(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 906 997**

(51) Int. CI.:

C12N 15/31 (2006.01)
C12N 15/81 (2006.01)
C07K 14/395 (2006.01)
C12N 9/00 (2006.01)
C12N 9/10 (2006.01)
C12P 21/02 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **20.06.2018 PCT/EP2018/066344**

(87) Fecha y número de publicación internacional: **27.12.2018 WO18234349**

(96) Fecha de presentación y número de la solicitud europea: **20.06.2018 E 18730809 (3)**

(97) Fecha y número de publicación de la concesión europea: **08.12.2021 EP 3642343**

(54) Título: **Cepas de expresión de proteínas mejoradas**

(30) Prioridad:

20.06.2017 EP 17176932

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
21.04.2022

(73) Titular/es:

ALBUMEDIX LTD. (100.0%)
HQ, Mabel Street
Nottingham NG2 3ED, GB

(72) Inventor/es:

FINNIS, CHRISTOPHER, JOHN, ARTHUR;
NORDEIDE, PER, KRISTOFFER y
MCLAUGHLAN, JENNIFER, MARY

(74) Agente/Representante:

ISERN JARA, Jorge

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

Cepas de expresión de proteínas mejoradas

Campo de la invención

La presente invención se refiere principalmente al desarrollo de cepas fúngicas que expresan proteínas a niveles sustancialmente más elevados que las cepas precursoras.

Antecedentes de la invención

Desde hace unos 30 años, se producen proteínas heterólogas deseadas en microorganismos. Sin embargo, una vez introducida la secuencia codificante necesaria y obtenida la expresión, aún queda mucho por hacer para optimizar el proceso de producción comercial. Un área de interés se refiere a la mejora de cepas, es decir, encontrar o producir cepas de microorganismos hospedadores que permitan producir la proteína con mayores rendimientos o mejor pureza, por ejemplo.

Para aumentar el rendimiento, una vez que se ha ideado un buen sistema de expresión, se podría considerar intentar aumentar el número de copias de la secuencia codificante, o aumentar la cantidad o la estabilidad del ARNm, o mejorar el plegamiento y/o la secreción de la proteína o disminuir la degradación de la proteína. Sin embargo, el efecto deseado de una mayor expresión solo se verá si se dirige a los factores limitantes.

Por lo tanto, lo que se requiere es una cepa hospedadora que permita aumentar el rendimiento de una proteína deseada, tal como una proteína heteróloga.

Los inventores han identificado de manera sorprendente que la mutación de *GSH1* en una célula fúngica da como resultado dicho mayor rendimiento de la proteína heteróloga.

El producto del gen *GSH1* en *Saccharomyces cerevisiae* es γ-glutamilcisteína sintetasa (Gsh1p). Esta enzima cataliza la primera etapa, y limitante de la velocidad, en la síntesis de glutatión (L-γ-glutamilcisteinilglicina; GSH). La glutatión sintetasa (Gsh2p) cataliza la segunda etapa. El glutatión es una molécula esencial que tiene muchas funciones importantes, incluida la protección de las células contra el estrés oxidativo, como cofactor en varias rutas de biosíntesis y reacciones de desintoxicación, y el mantenimiento del genoma mitocondrial de la levadura (Ayer *et al.*, (2010) Free Radical Biology & Medicine 49 1956-1968)). Las células de levadura con el gen *GSH1* eliminado, a menos que se cultiven en medios que contengan glutatión, son hipersensibles a una variedad de condiciones de estrés, tales como condiciones oxidativas y exposición a metales pesados y, de manera eventual, experimentan una detención del crecimiento (Spector *et al.*, (2001) The Journal of Biological Chemistry 276 (10) 7011-7016). También se sabe que el glutatión protege el núcleo en momentos de estrés oxidativo (Hatem *et al.*, (2014) Free Radical Biology & Medicine 67 103-114). Biterova y Barycki (2009) (The Journal of Biological Chemistry 284, 32700-32708) analizaron las estructuras cristalinas de Ghs1 (también conocida como glutamato cisteína ligasa) para investigar el mecanismo de biosíntesis del glutatión y afirman que se han observado variantes de la glutamato cisteína ligasa humana (por ejemplo, R127C, P158L, H370L y P414L).

Previamente (PCT/US2016/068239), se identificó que la mutación de *NOT4* (también conocida como *MOT2)* da como resultado un mayor rendimiento de proteína heteróloga producida en una cepa hospedadora. De manera sorprendente, los inventores ahora han identificado que combinar las mutaciones *NOT4* y *GSH1* da como resultado un aumento adicional en el rendimiento.

Not4 es una enzima que se une a ubiquitina y forma parte del complejo Ccr4-Not. El complejo Ccr4-Not se conserva en las células eucariotas, y en la levadura el complejo consiste en 9 subunidades centrales: Ccr4, Caf1, Caf40, Caf130, Not1, Not2, Not3, Not4 y Not5 (Collar, 2003, Global control of gene expression in yeast by the Ccr4-Not complex. Gene 313: 1-16; Bai *et al.*, 1999, The CCR4 and Caf1 proteins of the Ccr4-Not complex are physically and functionally separated from Not2, Not4 y Not5. Mol. Cell. Biol. 19: 6642-6651). Se ha propuesto que el complejo funcione como una centralita que puede interpretar señales del entorno y coordinar todos los niveles de expresión génica para responder de manera económica a la señal (Collart, 2012, The Ccr4-Not complex. Gene 492(1): 42-53). Se cree que las proteínas Not (Not1, Not2, Not3 y Not4) son necesarias para el ensamblaje del complejo de ARN polimerasa II, lo que sugiere un papel general en la regulación de la transcripción (Collart, 1994, Not1(cdc39), Not2(cdc36), Not3, and Not4 encode a global-negative regulator of transcription that differentially affects tata-element utilization. Genes & Development 8(5): 525-537; Collart, 2012, como se cita anteriormente). Una estructura de cocristal sugirió cómo la región carboxiterminal de Not4 envuelve una región de repetición HEAT de Not1, la proteína de armazón en el complejo Ccr4-Not (Bhaskar, 2015, Architecture of the ubiquitylation module of the yeast Ccr4-Not complex. Structure 23(5): 921-8).

Gutierrez-Escobedo *et al.*, (2013), Current Genetics: Eukaryotes with emphasis on yeasts, fungi, protists, cell organelles (Springer, Berlín, Alemania) vol. 59(3): 91-106 divulga el papel del glutatión en la respuesta al estrés oxidativo en el patógeno fúngico *Candida glabrata.*

Maris *et al.*, (2000), Current Genetics, vol. 37(3): 175-182 divulga que el glutatión, pero no el factor de transcripción Yap1, se requiere para la resistencia al estrés oxidativo dependiente de la fuente de carbono en *Saccharomyces cerevisiae.*

Wheeler *et al.*, (2002), Molecular Microbiology, vol. 46(2): 545-556 divulga que el glutatión regula la expresión de γ-glutamilcisteína sintetasa *a través* del factor de transcripción Met4.

El documento CN102296033 (Universidad de Wuhan) divulga una nueva cepa mutante por eliminación de δ-γ-glutamilcisteína sintetasa 1 (gsh1) de *Saccharomyces cerevisiae* depositada bajo el n.º de CCTCC: M 2011130, útil para detectar la toxicidad del cloruro de cadmio y otros metales pesados.

Grant *et al.*, (1997), Molecular Biology of the Cell (American Society for Cell Biology, EE. UU.), vol. 8(9): 1699-1707, divulga que la glutatión sintetasa es prescindible para el crecimiento en condiciones tanto normales como de estrés oxidativo en la levadura *Saccharomyces cerevisiae* debido a una acumulación del dipéptido γ-glutamilcisteína.

Kistler *et al.*, (1990), Mutagene (IRL Press, Oxford, Reino Unido), vol. 5(1): 39-44, divulga un análisis genético y bioquímico de mutantes deficientes en glutatión de *Saccharomyces cerevisiae.*

Lee *et al.*, 2001), FEMS Yeast Research, vol. 1(1): 57-65 divulga un papel esencial y auxiliar del glutatión en *Saccharomyces cerevisiae* analizado con una cepa disruptiva grande gsh1.

Otake *et al.*, (1991), Yeast, 1991, vol. 7(9): 953-961 divulga la clonación molecular del gen de la γ-glutamilcisteína sintetasa de *Saccharomyces cerevisiae.*

Sumario de la invención

La invención proporciona un método para producir una proteína heteróloga que comprende:

a. proporcionar una célula hospedadora fúngica que tiene:

1. una proteína Gsh1 modificada que tiene un nivel reducido de actividad en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
2. un nivel reducido de actividad de la proteína Gsh1 en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
3. un gen *GSH1* modificado que tiene un nivel reducido de actividad en relación con la actividad de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
4. un nivel reducido de expresión del gen *GSH1* en relación con el nivel de expresión de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y

b. cultivar la célula hospedadora para producir la proteína heteróloga,
en donde

(i) la proteína heteróloga es una proteína vegetal o animal, y/o
(ii) el cultivo es a una escala de al menos 1 litro, y/o
(iii) el método comprende además purificar la proteína heteróloga, y/o
(iv) el método comprende además formular la proteína heteróloga con un vehículo o diluyente terapéuticamente aceptable para producir así un producto terapéutico adecuado para la administración a un ser humano o un animal.

La célula hospedadora fúngica puede tener además:

5. una proteína Not4 modificada, y/o
6. un nivel de actividad de la proteína Not4 modificado, y/o
7. un gen *NOT4* modificado, y/o
8. un nivel de expresión del gen *NOT4* modificado.

Los inventores también divulgan un cultivo de células hospedadoras fúngicas que contienen una secuencia polinucleotídica que codifica una proteína deseada, tal como una proteína heteróloga, caracterizado por que las células hospedadoras fúngicas tienen un nivel de actividad y/o un nivel de expresión reducidos de la proteína Gsh1 y, opcionalmente, tienen además un nivel de actividad y/o un nivel de expresión reducidos de la proteína Not4.

La invención proporciona además un método para producir una proteína deseada, tal como una proteína heteróloga, a partir de una célula hospedadora fúngica.

La invención proporciona un método para modificar el rendimiento de producción de un polipéptido deseado a partir

de una célula hospedadora fúngica.

La invención también proporciona una proteína deseada, tal como una proteína heteróloga. La albúmina o variante, fragmento y/o fusión de la misma es una proteína deseada preferida.

La invención proporciona además una composición, tal como una composición farmacéutica, que comprende la proteína deseada.

La invención también proporciona un método para tratar a un paciente que comprende administrar una cantidad eficaz de la composición al paciente.

La invención proporciona además un método para preparar una célula hospedadora fúngica que tiene la propiedad o propiedades mencionadas anteriormente.

La invención también proporciona una proteína Gsh1 que comprende al menos un 50 % de identidad con la SEQ ID NO: 2 y una mutación en una posición correspondiente a una o más posiciones seleccionadas de 47, 48, 49, 50, 51, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 409, 451, 452, 453, 454 y 455 de la SEQ ID NO: 2.

La invención proporciona además un polinucleótido que codifica una variante de Gsh1 de la presente invención.

Cualquier realización descrita en el presente documento, incluidas aquellas descritas solo en los ejemplos y/o en la sección de realizaciones preferidas, están destinadas a poder combinarse con cualquier otra realización, a menos que se rechace de manera explícita o la combinación sea inadecuada.

Breve descripción de los dibujos

La figura 1 es un diagrama de Venn que muestra las clases de veinte aminoácidos y las relaciones entre ellos.
La figura 2 muestra la construcción del plásmido pDB5438.
La figura 3 muestra la construcción del plásmido pDB2305, "HSA" significa albúmina humana recombinante, "mFL" es una secuencia líder.
La figura 4 muestra la construcción del plásmido pDB2244, "rHA" significa albúmina humana recombinante, "FL" es una secuencia líder.
La figura 5 muestra la construcción del plásmido YCp50.
La figura 6 muestra la construcción del plásmido pDB5862. "Variante de albúmina" es el ácido nucleico que codifica la SEQ ID NO: 45. "mFL" es una secuencia líder.
La figura 7 muestra la construcción del plásmido pDB3936.
La figura 8 muestra la construcción del plásmido pDB5912. "mFL" es una secuencia líder.
La figura 9 muestra la construcción del plásmido pDB3029. "Inv" es una secuencia líder.

Definiciones

Albúmina: El término "albúmina" significa una proteína que tiene la misma estructura terciaria y/o muy similar a la seroalbúmina humana (HSA) o dominios HSA y tiene propiedades similares a la HSA o a los dominios en cuestión. Las estructuras terciarias similares son, por ejemplo, las estructuras de las albúminas de las especies mencionadas en "albúmina precursora". Algunas de las principales propiedades de la albúmina son i) su capacidad para regular el volumen en plasma (actividad oncótica), ii) una semivida en plasma prolongada de alrededor de 19 días ± 5 días, iii) unión a gp60, también conocida como albondina, iv) unión a FcRn, v) unión a ligando, por ejemplo, unión de moléculas endógenas tales como compuestos ácidos y lipófilos que incluyen bilirrubina, ácidos grasos, hemina y tiroxina (ver también la tabla 1 de Kragh-Hansen *et al.*, 2002, Biol. Pharm. Bull. 25, 695), vi) unión de compuestos orgánicos pequeños con características ácidas o electronegativas, por ejemplo, fármacos tales como la warfarina, diazepam, ibuprofeno y paclitaxel (véase también la tabla 1 de Kragh-Hansen *et al.*, 2002, Biol. Pharm. Bull. 25, 695). No es necesario cumplir todas estas propiedades para caracterizar una proteína o fragmento como albúmina. Si un fragmento, por ejemplo, no comprende un dominio responsable de la unión de determinados ligandos o compuestos orgánicos, tampoco se espera que la variante de dicho fragmento tenga estas propiedades.

Variante alélica: El término "variante alélica" significa cualquiera de dos o más formas alternativas de un gen que ocupa el mismo locus cromosómico. La variación alélica surge de manera natural a través de la mutación y puede dar lugar al polimorfismo dentro de las poblaciones. Las mutaciones génicas pueden ser silenciosas (sin cambios en el polipéptido codificado) o pueden codificar polipéptidos que tienen secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

ADNc: El término "ADNc" significa una molécula de ADN que se puede preparar mediante transcripción inversa a partir de una molécula de ARNm madura y cortada y empalmada obtenida de una célula eucariota o procariota. El ADNc carece de secuencias de intrones que pueden estar presentes en el ADN genómico correspondiente. El transcrito de ARN primario inicial es un precursor del ARNm que se procesa a través de una serie de etapas, que incluyen corte y empalme, antes de aparecer como ARNm maduro y cortado y empalmado.

Secuencia codificante: La expresión "secuencia codificante" significa un polinucleótido, que especifica directamente la secuencia de aminoácidos de un polipéptido. Los límites de la secuencia codificante están determinados en general por un marco de lectura abierto, que comienza con un codón de inicio, tal como ATG, GTG o TTG y termina con un codón de terminación, tal como TAA, TAG o TGA. La secuencia codificante puede ser un ADN genómico, ADNc, ADN sintético o una combinación de los mismos.

Secuencias de control: La expresión "secuencias de control" significa secuencias de ácido nucleico necesarias para la expresión de un polinucleótido que codifica un polipéptido de la presente invención. Cada secuencia de control puede ser natural (es decir, del mismo gen) o extraña (es decir, de un gen diferente) al polinucleótido que codifica el polipéptido o natural o extraña entre sí. Dichas secuencias de control incluyen, pero sin limitación, un líder, una secuencia de poliadenilación, una secuencia propéptido, un promotor, una secuencia de péptido señal y un terminador de la transcripción. Como mínimo, las secuencias de control incluyen un promotor y señales de terminación de la transcripción y la traducción. Las secuencias de control pueden proporcionarse con enlazadores con el fin de introducir sitios de restricción específicos que faciliten la unión de las secuencias de control con la región codificante del polinucleótido que codifica un polipéptido.

Expresión: El término "expresión" incluye cualquier etapa implicada en la producción de un polipéptido, que incluye, pero sin limitación, transcripción, modificación postranscripcional, traducción, modificación postraduccional y secreción.

Casete de expresión: La expresión "casete de expresión" significa el polinucleótido que codifica un polipéptido y las secuencias de control cadena arriba y cadena abajo que proporcionan su expresión.

Hospedador de expresión: La expresión "hospedador de expresión" significa cualquier célula hospedadora que expresa una proteína deseada, particularmente una proteína heteróloga.

Vector de expresión: La expresión "vector de expresión" significa una molécula de ADN lineal o circular que comprende un polinucleótido que codifica un polipéptido y está unido de manera operativa a las secuencias de control que proporcionan su expresión.

Fragmento: El término "fragmento" significa un polipéptido que tiene uno o más (varios) aminoácidos eliminados del extremo aminoterminal y/o carboxiterminal de un polipéptido maduro y/o de una región interna de un polipéptido maduro. Los fragmentos pueden consistir en una secuencia ininterrumpida procedente de un polipéptido o pueden comprender dos o más secuencias procedentes de diferentes partes del polipéptido. Con respecto a la albúmina, un fragmento puede tener un tamaño de más de aproximadamente 20 restos de aminoácidos, preferentemente más de 30 restos de aminoácidos, más preferentemente más de 40 restos de aminoácidos, más preferentemente más de 50 restos de aminoácidos, más preferentemente más de 75 restos de aminoácidos, más preferentemente más de 100 restos de aminoácidos, más preferentemente más de 200 restos de aminoácidos, más preferentemente más de 300 restos de aminoácidos, incluso más preferentemente más de 400 restos de aminoácidos y más preferentemente más de 500 restos de aminoácidos. En una realización preferida, un fragmento corresponde a uno o más de los dominios de albúmina. Los dominios de albúmina preferidos de la invención son dominios que tienen al menos un 70, 75, 80, 85, 90, 95, 96, 97, 98, 99, 99,5 % o el 100 % de identidad con un dominio I de HSA que consiste en los restos de aminoácidos 1 a 194 ± 1 a 15 aminoácidos de la SEQ ID NO: 10; al menos un 70, 75, 80, 85, 90, 95, 96, 97, 98, 99, 99,5 % o el 100 % de identidad con el dominio II de HSA que consiste en los restos de aminoácidos 192 a 387 ± 1 a 15 aminoácidos de la SEQ ID NO: 10 y al menos un 70, 75, 80, 85, 90, 95, 96, 97, 98, 99, 99,5 % o el 100 % de identidad con el dominio III de HSA que consiste en los restos de aminoácidos 381 a 585 ± 1 a 15 aminoácidos de la SEQ ID NO: 10 o una combinación de uno o más (varios) de estos dominios, por ejemplo, los dominios I y II, los dominios II y III o los dominios I y III fusionados juntos. No existe una convención generalmente aceptada para los límites exactos de los dominios de albúmina y la superposición en los intervalos mencionados anteriormente y la admisión de una longitud variable de más o menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 o 15 aminoácidos, preferentemente de 1 a 15 aminoácidos, más preferentemente de 1 a 10 aminoácidos, lo más preferentemente de 1 y 5 aminoácidos, en el extremo aminoterminal y/o carboxiterminal de los dominios, que permite una variación total en la longitud de hasta 30 aminoácidos, preferentemente de hasta 20 aminoácidos, más preferentemente de hasta 10 aminoácidos para cada dominio refleja este hecho y que puede haber algunas opiniones divergentes sobre los restos de aminoácidos en el borde entre los dominios que pertenecen a uno u otro dominio. Por la misma razón, es posible encontrar referencias a los restos de aminoácidos de los dominios de albúmina que difieren de los números anteriores, sin embargo, el experto apreciará cómo identificar los dominios de albúmina en función de la enseñanza en la bibliografía y la enseñanza anterior. Los dominios correspondientes de albúminas no humanas se pueden identificar mediante alineación con HSA utilizando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, J. Mol. Biol. 48: 443-453) como se implementó en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, Trends Genet. 16: 276-277), preferentemente la versión 3.0.0 o posterior, más preferentemente la versión 5.0.0 o posterior. Los parámetros opcionales utilizados son penalización por apertura de hueco de 10, penalización por extensión de hueco de 0,5, y la matriz de sustitución EBLOSUM62 (versión EMBOSS de BLOSUM62). También se pueden utilizar herramientas de alineación alternativas, por ejemplo, MUSCLE como se describe en el presente documento. Los dominios también se pueden

definir de acuerdo con Dockal o Kjeldsen: Dockal *et al.*, (The Journal of Biological Chemistry, 1999, Vol. 274(41): 29303-29310) define los dominios de la HSA como: Dominio I: aminoácidos 1 a 197, Dominio II: aminoácidos 189 a 385 de la SEQ ID NO: 10, Dominio III: aminoácidos 381 a 585 de la SEQ ID NO: 10. Kjeldsen *et al.*, (Protein Expression and Purification, 1998, Vol 13: 163-169) define los dominios como: Dominio I: aminoácidos 1 a 192, Dominio II: aminoácidos 193 a 382, Dominio III: aminoácidos 383 a 585. Cada dominio está compuesto por dos subdominios homólogos, a saber, 1-105, 120-194, 195-291, 316-387, 388-491 y 512-585, con regiones enlazadoras flexibles entre subdominios que comprenden los restos Lys106 a Glu119, Glu292 a Va1315 y Glu492 a Ala511.

Por lo tanto, en la presente invención, se prefieren las siguientes definiciones de dominio. Los números de aminoácidos corresponden a los de la SEQ ID NO: 10 (HSA). Sin embargo, utilizando estos números, el experto en la materia puede identificar los dominios correspondientes en otras secuencias de albúmina. El dominio I puede o no comenzar en el aminoácido 1 y puede o no terminar en cualquiera de los aminoácidos 192, 193, 194, 195, 196 o 197, preferentemente en cualquiera de los aminoácidos 192, 194 o 197. El dominio II puede o no comenzar en el aminoácido 189, 190, 191, 192 o 193, preferentemente en cualquiera de los aminoácidos 189, 192 o 193, y puede o no terminar en el aminoácido 382, 383, 384, 385, 386 o 387, preferentemente en cualquiera de los aminoácidos 382, 285 o 387. El dominio III puede o no comenzar en el aminoácido 381, 382 o 383, preferentemente en el aminoácido 381 o 383, y puede o no terminar en el aminoácido 585. Los dominios en albúminas no humanas pueden tener longitudes y/o números de restos de aminoácidos iguales o diferentes que HSA. Por ejemplo, se puede preparar una alineación múltiple o una alineación por pares utilizando HSA y una o más (varias) de otras albúminas, fragmentos, derivados, variantes y/o fusiones para identificar dominios correspondientes a los dominios I, II y/o III de la HSA.

Compañero de fusión: A lo largo de la presente memoria descriptiva, un compañero de fusión es una fracción que no es albúmina que puede fusionarse de manera genética con una albúmina o una variante y/o un fragmento de la misma.

Proteína heteróloga: una proteína heteróloga es aquella que no se produce de manera natural por la célula hospedadora y, preferentemente, no incluye proteínas tales como marcadores de selección (por ejemplo, marcadores de resistencia a antibióticos, marcadores seleccionables auxótrofos), chaperonas, FLP, REP1 o REP2.

Célula hospedadora: La expresión "célula hospedadora" significa cualquier tipo de célula que sea susceptible de transformación, transfección, transducción o similar con una construcción de ácido nucleico o vector de expresión que comprende un polinucleótido de la presente invención. La expresión "célula hospedadora" abarca cualquier progenie de una célula precursora que no sea idéntica a la célula precursora debido a mutaciones que se producen durante la replicación.

Polipéptido maduro: La expresión "polipéptido maduro" significa un polipéptido en su forma final después de la traducción y cualquier modificación postraduccional, tal como procesamiento aminoterminal, truncamiento carboxiterminal, glucosilación, fosforilación, *etc.* La secuencia madura de la albúmina humana se proporciona en la SEQ ID NO: 10, mientras que se proporciona un ejemplo de una forma inmadura en la SEQ ID NO: 12.

Secuencia codificante de polipéptido maduro: La expresión "secuencia codificante de polipéptido maduro" significa un polinucleótido que codifica un polipéptido maduro. En la SEQ ID NO: 9 se proporciona un ejemplo de una secuencia codificante de polipéptido maduro de la albúmina humana, mientras que en la SEQ ID NO: 11 se proporciona un ejemplo de una secuencia codificante para una forma inmadura de la albúmina humana.

Mutante: El término "mutante" significa un polinucleótido que codifica una variante.

Construcción de ácido nucleico: La expresión "construcción de ácido nucleico" significa una molécula de ácido nucleico, monocatenaria o bicatenaria, que se aísla de un gen de origen natural o se modifica para contener segmentos de ácidos nucleicos de una manera que de otro modo no existiría en la naturaleza o que es sintética, que comprende una o más secuencias de control.

Unido de manera operativa: La expresión "unido de manera operativa" significa una configuración en la que se coloca una secuencia de control en una posición adecuada en relación con la secuencia codificante de un polinucleótido, de tal manera que la secuencia de control dirige la expresión de la secuencia codificante.

Precursor o albúmina precursora: El término "precursor" o "albúmina precursora" significa una albúmina a la que se le hace una alteración para producir las variantes de albúmina de la presente invención. El precursor puede ser un polipéptido de origen natural (tipo silvestre) o un alelo del mismo o una variante del mismo. En una realización preferida, la albúmina precursora es una albúmina de tipo silvestre, más preferentemente una albúmina de tipo silvestre de *Homo sapiens* como se divulga en la SEQ ID NO: 12 (UNIPROT: P02768.2) o la secuencia madura de la misma (SEQ ID NO: 10). Las albúminas de tipo silvestre alternativas se pueden seleccionar de la lista no detallada que se muestra en la tabla 1.

Tabla 1: Albúminas de tipo silvestre de varias especies.

Nombre común	Especie	N.º de referencia de SwissProt o GenBak	% de identidad respecto a la SEQ ID NO: 10*	Longitud (aa)
Ser humano	*Homo sapiens*	P02768.2	100,0	609
Chimpancé	*Pan troglodytes*	XP_517233 (secuencia prevista)	98,8	609
Orangután de Sumatra	*Pongo abelii*	Q5NVH5.2	98,5	609
Macaco (Macaco de la India)	*Macaca mulatta*	Q28522.1	93,3	600
Gato	*Felis catus*	P49064.1	81,9	608
Perro	*Canis lupus familiaris*	P49822.3	80,0	608
Burro	*Equus asinus*	Q5XLE4.1	76,7	607
Caballo	*Equus caballus*	P35747.1	76,3	607
Esquistosoma	*Schistosoma mansoni*	Q95VB7	76,2	608
Bovino	*Bos taurus*	P02769.4 (NP_851335.1)	75,6	607
Cerdo	*Sus scrofa*	P08835.2	75,1	607
Oveja	*Ovis aries*	P14639.1	75,0	607
Cabra	*Capra hircus*	ACF10391.1	74,8	607
Conejo	*Oryctolagus cuniculus*	P49065.2	74,3	608
Jerbo de Mongolia	*Meriones unguiculatus*	035090,1	73,6	609
Rata	*Rattus norvegicus*	P02770. 2.	73,3	608
Ratón	*Mus musculus*	P07724.3.	72,3	608
Cobaya	*Cavia porcellus*	Q6WDN9	72,1	608
Pollo	*Gallus gallus*	P19121.2	47,0	615
* La identidad de la secuencia se calculó utilizando el algoritmo Needleman-Wunsch implementado en el programa Needle de EBLOSUM62 (paquete de programas EMBOSS, versión 6.1.0) utilizando una penalización por apertura de hueco de 10, penalización por extensión de hueco de 0,5 como se describe en el presente documento.				

Preferentemente, la albúmina precursora es una albúmina madura. En otra realización, la albúmina precursora es al menos un 70 %, más preferentemente el 75 %, más preferentemente al menos un 80 %, más preferentemente al menos un 85 %, incluso más preferentemente al menos un 90 %, lo más preferente al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, al menos un 99,5 % o al menos un 99,8 % idéntica a la SEQ ID NO: 10, y mantiene al menos una de las propiedades principales de la albúmina o una estructura terciaria similar a la de la albúmina, tal como la HSA. Las propiedades principales de la albúmina se resumen en Sleep, 2015, "Albumin and its application in drug delivery", Expert Opinion on Drug Delivery 12(5): 793-812.

Identidad de secuencia: La relación entre dos secuencias de aminoácidos o entre dos secuencias de nucleótidos se describe mediante el parámetro "identidad de secuencia".

Para los fines de la presente invención, la identidad de secuencia entre dos secuencias de aminoácidos se determina usando el algoritmo Needleman-Wunsch (Needleman y Wunsch, 1970, J. Mol. Biol. 48: 443-453) como se implementó en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, Trends Genet. 16: 276-277), preferentemente la versión 5.0.0 o posterior. Los parámetros utilizados son penalización por apertura de hueco de 10, penalización por extensión de hueco de 0,5, y la matriz de sustitución EBLOSUM62 (versión EMBOSS de BLOSUM62). El resultado de Needle marcado como "identidad más larga" (obtenida usando la opción *-nobrief*) se usa como porcentaje de identidad y se calcula de la siguiente manera:

(Restos idénticos × 100)/(Longitud de alineación - número total de huecos en la alineación)

Para los fines de la presente invención, la identidad de secuencia entre dos secuencias de desoxirribonucleótidos se *determina utilizando el algoritmo Needleman-Wunsch (Needleman y Wunsch, 1970, citado anteriormente)* implementado en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, citado anteriormente), preferentemente la versión 5.0.0 o posterior. Los parámetros utilizados son penalización por apertura de hueco de 10, penalización por extensión de hueco de 0,5, y la matriz de sustitución EDNAFULL (versión EMBOSS de NCBI NUC4.4). El resultado de Needle marcado como "identidad más larga" (obtenida usando la opción *-nobrief*) se usa como porcentaje de identidad y se calcula de la siguiente manera:

(Desoxirribonucleótidos idénticos × 100)/(Longitud de alineación - Número total de huecos en la alineación)

Variante: El término "variante" significa un polipéptido procedente de un polipéptido precursor, por ejemplo,

albúmina, que comprende una alteración, es decir, una sustitución, una inserción y/o una eliminación, en una o más (varias) posiciones. Una sustitución significa un reemplazo de un aminoácido que ocupa una posición por un aminoácido diferente; una eliminación significa la eliminación de un aminoácido que ocupa una posición; y una inserción significa añadir de 1 a 3 aminoácidos adyacentes a un aminoácido que ocupa una posición. El polipéptido alterado (variante) se puede obtener a través de la intervención humana mediante la modificación de la secuencia polinucleotídica que codifica el polipéptido precursor, por ejemplo, albúmina. La albúmina variante es preferentemente al menos un 70 %, preferentemente al menos un 75 %, más preferentemente al menos un 80 %, más preferentemente al menos un 85 %, incluso más preferentemente al menos un 90 %, lo más preferente al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 %, al menos un 99,5 % o al menos un 99,8 % idéntica a la SEQ ID NO: 10 y puede o no mantener al menos una de las propiedades principales de la albúmina precursora o una estructura terciaria similar tal como la HSA. En general, las variantes o fragmentos de la HSA tendrán al menos un 10 % (preferentemente al menos un 50 %, un 60 %, un 70 %, un 80 %, un 90 % o un 95 %) de actividad de unión al ligando de HSA (por ejemplo, unión a bilirrubina) y al menos un 50 % (preferentemente al menos un 70 %, un 80 %, un 90 % o un 95 %) de la actividad oncótica de la HSA, peso por peso. La actividad oncótica, también conocida como presión osmótica coloidal, de la albúmina, las variantes de la albúmina o los fragmentos de la albúmina pueden determinarse mediante el método descrito por Hoefs, J. C. (1992) Hepatology 16:396-403. La unión a bilirrubina se puede medir mediante aumento de fluorescencia a 527 nm en relación con la HSA. Se disuelve bilirrubina (1,0 mg) en 50 µl de NaOH 1 M y se diluye a 1,0 ml con agua desmineralizada. La reserva de bilirrubina se diluye en Tris-HCl 100 mM pH 8,5, EDTA 1 mM para dar 0,6 nmol de bilirrubina/ml en una cubeta de fluorómetro. La fluorescencia se mide mediante excitación a 448 nm y emisión a 527 nm (anchos de ranura de 10 nm) durante la valoración con HSA en un intensidad de proporciones de HSA:bilirrubina de 0 a 5 mol:mol. La variante puede tener una afinidad de unión alterada a FcRn y/o una semivida en plasma alterada en comparación con la albúmina precursora.

Con respecto a una variante de la proteína Gsh1, se aplican los mismos principios, con la excepción de que la actividad es actividad de Gsh1 en lugar de actividad de albúmina. La proteína Gsh1 precursora puede tener al menos un 50, 60, 70, 80 o 90 % de identidad con la SEQ ID NO: 2, más preferentemente el 100 % de identidad con la SEQ ID NO: 2. La variante de la proteína Gsh1 puede tener al menos un 50, 60, 70, 80, 85, 90, 95, 96, 97, 98 o al menos un 99 % de identidad con la SEQ ID NO: 2.

Con respecto a una variante de la proteína Not4, se aplican los mismos principios, con la excepción de que la actividad es actividad de Not4 en lugar de actividad de albúmina. La proteína Not4 precursora puede tener al menos un 50, 60, 70, 80 o 90 % de identidad con la SEQ ID NO: 6, más preferentemente el 100 % de identidad con la SEQ ID NO: 6. La variante de la proteína Not4 puede tener al menos un 50, 60, 70, 80, 85, 90, 95, 96, 97, 98 o al menos un 99 % de identidad con la SEQ ID NO: 6.

La secuencia polipeptídica variante es preferentemente una que no se encuentra en la naturaleza.

Vector: El término "vector" significa una molécula de ADN lineal o circular que comprende un polinucleótido que codifica un polipéptido y está unido de manera operativa a las secuencias de control que proporcionan su expresión. Los vectores incluyen plásmidos. Los vectores incluyen vectores de expresión.

Tipo silvestre: La expresión albúmina de "tipo silvestre" (TS) significa una albúmina que tiene la misma secuencia de aminoácidos de las albúminas que se encuentran de manera natural en un animal o en un ser humano. La SEQ ID NO: 10 es un ejemplo de una albúmina de tipo silvestre de *Homo sapiens.* La secuencia de la albúmina humana (HSA) de "tipo silvestre" está dada por el número de referencia de GenBank AAA98797.1 (Minghetti *et al.* "Molecular structure of the human albumin gene is revealed by nucleotide sequence within q11-22 of chromosome 4", J. Biol. Chem. 261 (15), 6747-6757 (1986). En la tabla 1 (anterior) se proporcionan ejemplos de albúminas de tipo silvestre.

Convenciones para la designación de las posiciones de aminoácidos

Para los fines de la presente invención, el polipéptido divulgado en la SEQ ID NO: 2 se utiliza para determinar el resto de aminoácido correspondiente en un homólogo de la proteína Gsh1. La secuencia de aminoácidos de un homólogo de la proteína Gsh1 se alinea con el polipéptido divulgado en la SEQ ID NO: 2, y en base a la alineación, el número de la posición de un aminoácido correspondiente a cualquier resto de aminoácido en el polipéptido divulgado en la SEQ ID NO: 2 se determina con el algoritmo Needleman-Wunsch (Needleman y Wunsch, 1970, J. Mol. Biol. 48: 443-453) como se implementó en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice *et al.*, 2000, Trends Genet. 16: 276-277), preferentemente la versión 5.0.0 o posterior. Los parámetros utilizados son penalización por apertura de hueco de 10, penalización por extensión de hueco de 0,5, y la matriz de sustitución EBLOSUM62 (versión EMBOSS de BLOSUM62). Los mismos principios se pueden utilizar en relación con Not4, basado en la SEQ ID NO: 6.

La identificación del resto de aminoácido correspondiente en un homólogo de la proteína Gsh1 se puede determinar mediante una alineación de múltiples secuencias polipeptídicas utilizando varios programas informáticos, incluidos, pero sin limitación, MUSCLE (comparación de secuencias múltiples mediante expectativa logarítmica; versión 3.5 o posterior; Edgar, 2004, Nucleic Acids Research 32: 1792-1797), MAFFT (versión 6.857 o posterior; Katoh y Kuma,

2002, Nucleic Acids Research 30: 3059-3066; Katoh *et al.*, 2005, Nucleic Acids Research 33: 511-518; Katoh y Toh, 2007, Bioinformatics 23: 372-374; Katoh *et al.*, 2009, Methods in Molecular Biology 537: 39-64; Katoh y Toh, 2010, Bioinformatics 26: 1899-1900), y EMBOSS EMMA employing ClustalW (1.83 o posterior; Thompson *et al.*, 1994, Nucleic Acids Research 22: 4673-4680), utilizando sus respectivos parámetros predeterminados. Los mismos principios se pueden utilizar en relación con Not4, basado en la SEQ ID NO: 6.

En la descripción de los polipéptidos de la presente invención, la nomenclatura descrita a continuación está adaptada para facilitar la referencia. Se emplea la abreviatura de aminoácidos de una o tres letras aceptada por la IUPAC.

Sustituciones. Para una sustitución de aminoácido, se utiliza la siguiente nomenclatura: Aminoácido originario, posición, aminoácido sustituido. Por consiguiente, la sustitución de treonina en la posición 226 por alanina se denomina "Thr226Ala" o "T226A". Las mutaciones múltiples se separan por marcas de adición ("+"), por ejemplo, "Gly205Arg + Ser411Phe" o "G205R + S411F", lo que representa sustituciones en las posiciones 205 y 411 de glicina (G) por arginina (R) y serina (S) por fenilalanina (F), respectivamente.

Eliminaciones. Para una eliminación de un aminoácido, se utiliza la siguiente nomenclatura: Aminoácido originario, posición, *. Por consiguiente, la eliminación de glicina en la posición 195 se denomina "Gly195*" o "G195*". Las eliminaciones múltiples están separadas por marcas de adición ("+"), por ejemplo, "Gly195* + Ser411*" o "G195* + S411*".

Inserciones. Como se ha divulgado anteriormente, una inserción puede ser en el lado amino ("cadena arriba", 'X-1') o en el lado carboxilo ('cadena abajo', 'X+1') del aminoácido que ocupa una posición ('el aminoácido nombrado (u originario)', 'X').

Para la inserción de un aminoácido en el lado carboxilo ("cadena abajo", 'X+1') del aminoácido originario ('X'), se utiliza la siguiente nomenclatura: Aminoácido originario, posición, aminoácido originario, aminoácido insertado. Por consiguiente, la inserción de lisina después de glicina en la posición 195 se denomina "Gly195GlyLys" o "G195GK". Una inserción de múltiples aminoácidos se denomina [aminoácido originario, posición, aminoácido originario, aminoácido insertado n.º 1, aminoácido insertado n.º 2; *etc*.]. Por ejemplo, la inserción de lisina y alanina después de la glicina en la posición 195 se indica como "Gly195GlyLysAla" o "G195GKA".

En dichos casos, los restos de aminoácidos insertados se numeran mediante la adición de letras minúsculas al número de posición del resto de aminoácido que precede a los restos de aminoácidos insertados. En el ejemplo anterior, la secuencia sería así:

Precursor:	Variante:
195	195 195a 195b
G	G - K - A

Para una inserción de aminoácido en el lado amino ("cadena arriba", 'X-1') del aminoácido originario (X), se utiliza la siguiente nomenclatura: Aminoácido originario, posición, aminoácido insertado, aminoácido originario. Por consiguiente, la inserción de lisina (K) antes de glicina (G) en la posición 195 se denomina "Gly195LysGly" o "G195KG". Una inserción de múltiples aminoácidos se denomina [aminoácido originario, posición, aminoácido insertado n.º 1, aminoácido insertado n.º 2; *etc*., aminoácido originario]. Por ejemplo, la inserción de lisina (K) y alanina (A) antes de la glicina en la posición 195 se indica como "Gly195LysAlaGly" o "G195KAG". En dichos casos, los restos de aminoácidos insertados se numeran mediante la adición de letras minúsculas al número de posición del resto de aminoácido que sigue a los restos de aminoácidos insertados. En el ejemplo anterior, la secuencia sería así:

Precursor:	Variante:
195	195a' 195b' 195
G	K - A - G

Múltiples alteraciones. Los polipéptidos que comprenden múltiples alteraciones están separados por marcas de adición ("+"), por ejemplo, "Arg170Tyr + Gly195Glu" o "R170Y + G195E" que representa una sustitución de arginina y glicina en las posiciones 170 y 195 por tirosina y ácido glutámico, respectivamente.

Diferentes alteraciones. Cuando se pueden introducir diferentes alteraciones en una posición, las diferentes alteraciones están separadas por una coma, por ejemplo, "Arg170Tyr,Glu" representa una sustitución de arginina en la posición 170 por tirosina o ácido glutámico. Por lo tanto, "Tyr167Gly,Ala + Arg170Gly,Ala" indica las siguientes variantes:
"Tyr167Gly + Arg170Gly", "Tyr167Gly + Arg170Ala", "Tyr167Ala + Arg170Gly" y "Tyr167Ala + Arg170Ala".

Descripción detallada de la invención

Los inventores divulgan una célula hospedadora fúngica que tiene:

a. una proteína Gsh1 modificada que tiene un nivel reducido de actividad en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
b. un nivel de actividad de la proteína Gsh1 modificado en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
c. un gen *GSH1* modificado que tiene un nivel reducido de actividad en relación con la actividad de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
d. un nivel de expresión modificado del gen *GSH1* en relación con el nivel de expresión de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica.

Dicha célula hospedadora fúngica es de interés para los aspectos de la invención que se divulgan a continuación.

La proteína Gsh1 modificada puede modificarse con respecto a una proteína Gsh1 de referencia, tal como una proteína Gsh1 de tipo silvestre, por ejemplo, de SEQ ID NO: 2. Preferentemente, la proteína Gsh1 modificada tiene al menos un 50 % de identidad con la SEQ ID NO: 2, más preferentemente al menos un 55, 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 99, 99,1, 99,2, 99,3, 99,4, 99,5, 99,6, 99,7, 99,8 o al menos 99,9 % de identidad con la SEQ ID NO: 2. La proteína Gsh1 modificada puede o no tener como máximo un 99.9, 99.8, 99.7, 99.6, 99.5, 99.4, 99.3, 99.2, 99.1, 99, 98, 97, 96, 95, 90, 85, 80, 75, 70, 65 o como máximo un 60 % de identidad con la SEQ ID NO: 2. Más preferentemente, la proteína Gsh1 modificada comprende o consiste en la SEQ ID NO: 4.

El nivel de proteína Gsh1 modificado es un nivel reducido de expresión de la proteína Gsh1 y/o un nivel reducido de actividad de la proteína Gsh1. Preferentemente, el nivel reducido es relativo al nivel en una célula hospedadora fúngica de referencia, tal como una célula hospedadora fúngica en la que la proteína Gsh1 comprende o consiste en la SEQ ID NO: 2. La proteína Gsh1 del hospedador fúngico de referencia puede ser una secuencia de Gsh1 de tipo silvestre, tal como la SEQ ID NO: 2. Una célula hospedadora fúngica de referencia adecuada es S288C de *S. cerevisiae* o DXY1 de *S. cerevisiae*. S288C tiene el genotipo *MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6.* DXY1 tiene el genotipo *leu2-3, leu2-122, can1, pra1, ubc4, ura3:yap3* (Kerry-Williams *et al.* (1998) Yeast 14:161-169). Otras células hospedadoras fúngicas de referencia adecuadas incluyen células que son idénticas a la célula hospedadora con la excepción del gen *GSH1* o la proteína Gsh1. Por ejemplo, el gen *GSH1* de la referencia puede ser de tipo silvestre (por ejemplo, SEQ ID NO: 1) o el gen *GSH1* de la referencia puede codificar la proteína Gsh1 de tipo silvestre (por ejemplo, SEQ ID NO: 2) o la proteína Gsh1 codificada por la referencia puede ser de tipo silvestre (por ejemplo, SEQ ID NO: 2). Preferentemente, la célula hospedadora de la invención es idéntica a una cepa precursora con la excepción del gen *GSH1* o la proteína Gsh1. Un hospedador fúngico de referencia también puede denominarse hospedador fúngico "correspondiente". Un hospedador fúngico de referencia puede ser un hospedador fúngico precursor.

Se puede lograr un nivel reducido de proteína Gsh1 y/o nivel de actividad de la proteína Gsh1, por ejemplo, mediante la mutación o eliminación del gen *GSH1*, dando lugar así a una proteína Gsh1 mutada o a una ausencia completa de la proteína Gsh1; mediante la eliminación o cambio del marco de lectura abierto del gen; mediante la mutación o cambio de las secuencias de control del gen *GSH1* tal como una secuencia promotora y/o una secuencia terminadora; mediante el bloqueo o reducción de la transcripción del gen *GSH1*, por ejemplo, mediante la introducción de ARN de interferencia adecuado tal como ARNm antisentido; mediante la introducción, control o modificación de genes activadores transcripcionales adecuados o mediante la introducción de un agente que bloquea el nivel de actividad de la proteína Gsh1. Los métodos para medir los niveles de proteínas y la actividad de las proteínas son bien conocidos en la materia.

El nivel reducido de actividad de la proteína Gsh1 puede dar lugar a de un 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 96, 97 o 98 a un 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 96, 97, 98 o 99 % del nivel de actividad de la proteína Gsh1 de una célula hospedadora fúngica precursora o de referencia, tal como una célula hospedadora fúngica de tipo silvestre. El nivel reducido de actividad de la proteína Gsh1 en una célula hospedadora fúngica puede ser relativo al nivel de actividad de la proteína Gsh1 de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica precursora o una célula hospedadora fúngica de tipo silvestre como se describe anteriormente. Por consiguiente, el nivel de actividad de la proteína Gsh1 en la célula hospedadora puede ser como máximo el 99 % del nivel de actividad de la proteína Gsh1 en una célula hospedadora fúngica de referencia, por ejemplo, como máximo el 98, 97, 96, 95, 90, 80, 70, 60, 50, 40, 30, 20 o como máximo el 10 % del nivel de actividad de la proteína Gsh1 en la célula hospedadora fúngica de referencia. El nivel de actividad de la proteína Gsh1 puede reducirse a cero o sustancialmente cero.

El nivel reducido de expresión (cantidad) de proteína Gsh1 puede dar lugar a de un 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 96, 97, 98 a un 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 96, 97, 98, 99 % del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica de referencia, tal como una célula hospedadora fúngica de tipo silvestre. El nivel reducido de expresión de la proteína Gsh1 en una célula hospedadora fúngica puede ser relativo al nivel de expression de la proteína Gsh1 de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica precursora o una célula hospedadora fúngica de tipo silvestre como se describe anteriormente. Por consiguiente, el nivel de expresión de la proteína Gsh1 en la célula hospedadora puede ser como máximo el 99 %

del nivel de expresión de la proteína Gsh1 en una célula hospedadora fúngica de referencia, por ejemplo, como máximo el 98, 97, 96, 95, 90, 80, 70, 60, 50, 40, 30, 20 o como máximo el 10 % del nivel de expresión de la proteína Gsh1 en la célula hospedadora fúngica de referencia. El nivel de expresión de la proteína Gsh1 puede reducirse a cero o sustancialmente cero.

La célula hospedadora fúngica puede carecer de un gen *GSH1* o proteína Gsh1 funcional. Por ejemplo, la célula hospedadora fúngica puede contener un gen *GSH1* modificado que puede dar como resultado un nivel reducido de expresión de la proteína Gsh1 o un nivel reducido de actividad de la proteína Gsh1. La célula hospedadora fúngica puede carecer de un gen *GSH1*, por ejemplo, debido a una eliminación y/o puede carecer de la proteína Gsh1.

La actividad de *GSH1* puede medirse utilizando un ensayo tal como el divulgado por Volohonsky *et al.*, Chemico-Biological Interactions 140 (2002) 49-65, particularmente la sección 2.8.1. El nivel de expresión de Gsh1 puede medirse, por ejemplo, mediante ELISA para determinar la cantidad de proteína o mediante RT-PCR cuantitativa para medir el nivel de ARN.

La célula hospedadora fúngica puede tener una proteína Gsh1 modificada que comprende una mutación en una posición correspondiente a una posición seleccionada de 47, 48, 49, 50, 51, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 409, 451, 452, 453, 454 y 455 de la SEQ ID NO: 2, preferentemente una posición seleccionada de:

una posición correspondiente a cualquiera de las posiciones 47, 48, 49, 50 y 51 de la SEQ ID NO: 2, preferentemente correspondiente a la posición 49;
una posición correspondiente a cualquiera de las posiciones 120, 121, 122, 123, 124, 125, 126, 127, 128, 129 o 130 de la SEQ ID NO: 2, preferentemente correspondiente a la posición 123, 124, 125, 126 o 127, más preferentemente correspondiente a la posición 125; o una posición correspondiente a la posición 409 de la SEQ ID NO: 2;
una posición correspondiente a cualquiera de las posiciones 451, 452, 453, 454 o 455 de la SEQ ID NO: 2, preferentemente correspondiente a la posición 453.

Mucho más preferentemente, la célula hospedadora fúngica tiene una proteína Gsh1 modificada que comprende una mutación en una posición correspondiente a la posición 125 de la SEQ ID NO: 2.

La mutación puede ser una sustitución, inserción y/o eliminación en una o más (por ejemplo, varias) posiciones. Se prefieren las sustituciones.

La célula hospedadora fúngica puede comprender una secuencia polinucleotídica que codifica la proteína Gsh1 modificada, por ejemplo, la SEQ ID NO: 3. Debido a la degeneración del código genético, otras secuencias polinucleotídicas también pueden codificar proteínas Gsh1 modificadas adecuadas.

Los aminoácidos pertenecen a varias clases bien conocidas. Por lo tanto, algunos aminoácidos están más estrechamente relacionados que otros. Tal como se utiliza en el presente documento, "sustituciones conservadoras de aminoácidos" se refiere a sustituciones realizadas dentro del mismo grupo y que, normalmente, no afectan de manera sustancial a la función de la proteína. Por "sustitución conservadora" se entiende dentro de grupos tales como los que se muestran en la figura 1, que es un diagrama de Venn que proporciona un sistema mediante el cual se puede visualizar el nivel de conservación. En general, las sustituciones de baja conservación son aquellas para las que existen muchos límites (líneas) entre el aminoácido inicial y la sustitución resultante. La "sustitución conservadora de aminoácidos" incluye una sustitución realizada dentro del mismo grupo, tal como dentro de:

aminoácidos aromáticos: F, H, W, Y;
aminoácidos alifáticos: I, L, V;
aminoácidos hidrófobos: A, C, F, H, I, K, L, M, T, V, W, Y;
aminoácidos cargados: D, E, H, K, R, por ejemplo:
aminoácidos con carga positiva: H, K, R; o
aminoácidos con carga negativa: D, E;
aminoácidos polares: C, D, E, H, K, N, Q, R, S, T, W, Y;
aminoácidos pequeños: A, C, D, G, N, P, S, T, V, por ejemplo:
aminoácidos minúsculos: A, C, G, S.

De manera alternativa, "sustitución conservadora" puede estar dentro de los siguientes grupos:

aminoácidos que tienen cadenas laterales alifáticas: G, A, V, L, I;
aminoácidos que tienen cadenas laterales aromáticas: F, Y, W;
aminoácidos que tienen cadenas laterales que contienen azufre: C, M;
aminoácidos que tienen cadenas laterales de hidroxilos alifáticos: S, T;
aminoácidos que contienen cadenas laterales básicas: K, R, H;
aminoácidos ácidos y sus derivados amídicos: D, E, N, Q.

Las sustituciones pueden hacerse mediante técnicas conocidas en la materia, tal como mediante mutagénesis dirigida al sitio como se divulga en la patente de EE. UU. N.º 4.302.386.

Las sustituciones de amino no conservadoras pueden referirse a sustituciones realizadas de un grupo a otro grupo, por ejemplo, del grupo que tiene cadenas laterales aromáticas al grupo que tiene cadenas laterales alifáticas.

La célula hospedadora fúngica puede comprender una proteína Gsh1 modificada en la que, en relación a la SEQ ID NO: 2, la mutación es una sustitución de un aminoácido, preferentemente un aminoácido no conservado, seleccionado de A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y.

La mutación en una posición correspondiente a la posición 125 de la SEQ ID NO: 2 puede ser una sustitución del aminoácido natural, tal como R, por un aminoácido no natural, tal como A, C, D, E, F, G, H, I, L, M, N, P, Q, S, T, V, W o Y, preferentemente por C, D, E o G, más preferentemente a G. La sustitución puede ser de un aminoácido con carga positiva por un aminoácido alifático, aromático, hidrófobo, pequeño, minúsculo, polar o con carga negativa, preferentemente un aminoácido con carga negativa por un aminoácido minúsculo. Una sustitución particularmente preferida es R por G. La sustitución por K es menos preferida.

La mutación en una posición correspondiente a la posición 49 de la SEQ ID NO: 2 puede ser una sustitución del aminoácido natural, tal como D, por un aminoácido no natural, tal como A, C, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y, preferentemente por H, K o R, más preferentemente por K o R. La sustitución puede ser de un aminoácido con carga negativa por un aminoácido con carga positiva. La sustitución por E es menos preferida.

La mutación en una posición correspondiente a la posición 409 de la SEQ ID NO: 2 puede ser una sustitución del aminoácido natural, tales como H, por un aminoácido no natural, tal como A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W o Y, preferentemente por A, I, L, V, M, F, W, Y, más preferentemente por A, I, L, V, lo más preferentemente por L. La sustitución puede ser de un aminoácido aromático por un aminoácido alifático. Una sustitución particularmente preferida es H por L.

La mutación en una posición correspondiente a la posición P453 de la SEQ ID NO: 2 puede ser una sustitución del aminoácido natural, tal como P, por un aminoácido no natural, tal como A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W o Y, preferentemente por A, I, L, V, M, F, W, Y, más preferentemente por A, I, L, V, lo más preferentemente por L. La sustitución puede ser de un aminoácido aromático por un aminoácido alifático. Una sustitución particularmente preferida es P por L.

Una proteína Gsh1 modificada preferida incluye una mutación en una posición correspondiente a R125 de la SEQ ID NO: 2.

Una proteína Gsh1 modificada preferida comprende o consiste en la SEQ ID NO: 4, es decir, que incluye la mutación R125G. Dicha proteína Gsh1 modificada puede proporcionarse utilizando una secuencia de nucleótidos que codifica una proteína Gsh1 (por ejemplo, la SEQ ID NO: 1), incluyendo la secuencia de nucleótidos el polimorfismo de un solo nucleótido (SNP, del inglés "single nucleotide polymorphism") A373G y que puede proporcionarse utilizando, por ejemplo, la SEQ ID NO: 3.

Se propone que la posición R125 participe en una interacción de puente salino con la posición D49. La mutación de cualquiera de los aminoácidos se podría utilizar para interrumpir esta interacción. La mutación de R125 a D o E o de D49 a R o K podría dar como resultado efectos de repulsión electrostática que se esperaría que desestabilizaran la interacción entre las posiciones 125 y 49. Sin embargo, la mutación de R125 a K y/o D49 a E (glutamato) puede mantener la interacción.

La proteína Gsh1 puede ser exógena a la célula hospedadora o puede ser endógena a la célula hospedadora. Cuando la proteína Gsh1 es exógena a la célula hospedadora, la célula hospedadora puede mantener o carecer de una proteína Gsh1 que es endógena a la célula hospedadora.

El hospedador fúngico puede tener además:

una proteína Not4 modificada, y/o
un nivel modificado de actividad de la proteína Not4, y/o
un gen *NOT4* modificado, y/o
un nivel modificado de expresión del gen *NOT4*.

NOT4 también se conoce como *MOT2*. La proteína Not4 modificada puede modificarse en relación con una proteína Not4 de referencia, tal como una proteína Not4 de tipo silvestre, por ejemplo, la SEQ ID NO: 6. Preferentemente, la proteína Not4 modificada tiene al menos un 70 % de identidad con la SEQ ID NO: 6, más preferentemente al menos un 75, 80, 85, 90, 95, 96, 97, 98, 99, 99,1, 99,2, 99,3, 99,4, 99,5, 99,6, 99,7, 99,8 o al menos un 99,9 % de identidad con la SEQ ID NO: 6. La proteína Not4 modificada puede o no tener como máximo un 99.9, 99.8, 99.7, 99.6, 99.5, 99.4, 99.3, 99.2, 99.1, 99, 98, 97, 96, 95, 90, 85, 80, 75, 70, 65 o como máximo un 60 % de identidad con la SEQ ID

NO: 6. Más preferentemente, la proteína Not4 modificada comprende o consiste en la SEQ ID NO: 8.

Se prefiere que el nivel modificado de proteína Not4 sea un nivel reducido de expresión de la proteína Not4 o un nivel reducido de actividad de la proteína Not4. Preferentemente, el nivel modificado, por ejemplo, reducido, es relativo al nivel en una célula hospedadora fúngica de referencia, tal como una célula hospedadora fúngica en la que la proteína Not4 comprende o consiste en la SEQ ID NO: 6. La proteína Not4 del hospedador fúngico de referencia puede ser una secuencia de Not4 de tipo silvestre, tal como la SEQ ID NO: 6. Una célula hospedadora fúngica de referencia adecuada es S288C de *S. cerevisiae* o DXY1 de *S. cerevisiae*, como se describe anteriormente. Otras células hospedadoras fúngicas de referencia adecuadas incluyen células que son idénticas a la célula hospedadora con la excepción del gen *NOT4* o la proteína Not4. Por ejemplo, el gen *NOT4* de la referencia puede ser de tipo silvestre (por ejemplo, SEQ ID NO: 5) o el gen *NOT4* de la referencia puede codificar la proteína Not4 de tipo silvestre (*por ejemplo*, SEQ ID NO: 6) o la proteína Not4 codificada por la referencia puede ser de tipo silvestre (*por ejemplo*, SEQ ID NO: 6). Preferentemente, la célula hospedadora de la invención es idéntica a una cepa precursora con la excepción del gen *NOT4* o la proteína Not4. Un hospedador fúngico de referencia también puede denominarse hospedador fúngico "correspondiente". Un hospedador fúngico de referencia puede ser un hospedador fúngico precursor.

Se puede lograr un nivel de proteína Not4 o un nivel de actividad de la proteína Not4 reducido, por ejemplo, mediante la mutación o eliminación del gen *NOT4*, dando lugar así a una proteína Not4 mutada o a una ausencia completa de la proteína Not4; mediante la eliminación o cambio del marco de lectura abierto del gen, mediante la mutación o cambio de las secuencias de control del gen *NOT4* tal como una secuencia promotora y/o una secuencia terminadora; mediante el bloqueo o reducción de la transcripción del gen *NOT4*, por ejemplo, mediante la introducción de ARN de interferencia adecuado tal como ARNm antisentido, mediante la introducción, control o modificación de genes activadores transcripcionales adecuados o mediante la introducción de un agente que bloquea el nivel de actividad de la proteína Not4. Los métodos para medir los niveles de proteína son bien conocidos en la materia.

El nivel de actividad modificado de la proteína Not4 puede reducirse, dando así como resultado de un 0, 10, 20, 30, 40, 50, 60, 70, 80 o 90 a un 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 96, 97 o 98 % del nivel de actividad de la proteína Not4 de una célula hospedadora fúngica precursora o de referencia, tal como una célula hospedadora fúngica de tipo silvestre. El nivel de actividad modificado, por ejemplo, reducido, de la proteína Not4 en una célula hospedadora fúngica puede ser relativo al nivel de actividad de la proteína Not4 de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica precursora o una célula hospedadora fúngica de tipo silvestre como se describe anteriormente. Por consiguiente, el nivel de actividad de la proteína Not4 en la célula hospedadora puede ser como máximo el 99 % del nivel de actividad de la proteína Not4 en una célula hospedadora fúngica de referencia, por ejemplo, como máximo el 98, 97, 96, 95, 90, 80, 70, 60, 50, 40, 30, 20 o como máximo el 10 % del nivel de actividad de la proteína Not4 en la célula hospedadora fúngica de referencia. El nivel de actividad de la proteína Not4 puede reducirse a cero o sustancialmente cero.

El nivel de expresión modificado (cantidad) de la proteína Not4 puede reducirse, dando así como resultado de un 0, 10, 20, 30, 40, 50, 60, 70, 80 o 90 a un 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 96, 97 o 98 % del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica de referencia, tal como una célula hospedadora fúngica de tipo silvestre. El nivel de expresión modificado, por ejemplo, reducido, de la proteína Not4 en una célula hospedadora fúngica puede ser relativo al nivel de expresión de la proteína Not4 de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica precursora o una célula hospedadora fúngica de tipo silvestre como se describe anteriormente. Por consiguiente, el nivel de expresión de la proteína Not4 en la célula hospedadora puede ser como máximo el 99 % del nivel de expresión de la proteína Not4 en una célula hospedadora fúngica de referencia, por ejemplo, como máximo el 98, 97, 96, 95, 90, 80, 70, 60, 50, 40, 30, 20 o como máximo el 10 % del nivel de expresión de la proteína Not4 en la célula hospedadora fúngica de referencia. El nivel de expresión de la proteína Not4 puede reducirse a cero o sustancialmente cero.

La célula hospedadora fúngica puede carecer de un gen *NOT4* o proteína Not4 funcional. Por ejemplo, la célula hospedadora fúngica puede contener un gen *NOT4* modificado que puede dar como resultado un nivel reducido de expresión de la proteína Not4 o un nivel reducido de actividad de la proteína Not4. La célula hospedadora fúngica puede carecer de un gen *NOT4*, por ejemplo, debido a una eliminación y/o puede carecer de la proteína Not4.

El nivel de expresión de Not4 puede medirse, por ejemplo, mediante ELISA para determinar la cantidad de proteína o mediante RT-PCR cuantitativa para medir el nivel de ARN.

La proteína Not4 modificada puede mutarse de modo que se altere su interacción con la proteína Not1. Por ejemplo, la región aminoterminal de la proteína Not4 puede estar mutada, tal como los aminoácidos que contienen hélice α correspondientes a las posiciones 426 a 439 de la SEQ ID NO: 6.

Por lo tanto, la invención también proporciona una célula hospedadora fúngica que además tiene una proteína Not4 que tiene una interacción más débil, tal como una interacción hidrófoba, con Not1 que la interacción entre una proteína Not4 de tipo silvestre (por ejemplo, SEQ ID NO: 6) y una proteína Not1 de tipo silvestre (*por ejemplo*, SEQ

ID NO: 13).

La célula hospedadora fúngica puede tener una proteína Not4 modificada que comprende una mutación en una posición correspondiente a una posición seleccionada de 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469 o 470 de la SEQ ID NO: 6, preferentemente una posición seleccionada de:

una posición correspondiente a 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438 o 439 de la SEQ ID NO: 6, preferentemente la posición 429, 430, 434 o 437, más preferentemente la posición 429;
una posición correspondiente a 460, 461, 462, 463, 464, 465, 466, 467, 468, 469 o 470 de la SEQ ID NO: 6, preferentemente la posición 463, 464 o 466; o
una posición correspondiente a 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455 o 456 de la SEQ ID NO: 6, preferentemente la posición 442, 445, 447 o 452.

La mutación puede ser una sustitución, inserción y/o eliminación en una o más (por ejemplo, varias) posiciones. Se prefieren las sustituciones.

La célula hospedadora fúngica puede comprender una secuencia polinucleotídica que codifica la proteína Not4 modificada, por ejemplo, la SEQ ID NO: 7. Debido a la degeneración del código genético, otras secuencias polinucleotídicas también pueden codificar proteínas Not4 modificadas adecuadas.

La célula hospedadora fúngica puede comprender una proteína Not4 modificada en la que, en relación a la SEQ ID NO: 6, la mutación es una sustitución de un aminoácido, preferentemente un aminoácido no conservado, seleccionado de A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W e Y.

La mutación en una posición correspondiente a la posición 429 de la SEQ ID NO: 6 puede ser una sustitución del aminoácido natural, tales como F, por un aminoácido no natural, tal como A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y, preferentemente por G, A, V, L o I, más preferentemente por V, L o I, lo más preferentemente por I. La sustitución puede ser por un aminoácido no conservado. La sustitución puede ser por un aminoácido alifático. Una sustitución particularmente preferida es F por I.

La mutación en una posición correspondiente a la posición 430 de la SEQ ID NO: 6 puede ser una sustitución del aminoácido natural, tal como L, por cualquier aminoácido no natural tal como A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y. La sustitución puede ser por un aminoácido no conservado.

La mutación en una posición correspondiente a la posición 434 puede ser una sustitución del aminoácido natural, tal como L, por cualquier aminoácido no natural tal como A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y. La sustitución puede ser por un aminoácido no conservado.

La mutación en una posición correspondiente a la posición 437 de la SEQ ID NO: 6 puede ser una sustitución del aminoácido natural, tal como L, por cualquier aminoácido no natural tal como A, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y. La sustitución puede ser por un aminoácido no conservado.

Una proteína Not4 modificada preferida incluye una mutación en una posición correspondiente a F429 de la SEQ ID NO: 6.

Una proteína Not4 modificada preferida comprende o consiste en la SEQ ID NO: 8, es decir, que incluye la mutación F429I. Dicha proteína Not4 modificada puede proporcionarse utilizando una secuencia de nucleótidos que codifica una proteína Not4 (por ejemplo, SEQ ID NO: 5), incluyendo la secuencia de nucleótidos el SNP T1285A, por ejemplo, la SEQ ID NO: 7.

De manera alternativa, el nivel modificado puede aumentarse. Es probable que un nivel aumentado o una actividad aumentada de la proteína Not4 disminuyan el rendimiento de la proteína deseada (tal como una proteína heteróloga). Dicha disminución del rendimiento puede ser deseable cuando, por ejemplo, la proteína deseada es perjudicial para la viabilidad de la célula hospedadora. Un nivel aumentado puede ser al menos un 101, 102, 103, 104, 105, 110, 120, 130, 140, 150, 175 o 200 % del nivel en un hospedador de referencia, tal como un hospedador precursor.

La proteína Not4 puede ser exógena a la célula hospedadora o puede ser endógena a la célula hospedadora. Cuando la proteína Not4 es exógena a la célula hospedadora, la célula hospedadora puede mantener o carecer de una proteína Not4 que es endógena a la célula hospedadora.

El hospedador fúngico puede tener (1) un nivel reducido de la proteína Gsh1 y/o un nivel reducido de actividad de la proteína Gsh1 y/o un gen *GSH1* modificado que reduce el nivel de actividad del gen *GSH1*, y/o un nivel reducido de expresión del gen *GSH1* y (2) un nivel modificado de proteína Not4 y/o un nivel modificado de actividad de la

proteína Not4, y/o un gen *NOT4* modificado y/o un nivel modificado de expresión del gen *NOT4*. Preferentemente, el hospedador fúngico tiene (1) un nivel reducido de la proteína Gsh1 o un nivel reducido de actividad de la proteína Gsh1 y/o un nivel reducido de expresión del gen *GSH1* y (2) un nivel reducido de la proteína Not4 o un nivel reducido de actividad de la proteína Not4 y/o un nivel reducido de expresión del gen *NOT4*. De manera alternativa, el hospedador fúngico puede tener (1) un nivel reducido de proteína Gsh1 o un nivel reducido de actividad de la proteína Gsh1 y/o un nivel reducido de expresión del gen *GSH1* y (2) un nivel aumentado de la proteína Not4 o un nivel aumentado de actividad de la proteína Not4 y/o un nivel aumentado de expresión del gen *NOT4*. Las mismas opciones se aplican con respecto a los niveles de expresión de *GSH1* y *NOT4.* 'Aumentado' y 'reducido' son como se describe en el presente documento.

Preferentemente, el nivel reducido de Gsh1 es relativo al nivel en una célula hospedadora fúngica de referencia, tal como una célula hospedadora fúngica en la que la proteína Gsh1 comprende o consiste en la SEQ ID NO: 2 y el nivel modificado, por ejemplo, reducido, de Not4 es relativo al nivel en una célula hospedadora fúngica de referencia, tal como una célula hospedadora fúngica en la que la proteína Not4 comprende o consiste en la SEQ ID NO: 6. La proteína Gsh1 del hospedador fúngico de referencia puede ser una secuencia de Gsh1 de tipo silvestre, tal como la SEQ ID NO: 2. La proteína Not4 del hospedador fúngico de referencia puede ser una secuencia de Not4 de tipo silvestre, tal como la SEQ ID NO: 6. Una célula hospedadora fúngica de referencia adecuada es S288C de *S. cerevisiae* o DXY1 de *S. cerevisiae* como se describe anteriormente. Otras células hospedadoras fúngicas de referencia adecuadas incluyen células que son idénticas a la célula hospedadora con la excepción del gen *GSH1* gen o la proteína Gsh1 y el gen *NOT4* o la proteína Not4. Por ejemplo, el gen *GSH1* de la referencia puede ser de tipo silvestre (*por ejemplo*, SEQ ID NO: 1) o el gen *GSH1* de la referencia puede codificar la proteína Gsh1 de tipo silvestre (*por ejemplo*, SEQ ID NO: 2) o la proteína Gsh1 codificada por la referencia puede ser de tipo silvestre (por ejemplo, SEQ ID NO: 2). El gen *NOT4* de la referencia puede ser de tipo silvestre (por ejemplo, SEQ ID NO: 5) o el gen *NOT4* de la referencia puede codificar la proteína Not4 de tipo silvestre (por ejemplo, SEQ ID NO: 6) o la proteína Not4 codificada por la referencia puede ser de tipo silvestre (*por ejemplo*, SEQ ID NO: 6). Preferentemente, la célula hospedadora de la invención es idéntica a una cepa precursora con la excepción del gen *GSH1* o la proteína Gsh1 y el gen *NOT4* o la proteína Not4. Un hospedador fúngico de referencia también puede denominarse hospedador fúngico "correspondiente". Un hospedador fúngico de referencia puede ser un hospedador fúngico precursor.

La célula hospedadora fúngica puede ser una célula de hongo filamentoso.

La célula hospedadora fúngica puede ser una levadura o un hongo filamentoso. "Hongos", como se usa en el presente documento, incluye los filos *Ascomycota*, *Basidiomycota*, *Chytridiomycota* y *Zygomycota* (según define Hawksworth *et al.*, en, Ainsworth and Bisby's Dictionary of The Fungi, 8.ª edición, 1995, CAB International, University Press, Cambridge, Reino Unido) así como el Oomycota (como se cita en Hawksworth *et al.*, 1995, *citado anteriormente,* página 171) y todos los hongos mitospóricos (Hawksworth *et al.*, 1995, citado anteriormente).

Preferentemente, la célula hospedadora fúngica es una célula de levadura. "Levadura", como se usa en el presente documento, incluye levadura ascoesporógena (Endomycetales), levadura basidiosporógena y levadura que pertenece a los hongos imperfectos (Blastomicetos). Puesto que la clasificación de las levaduras puede cambiar en el futuro, para los fines de la presente invención, la levadura se definirá como se describe en *Biology and Activities of Yeast* (Skinner, F. A., Passmore, S. M. y Davenport, R. R., eds, Soc. App. Bacteriol. n.º de serie de simposios: 9, 1980, páginas 1 a 27).

Más preferentemente, la célula hospedadora de levadura es una célula de *Candida*, *Hansenula, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces* o *Yarrowia*.

Incluso más preferentemente, la célula hospedadora de levadura es una célula de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis, Saccharomyces oviformis, Kluyveromyces lactis* o *Yarrowia lipolytica*. Se prefiere particularmente un hospedador de *Saccharomyces cerevisiae*.

El hospedador *S. cerevisiae* puede comprender o no una o más de las siguientes características genotípicas: *leu2-3, leu2-122, can1, pra1, ubc4, ura3, yap3::URA3, lys2, hsp150::LYS2, pmt1::URA3* (como se define en el documento WO2014/138371), por ejemplo, BXP10 de *S. cerevisiae*. Preferentemente el hospedador *S. cerevisiae* incluye *MATa.*

El hospedador *S. cerevisiae* puede comprender o no uno o más de los siguientes genotipos, *MATa, leu2-3, leu2-112, ubc4, ura3, yap3::URA3, lys2, hsp150::LYS2;* con *PDI1*, *URA3 y* Ylplac211 integrados en el locus *PDI1* (Finnis *et al.*, 2010, Microbial Cell Factories 9:87), por ejemplo, DP9 de *S. cerevisiae*.

El hospedador *S. cerevisiae* puede comprender o no uno o más de los siguientes genotipos, *MATα, leu2, pep4-3,* por ejemplo, MT302/28B de *S. cerevisiae* como se describe en Finnis *et al.*, 1993, Eur. J. Biochem, 212: 201-210.

El hospedador *S. cerevisiae* puede o no comprender el siguiente genotipo: *MATα, SUC2, gal2, mal2, mel, flo1, flo8-1, hap1, ho, bio1, bio6* (Mortimer y Johnston (1986) Genetics 113:35-43), por ejemplo, S288C de *S. cerevisiae*.

La cepa hospedadora preferida de *S. cerevisiae* comprende o consiste en todos los *MATa, leu2-3, leu2-122, can1, pra1, ubc4, ura3, yap3::URA3, lys2, hsp150::LYS2 y pmt1::URA3.*

Otro hospedador preferido *S. cerevisiae* comprende o consiste en todo: *MATα, leu2-3, leu2-112, ubc4, ura3, yap3::URA3, lys2, hsp150::LYS2,* con *PDI1*, *URA3 y* Ylplac211 integrados en el locus *PDI1*.

Otro hospedador preferido *S. cerevisiae* comprende o consiste en todo: *MATα, SUC2, gal2, mal2, mel, flo1, flo8-1, hap1, ho, bio1, bio6.*

Otro hospedador preferido *S. cerevisiae* comprende o consiste en todo: *MATα, leu2, pep4-3.*

El hospedador puede ser poliploide, diploide o haploide. Se prefiere una levadura hospedadora haploide o diploide, preferentemente haploide.

El tipo de apareamiento del hospedador puede ser, por ejemplo, *MATa* o *MATα* (Mat-alfa). Preferentemente el hospedador *S. cerevisiae* contiene un plásmido que codifica la albúmina humana o una variante, fragmento y/o fusión de la misma.

Los "hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión Eumycota y Oomycota (según define Hawksworth *et al.*, 1995, citado anteriormente). Los hongos filamentosos se caracterizan por una pared micelial compuesta por quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos. El crecimiento vegetativo es por alargamiento hifal y el catabolismo del carbono es obligatoriamente aeróbico. Por el contrario, el crecimiento vegetativo de levaduras tales como *Saccharomyces cerevisiae* es por gemación de un talo unicelular y el catabolismo del carbono puede ser fermentativo.

Las células hospedadoras de hongos filamentosos preferidas pueden o no incluir *Acremonium, Aspergillus, Aureobasidium, Bjerkandera, Ceriporiopsis, Chrysosporium, Coprinus, Coriolus, Cryptococcus, Filibasidium, Fusarium, Humicola, Magnaporthe, Mucor, Myceliophthora, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Phanerochaete, Phlebia, Piromyces, Pleurotus, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, Trametes* o *Trichoderma.*

La célula hospedadora fúngica, puede comprender una secuencia de nucleótidos que codifica una proteína deseada. Preferentemente, la proteína deseada es una proteína heteróloga. Una proteína heteróloga es aquella que no se produce de manera natural por la célula hospedadora y, preferentemente, no incluye proteínas tales como marcadores seleccionables, por ejemplo, marcadores de resistencia a antibióticos o marcadores auxótrofos, chaperonas, FLP o FRT.

La célula hospedadora fúngica puede ser un hospedador de expresión. La célula hospedadora fúngica puede comprender un casete de expresión, por ejemplo, que codifica una proteína deseada tal como una proteína heteróloga. El casete de expresión puede estar, por ejemplo, dentro de un vector tal como un plásmido. La célula hospedadora fúngica puede comprender un vector de expresión.

La proteína deseada puede ser o no una proteína vegetal o animal o una variante de la misma. La proteína deseada puede, o no, comprender la secuencia de la albúmina, un anticuerpo monoclonal, un etopósido, una proteína sérica (tal como un factor de coagulación de la sangre), antiestasina, un péptido anticoagulante de garrapatas, transferrina, lactoferrina, endostatina, angiostatina, colágenos, inmunoglobulinas o moléculas basadas en inmunoglobulinas o fragmentos de las mismas (*por ejemplo*, un Small Modular ImmunoPharmaceutical™ ("SMIP") o dAb, fragmentos Fab', F(ab')2, scAb, scFv o fragmento scFv), una proteína con dominio Kunitz (tal como las descritas en el documento WO03/066824, con o sin fusiones de albúmina), interferones, interleucinas, IL-10, IL-11, IL-2, especies y subespecies de interferón α (alfa), especies y subespecies de interferón β (beta), especies y subespecies de interferón γ (gamma), leptina, CNTF, $CNTF_{Ax15}$, antagonista del receptor de IL-1, eritropoyetina (EPO) e imitadores de EPO, trombopoyetina (TPO) e imitadores de TPO, prosaptida, cianovirina-N, 5 hélices, péptido T20, péptido T1249, gp41 del VIH, gp120 del VIH, urocinasa, prourocinasa, tPA, hirudina, factor de crecimiento procedente de plaquetas, hormona paratiroidea, proinsulina, insulina, glucagón, péptidos similares al glucagón tales como exendina-4, GLP-1 o GLP-2, factor de crecimiento de tipo insulina, calcitonina, hormona del crecimiento, factor de crecimiento transformante β (beta), factor de necrosis tumoral, G-CSF, GM-CSF, M-CSF, FGF, factores de coagulación en formas tanto pre como activas, que incluyen, pero sin limitación, plasminógeno, fibrinógeno, trombina, pretrombina, protrombina, factor von Willebrand, antitripsina α_1, activadores de plasminógeno, factor VII, factor VIII, factor IX, factor X y factor XIII, factor de crecimiento nervioso, LACI, factor de crecimiento de células endoteliales procedente de plaquetas (PD-ECGF), glucosa oxidasa, colinesterasa sérica, aprotinina, proteína precursora de amiloide, inhibidor de intertripsina α, antitrombina III, especies de apo-lipoproteínas, proteína C, proteína S, un metabolito, un antibiótico o una variante o fragmento de cualquiera de los anteriores.

Preferentemente, la variante tiene al menos un 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99 % de identidad con una o más de las proteínas divulgadas anteriormente.

Una proteína deseada preferida puede o no ser una proteína sérica tal como una albúmina o una variante, fragmento y/o fusión de la misma. Preferentemente, la albumina tiene del 60, 65, 70, 75, 80, 85, 90, 95, 96, 97, 98, 98.2, 98.4, 98.6, 98.8, 99, 99.2, 99.3, 99.4, 99.5, 99.6, 99.7, 99.8, o 99,9 al 70, 75, 80, 85, 90, 95, 96, 97, 98, 98,2, 98,4, 98,6, 98,8, 99, 99,2, 99,3, 99,4, 99,5, 99,6, 99,7, 99,8, 99,9 o 100 % de identidad de secuencia con la SEQ ID NO: 10. Mucho más preferentemente, la albúmina comprende o consiste en la SEQ ID NO: 10.

La variante de albúmina, fragmento y/o fusión del mismo puede o no comprender de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18 o 19 a 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19 o 20 mutaciones en relación con la SEQ ID NO: 10. Las variantes de albúmina preferidas comprenden de 1 a 10 mutaciones en relación con la SEQ ID NO: 10, más preferentemente de 1 a 5 mutaciones. Las mutaciones preferidas incluyen sustituciones.

La variante de albúmina, fragmento y/o fusión de la misma puede o no comprender A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W o Y en una posición correspondiente a la posición K573 de la SEQ ID NO: 10, más preferentemente una P, H, W o Y en una posición correspondiente a la posición K573 de la SEQ ID NO: 10. Las variantes de albúmina particularmente preferidas tienen al menos un 95 % de identidad con la SEQ ID NO: 10 (más preferentemente al menos un 96, 97, 98 o 99 % de identidad) y comprenden P en una posición correspondiente a 573 de la SEQ ID NO: 10.

La variante de albúmina, fragmento y/o fusión de la misma puede o no comprender A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y en una posición correspondiente a la posición E492 de la SEQ ID NO: 10, más preferentemente una G, D, F, H, M o R en una posición correspondiente a la posición E492 de la SEQ ID NO: 10, incluso más preferentemente una G o D en una posición correspondiente a la posición E492 de la SEQ ID NO: 10. Las variantes de albúmina particularmente preferidas tienen al menos un 95 % de identidad con la SEQ ID NO: 10 (más preferentemente al menos un 96, 97, 98 o 99 % de identidad) y comprenden G en una posición correspondiente a E492 de la SEQ ID NO: 10.

La variante de albúmina, fragmento y/o fusión de la misma puede o no comprender A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W o Y en una posición correspondiente a la posición K574 de la SEQ ID NO: 10, más preferentemente una H, G, D, F, N, S o Y en una posición correspondiente a la posición K574 de la SEQ ID NO: 10, aún más preferentemente una D, F, G o H en una posición correspondiente a la posición K574 de la SEQ ID NO: 10. Las variantes de albúmina particularmente preferidas tienen al menos un 95 % de identidad con la SEQ ID NO: 10 (más preferentemente al menos un 96, 97, 98 o 99 % de identidad) y comprenden H en una posición correspondiente a K574 de la SEQ ID NO: 10.

La variante de albúmina, fragmento y/o fusión de la misma puede o no comprender A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W o Y en una posición correspondiente a la posición Q580 de la SEQ ID NO: 10, más preferentemente una K o R en una posición correspondiente a la posición Q580 de la SEQ ID NO: 10. Las variantes de albúmina particularmente preferidas tienen al menos un 95 % de identidad con la SEQ ID NO: 10 (más preferentemente al menos un 96, 97, 98 o 99 % de identidad) y comprenden K en una posición correspondiente a Q580 de la SEQ ID NO: 10.

Las variantes de albúmina preferidas pueden comprender o no mutaciones en una o más posiciones seleccionadas entre las correspondientes a las posiciones 492, 573, 574 y 580, por ejemplo, las sustituciones descritas anteriormente. Por ejemplo, la variante de albúmina puede comprender o consistir en la SEQ ID NO: 45.

Otras variantes de albúmina preferidas, fragmentos y/o fusiones de la misma incluyen las divulgadas en los documentos WO2011/051489, WO2011/124718, WO2012/059486, WO2012/150319, WO2014/072481, WO2013/135896, WO2015/036579, WO2010/092135, WO2013/075066, WO2014/179657, WO2009/126920, WO2010/059315, WO2011/103076, WO2012/112188, WO2015/063611 y WO 2017/029407 o fragmentos o fusiones de la misma.

La albúmina puede ser o no un fragmento de una albúmina o una variante de la misma.

La variante de albúmina, fragmento y/o fusión de la misma puede tener una afinidad de unión a FcRn que es más fuerte o más débil (y, preferentemente, es más fuerte) que la de la albúmina, fragmento y/o fusión de la misma precursora.

La variante de albúmina, fragmento y/o fusión de la misma puede tener una KD para FcRn (por ejemplo, shFcRn) que es menor que la KD correspondiente para la HSA o conjugado de la misma. Preferentemente, la KD para la variante de albúmina, fragmento y/o fusión de la misma es inferior a 0,9× la KD para la HSA a FcRn, más preferente inferior a 0,5× la KD para la HSA a FcRn, más preferente inferior a 0,1× la KD para la HSA a FcRn, aún más preferente inferior a 0,05× la KD para la HSA a FcRn, aún más preferente inferior a 0,02× la KD para la HSA a FcRn, aún más preferente inferior a 0,01× la KD para la HSA a FcRn y lo más preferente inferior a 0,001× la KD para la HSA a FcRn (donde × significa "multiplicado por"). Una KD inferior corresponde a una afinidad de unión más fuerte.

La variante de albúmina, fragmento y/o fusión de la misma puede tener una KD para FcRn que sea más elevada que la correspondiente KD para la HSA o conjugado de la misma a FcRn. Preferentemente, la KD para la variante de albúmina, fragmento y/o fusión de la misma es superior a 2× la KD para la HSA a FcRn, más preferente superior a 5× la KD para la HSA a FcRn, más preferente superior a 10× la KD para la HSA a FcRn, aún más preferente superior a 25× la KD para la HSA a FcRn, lo más preferente superior a 50× la KD para la HSA a FcRn. La variante de albúmina, fragmento y/o fusión de la misma puede ser un enlazador nulo para FcRn. Una KD más elevada corresponde a una afinidad de unión más débil.

Cuando se determina y/o compara la KD, se pueden utilizar uno o más (por ejemplo, varios) (y preferentemente todos) de los siguientes parámetros:

Instrumento: Instrumento Biacore 3000 (GE Healthcare)
Cubeta de lectura: Chip sensor CM5
FcRn: FcRn humano, preferentemente FcRn humano soluble, de manera opcional acoplado a un marcador tal como glutatión S transferasa (GST) o histidina (His), lo más preferentemente His tal como 6 restos de histidina en el extremo carboxiterminal de la microglobulina β_2.
Cantidad de FcRn: 1200 a 2500 UR
Química de acoplamiento: química de acoplamiento de amina (por ejemplo, como se describe en el protocolo proporcionado por el fabricante del instrumento).
Método de acoplamiento: El acoplamiento se puede realizar mediante la inyección de 20 µg/ml de la proteína en acetato de sodio 10 mM pH 5,0 (GE Healthcare). El tampón de fosfato (tampón de fosfato 67 mM, NaCl 0,15 M, Tween 20 al 0,005 %) a pH 5,5 se puede utilizar como tampón de ejecución y tampón de dilución. La regeneración de las superficies se puede realizar mediante inyecciones de tampón HBSEP (HEPES 0,01 M, NaCl 0,15 M, EDTA 3 mM, tensioactivo P20 al 0,005 %) a pH 7,4 (Biacore AB).
Cantidad de inyección de la molécula de prueba (por ejemplo, HSA o variante) de 20 a 0,032 µM.
Caudal de inyección: constante, por ejemplo, 30 µl/ml.
Temperatura de inyección: 25 °C.
Programa informático para la evaluación de datos: Programa informático BIAevaluation 4.1 (BIAcore AB).

La variante de albúmina, fragmento y/o fusión de la misma puede tener una semivida en plasma que es más larga o más corta, preferentemente más larga, que la de la albúmina, fragmento y/o fusión de la misma precursora.

La semivida en plasma se determina de manera ideal *in vivo* en individuos adecuados. Sin embargo, dado que lleva mucho tiempo y es costoso e inevitablemente existen preocupaciones éticas relacionadas con la realización de experimentos en animales o en el ser humano, es conveniente utilizar un ensayo *in vitro* para determinar si la semivida en plasma se prolonga o se reduce. Se sabe que la unión de la albúmina a su receptor (FcRn) es importante para la semivida en plasma y la correlación entre la unión al receptor y la semivida en plasma es que una mayor afinidad de la albúmina por su receptor conduce a una semivida en plasma más larga. Por lo tanto, para la invención, una mayor afinidad de la albúmina por FcRn se considera indicativa de una semivida en plasma aumentada y una afinidad menor de la albúmina por su receptor se considera indicativa de una semivida en plasma reducida.

La unión de la albúmina a su receptor FcRn puede describirse utilizando el término afinidad y las expresiones "más fuerte" o "más débil". Por lo tanto, debería comprenderse que, se considera que una molécula que tiene una afinidad más elevada por FcRn que la afinidad de la HSA por FcRn se une de manera más fuerte a FcRn que lo que la HSA se une a FcRn y una molécula que tiene una afinidad más baja por FcRn que la afinidad de la HSA por FcRn se considera que se une de manera más débil a FcRn que lo que la HSA se une a FcRn. El término "coeficiente de unión" se puede utilizar en lugar del término "afinidad de unión".

Se entiende que las expresiones "semivida en plasma más larga" o "semivida en plasma más corta" y expresiones similares están en relación con la correspondiente molécula de albúmina precursora o de referencia o correspondiente. Por lo tanto, una semivida en plasma más larga con respecto a una variante de albúmina de la invención significa que la variante tiene una semivida en plasma más larga que la de la albúmina correspondiente que tiene las mismas secuencias excepto por la(s) alteración(es) descrita(s) en el presente documento.

La albúmina o variante, y/o fragmento de la misma puede o no estar fusionada de manera genética con un compañero de fusión. Preferentemente, el compañero de fusión es una proteína que no es albúmina. El compañero de fusión puede fusionarse en el extremo N' o C' de la albúmina. Puede haber o no uno o más aminoácidos espaciadores ubicados entre la fracción de albúmina y la fracción asociada. Los compañeros de fusión pueden insertarse dentro de la secuencia de albúmina. El compañero de fusión puede tener al menos 5 aminoácidos de longitud, por ejemplo, al menos 5, 10, 15, 20, 25, 30, 35, 40, 50, 60, 70, 80, 90 o al menos 100 aminoácidos de longitud. El compañero de fusión puede tener o no una longitud máxima de 35, 30, 40, 50, 60, 70, 80, 90, 100, 125, 150, 175, 200, 300, 400, 500, 600, 700, 800, 900 o 1000 aminoácidos de largo. La proteína de fusión puede comprender una o más compañeros de fusión, por ejemplo, fusionados en el extremo N' o C' de la albúmina o insertados dentro de la secuencia de la albúmina. La proteína de fusión puede comprender uno o más (por ejemplo, varios, tal como 2, 3, 4 o 5) copias del mismo compañero de fusión o dos o más compañeros diferentes. El

compañero de fusión se puede seleccionar de proteínas deseadas o heterólogas como se divulga anteriormente.

Una proteína de fusión preferida puede comprender un polipéptido que tenga actividad GLP-1 tal como los descritos en el documento WO2014/138371 (con referencia particular a las páginas 13, 14, 26 y 34 a 37). Por ejemplo, una proteína de fusión preferida puede comprender HSA (SEQ ID NO: 10) o una variante (por ejemplo, SEQ ID NO: 45) y/o un fragmento de HSA genéticamente fusionado en serie a una copia de un análogo de GLP (*por ejemplo*, SEQ ID NO: 14 o 51) o HSA (SEQ ID NO: 10) o una variante (por ejemplo SEQ ID NO: 45) y/o fragmento de HSA genéticamente fusionado en serie a una repetición en tándem de un análogo de GLP (*por ejemplo*, SEQ ID NO: 15, 52 o 53). Por ejemplo, la proteína de fusión puede comprender o consistir en la SEQ ID NO: 16 (albiglutida).

Las células hospedadoras fúngicas particularmente adecuadas para la producción de albúminas, variantes, fragmentos y/o fusiones de las mismas incluyen, pero sin limitación, *Aspergillus* (Documento WO06/066595), *Kluyveromyces* (Fleer, 1991, Bio/technology 9: 968-975), *Pichia* (Kobayashi, 1998, Therapeutic Apheresis 2: 257-262) y *Saccharomyces* (Sleep, 1990, Bio/technology 8: 42-46)).

La proteína deseada (tal como una proteína heteróloga) puede ser o no una proteína secretada. Por lo tanto, la proteína codificada por la célula hospedadora puede comprender o no un péptido señal (que en alguna bibliografía también puede denominarse "secuencia líder"). Normalmente, la secuencia del péptido señal se escinde de la proteína durante la secreción de la célula hospedadora, por lo tanto, la proteína resultante (madura) no comprende una secuencia de péptido señal. A continuación se proporcionan ejemplos de secuencias de péptidos señal adecuadas. Un péptido señal puede comprender o no un propéptido.

De manera alternativa, la proteína deseada puede o no ser intracelular.

La proteína deseada puede o no estar codificada por un plásmido.

La proteína deseada puede o no estar codificada por un ácido nucleico cromosómico.

Los plásmidos adecuados incluyen plásmidos de la familia de 2 µm tales como los descritos en el documento WO2006/067511 (con especial énfasis en la sección titulada "The 2µm-family plasmids": en las páginas 46 a 61). Dichos plásmidos, denominados de manera colectiva "plásmidos de la familia de 2 µm", incluyen pSR1, pSB3 y pSB4 de *Zygosaccharomyces rouxii* (anteriormente clasificado como *Zygosaccharomyces bisporus),* plásmidos pSB1 y pSB2 de *Zygosaccharomyces bailii,* plásmido pSM1 de *Zygosaccharomyces fermentati,* plásmido pKD1 de *Kluyveromyces drosphilarum,* un plásmido sin nombre de *Pichia membranaefaciens* ("pPM1") y el plásmido de 2 µm (como se muestra en la figura 1 del documento WO2006/067511) y variantes (tales como Scp1, Scp2 y Scp3) de *Saccharomyces cerevisiae* (Volkert, *et al.*, 1989, Microbiological Reviews 53: 299; Murray *et al.*, 1988, J. Mol. Biol. 200: 601; Painting, *et al.*, 1984, J. Applied Bacteriology 56: 331).

Un plásmido de la familia de 2 µm normalmente comprende al menos tres marcos de lectura abiertos ("ORF", del inglés "open reading frames"), cada uno de los cuales codifica una proteína que funciona en el mantenimiento estable del plásmido de la familia de 2 µm como un plásmido multicopia. Las proteínas codificadas por los tres ORF pueden denominarse *FLP, REP1* y *REP2.* Cuando un plásmido de la familia de 2 µm no comprende los tres ORF que codifican *FLP, REP1* y *REP2* entonces los ORF que codifican la(s) proteína(s) que falta(n) deben suministrarse en *trans,* ya sea en otro plásmido o mediante integración cromosómica.

Un plásmido preferido es el plásmido de 2 µm de *S. cerevisiae,* que codifica preferentemente una proteína deseada tal como una proteína heteróloga.

La proteína Gsh1, la proteína Not4 y/o la proteína deseada, por ejemplo, heteróloga, pueden estar codificadas por una secuencia de nucleótidos unida de manera operativa a una o más secuencias de control que dirigen la expresión de la secuencia codificante en una célula hospedadora adecuada en condiciones compatibles con las secuencias de control.

El polinucleótido puede manipularse de diversas de maneras para proporcionar la expresión de un polipéptido. La manipulación del polinucleótido antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para la modificación de polinucleótidos utilizando métodos de ADN recombinantes son conocidas en la materia.

La secuencia de control puede ser un promotor, un polinucleótido que es reconocido por una célula hospedadora para la expresión del polinucleótido. El promotor contiene secuencias de control de la transcripción que median en la expresión del polipéptido. El promotor puede ser cualquier polinucleótido que muestra actividad transcripcional en la célula hospedadora, incluyendo promotores mutantes, truncados e híbridos, y puede obtenerse a partir de genes que codifican polipéptidos extracelulares o intracelulares, bien homólogos o heterólogos de la célula hospedadora.

Ejemplos de promotores adecuados para dirigir la transcripción de las construcciones de ácido nucleico de la presente invención en una célula hospedadora de hongos filamentosos son los promotores obtenidos de los genes

de la acetamidasa de *Aspergillus nidulans*, α-amilasa neutra de *Aspergillus niger*, α-amilasa estable a los ácidos de *Aspergillus niger*, glucoamilasa (glaA) de *Aspergillus niger* o *Aspergillus awamori*, TAKA amilasa de *Aspergillus oryzae*, proteasa alcalina de *Aspergillus oryzae*, triosa fosfato isomerasa de *Aspergillus oryzae*, proteasa similar a tripsina de *Fusarium oxysporum* (Documento WO96/00787), amiloglucosidasa de *Fusarium venenatum* (Documento WO00/56900), Daria de *Fusarium venenatum* (Documento WO00/56900), Quinn de *Fusarium venenatum* (Documento WO00/56900), lipasa de *Rhizomucor miehei*, proteinasa aspártica de *Rhizomucor miehei*, β-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa IV de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasa I de *Trichoderma reesei*, xilanasa II de *Trichoderma reesei*, β-xilosidasa de *Trichoderma reesei*, así como el promotor NA2-tpi (un promotor modificado a partir de un gen de α-amilasa neutra de *Aspergillus* en el que la secuencia líder no traducida se ha reemplazado por una secuencia líder no traducida de un gen de triosa fosfato isomerasa de *Aspergillus*; los ejemplos no limitantes incluyen promotores a partir de un gen de α-amilasa neutra de *Aspergillus niger* en el que la secuencia líder no traducida se ha reemplazado por una secuencia líder no traducida de un gen de triosa fosfato isomerasa de *Aspergillus nidulans* o *Aspergillus oryzae*; y promotores mutantes, truncados e híbridos de los mismos.

En un hospedador de levadura, los promotores útiles se obtienen de los genes para la enolasa (ENO1) de *Saccharomyces cerevisiae*, galactocinasa (GAL1) de *Saccharomyces cerevisiae*, alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa (ADH1, ADH2/GAP) de *Saccharomyces cerevisiae*, triosa fosfato isomerasa (TPI) de *Saccharomyces cerevisiae*, metalotioneína (CUP1) de *Saccharomyces cerevisiae* y 3-fosfoglicerato cinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para las células hospedadoras de levadura se describen en Romanos *et al.*, 1992, *citado anteriormente.*

La secuencia de control también puede ser un terminador de la transcripción, que es reconocido por una célula hospedadora para terminar la transcripción. La secuencia terminadora está unida de manera operativa al extremo 3' del polinucleótido que codifica el polipéptido. Se puede utilizar cualquier terminador que sea funcional en la célula hospedadora.

Los terminadores preferentes para células hospedadoras de hongos filamentosos se obtienen a partir de los genes de antranilato sintasa de *Aspergillus nidulans*, glucoamilasa de *Aspergillus niger*, α-glucosidasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae* y proteasa de tipo tripsina de *Fusarium oxysporum*.

Los terminadores preferentes para las células hospedadoras de levadura se obtienen de los genes de enolasa de *Saccharomyces cerevisiae*, citocromo C (CYC1) de *Saccharomyces cerevisiae* y gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae.* Otros terminadores útiles para las células hospedadoras de levadura se describen en Romanos *et al.*, 1992, *citado anteriormente.*

La secuencia de control también puede ser una región estabilizadora de ARNm cadena abajo de un promotor y cadena arriba de la secuencia codificante de un gen que aumenta la expresión del gen.

La secuencia de control también puede ser una región líder, no traducida de un ARNm que es importante para la traducción por parte de la célula hospedadora. La secuencia líder está unida de manera operativa al extremo 5' del polinucleótido que codifica el polipéptido. Se puede utilizar cualquier líder que sea funcional en la célula hospedadora.

Las líderes preferentes para las células hospedadoras de hongos filamentosos se obtienen a partir de los genes de TAKA amilasa de *Aspergillus oryzae* y triosa fosfato isomerasa de *Aspergillus nidulans*.

Las líderes adecuadas para las células hospedadoras de levadura se obtienen de los genes de enolasa (ENO-1) de *Saccharomyces cerevisiae*, 3-fosfoglicerato cinasa de *Saccharomyces cerevisiae*, factor α de *Saccharomyces cerevisiae* y alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa (ADH2/GAP) de *Saccharomyces cerevisiae*.

La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia unida de manera operativa al extremo 3' de la secuencia que codifica el polipéptido y, cuando se transcribe, es reconocida por la célula hospedadora como una señal para añadir restos de poliadenosina al ARNm transcrito. Se puede utilizar cualquier secuencia de poliadenilación que sea funcional en la célula hospedadora.

Las secuencias de poliadenilación preferentes para células hospedadoras de hongos filamentosos se obtienen de los genes de antranilato sintasa de *Aspergillus nidulans*, glucoamilasa de *Aspergillus niger*, α-glucosidasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae* y proteasa de tipo tripsina de *Fusarium oxysporum*.

Las secuencias de poliadenilación útiles para células hospedadoras de levadura están descritas en Guo y Sherman, 1995, Mol. Cellular Biol. 15: 5983-5990.

La secuencia de control también puede ser una región codificante del péptido señal que codifica un péptido señal unido al extremo N de un polipéptido y dirige el polipéptido dentro de la ruta secretora de la célula. El extremo 5' de la secuencia codificante del polinucleótido puede contener de manera inherente una secuencia codificante del péptido señal unida de manera natural en el marco de lectura de traducción con el segmento de la secuencia codificante que codifica el polipéptido. De manera alternativa, el extremo 5' de la secuencia codificante puede contener una secuencia codificante del péptido señal que es extraña a la secuencia codificante. Puede requerirse una secuencia codificante del péptido señal extraño cuando la secuencia codificante no contiene de manera natural una secuencia codificante del péptido señal. De manera alternativa, una secuencia codificante del péptido señal extraño puede simplemente reemplazar la secuencia codificante del péptido señal natural para potenciar la secreción del polipéptido. Sin embargo, puede utilizarse cualquier secuencia codificante del péptido señal que dirija el polipéptido expresado a la ruta secretora de una célula hospedadora.

Las secuencias codificantes del péptido señal eficaces para células hospedadoras de hongos filamentosos son las secuencias codificantes del péptido señal obtenidas de los genes de amilasa neutra de *Aspergillus niger*, glucoamilasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae*, celulosa de *Humicola insolens*, endoglucanasa V de *Humicola insolens*, lipasa de *Humicola lanuginosa* y proteinasa aspártica de *Rhizomucor miehei*.

Los péptidos señal preferidos para células hospedadoras de levadura, por ejemplo, células hospedadoras de levadura para la producción de albúmina, o variante, fragmento y/o fusión de la misma, incluyen:

un péptido señal obtenido del gen del factor α de *Saccharomyces cerevisiae*,
un péptido señal obtenido del gen de invertasa de *Saccharomyces cerevisiae*,
un péptido señal obtenido del gen de KEX2 de *Saccharomyces cerevisiae*, por ejemplo, que comprende o consiste en la SEQ ID NO: 17 o una secuencia de péptido señal KEX2 modificada, por ejemplo, que comprende o consiste en la SEQ ID NO: 18.

Los péptidos señal particularmente preferidos incluyen:

un péptido señal que comprende una fusión de la secuencia del péptido señal del factor α de apareamiento y la secuencia del péptido señal de la albúmina humana como se indica en el documento WO90/01063, un ejemplo de dicha secuencia del péptido señal se proporciona en la SEQ ID NO: 19;
un péptido señal que comprende el motivo pentapeptídico de SEQ ID NO: 20, en donde el motivo pentapeptídico está ubicado en el dominio hidrófobo de la secuencia del péptido señal, por ejemplo, de las posiciones -10 a -25 de una proteína inmadura, donde la posición -1 se refiere al aminoácido de la secuencia del péptido señal que está inmediatamente adyacente al extremo N del primer aminoácido de la secuencia madura, o para las secuencias del péptido señal que comprenden un propéptido, la posición -1 se refiere al aminoácido de la secuencia del péptido señal que está inmediatamente adyacente al extremo N del primer aminoácido del propéptido, ejemplos de dichas secuencias de péptidos señal se divulgan en el documento WO2004/009819;
un péptido señal de albúmina que se modifica para comprender el motivo pentapeptídico de SEQ ID NO: 20, el motivo pentapeptídico puede estar ubicado en el dominio hidrófobo de la secuencia del péptido señal, se proporciona un ejemplo de dicha secuencia del péptido señal modificada en la SEQ ID NO: 21. El motivo pentapeptídico puede insertarse en un péptido señal de invertasa para generar un péptido señal de invertasa modificado, se proporcionan ejemplos de péptidos señal de invertasa modificados en las SEQ ID NO: 41 y SEQ ID NO: 42; o un péptido señal de albúmina que se modifica para comprender el motivo pentapeptídico de SEQ ID NO: 20 y comprende un propéptido en el extremo C' de la secuencia del péptido señal, el motivo pentapeptídico puede estar ubicado en el dominio hidrófobo de la secuencia del péptido señal, se proporcionan ejemplos de dicha secuencia del péptido señal modificada en la SEQ ID NO: 22, la SEQ ID NO: 23 y la SEQ ID NO: 24.

Los péptidos señal que comprenden o consisten en la SEQ ID NO: 19, SEQ ID NO: 24 y SEQ ID NO: 42 son especialmente preferidos, por ejemplo, para la expresión de albúmina o una variante, fragmento y/o fusión de la misma.

Otras secuencias codificantes de péptidos señal útiles se describe en Romanos *et al.*, 1992, *citado anteriormente*.

La secuencia de control también puede ser una secuencia codificante de propéptido que codifica un propéptido posicionado en el extremo aminoterminal de un polipéptido. El polipéptido resultante se conoce como proenzima o propolipéptido (o zimógeno en algunos casos). En general, un propolipéptido es inactivo y se puede convertir en un polipéptido activo mediante la escisión catalítica o autocatalítica del propéptido del propolipéptido. La secuencia codificante del propéptido puede obtenerse de los genes del factor α de *Saccharomyces cerevisiae*.

Cuando están presentes tanto las secuencias del péptido señal como las del propéptido, la secuencia de propéptido se ubica junto al extremo aminoterminal del polipéptido y la secuencia del péptido señal se ubica junto al extremo aminoterminal de la secuencia del propéptido.

También puede ser deseable añadir secuencias reguladoras que regulan la expresión del polipéptido en relación con

el crecimiento de la célula hospedadora. Los ejemplos de sistemas reguladores son aquellos que hacen que la expresión del gen se active o desactive en respuesta a un estímulo químico o físico, incluida la presencia de un compuesto regulador. En levaduras, se puede usar el sistema ADH2 o el sistema GAL1. En hongos filamentosos, pueden usarse el promotor de la glucoamilasa de *Aspergillus niger*, el promotor de la α-amilasa TAKA de *Aspergillus oryzae* y el promotor de la glucoamilasa de *Aspergillus oryzae*. Otros ejemplos de secuencias reguladoras son aquellas que permiten la amplificación génica. En los sistemas eucariotas, estas secuencias reguladoras incluyen el gen de dihidrofolato reductasa que se amplifica en presencia de metotrexato y los genes de metalotioneína que se amplifican con metales pesados. En estos casos, el polinucleótido que codifica el polipéptido estaría unido de manera operativa a la secuencia reguladora.

La cepa hospedadora puede o no expresar o sobreexpresar una o más proteínas chaperonas tales como las descritas en los documentos WO2005/061718, WO2006/067511, WO2006/136831 o WO2014/138371. Por ejemplo, la cepa hospedadora puede o no sobreexpresar uno o más de: *AHA1, CCT2, CCT3, CCT4, CCT5, CCT6, CCT7, CCT8, CNS1, CPR3, CPR6, ERO1, EUG1, FMO1, HCH1, HSP10, HSP12, HSP104, HSP26, HSP30, HSP42, HSP60, HSP78, HSP82, JEM1, MDJ1, MDJ2, MPD1, MPD2, PDI1, PFD1, ABC1, APJ1, ATP11, ATP12, BTT1, CDC37, CPR7, HSC82, KAR2, LHS1, MGE1, MRS11, NOB1, ECM10, SSA1, SSA2, SSA3, SSA4, SSC1, SSE2, SIL1, SLS1, ORM1, ORM2, PER1, PTC2, PSE1, UBI4* y *HAC1* o un *HAC1* truncado sin intrones (Valkonen *et al.*, 2003, Applied Environ. Micro., 69: 2065), así como *TIM9, PAM18* (también conocido como *TIM14)* y *TCP1* (también conocido como *CCT1*) o una variante de los mismos. Se prefiere la sobreexpresión de *PDI1* (SEQ ID NO: 25) o variante o fragmento del mismo y/o *ERO1* (SEQ ID NO: 26) o variante o fragmento del mismo. La sobreexpresión incluye el aumento de la expresión de la chaperona en al menos un 25, 50, 75, 100, 200, 300, 400 o 500 % en relación con el nivel de expresión natural de la chaperona en la célula hospedadora. La sobreexpresión puede corresponder a un aumento en la cantidad de chaperonas o un aumento en la actividad de las chaperonas. La sobreexpresión se puede lograr mediante el aumento del número de copias del gen que codifica la chaperona, por ejemplo, proporcionando una célula hospedadora que comprende 2, 3, 4, 5, 6, 7, 8, 9, 10 o más copias del gen. Preferentemente, la chaperona variante tiene al menos un 70, 75, 80, 85, 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99 % de identidad de secuencia con la chaperona. Preferentemente, la variante mantiene la actividad funcional de la chaperona.

La célula hospedadora puede comprender o no al menos un ácido nucleico heterólogo que codifica una proteasa o un fragmento y/o variante de la misma. La célula hospedadora puede comprender o no al menos un ácido nucleico que codifica una proteasa tal como una serina proteasa dependiente de calcio tal como la proteasa que expresa toxina asesina (Kex2p) o un fragmento y/o variante de la misma. Preferentemente, la variante o fragmento de proteasa es funcional, por ejemplo, tiene la capacidad de escindir polipéptidos en el extremo carboxilo de la secuencia de reconocimiento Arg-Arg/X o Lys-Arg/X. A secuencia de nucleótidos *KEX2* puede comprender o consistir en la SEQ ID NO: 27, una proteína Kex2p puede comprender o consistir en la SEQ ID NO: 28. Las variantes de *KEX2* y Kex2p puede tener al menos un 70, 75, 80, 85, 90, 95, 96, 97, 98 o 99 % de identidad con la SEQ ID NO: 27 y la SEQ ID NO: 28, respectivamente. *KEX2* puede o no estar sobreexpresado.

Una célula hospedadora preferida, lo más preferentemente *S. cerevisiae,* sobreexpresa *PDI1* y/o *ERO1* y comprende al menos un ácido nucleico que codifica Kex2p.

Las secuencias de nucleótidos que codifican la proteína Gsh1, la proteína Not4 y las proteínas deseadas se pueden preparar utilizando cualquier procedimiento de mutagénesis conocido en la materia, tal como mutagénesis dirigida al sitio, construcción de genes sintéticos, construcción de genes semisintéticos, mutagénesis aleatoria, reordenamiento, *etc.*

La mutagénesis dirigida al sitio es una técnica en la que se introducen una o más (por ejemplo, varias) mutaciones en uno o más sitios definidos en un polinucleótido que codifica el precursor.

La mutagénesis dirigida al sitio se puede lograr *in vitro* mediante PCR que implica el uso de cebadores de oligonucleótidos que contienen la mutación deseada. También se puede realizar mutagénesis dirigida al sitio *in vitro* mediante mutagénesis de casete que implica la escisión mediante una enzima de restricción en un sitio en el plásmido que comprende un polinucleótido que codifica el precursor y la posterior unión de un oligonucleótido que contiene la mutación en el polinucleótido. Por lo general, la enzima de restricción que digiere el plásmido y el oligonucleótido es la misma, permitiendo que los extremos cohesivos del plásmido y el inserto se unan entre sí. Véanse, por ejemplo, Scherer y Davis, 1979, Proc. Natl. Acad. Sci. EE. UU. 76: 4949-4955; y Barton *et al.*, 1990, Nucleic Acids Res. 18: 7349-4966.

La mutagénesis dirigida al sitio también se puede lograr *in vivo* mediante métodos conocidos en la materia. Véanse, por ejemplo, la publicación de solicitud de patente de EE. UU. n.º 2004/0171154; Storici *et al.*, 2001, Nature Biotechnol. 19: 773-776; Kren *et al.*, 1998, Nat. Med. 4: 285-290; y Calissano y Macino, 1996, Fungal Genet. Newslett. 43: 15-16.

Cualquier procedimiento de mutagénesis dirigida al sitio se puede utilizar en la presente invención. Hay muchos kits comerciales disponibles que se pueden utilizar para preparar polipéptidos.

La construcción de genes sintéticos implica la síntesis *in vitro* de una molécula polinucleotídica diseñada para codificar un polipéptido de interés. La síntesis de genes se puede realizar utilizando una serie de técnicas, tales como la tecnología múltiple basada en microchip descrita por Tian *et al.* (2004, Nature 432: 1050-1054) y tecnologías similares en donde los oligonucleótidos se sintetizan y ensamblan en chips de microfluidos fotoprogramables.

Se pueden realizar y analizar sustituciones, eliminaciones y/o inserciones de aminoácidos simples o múltiples utilizando métodos conocidos de mutagénesis, recombinación y/o reordenamiento, seguido de un procedimiento de selección pertinente, tal como los divulgados por Reidhaar-Olson y Sauer, 1988, *Science* 241: 53-57; Bowie y Sauer, 1989, Proc. Natl. Acad. Sci. EE. UU. 86: 2152-2156; o los documentos WO95/17413; o WO95/22625. Otros métodos que se pueden utilizar incluyen PCR propensa a errores, presentación de fagos (por ejemplo, Lowman *et al.*, 1991, Biochemistry 30: 10832-10837; patente de EE. UU. n.º 5.223.409; documento WO92/06204) y mutagénesis dirigida a región (Derbyshire *et al.*, 1986, Gene 46: 145; Ner *et al.*, 1988, DNA 7: 127).

Los métodos de mutagénesis/reordenamiento se pueden combinar con métodos de detección automatizados de alto rendimiento para detectar la actividad de polipéptidos clonados y mutagenizados expresados por células hospedadoras (Ness *et al.*, 1999, Nature Biotechnology 17: 893-896). Las moléculas de ADN mutagenizado que codifican polipéptidos activos se pueden recupera de las células hospedadoras y secuenciarse de manera rápida utilizando métodos convencionales en la materia. Estos métodos permiten la determinación rápida de la importancia de los restos de aminoácidos individuales en un polipéptido.

La construcción de genes semisintéticos se logra mediante la combinación de aspectos de la construcción de genes sintéticos y/o mutagénesis dirigida al sitio y/o mutagénesis aleatoria y/o reordenamiento. La construcción semisintética se caracteriza por un proceso que utiliza fragmentos de polinucleótidos que se sintetizan, en combinación con técnicas de PCR. Por lo tanto, pueden sintetizarse *de novo* regiones definidas de genes, mientras que otras regiones pueden amplificarse utilizando cebadores mutagénicos específicos del sitio, mientras que otras regiones pueden someterse a amplificación por PCR propensa a errores o no propensa a errores. A continuación, las subsecuencias de polinucleótidos se pueden reordenar.

También se divulga por los inventores, y es relevante para la invención divulgada a continuación, un cultivo de células hospedadoras fúngicas que contienen una secuencia de polinucleótidos que codifica una proteína deseada, tal como una proteína heteróloga, caracterizado porque las células hospedadoras fúngicas tienen un nivel reducido de actividad de la proteína Gsh1 y/o un nivel reducido de expresión de la proteína Gsh1.

El cultivo de células hospedadoras fúngicas también puede tener un nivel de actividad modificado, tal como reducido, de la proteína Not4 y/o un nivel de expresión modificado, tal como reducido, de la proteína Not4. Las células hospedadoras fúngicas son como se ha descrito anteriormente.

El método puede comprender el cultivo en presencia de glutatión. Esto es particularmente útil cuando el hospedador fúngico carece de una proteína Gsh1 funcional, por ejemplo, debido a mutación o eliminación. También puede ser útil cuando la célula hospedadora fúngica contiene una proteína Gsh1 funcional o parcialmente funcional. El glutatión puede estar presente en los medios de fermentación en al menos 0,05 mM, por ejemplo de aproximadamente 0,05, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450 mM a aproximadamente 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500 mM, preferentemente de aproximadamente 0,5 a aproximadamente 50 mM, más preferentemente de aproximadamente 1 a aproximadamente 10 mM, más preferentemente de aproximadamente 3 a aproximadamente 7 mM, lo más preferentemente a aproximadamente 5 mM.

Un primer aspecto de la invención proporciona un método para producir una proteína heteróloga que comprende:

a. proporcionar una célula hospedadora fúngica que tiene:

1. una proteína Gsh1 modificada que tiene un nivel reducido de actividad en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
2. un nivel reducido de actividad de la proteína Gsh1 en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
3. un gen *GSH1* modificado que tiene un nivel reducido de actividad en relación con la actividad de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
4. un nivel reducido de expresión del gen *GSH1* en relación con el nivel de expresión de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y

b. cultivar la célula hospedadora para producir la proteína heteróloga,
en donde

(i) la proteína heteróloga es una proteína vegetal o animal, y/o
(ii) el cultivo es a una escala de al menos 1 litro, y/o

(iii) el método comprende además purificar la proteína heteróloga, y/o
(iv) el método comprende además formular la proteína heteróloga con un vehículo o diluyente terapéuticamente aceptable para producir así un producto terapéutico adecuado para la administración a un ser humano o un animal.

El método puede utilizarse para aumentar el rendimiento de producción de un polipéptido deseado a partir de una célula hospedadora fúngica.

La proteína deseada puede o no secretarse por la célula hospedadora, siendo preferida una proteína secretada.

Las células hospedadoras pueden cultivarse en un medio nutritivo adecuado para la producción de la proteína deseada utilizando métodos conocidos en la materia. Por ejemplo, la célula se puede cultivar mediante cultivo en matraz con agitación, o fermentación a pequeña o gran escala (incluida la fermentación continua, discontinua, alimentada-discontinua o en estado sólido) en laboratorio o fermentadores industriales realizadas en un medio adecuado y en condiciones que permitan expresar y/o aislar el polipéptido. El cultivo puede tener lugar en medios de nutrientes adecuados que comprenden fuentes de carbono y nitrógeno y sales inorgánicas, usando procedimientos conocidos en la materia. Los medios adecuados están disponibles en proveedores comerciales o se pueden preparar de acuerdo con las composiciones publicadas (por ejemplo, en los catálogos de la American Type Culture Collection). Los medios preferidos incluyen MW11D como se describe en el ejemplo 5. Si la proteína deseada se secreta en el medio nutritivo, la proteína deseada puede recuperarse directamente del medio. Si la proteína deseada no se secreta, puede recuperarse de lisados celulares.

El cultivo puede ser a pequeña o gran escala, por ejemplo, a escala de placa de microvaloración (*por ejemplo*, de 10 a 500 microlitros de medios de cultivo en volumen), a escala de matraz con agitación (*por ejemplo*, de 5 a 1000 mililitros (ml) de volumen de cultivo) o a escala de fermentador o sistemas equivalentes (*por ejemplo*, al menos a partir de 5 ml de volumen de cultivo, más preferentemente al menos 1, 2, 3, 4 o 5 litros (l), más preferentemente al menos 10, 50, 100 l, por ejemplo, al menos 500, 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 20000, 30000 l de volumen de cultivo).

El cultivo puede ser a un pH adecuado para la célula hospedadora. Para S. *cerevisiae,* el pH es, preferentemente, de 5 a 7, por ejemplo, de 5, 5,1, 5,2, 5,3, 5,4, 5,5, 5,6, 5,7, 5,8, 5,9, 6, 6,1, 6,2, 6,3, 6,4, 6,5, 6,6, 6,7, 6,8 o 6,9 a 5,1, 5,2, 5,3, 5,4, 5,5, 5,6, 5,7, 5,8, 5,9, 6, 6,1, 6,2, 6,3, 6,4, 6,5, 6,6, 6,7, 6,8, 6,9 o 7. Un intervalo de pH preferido es de aproximadamente 6,0 a aproximadamente 6,4.

El cultivo puede ser a una temperatura de aproximadamente 20 °C a aproximadamente 35 °C, por ejemplo de aproximadamente 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 °C a aproximadamente 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34 o 35 °C, preferentemente de aproximadamente 28 a aproximadamente 32 °C, más preferentemente aproximadamente 30 °C.

El cultivo puede o no implicar agitación, por ejemplo, la rotación del recipiente de cultivo o mediante el uso de un agitador dentro del recipiente de cultivo. Se prefiere la agitación.

La proteína deseada puede detectarse utilizando métodos conocidos en la materia que son específicos para la proteína deseada. Estos métodos de detección incluyen, pero sin limitación, el uso de anticuerpos específicos o cromatografía líquida de alta resolución (HPLC).

Una HPLC preferida es la HPLC de permeación en gel (GP-HPLC). El equipo adecuado incluye un sistema de HPLC LC2010 (Shimadzu) equipado con detección UV bajo el control del programa informático del servidor cliente Shimadzu VP7.3. Se pueden realizar inyecciones de 75 µl en una columna TSK G3000SWXL de 7,8 mm de d.i. × 300 mm de longitud (Tosoh Bioscience), con una precolumna TSK SW de 6,0 mm de d.i. × 40 mm de longitud (Tosoh Bioscience). Las muestras se pueden cromatografiar en fosfato de sodio 25 mM, sulfato de sodio 100 mM, azida sódica al 0,05 % (p/v), pH 7,0 a 1 $ml \cdot min^{-1}$, con un tiempo de ejecución de 20 minutos. Las muestras pueden cuantificarse mediante detección de UV a 280 nm, por área de máximos, en relación con un patrón de albúmina humana recombinante de concentración conocida (por ejemplo, 10 mg/ml) y corregido por sus coeficientes de extinción relativos.

De manera opcional, el método comprende recuperar la proteína deseada, por ejemplo, mediante el aislamiento de la proteína deseada de la célula hospedadora o cultivo de células hospedadoras, por ejemplo, medios celulares o lisado celular.

La proteína deseada puede recuperarse utilizando métodos conocidos en la materia. Por ejemplo, la proteína deseada puede recuperarse del medio nutritivo mediante procedimientos convencionales, que incluyen, pero sin limitación, recogida, centrifugación, filtración, extracción, secado por pulverización, evaporación o precipitación.

De manera opcional, el método comprende purificar la proteína deseada. La proteína deseada se puede purificar mediante diversos procedimientos conocidos en la materia, que incluyen, pero sin limitación, cromatografía (*por*

ejemplo, de intercambio iónico, de afinidad, hidrófoba, cromatoenfoque y de exclusión por tamaños), procedimientos electroforéticos (*por ejemplo*, enfoque isoeléctrico preparativo), solubilidad diferencial (por ejemplo, precipitación con sulfato amónico), SDS-PAGE o extracción (véase, por ejemplo, Protein Purification, Janson y Ryden, editores, VCH Publishers, Nueva York, 1989) para obtener proteínas deseadas sustancialmente puras.

En un aspecto alternativo, la proteína deseada no se recupera, sino que se utiliza una célula hospedadora de la presente invención que expresa la proteína deseada como fuente de la proteína deseada.

La etapa de purificar la proteína deseada (tal como una proteína heteróloga deseada) a partir de la célula hospedadora cultivada o del medio de cultivo comprende de manera opcional inmovilización celular, separación celular y/o rotura celular, pero siempre comprende al menos otra etapa de purificación diferente a la etapa o etapas de inmovilización, separación y/o rotura celular.

Las técnicas de inmovilización celular, tales como recubrir las células con perlas de alginato de calcio, se conocen en la materia. De manera análoga, las técnicas de separación celular, tales como la centrifugación, la filtración (*por ejemplo*, filtración de flujo cruzado), la cromatografía de lecho expandido y similares son conocidas en la materia. Del mismo modo, los métodos de rotura celular, que incluyen, molienda con perlas, ultrasonidos, exposición enzimática y similares son conocidos en la materia.

La al menos otra etapa de purificación puede ser cualquier otra etapa adecuada para la purificación de proteínas conocida en la materia. Por ejemplo, se han divulgado técnicas de purificación para la recuperación de albúmina expresada de forma recombinante en: el documento WO2010/128142, purificación por afinidad utilizando un ligando específico de albúmina tal como 2-cloro-4,6-di(2'-sulfoanilino)-S-triazina, el documento WO92/04367, eliminación de colorante procedente de la matriz; el documento EP464590, eliminación de colorantes procedentes de levadura; el documento EP319067, precipitación alcalina y posterior aplicación de la albúmina a una fase lipófila; y los documentos WO96/37515, US5728553 y WO00/44772, que describen procesos completos de purificación.

Las proteínas deseadas distintas de la albúmina pueden purificarse del medio de cultivo mediante cualquier técnica que se haya encontrado útil para purificar dichas proteínas.

Los métodos adecuados incluyen precipitación con sulfato de amonio o etanol, extracción con ácido o disolvente, cromatografía de intercambio aniónico o catiónico, cromatografía de fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía de afinidad, cromatografía de hidroxilapatita, cromatografía de lectina, concentración, dilución, ajuste de pH, diafiltración, ultrafiltración, cromatografía líquida de alto rendimiento ("HPLC"), HPLC en fase inversa, ajuste de conductividad y similares.

De manera opcional, el método puede comprender purificar la proteína aislada hasta un nivel de pureza comercial o industrialmente aceptable. Por el nivel de pureza comercial o industrialmente aceptable, se incluye el aporte de la proteína a una concentración de al menos 0,01 $g \cdot l^{-1}$, 0,02 $g \cdot l^{-1}$, 0,03 $g \cdot l^{-1}$, 0,04 $g \cdot l^{-1}$, 0,05 $g \cdot l^{-1}$, 0,06 $g \cdot l^{-1}$, 0,07 $g \cdot l^{-1}$, 0,08 $g \cdot l^{-1}$, 0,09 $g \cdot l^{-1}$, 0,1 $g \cdot l^{-1}$, 0,2 $g \cdot l^{-1}$, 0,3 $g \cdot l^{-1}$, 0,4 $g \cdot l^{-1}$, 0,5 $g \cdot l^{-1}$, 0,6 $g \cdot l^{-1}$, 0,7 $g \cdot l^{-1}$, 0,8 $g \cdot l^{-1}$, 0,9 $g \cdot l^{-1}$, 1 $g \cdot l^{-1}$, 2 $g \cdot l^{-1}$, 3 $g \cdot l^{-1}$, 4 $g \cdot l^{-1}$, 5 $g \cdot l^{-1}$, 6 $g \cdot l^{-1}$, 7 $g \cdot l^{-1}$, 8 $g \cdot l^{-1}$, 9 $g \cdot l^{-1}$, 10 $g \cdot l^{-1}$, 15 $g \cdot l^{-1}$, 20 $g \cdot l^{-1}$, 25 $g \cdot l^{-1}$, 30 $g \cdot l^{-1}$, 40 $g \cdot l^{-1}$, 50 $g \cdot l^{-1}$, 60 $g \cdot l^{-1}$, 70 $g \cdot l^{-1}$, 80 $g \cdot l^{-1}$, 90 $g \cdot l^{-1}$, 100 $g \cdot l^{-1}$, 150 $g \cdot l^{-1}$, 200 $g \cdot l^{-1}$, 250 $g \cdot l^{-1}$, 300 $g \cdot l^{-1}$, 350 $g \cdot l^{-1}$, 400 $g \cdot l^{-1}$, 500 $g \cdot l^{-1}$, 600 $g \cdot l^{-1}$, 700 $g \cdot l^{-1}$, 800 $g \cdot l^{-1}$, 900 $g \cdot l^{-1}$, 1000 $g \cdot l^{-1}$ o más. Por el nivel de pureza comercial o industrialmente aceptable, se incluye el aporte de la proteína aislada en la que otro material (por ejemplo, uno o más (por ejemplo, varios) contaminantes) están presentes en un nivel de menos de un 50 %, un 40 %, un 30 %, un 20 %, un 10 %, un 5 %, un 4 %, un 3 %, un 2 %, un 1 %, un 0,5 %, un 0,1 %, un 0,01 %, un 0,001 %, un 0,0001 %, un 0,00001 % o un 0,000001 % y, lo más preferentemente a un nivel del 0 %.

La proteína se puede proporcionar a una concentración de al menos 0,01 $g \cdot l^{-1}$, 0,02 $g \cdot l^{-1}$, 0,03 $g \cdot l^{-1}$, 0,04 $g \cdot l^{-1}$, 0,05 $g \cdot l^{-1}$, 0,06 $g \cdot l^{-1}$, 0,07 $g \cdot l^{-1}$, 0,08 $g \cdot l^{-1}$, 0,09 $g \cdot l^{-1}$, 0,1 $g \cdot l^{-1}$, 0,2 $g \cdot l^{-1}$, 0,3 $g \cdot l^{-1}$, 0,4 $g \cdot l^{-1}$, 0,5 $g \cdot l^{-1}$, 0,6 $g \cdot l^{-1}$, 0,7 $g \cdot l^{-1}$, 0,8 $g \cdot l^{-1}$, 0,9 $g \cdot l^{-1}$, 1 $g \cdot l^{-1}$, 2 $g \cdot l^{-1}$, 3 $g \cdot l^{-1}$, 4 $g \cdot l^{-1}$, 5 $g \cdot l^{-1}$, 6 $g \cdot l^{-1}$, 7 $g \cdot l^{-1}$, 8 $g \cdot l^{-1}$, 9 $g \cdot l^{-1}$, 10 $g \cdot l^{-1}$, 15 $g \cdot l^{-1}$, 20 $g \cdot l^{-1}$, 25 $g \cdot l^{-1}$, 30 $g \cdot l^{-1}$, 40 $g \cdot l^{-1}$, 50 $g \cdot l^{-1}$, 60 $g \cdot l^{-1}$, 70 $g \cdot l^{-1}$, 80 $g \cdot l^{-1}$, 90 $g \cdot l^{-1}$, 100 $g \cdot l^{-1}$, 150 $g \cdot l^{-1}$, 200 $g \cdot l^{-1}$, 250 $g \cdot l^{-1}$, 300 $g \cdot l^{-1}$, 350 $g \cdot l^{-1}$, 400 $g \cdot l^{-1}$, 500 $g \cdot l^{-1}$, 600 $g \cdot l^{-1}$, 700 $g \cdot l^{-1}$, 800 $g \cdot l^{-1}$, 900 $g \cdot l^{-1}$, 1000 $g \cdot l^{-1}$ o más.

Se prefiere que la proteína deseada se purifique para lograr un nivel de pureza farmacéuticamente aceptable. Una proteína tiene un nivel de pureza farmacéuticamente aceptable si está esencialmente libre de pirógenos y se puede administrar en una cantidad farmacéuticamente eficaz sin provocar efectos médicos no asociados con la actividad de la proteína.

De manera opcional, el método comprende además formular la proteína deseada con un vehículo o diluyente terapéuticamente aceptable para producir así un producto terapéutico adecuado para la administración a un ser humano o un animal.

La proteína deseada resultante puede, o no, utilizarse para cualquiera de sus utilidades conocidas, que, en el caso de la albúmina, incluye administración *intravenosa* (i.v.) a pacientes para tratar quemaduras graves, conmoción y

pérdida de sangre, como complemento de medios de cultivo y como excipiente en formulaciones de otras proteínas.

Aunque es posible que una proteína deseada de utilidad terapéutica, diagnóstica, industrial, doméstica o nutricional obtenida mediante un proceso de la invención se presente o administre sola, es preferible presentarla como una formulación (tal como una formulación farmacéutica, en particular, en el caso de proteínas útiles de manera terapéutica y/o diagnóstica), junto con uno o más vehículos o diluyentes aceptables. Los vehículos o diluyentes deben ser "aceptables" en el sentido de que sean compatibles con la proteína deseada y, cuando la formulación esté destinada a la administración a un receptor, que no sea perjudicial para el destinatario de la misma. Normalmente, los vehículos o diluyentes serán agua o solución salina que serán estériles y sin pirógenos.

De manera opcional, la proteína así formulada se presentará en forma de dosificación unitaria, tal como en forma de un comprimido, cápsula, solución inyectable o similar.

De manera opcional, el método comprende además proporcionar la proteína deseada en forma de dosificación unitaria.

Un segundo aspecto de la invención proporciona un método para aumentar el rendimiento de una proteína deseada (tal como una proteína heteróloga) que comprende el método de acuerdo con el primer aspecto de la invención.

El segundo aspecto de la invención también proporciona el uso de una célula hospedadora como se describe anteriormente o un cultivo como se describe anteriormente para aumentar el rendimiento de una proteína deseada (tal como una proteína heteróloga).

El rendimiento se refiere a la cantidad de producto, por ejemplo, proteína deseada, en solución, por ejemplo, caldo de cultivo o mezcla de lisis celular. El rendimiento puede expresarse en términos relativos, por ejemplo, siendo el rendimiento a partir de una cepa hospedadora de referencia del 100 %. Cuando se comparan las cepas hospedadoras, se prefiere que el rendimiento se mida en un conjunto de condiciones definido. El rendimiento absoluto se puede expresar como nanogramos por microlitro (ng/µl) o gramos por litro (g/l). El rendimiento puede expresarse de manera alternativa como tasa de productividad celular específica (Y_{PXT}).

Preferentemente, el rendimiento de la proteína deseada es al menos un 2 % superior al rendimiento (g/l o Y_{PXT}) de una célula hospedadora fúngica de referencia, tal como una célula hospedadora fúngica que tiene una proteína Gsh1 de tipo silvestre, tal como la SEQ ID NO: 2, preferentemente al menos un 3, 4, 5, 6, 7, 8, 9, 10, 12,5, 15, 17,5, 20, 22,5, 25, 27,5, 30, 35, 40, 45 o al menos un 50 % superior. Una célula hospedadora fúngica de referencia preferida tiene una proteína Gsh1 de SEQ ID NO: 2.

Preferentemente, el rendimiento de la proteína deseada es al menos un 2 % superior al rendimiento (g/l o Y_{PXT}) de una célula hospedadora fúngica de referencia, tal como una célula hospedadora fúngica que tiene una proteína Not4 de tipo silvestre, tal como la SEQ ID NO: 6, preferentemente al menos un 3, 4, 5, 6, 7, 8, 9, 10, 12,5, 15, 17,5, 20, 22,5, 25, 27,5, 30, 35, 40, 45 o al menos un 50 % superior. Una célula hospedadora fúngica de referencia preferida tiene una proteína Not4 de SEQ ID NO: 6.

Preferentemente, el rendimiento de la proteína deseada es al menos un 2 % superior al rendimiento (g/l o Y_{PXT}) de una célula hospedadora fúngica de referencia, tal como una célula hospedadora fúngica que tiene una proteína Gsh1 de tipo silvestre, tal como la SEQ ID NO: 2, y una proteína Not4 de tipo silvestre, tal como la SEQ ID NO: 6, preferentemente al menos un 3, 4, 5, 6, 7, 8, 9, 10, 12,5, 15, 17,5, 20, 22,5, 25, 27,5, 30, 35, 40, 45 o al menos un 50 % superior. Una célula hospedadora fúngica de referencia preferida tiene una proteína Gsh1 de SEQ ID NO: 2 y una proteína Not4 de SEQ ID NO: 6.

La proteína deseada puede ser como se describe anteriormente, especialmente una albúmina o variante, fragmento y/o fusión de la misma.

Un tercer aspecto de la invención proporciona una proteína deseada (tal como una proteína heteróloga) producida por el método de acuerdo con el primer o el segundo aspecto de la invención.

La invención también proporciona una composición, tal como una composición farmacéutica, que comprende la proteína deseada del tercer aspecto de la invención. La composición farmacéutica puede comprender uno o más vehículos farmacéuticamente aceptables, tales como los aprobados por una autoridad reguladora, tal como la US Food and Drug Administration o la European Medicines Agency. La invención proporciona además un método para tratar a un paciente que comprende administrar una cantidad eficaz de la composición farmacéutica al paciente.

Como un cuarto aspecto de la invención, el primer y/o segundo aspecto pueden incluir la preparación de una célula hospedadora fúngica como se describe anteriormente o un cultivo como se describe anteriormente en el que el método comprende modificar de manera genética una célula hospedadora fúngica (precursora), modificar la proteína Gsh1 resultante, reducir el nivel de actividad de la proteína Gsh1, modificar un gen *GSH1* o una secuencia de control del mismo o reducir el nivel de expresión de un gen *GSH1*. De manera opcional, el método también puede

comprender modificar de manera genética la célula hospedadora fúngica (precursora) para modificar la proteína Not4 resultante, modificar, por ejemplo, reducir, el nivel de actividad de la proteína Not4, modificar un gen *NOT4* o una secuencia de control del mismo o modificar el nivel de expresión de un gen *NOT4*. Las mutaciones, eliminaciones y la modificación de los niveles de actividad y/o expresión pueden ser como se describe anteriormente para la célula hospedadora fúngica o para el cultivo, o como se describe para el primer aspecto de la invención. En la materia se conocen métodos para genomanipular células hospedadoras. Como alternativa a la modificación genética de la célula hospedadora, puede modificarse el nivel o actividad de la proteína Gsh1 mediante la adición de un inhibidor o potenciador al medio de crecimiento. Del mismo modo, como una alternativa a la modificación genética de la célula hospedadora, puede modificarse el nivel o actividad de la proteína Not4 mediante la adición de un inhibidor o potenciador al medio de crecimiento.

Los inventores también divulgan una proteína Gsh1 que comprende al menos un 50 % de identidad con la SEQ ID NO: 2 y una mutación en una posición correspondiente a una o más posiciones seleccionadas de 47, 48, 49, 50, 51, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 409, 451, 452, 453, 454 y 455 de la SEQ ID NO: 2, preferentemente una posición seleccionada de (a) 57 a 51; (b) 120 a 130; (c) 409 o (d) 451 a 455. En particular, se prefiere una mutación en una posición correspondiente a la posición 125 de la SEQ ID NO: 2.

La proteína Gsh1 divulgada puede ser como se describe en relación con las células hospedadoras fúngicas descritas anteriormente. Preferentemente, la proteína Gsh1 comprende o consiste en la SEQ ID NO: 4. La proteína Gsh1 puede ser o no una proteína aislada.

Los inventores también divulgan un polinucleótido que codifica una variante de Gsh1 de la presente invención, tal como una variante de la SEQ ID NO: 2 que da como resultado un nivel menor de expresión de la proteína Gsh1 y/o un nivel menor de actividad de la proteína Gsh1, que una célula hospedadora que codifica una proteína Gsh1 de tipo silvestre tal como la SEQ ID NO: 2. Dichas proteínas Gsh1 se describen en relación con la célula hospedadora fúngica como se describe anteriormente, en relación con el cultivo como se describe anteriormente y en relación con los aspectos primero a cuarto de la invención.

Un polinucleótido preferido codifica una proteína Gsh1 con la mutación R125G (SEQ ID NO: 4), un ejemplo de dicha secuencia de polinucleótidos se proporciona mediante la SEQ ID NO: 3.

Por ejemplo, la presente invención también se refiere a construcciones de ácidos nucleicos que comprenden un polinucleótido que codifica una variante de Gsh1 de la presente invención unida de manera operativa a una o más secuencias de control que dirigen la expresión de la secuencia codificante en una célula hospedadora adecuada en condiciones compatibles con las secuencias de control. Las secuencias de control adecuadas se describen en relación con la célula hospedadora fúngica, en relación con el cultivo, en relación con los aspectos primero a tercero de la invención, en relación con el método de preparación de un hospedador como se describe anteriormente y en relación con la proteína Gsh1 como se describe anteriormente.

La construcción de ácido nucleico puede comprender adicionalmente un polinucleótido que codifica una variante de Not4 de la presente invención unido de manera operativa a una o más secuencias de control que dirigen la expresión de la secuencia codificante en una célula hospedadora adecuada en condiciones compatibles con las secuencias de control. Las secuencias de control adecuadas se describen en relación con la célula hospedadora fúngica, en relación con el cultivo y en relación con los aspectos primero a cuarto de la invención.

El(los) polinucleótido(s) puede(n) ubicarse en un vector o en el genoma de la célula hospedadora.

Por consiguiente, la presente invención también se refiere a vectores recombinantes que comprenden un polinucleótido que codifica una variante de Gsh1 de la presente invención, un promotor y señales de terminación de la transcripción y la traducción. De manera opcional, el vector comprende además un polinucleótido que codifica una variante de Not4 de la presente invención, un promotor y señales de terminación de la transcripción y la traducción. La Gsh1 puede estar codificada en el mismo polinucleótido (por ejemplo, vector) o en un polinucleótido diferente (*por ejemplo*, vector) a Not4. La invención también se refiere a vectores que comprenden un polinucleótido que codifica Gsh1 y una o más (*por ejemplo*, varias) secuencias de control que provocan que se reduzca el nivel de Gsh1 o la actividad de Gsh1. De manera opcional, los polinucleótidos (por ejemplo, vectores) comprenden además un polinucleótido que codifica Not4 y una o más (por ejemplo, varias) secuencias de control que provocan que se modifique, por ejemplo, que se reduzca, el nivel de Not4 o la actividad de Not4. Las diversas secuencias de nucleótidos y de control pueden estar unidas entre sí para producir un vector recombinante que puede incluir uno o más sitios de restricción convenientes para permitir la inserción o sustitución del polinucleótido que codifica la variante en dichos sitios. De manera alternativa, el polinucleótido puede expresarse mediante la inserción del polinucleótido o una construcción de ácido nucleico que comprende el polinucleótido en un vector adecuado para la expresión. En la creación del vector, la secuencia codificante se ubica en el vector de manera que la secuencia codificante esté unida de manera operativa a las secuencias de control adecuadas para la expresión.

El vector recombinante puede ser cualquier vector (*por ejemplo*, un plásmido o virus) que pueda someterse de manera conveniente a procedimientos de ADN recombinante y pueda provocar la expresión del polinucleótido. La

elección del vector normalmente dependerá de la compatibilidad del vector con la célula hospedadora en la que ha de introducirse el vector. El vector puede ser un plásmido lineal o circular cerrado.

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, un elemento extracromosómico, un minicromosoma o un cromosoma artificial. El vector puede contener cualquier medio para garantizar la autorreplicación. De manera alternativa, el vector puede ser uno que, cuando se introduzca en la célula hospedadora, se integre en el genoma y se replique junto con el cromosoma o cromosomas en los que se ha integrado. Además, se pueden usar un único vector o plásmido o dos o más vectores o plásmidos que juntos contienen el ADN total a introducir en el genoma de la célula huésped o un transposón.

El vector contiene preferentemente uno o más marcadores seleccionables que permiten la selección de células transformadas, transfectadas, transducidas o similares. Un marcador seleccionable es un gen, cuyo producto proporciona biocidas o resistencia vírica, resistencia a metales pesados, prototrofia a auxótrofos y similares.

El vector contiene, preferentemente, un elemento o elementos que permiten la integración del vector en el genoma de la célula hospedadora o la replicación autónoma del vector en la célula de manera independiente del genoma.

Para la integración en el genoma de la célula hospedadora, el vector puede basarse en la secuencia del polinucleótido que codifica la variante o cualquier otro elemento del vector para su integración en el genoma mediante recombinación homóloga o no homóloga. De manera alternativa, el vector puede contener polinucleótidos adicionales para dirigir la integración mediante recombinación homóloga en el genoma de la célula hospedadora en una o más ubicaciones precisas en el o los cromosomas. Para aumentar la probabilidad de integración en una ubicación precisa, los elementos integracionales deberían contener un número suficiente de ácidos nucleicos, tal como de 100 a 10.000 pares de bases, de 400 a 10.000 pares de bases, y de 800 a 10.000 pares de bases, que tienen un alto grado de identidad de secuencia con la secuencia diana correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos integracionales pueden ser cualquier secuencia que sea homóloga con la secuencia diana en el genoma de la célula hospedadora. Además, los elