-735640 (-)
Proteína: XP_003506588.1, XP_007626334.1, ERE63792.1, EGV92105.1

10 Identificadores CTSD del NCBI:

ID del gen: 100766628; ARNm, XM_003510234.2, XM_007611606.1, JI884212.1, JP051918.1,
JH001272.1:176080-187386 (-), KE672442.1:201083-212389 (+)
15 Proteína, EGW04839.1, ERE79325.1, XP_007609796.1, XP_003510282.1

Realizaciones específicas de la invención

20 La presente invención se refiere a una célula CHO modificada para interrumpir la expresión de una o más proteínas endógenas seleccionadas entre el grupo que consiste en TIMP1, NID1-1, NID1-2 y LGALS3BP.

En algunas realizaciones, la célula CHO se modifica para interrumpir la expresión de TIMP1.

25 En algunas realizaciones, la célula CHO se modifica para interrumpir la expresión de LGALS3BP.

En algunas realizaciones, la célula CHO se modifica para interrumpir la expresión de NID1-1, NID1-2 o ambos.

30 En algunas realizaciones, la célula CHO se modifica para interrumpir la expresión de las proteínas endógenas TIMP1, NID1-1, NID1-2 y LGALS3BP.

En algunas realizaciones, la célula CHO se modifica para interrumpir la expresión de

- 35 (a) al menos 5 proteínas endógenas;
(b) como máximo 50 proteínas endógenas; o
(c) entre 5 y 50 proteínas endógenas.

En algunas realizaciones, la célula CHO se modifica para interrumpir la expresión de TIMP1, LGALS3BP, NID1-1 y NID1-2.

40 En algunas realizaciones, la expresión de al menos una proteína endógena se ve interrumpida por la atenuación del gen que codifica la proteína endógena.

45 En algunas realizaciones, la expresión de al menos una proteína endógena se ve interrumpida por la inactivación del gen que codifica la proteína endógena.

En algunas realizaciones, la célula CHO de la invención comprende una secuencia de ácido nucleico que codifica una proteína recombinante de interés, estando la secuencia de ácido nucleico extracromosómica o cromosómicamente integrada y bajo el control de un promotor inducible o de un promotor constitutivo.

50 En algunas realizaciones, la célula de la invención produce niveles más altos de dicha proteína recombinante de interés y/o tiene una mayor densidad celular en comparación con una célula de control relevante sin expresión interrumpida de una o más proteínas endógenas, tal como una célula CHO de tipo salvaje.

55 En algunas realizaciones, la célula CHO según la invención se selecciona entre células CHO-K1, CHO-S y CHO DG44.

De acuerdo con las células y métodos de la invención, la expresión génica de ciertos genes está "interrumpida" por al menos una etapa seleccionada entre el grupo que consiste en silenciamiento génico, atenuación génica, inactivación génica, liberación de una construcción negativa dominante, inactivación génica condicional y/o mediante alteración génica con respecto a un gen específico.

60 La expresión "expresión génica", tal como se usa en el presente documento, está destinado a abarcar al menos una etapa seleccionada entre el grupo que consiste en transcripción de ADN en ARNm, procesamiento de ARNm, maduración de ARNm no codificante, exportación de ARNm, traducción, plegamiento de proteínas y/o transporte de proteínas.

65 La inhibición o reducción de la expresión génica de un gen se refiere a métodos que interfieren directamente con la

expresión génica, que abarcan, pero sin limitaciones, inhibición o reducción de la transcripción de DNA, por ejemplo, mediante el uso de represores específicos relacionados con el promotor, mediante mutagénesis específica del sitio de un promotor dado, por cambio de promotor o inhibición o reducción de la traducción, por ejemplo, mediante silenciamiento génico postranscripcional inducido por ARNi. La expresión de un producto génico disfuncional o inactivo con actividad reducida, puede, por ejemplo, lograrse mediante mutagénesis aleatoria o específica del sitio, inserciones o deleciones dentro del gen codificante.

La inhibición o reducción de la actividad del producto génico puede, por ejemplo, lograrse mediante la administración de, o la incubación con, un inhibidor de la enzima respectiva, antes o simultáneamente con la expresión de proteínas. Los ejemplos para tales inhibidores incluyen, pero sin limitación, un péptido inhibidor, un anticuerpo, un aptámero, una proteína de fusión o un anticuerpo mimético contra dicha enzima, o un ligando o receptor del mismo, o un péptido inhibidor o ácido nucleico, o una molécula pequeña con actividad de unión similar.

Otras formas de inhibir la enzima son la reducción de cofactores específicos de la enzima en el medio, como cobre, que es un cofactor de iones específicos de PAM (por ejemplo, en forma de CuSO_4), ascorbato, que actúa como donante de electrones para PAM, oxígeno molecular, catalasa y otras conocidas hoy en día por el experto en la materia, o aún por descubrir en el futuro.

El silenciamiento génico, la atenuación génica y la inactivación génica se refieren a técnicas mediante las cuales se reduce la expresión de un gen, ya sea a través de modificación genética o por tratamiento con un oligonucleótido con una secuencia complementaria a un transcrito de ARNm o un gen. Si se realiza una modificación genética del ADN, el resultado es un organismo atenuado o inactivado. Si el cambio en la expresión génica es causado por un oligonucleótido que se une a un ARNm o se une temporalmente a un gen, esto da como resultado un cambio temporal en la expresión génica sin modificación del ADN cromosómico y se denomina atenuación transitoria.

En una atenuación transitoria, que también está comprendida en el término anterior, la unión de este oligonucleótido al gen activo o a sus transcritos provoca una disminución de la expresión mediante el bloqueo de la transcripción (en el caso de la unión de genes), degradación de la transcripción de ARNm (por ejemplo, por ARN de interferencia pequeño (ARNip) o antisentido dependiente de ARNasa-H) o bloqueo de la traducción de ARNm, sitios de corte y empalme de pre-ARNm o sitios de escisión de nucleasas utilizados para la maduración de otros ARN funcionales como miARN (por ejemplo, por oligos de morfolino u otro antisentido independiente de ARNasa-H). Otros enfoques implican el uso de ARNhc (ARN de horquilla pequeña, que es una secuencia de ARN que hace un giro de horquilla cerrado que se puede usar para silenciar la expresión génica a través de la interferencia de ARN), eARNip (ARNip preparados con endorribonucleasa, que son una mezcla de oligos de ARNip resultantes de la escisión de ARN de doble cadena largo (ARNbc) con una endorribonucleasa), o la activación del complejo de silenciamiento inducido por ARN (RISC).

Otros enfoques para llevar a cabo el silenciamiento génico, los expertos en la técnica conocen la atenuación e inactivación por la bibliografía respectiva, y su aplicación en el contexto de la presente invención se considera rutinaria.

La inactivación génica se refiere a las técnicas mediante las cuales la expresión de un gen se bloquea por completo, es decir, el gen respectivo no funciona o incluso se ha eliminado. Los enfoques metodológicos para lograr este objetivo son múltiples y conocidos por el experto en la materia. Los ejemplos son la producción de un mutante que es predominantemente negativo para el gen dado. Dicho mutante se puede producir mediante mutagénesis dirigida al sitio (por ejemplo, delección, delección parcial, inserción o sustitución de ácido nucleico), mediante el uso de transposones adecuados o mediante otros enfoques que son conocidos por el experto en la literatura respectiva, cuya aplicación en el contexto de la presente invención se considera, por tanto, rutinaria. Un ejemplo de una técnica recientemente desarrollada que el experto consideraría útil en el contexto de la presente invención es la desactivación mediante el uso de nucleasas con dedos de cinc dirigidas. Sigma Aldrich proporciona un kit respectivo como "CompoZR knockout ZFN". Otro enfoque abarca el uso de nucleasas efectoras similares a activadores de transcripción (TALEN).

La administración de una construcción negativa dominante implica la introducción de una secuencia que codifica una enzima disfuncional, por ejemplo, mediante transfección. Dicha secuencia codificante está acoplada funcionalmente a un promotor fuerte, de tal manera que la expresión génica de la enzima disfuncional anula la expresión natural de la enzima de tipo salvaje, que, a su vez, conduce a un defecto fisiológico efectivo de la actividad enzimática respectiva.

Una inactivación génica condicional permite bloquear la expresión génica de una manera específica de tejido o de tiempo. Esto se realiza, por ejemplo, introduciendo secuencias cortas llamadas sitios loxP alrededor del gen de interés. De nuevo, otros enfoques son conocidos por el experto en la literatura respectiva y su aplicación en el contexto de la presente invención se considera rutinaria.

Otro enfoque es la alteración génica que puede conducir a un producto génico disfuncional o a un producto génico con actividad reducida. Este enfoque implica la introducción de mutaciones de cambio de marco, mutaciones sin sentido (es decir, introducción de un codón de terminación prematuro) o mutaciones que conducen a una sustitución de aminoácidos que hace que todo el producto génica sea disfuncional o provoque una actividad reducida. Tal alteración génica puede ser producida, por ejemplo, por mutagénesis (por ejemplo, delección, delección parcial, inserción

o sustitución de ácido nucleico), ya sea mutagénesis inespecífica (aleatoria) o mutagénesis dirigida al sitio.

Protocolos que describen la aplicación práctica del silenciamiento génico, atenuación génica, inactivación génica, liberación de una construcción negativa dominante, inactivación condicional del gen y/o la alteración del gen están comúnmente disponibles para el experto en la materia y están dentro de su rutina. La enseñanza técnica proporcionada en el presente documento está totalmente habilitada con respecto a todos los métodos imaginables que conducen a una inhibición o reducción de la expresión génica de un producto génico, o a la expresión de un producto génico disfuncional o inactivo, o con actividad reducida.

10 Preparación de proteína de interés

La invención también se refiere a un método para preparar una proteína de interés como se ha mencionado anteriormente. La proteína de interés descrita en el presente documento se puede producir mediante técnicas de ácido nucleico recombinante. En general, una secuencia de ácido nucleico clonada se modifica para codificar la proteína deseada. Esta secuencia modificada se inserta luego en un vector de expresión, que a su vez se transforma o transfecta en células huésped. Las células eucariotas superiores, en particular células de mamífero cultivadas, se prefieren como células huésped. Las secuencias completas de nucleótidos y aminoácidos para seres humanos

Las alteraciones opcionales de la secuencia de aminoácidos pueden lograrse mediante diversas técnicas. La modificación de la secuencia de ácido nucleico puede ser por mutagénesis específica del sitio. Las técnicas para la mutagénesis específica del sitio son bien conocidas en la técnica y se describen en, por ejemplo, Zoller y Smith (DNA 3: 479-488, 1984) o "Splicing by extension overlap", Horton *et al.*, Gene 77, 1989, pp. 61-68. Por lo tanto, utilizando las secuencias de nucleótidos y aminoácidos de la proteína de interés, se puede(n) introducir la(s) alteración(es) de elección. Análogamente, los procedimientos para preparar una construcción de ADN mediante la reacción en cadena de la polimerasa utilizando cebadores específicos son bien conocidos por los expertos en la materia (véase PCR Protocols, 1990, Academic Press, San Diego, California, EE.UU.).

La construcción de ácido nucleico que codifica la proteína de interés puede ser adecuadamente de origen genómico o de ADNc, por ejemplo, obtenida mediante la preparación de una biblioteca genómica o de ADNc y la detección de secuencias de ADN que codifican todo o parte del polipéptido mediante hibridación utilizando sondas de oligonucleótidos sintéticos de acuerdo con técnicas estándar (véase Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª Ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York, 1989).

La construcción de ácido nucleico que codifica la proteína de interés también se puede preparar sintéticamente mediante métodos estándar establecidos, por ejemplo, el método de la fosfoamidita descrito por Beaucage y Caruthers, Tetrahedron Letters 22 (1981), 1859-1869, o el método descrito por Matthes *et al.*, EMBO Journal 3 (1984), 801-805. Según el método de la fosfoamidita, se sintetizan oligonucleótidos, por ejemplo, en un sintetizador automático de ADN, se purifican, se hibridan, se ligan y se clonan en vectores adecuados. La proteína de interés también se puede preparar por reacción en cadena de la polimerasa utilizando cebadores específicos, por ejemplo, como se describe en el documento US 4.683.202, Saiki *et al.*, Science 239 (1988), 487-491 o Sambrook *et al.*, citado anteriormente.

Adicionalmente, la construcción de ácido nucleico puede ser de mezcla sintética y genómica, ADNc y sintético mixto o de origen genómico y ADNc mixto preparado ligando fragmentos de ADN sintético, origen genómico o ADNc (según corresponda), los fragmentos correspondientes a varias partes de la construcción completa de ácido nucleico, de acuerdo con las técnicas convencionales.

La construcción de ácido nucleico es preferiblemente una construcción de ADN. Las secuencias de ADN para su uso en la producción de la proteína de interés normalmente codificarán un polipéptido pre-pro en el extremo amino terminal de la proteína de interés para obtener un procesamiento postraducciona adecuado (por ejemplo, carboxilación gamma de residuos de ácido glutámico) y secreción de la célula huésped.

Las secuencias de ADN que codifican la proteína de interés generalmente se insertan en un vector recombinante que puede ser cualquier vector, que puede someterse convenientemente a procedimientos de ADN recombinante, y la elección del vector dependerá a menudo de la célula huésped en la que se va a introducir. Por lo tanto, el vector puede ser un vector de replicación autónoma, es decir, un vector, que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo, un plásmido. Como alternativa, el vector puede ser uno que, cuando se introduce en la célula hospedadora, se integra en el genoma de la célula hospedadora y se replique junto con el cromosoma o cromosomas en los que se ha integrado.

El vector es, preferentemente, un vector de expresión en el que la secuencia de ADN que codifica la proteína de interés se une de manera operativa a segmentos adicionales necesarios para la transcripción del ADN. En general, el vector de expresión deriva de plásmido o ADN viral, o puede contener elementos de ambos. La expresión "unido de manera operativa" indica que los segmentos están dispuestos de modo que funcionan en conjunto para sus propósitos previstos, por ejemplo, la transcripción se inicia en un promotor y continúa a través de la secuencia de ADN que codifica el polipéptido.

Los vectores de expresión para su uso en la expresión de una proteína de interés comprenderán un promotor capaz de dirigir la transcripción de un gen o ADNc clonado. El promotor puede ser cualquier secuencia de ADN, que muestra actividad transcripcional en la célula huésped de elección y puede derivar de genes que codifican proteínas homólogas o heterólogas a la célula huésped.

Ejemplos de promotores adecuados para dirigir la transcripción del ADN que codifica una proteína de interés en células de mamíferos son el promotor SV40 (Subramani *et al.*, Mol. Cell Biol. 1 (1981), 854-864), el promotor MT-1 (gen de la metalotioneína) (Palmiter *et al.*, Science 222 (1983), 809-814), el promotor de CMV (Boshart *et al.*, Cell 41: 521-530, 1985) o el principal promotor tardío del adenovirus 2 (Kaufman y Sharp, Mol. Cell. Biol. 2: 1304-1319, 1982).

Las secuencias de ADN que codifican la proteína de interés también pueden, en caso necesario, estar conectadas de manera operativa a un terminador adecuado, tal como el terminador de la hormona del crecimiento humano (Palmiter *et al.*, Science 222, 1983, pp. 809-814) o los terminadores TPI 1 (Alber y Kawasaki, J. Mol. Appl. Gen. 1, 1982, pp. 419-434) o ADH3 (McKnight *et al.*, The EMBO J. 4, 1985, pp. 2093-2099). Los vectores de expresión también pueden contener un conjunto de sitios de empalme de ARN ubicados aguas abajo del promotor y aguas arriba del sitio de inserción para la propia secuencia que codifica la proteína de interés. Los sitios de corte y empalme de ARN preferidos pueden obtenerse a partir de genes de adenovirus y/o inmunoglobulina. También contenida en los vectores de expresión hay una señal de poliadenilación ubicada cadena abajo del sitio de inserción. Las señales de poliadenilación particularmente preferidas incluyen la señal de poliadenilación temprana o tardía de SV40 (Kaufman y Sharp, *ibid.*), la señal de poliadenilación de la región 5' del adenovirus, el terminador del gen de la hormona del crecimiento humano (DeNoto *et al.* Nucl. Acids Res. 9: 3719-3730, 1981) o la señal de poliadenilación del gen que codifica la proteína de interés. Los vectores de expresión también pueden incluir una secuencia líder viral no codificante, como el líder de adenovirus tripartito 2, ubicado entre el promotor y los sitios de corte y empalme de ARN; y secuencias potenciadoras, tal como el potenciador SV40.

Para dirigir la proteína de interés hacia la vía secretora de las células huésped, puede proporcionarse una secuencia señal secretora (también conocida como secuencia líder, preprosecuencia o presecuencia) en el vector recombinante. La secuencia señal secretora se une a las secuencias de ADN que codifican la proteína en el marco de lectura correcto. Las secuencias señal secretoras se colocan comúnmente en 5' con respecto a la secuencia de ADN que codifica el péptido. La secuencia señal secretora puede ser que normalmente esté asociada a la proteína o puede ser de un gen que codifica otra proteína secretada.

Los procedimientos utilizados para ligar las secuencias de ADN que codifican la proteína de interés, el promotor y, opcionalmente, el terminador y/o la secuencia señal secretora, respectivamente, e insertarlos en vectores adecuados que contengan la información necesaria para la replicación, son bien conocidas por los expertos en la técnica (véase, por ejemplo, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, Nueva York, 1989).

Los métodos para transfectar células de mamífero y expresar secuencias de ADN introducidas en las células se describen, por ejemplo, en Kaufman y Sharp, J. Mol. Biol. 159 (1982), 601-621; Southern y Berg, J. Mol. Appl. Genet. 1 (1982), 327-341; Loyer *et al.*, Proc. Natl. Acad. Sci. USA 79 (1982), 422-426; Wigler *et al.*, Cell 14 (1978), 725; Corsaro y Pearson, Somatic Cell Genetics 7 (1981), 603; Graham y van der Eb, Virology 52 (1973), 456; y Neumann *et al.*, EMBO J. 1 (1982), 841-845. Las secuencias de ADN clonadas se introducen en células de mamífero cultivadas mediante, por ejemplo, transfección mediada por fosfato de calcio (Wigler *et al.*, Cell 14: 725-732, 1978; Corsaro y Pearson, Somatic Cell Genetics 7: 603-616, 1981; Graham y Van der Eb, Virology 52d: 456-467, 1973) o electroporación (Neumann *et al.*, EMBO J. 1: 841-845, 1982).

Para identificar y seleccionar células que expresan el ADN exógeno, un gen que confiere un fenotipo seleccionable (un marcador seleccionable) generalmente se introduce en las células junto con el gen o el ADNc de interés. Los marcadores seleccionables preferidos incluyen genes que confieren resistencia a fármacos, tales como neomicina, higromicina y metotrexato. El marcador seleccionable puede ser un marcador seleccionable amplificable. Un marcador seleccionable amplificable preferido es una secuencia de dihidrofolato reductasa (DHFR). Los marcadores seleccionables son revisados por Thilly (Mammalian Cell Technology, Butterworth Publishers, Stoneham, MA). El experto en la materia podrá elegir fácilmente los marcadores seleccionables adecuados.

Los marcadores seleccionables pueden introducirse en la célula en un plásmido separado al mismo tiempo que el gen de interés, o pueden introducirse en el mismo plásmido. Si, en el mismo plásmido, el marcador seleccionable y el gen de interés pueden estar bajo el control de diferentes promotores o del mismo promotor, esta última disposición produce un mensaje dicistrónico. Las construcciones de este tipo son conocidas en la técnica (por ejemplo, Levinson y Simonsen, documento U.S. 4.713.339). También puede ser ventajoso añadir ADN adicional, conocido como "ADN vehículo," a la mezcla que se introduce en las células.

Una vez que las células han absorbido el ADN, se cultivan en un medio de crecimiento apropiado, normalmente 1-2 días, para comenzar a expresar el gen de interés. Como se usa en el presente documento, el término "medio de crecimiento apropiado" significa un medio que contiene nutrientes y otros componentes necesarios para el crecimiento de las células y la expresión de la proteína humana de interés. Los medios generalmente incluyen una fuente de

carbono, una fuente de nitrógeno, aminoácidos esenciales, azúcares esenciales, vitaminas, sales, fosfolípidos, proteínas y factores de crecimiento. Para la producción de proteínas gamma-carboxiladas, el medio contendrá vitamina K, preferentemente a una concentración de aproximadamente 0,1 lg/ml a aproximadamente 5 Lg/ml. A continuación, se aplica la selección de fármacos para seleccionar el crecimiento de células que expresan el marcador seleccionable de forma estable. Para las células que han sido transfectadas con un marcador seleccionable amplificable, la concentración del fármaco puede aumentarse para seleccionar un mayor número de copias de las secuencias clonadas, aumentando así los niveles de expresión. A continuación, los clones de células transfectadas de forma estable se analizan para detectar la expresión de la proteína de interés.

La célula huésped en la que se introduce la secuencia de ADN que codifica la proteína de interés puede ser cualquier célula, tal como una que sea capaz de producir las proteínas modificadas postraduccionales, tales como células eucariotas superiores.

Ejemplo 1

Identificación de las 6 principales proteínas secretadas en nuestra línea celular CHO

Los primeros datos de espectroscopia de masas de proteínas de los medios extracelulares de varios cultivos de células CHO se analizaron mediante una serie de procesos implementados y ejecutados en python (<https://www.python.org/>): Los datos relevantes se extrajeron de los archivos de salida de datos del software de especificaciones de masas, por ejemplo, identificadores cortos y largos, 3 principales sumas de intensidad de péptido coincidentes e identificadores para posibles homólogos.

Los homólogos se colapsaron por suma y las sumas de intensidad se normalizaron convirtiéndolas en porcentaje de sumas de intensidad de péptido total.

Se añadió información externa sobre las condiciones de ejecución a los datos.

A continuación, se seleccionaron las 20 proteínas principales en función del criterio de que estaban presentes en al menos la mitad de nuestras muestras y tenían los 20 valores de suma de intensidad normalizada media más altos.

Luego, estas 20 proteínas se analizaron más a fondo mediante ontología génica para excluir proteínas por varias razones:

La ontología génica, compilada por, por ejemplo, el Gene Ontology Consortium (GO, <http://geneontology.org/>), se puede utilizar para inferir el origen de los productos génicos. Los medios por los que las proteínas alcanzan los medios de un cultivo celular pueden ser, por ejemplo, secreción controlada o lisis celular. Los inventores estaban interesados principalmente en las proteínas que habían llegado a los medios por secreción y, específicamente, no nos interesaban las proteínas que habían llegado por lisis celular. Si, por ontología génica, se observó que una proteína residía principalmente intracelularmente, se supuso que su presencia en los medios no se debía a una secreción organizada, sino que es poco probable que la lisis celular y, en consecuencia, la eliminación o eliminación de células libere la capacidad secretora.

Se asignaron términos de ontología génica a nuestras 20 proteínas principales mediante el uso de la base de datos de homología entre especies GO (GO CHO, <https://ebdrupor.ejemplo.biosustain.dtu.dk/gocho/>) que extrae términos para un producto génico buscando un producto génico con un nombre similar en una selección de organismos relacionados bajo el supuesto de que la función del gen y, por lo tanto, los términos GO, se conservan en gran medida entre especies estrechamente relacionadas.

Usando la interfaz GO CHO, a continuación se excluyó las proteínas de la lista de las 20 principales que no se sabe que estén presentes en los medios extracelulares.

Se conoce a partir de la base de datos GeneOntology. Esta es una base de datos de anotación manual (literatura) y electrónica (bibliografía y varios métodos) para productos genéticos.

Además, se excluyó a las proteínas de la lista de las 20 principales, que se sabe que son esenciales o importantes para el funcionamiento de la célula.

Estos procesos finalmente dieron como resultado la identificación de 6 proteínas secretadas. (TIMP1, BGN, CTSD; NID1-1, NID1-2 y LGALS3BP).

El ARNg seleccionado se sintetizó como oligos sintéticos, se hibridó para construir ADNbc y se clonó en un vector de expresión utilizando la clonación de USER (New England Biolabs) de acuerdo con el protocolo del fabricante.

Primero, los 6 genes se eliminaron individualmente en células CHO-S. A continuación, se inactivaron 5 genes en la línea celular defectiva que ya contenía TIMP1 para crear líneas celulares que contenían secuencias codificantes interrumpidas para los 6 genes.

Los oligos hibridados se clonaron en vectores de expresión utilizando la clonación del USER, según el protocolo de los fabricantes (New England Biolabs).

- 5 Los plásmidos con los oligos de ARNg clonados se extrajeron y transfectaron con un plásmido que codifica Cas9 marcado con GFP (Ronda C, Pedersen LE, et al Biotechnol Bioeng. 2014 Aug;111(8):1604-16. y Grav LM, et al, Biotechnol J. 2015 Sep;10(9):1446-56.) en células CHO-S de Thermo, se cultivaron en medios CD-CHO con glutamina 8 mM, usando el agente de transfección FreeStyle (Thermo) de acuerdo con las recomendaciones del fabricante.

10 Ejemplo 2. Preparación de células CHO KO:

Las construcciones de expresión de ARNsg se diseñaron fusionando tracrARN y ARNcr en un ARNsg quimérico (Jinek). *et al.*, 2012) y se localizaron inmediatamente aguas abajo de un promotor U6 (Chang *et al.*, 2013). Inicialmente, el casete de expresión de ARNsg se sintetizó como gBlock (Integrated DNA Technologies) y se subclonó en el vector pRSFDuet-1 (Novagen, Merck) usando los sitios de restricción KpnI y HindIII. Este vector de expresión pRSFDuet-1/ARNsg se usó como columna vertebral en un método de clonación de reactivo de escisión específico de uracilo (USER) basado en PCR. Este método fue diseñado para cambiar fácil y rápidamente la región variable de 19 pb de longitud (N19) del ARNsg para generar nuestras construcciones de ARNsg. Del vector de expresión pRSFDuet-1/ARNsg, se generó un amplicón de 4221 pb de longitud (estructura del vector de expresión mediante PCR (1x: 98°C durante 2 min; 30x: 98°C durante 10 s, 57°C durante 30 s, 72°C durante 4 min 12 s; 1x: 72°C durante 5 min) utilizando dos cebadores que contienen uracilo (ARNsg Estructura dir y ARNsg Estructura inv, Integrated DNA Technologies, Tabla 2) y la ADN polimerasa X7. Después del tratamiento enzimático FastDigest DpnI (Thermo Fisher Scientific), el amplicón se purificó a partir de un gel de TBE de agarosa al 2 % utilizando el kit de extracción de gel QIAEX II (Qiagen). En paralelo, se sintetizaron oligos monocatenarios de 54 pb de longitud y 53 pb de longitud, (cadena sentido y antisentido, respectivamente) que comprende la región variable del ARNsg, (TAG Copenhagen, Tabla 2). La hibridación de los oligos monocatenarios sentido y antisentido (100 pM) se realizó con enfriamiento lento de 95 °C a 25 °C durante 15 min en un termociclador en tampón 4 de New England Biolabs. Luego, los oligos hibridados se mezclaron con el vector de expresión purificado en gel y se trataron con la enzima USER (New England Biolabs) de acuerdo con las recomendaciones del fabricante. Después del tratamiento enzimático USER, la mezcla de reacción se transformó en células competentes de *E. coli* Mach1 (Life Technologies) según procedimientos estándar. Los clones transformantes se seleccionaron en placas LB de kanamicina de 50 µg/ml (Sigma-Aldrich). Todas las construcciones se verificaron mediante secuenciación y se purificaron con NucleoBond Xtra Midi EF (Macherey-Nagel) de acuerdo con las pautas del fabricante.

35 El vector de expresión Cas9 optimizado por codón de CHO aplicado en el estudio se describe en

(Ronda C, Pedersen LE, et al Biotechnol Bioeng. 2014 Aug;111(8):1604-16. y Grav LM, et al, Biotechnol J. 2015 Sep;10(9):1446-56.). La herramienta bioinformática CRISPy (<http://staff.biosustain.dtu.dk/laeb/crispy/>) se aplicó para generar ARNsg dirigidos a los 6 genes elegidos para KO. El ARNsg dirigido a los 6 genes se describe en la tabla 1.

40

Tabla 1 Nombres y secuencias de oligos.

ID	Seq (5'-3')
gRNA_BGN_1140221_fwd	GGAAAGGACGAAACACCGCGGTCA GTGACGCAGCGGAGTTTTAGAGCTA GAAAT
gRNA_TIMP1_1673356_fwd	GGAAAGGACGAAACACCGCCTTCTGCAACTCCGACCTGTTTTAGAGCTA GAAAT
gRNA_CTSD_1486038_fWd	GGAAAGGACGAAACACCGCAAGTTCACGTCTATCCGTGTTTTAGAGCTA GAAAT
gRNA_LGALS3BP_390023_fwd	GGAAAGGACGAAACACCGATCTCCACGCGGCCCTCACGTTTTAGAGCTA GAAAT
gRNA_NID1- ma13382_1237189_fWd	GGAAAGGACGAAACACCGTGACTGAGCCTGATCCGGAGTTTTAGAGCTA GAAAT
gRNA_NID1- ma15820_1455331_fWd	GGAAAGGACGAAACACCGTTGCGGCCGAGATGGTCAAGTTTTAGAGCTA GAAAT
gRNA_BGN_1140221_rev	CTAAAACTCCGCTGCGTCACTGACCGCGGTGTTTCGTCCTTTCCACAAGA TAT

(continuación)

ID	Seq (5'-3')
gRNA_TIMP1_1673356_rev	CTAAAACAGGTCGGAGTTGCAGAAGGCGGTGTTTCGTCCTTTCCACAAG ATAT
gRNA_CTSD_1486038_rev	CTAAAACACGGATAGACGTGAACTTGCAGGTGTTTCGTCCTTTCCACAAGA TAT
gRNA_LGALS3BP_390023_rev	CTAAAACGTGAGGGCCGCGTGGAGATCGGTGTTTCGTCCTTTCCACAAG ATAT
gRNA_NID1- rna13382_1237189_rev	CTAAAACCTCCGGATCAGGCTCAGTCACGGTGTTCGTCCTTTCCACAAGA TAT
gRNA_NID1- rna15820_1455331_rev	CTAAAACCTTGACCATCTCGGCCGCAACGGTGTTCGTCCTTTCCACAAGA TAT
ID	Seq (5'-3')
MiSeq_BGN_1140221_fWd	CATGGTGGCTACTACTGCTTTT
MiSeq_TIMP1_1673356_fWd	TCACTGATAACCTCCAGCAAGG
MiSeq_CTSD_1486038_fwd	CTCAGCAGGTGACTGATTGGAG
MiSeq_LGALS3BP_390023_fW d	GCACAAAGCATCGAAGTCCTG
MiSeq_NID1- rna13382_1237189_fWd	TGGGTAGGGACTGCTACCG
MiSeq_NID1- rna15820_1455331_fWd	TGCCATCCATTCTTGCAACTTC
ID	Seq (5'-3')
MiSeq_BGN_1140221_rev	CAAGGTGGGCATCAACGACT
MiSeq_TIMP1_1673356_rev	ATGGCCAAAGTGGATCTATGCAA
MiSeq_CTSD_1486038_rev	AGCCTGGAGTATCCTTGGGTAG
MiSeq_LGALS3BP_390023_rev	GCATCTAAAATGTTCCAGAGGTTGT
MiSeq_NID1- rna13382_1237189_rev	TGGAGTCATCATGGACACACTC
MiSeq_NID1- rna15820_1455331_rev	TCCCATTGTTTCTCAAGGTTT

Las células CHO, por ejemplo, las células CHO-S de Life Technologies, se cultivaron en un medio apropiado, por ejemplo, Medio CD CHO (Life Technologies) suplementado con L-Glutamina 8 mM y se cultivaron en matraces de agitación. Las células se incubaron a 37 °C, CO₂ al 5 %, con agitación de 120 rpm y pases cada 2-3 días. La transfección se realizó con los vectores de expresión que codifican Cas9 optimizado para CHO y ARNsg dirigidos a los respectivos sitios KO, anteriormente descritos. Para cada muestra, se transfectaron 3 X 10⁶ células con un total de 3,75 pg de ADN. 16 horas después de la transfección, las muestras se incubaron a 30 °C durante 32 horas antes de volver a transferirlas a 37 °C. Las células transfectadas se clasificaron de forma individual en placas de 96 pocillos, utilizando clasificación de células activadas por fluorescencia (FACS) para cosechar células positivas para GFP y se cultivaron a 37 °C, 5 % de CO₂ durante 3 semanas. Se evitó el secado de los pocillos rellenando con medio nuevo según fuera necesario. Los clones en crecimiento se colapsaron en placas individuales de 96 pocillos y se tomó una muestra del clon en crecimiento para análisis genómico para identificar genes KO. El ADN genómico se prepara utilizando el kit Quick Extract (Epicentre) de acuerdo con las recomendaciones del fabricante. Los amplicones que cubrían los sitios objetivo se crearon utilizando cebadores miSeq específicos para el objetivo (tableX) usando polimerasa Phusion en un termociclador usando el siguiente programa: 1x: 95 °C durante 7 min; 20x: 95 °C durante 45 s, 69,5 °C durante 30 s (Δ -0,5 °C/ciclo), 72 °C durante 30 s; 1x: 72 °C durante 7 min.

Los amplicones resultantes se purifican utilizando AMPure XP Beads (Beckman Coulter) según las recomendaciones del fabricante.

- 5 Los amplicones purificados se indexan con el kit de indexación Nextera XT de Illumina con el siguiente protocolo:

Mezcla	x1
ADN	5,0µl
Nextera XT Index 1	2,5µl
Nextera XT Index 2	2,5µl
2x KAPA HiFi HotStart Ready Mix	12,5µl
H2O	2,5µL
	25µl

y se ejecutó el proceso en un termociclador usando el siguiente programa: 1x: 95 °C durante 3 min; 8x: 95 °C durante 30 s, 55 °C durante 30 s 72 °C durante 30 s; 1x: 72 °C durante 5 min

- 10 Las bibliotecas indexadas resultantes se purifican utilizando AMPure XP Beads (Beckman Coulter) de acuerdo con las recomendaciones del fabricante.

- 15 Las bibliotecas están normalizadas a 10 nM, agrupadas y ejecutadas en un secuenciador de ADN miSeq (Illumina) utilizando el siguiente protocolo:

Lleve la biblioteca de 10 nM a 4 nM combinando:

- 20 - 4 µl de biblioteca 10 nM
- 6 µl de tampón de hibridación1 (HT1) (Illumina)

Combinar juntos:

- 25 - 5 µl de biblioteca 4 nM
- 5 µl de NaOH 0,2 N

Agitar en vórtex brevemente y centrifugar 1 minuto a 280 g.

- 30 Incubar durante 5 minutos a temperatura ambiente para desnaturalizar el ADN en cadenas sencillas.

Añadir 990 µl de HT1 preenfriado (la biblioteca ahora está a 20 pM y desnaturalizada en NaOH 1 mM).

Se conservan en hielo de ahora en adelante.

- 35 Las bibliotecas desnaturalizadas se diluyen a 10 pM con HT1 preenfriado y se añaden 30 µl de Illumina PhiX Control (Illumina) desnaturalizado, se carga inmediatamente en un cartucho miSeq y se ejecuta en un secuenciador Illumina miSeq Desktop.

- 40 Después de MiSeq, los datos de la secuencia se analizan ejecutándolos a través de una serie de etapas ejecutadas con python. Las primeras secuencias emparejadas se fusionan en una sola secuencia. A continuación, las secuencias se verifican y recortan la calidad o se descartan si están por debajo del umbral. Luego, las secuencias se alinean con la versión de tipo salvaje de los sitios de destino. Luego, cada clon se identifica sobre la base de su par de índice único y se analiza con respecto a la situación genómica en cada uno de los sitios objetivo. Si está claro que todos los sitios de destino se han interrumpido de una manera que resulta en un cambio de marco y los datos son generalmente plausibles y limpios, el clon se guarda para su posterior propagación.

- 45 Todas las KO se generaron primero como KO individuales y, a partir de entonces, la KO de TIMP1 se utilizó como base para la KO de los otros 5 objetivos genéticos en un solo procedimiento, donde los 5 vectores ARNg se cotransfectaron con el vector Cas9, para generar una línea de células CHO con los 6 genes eliminados (6XSeckO), utilizando los protocolos descritos anteriormente.

Ejemplo 3: Caracterización de células CHO KO

Se analizó el crecimiento y la expresión de rituximab de las células CHO utilizando el siguiente protocolo:

- 55 Descongelar la célula del banco de celdas

Mantener y propagar células durante 3 pases

El Día 1: Lavar las células y el cultivo o/n sin agente antiaglomerante

El Día 2: Contar las células y ajuste a 1.000.000 de células/ml (3 matraces de agitación de 125 ml con 30 ml para cada línea celular), transfectar células con el vector de expresión Rituximab según el protocolo descrito anteriormente

El día 3, extraer una muestra de 1 ml de cada matraz para el recuento de células y la determinación de la eficiencia de la transfección y reemplazar con 1 ml de medio fresco (que contenga 60 ul de agente antiaglomerante por ml)

El día 4, tomar una muestra de 1 ml de cada matraz para el recuento de células y la determinación de la eficiencia de la transfección y reemplazar con 1 ml de medio fresco

El día 5: Extraer una muestra de 1 ml de cada matraz para el recuento de células y la determinación de la eficiencia de la transfección y reemplazar con 1 ml de medio fresco y cosechar los sobrenadantes para el análisis del título de Rituximab.

Todas las incubaciones se realizan a 37 °C, 5 %de CO2 a 120 rpm en una Incubadora Multitron (Infors).

Los títulos de rituximab se determinaron usando el sistema Octet RED96 de Pall de acuerdo con las recomendaciones del fabricante para la detección de IgG.

Las células se contaron utilizando un Nucleocounter NC-200 de Chemometec de acuerdo con las recomendaciones del fabricante para el recuento de células CHO

En la figura 1 se ve un efecto positivo de eliminar todos los genes excepto quizás BGN.

En la figura 2 se ve que las 6XSeckO alcanzan densidades celulares máximas significativamente más altas que las células CHO-S salvaje. El gen KO que parece ser el principal contribuyente a este fenotipo es la eliminación de los genes NID1-1.

Ejemplo 4 Cuantificación del contenido de proteína de la célula huésped (HCP) en el sobrenadante

Medios usados de CHO-S, 6XseckO y 11XseckO, se cosechó y los HCP se precipitaron con TCA utilizando el siguiente protocolo:

1. Añadir 1 volumen de stock de TCA a 4 volúmenes de muestra de proteína. (250 ul TCA + 1000 ul SN) es decir en tubo de 1,5 ml con volumen máximo, añadir 250 µl de TCA a 1,0 ml de muestra.

2. Incubar 10 min a 4 °C.

3. Girar el tubo en microcentrífuga a 14K rpm, 5 min.

4. Retirar el sobrenadante, dejando el sedimento de proteína intacto. El sedimento debe formarse a partir de ppt blanquecino esponjoso.

5. Lavar el sedimento con 200 µl de acetona fría.

6. Girar sintonizar en microcentrífuga a 14K rpm, 5 min.

7. Repetir las etapas 4 a 6 para un total de 2 lavados con acetona.

8. Secar el sedimento colocando el tubo en un bloque térmico a 95 °C durante 5-10 minutos para eliminar la acetona.

9. Disolver los sedimentos en 200 ul SDS al 1 %, Tris 50 mM, pH 7,5 a 95 C

El contenido de proteína del HCP redissuelto se determinó usando el ensayo Bradford Protein (Pierce) de acuerdo con las instrucciones del proveedor y usando absorción a 280 nm en un Nanodrop (Thermo Fischer). La presencia de HCP también se determinó utilizando un kit HCP de células CHO basado en anticuerpos (ForteBIO).

El análisis del HCP presente en los medios usados muestra que los 2 KO contienen cantidades significativamente más bajas de HCP en los medios usados que la línea celular CHO-S (entre aproximadamente el 55 % y el 70 % en comparación con CHO-S) e incluso HCP menos inmunogénico (aprox. el 35 %), mostrado por el ensayo HCP basado

en anticuerpos. (Véase la figura 3).

5 Análisis del HCP presente en los medios usados de células con los genes eliminados individualmente, muestra que especialmente TIMP1, BGN, La eliminación de NID1-1 y NID1-2 reduce la señal inmunogénica observada en el ensayo basado en anticuerpos (véase la figura 4)

Listado de referencias

- 10 WO 2007/006808 A1
WO 2012/082509 A2
Baycin-Hizal *et al.*, J Proteome Res 2012;11(11):5265-5276
Chaudhuri *et al.*, J Bioprocess Biotech 2015;5:7
Datta *et al.*, Biotechnol Bioeng 2013;110(5):1255-1271
Kim *et al.*, Appl Microbiol Biotechnol 2012;93:917-930
15 Slade *et al.*, J Proteome Res 2012;11:6175-6186
Zhang, "Mammalian Cell Culture for Biopharmaceutical production", En: Manual of Industrial Microbiology and Biotechnology, Edición: 3.^a ed., Capítulo: 12, Editorial: ASM Press, Washington, DC., pp.157 - 178 (2010)
Lee *et al.*, Biotechnology Journal 2015; 10(7):979-994
Walsh, Nat Biotechnol 2014;32:992-1000.

REIVINDICACIONES

1. Una célula CHO modificada para interrumpir la expresión de una o más proteínas endógenas seleccionadas de entre el grupo que consiste en TIMP1, NID1-1, NID1-2 y LGALS3BP.
2. La célula CHO de la reivindicación 1, en donde la célula CHO es modificada para interrumpir la expresión de TIMP1.
3. La célula CHO de una cualquiera de las reivindicaciones anteriores, en donde la célula CHO es modificada para interrumpir la expresión de LGALS3BP.
4. La célula CHO de una cualquiera de las reivindicaciones anteriores, en donde la célula CHO es modificada para interrumpir la expresión de NID1-1, NID1-2 o de ambos.
5. La célula CHO de una cualquiera de las reivindicaciones anteriores, en donde la célula CHO es modificada para alterar la expresión de:
 - (a) al menos 5 proteínas endógenas;
 - (b) como máximo 50 proteínas endógenas; o
 - (c) entre 5 y 50 proteínas endógenas.
6. La célula CHO de una cualquiera de las reivindicaciones anteriores, en donde la célula CHO es modificada para interrumpir la expresión de TIMP1, LGALS3BP, NID1-1 y NID1-2.
7. La célula CHO de una cualquiera de las reivindicaciones anteriores, en donde la expresión de al menos una proteína endógena se ve interrumpida por la atenuación del gen que codifica la proteína endógena.
8. La célula CHO de una cualquiera de las reivindicaciones anteriores, en donde la expresión de al menos una proteína endógena se ve interrumpida por la inactivación del gen que codifica la proteína endógena.
9. La célula CHO de una cualquiera de las reivindicaciones anteriores, que comprende una secuencia de ácido nucleico que codifica una proteína recombinante de interés, estando la secuencia de ácido nucleico extracromosómica o cromosómicamente integrada y bajo el control de un promotor inducible o de un promotor constitutivo.
10. La célula CHO de acuerdo con la reivindicación 9, célula que produce niveles más altos de dicha proteína recombinante de interés y/o que tiene una mayor densidad celular en comparación con una célula de control relevante sin expresión interrumpida de una o más proteínas endógenas, tal como una célula CHO de tipo salvaje.
11. Un método para producir una proteína recombinante de interés, que comprende:
 - (a) cultivar la célula CHO de la reivindicación 9 en un medio de cultivo tal que la proteína recombinante de interés se exprese a partir de la secuencia de ácido nucleico;
 - (b) recolectar la proteína recombinante de interés del medio de cultivo; y
 - (c) opcionalmente, purificar adicionalmente la proteína recombinante recolectada de interés.
12. Un método para la preparación de la célula CHO de una cualquiera de las reivindicaciones 1 a 10, que comprende las etapas de:
 - (a) modificar la célula CHO para atenuar o para inactivar el gen endógeno o los genes endógenos que codifican uno o más de TIMP1, NID1-1, NID1-2 y LGALS3BP; y
 - (b) opcionalmente, transfectar la célula CHO modificada con una secuencia de ácido nucleico que codifica una proteína de interés.

Figura 1:

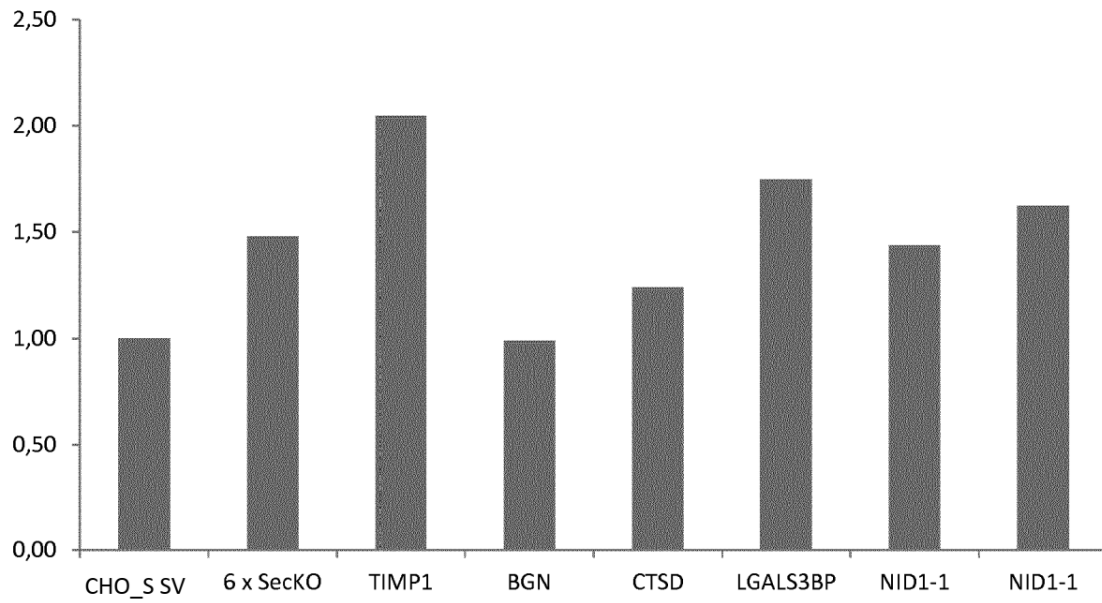


Figura 2:

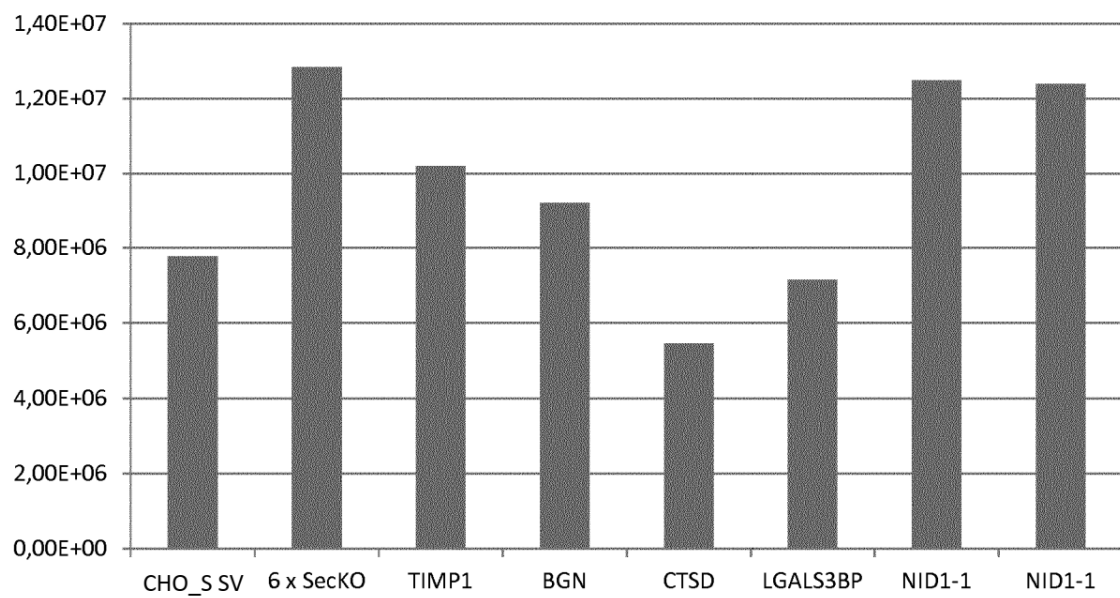


Figura 3:

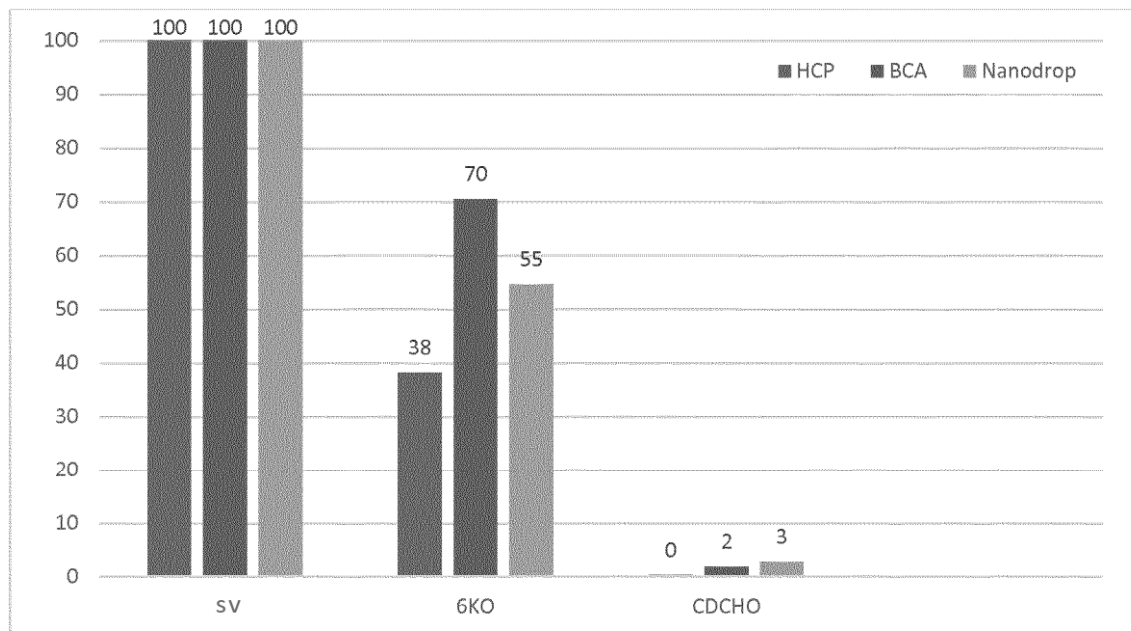


Figura 4:

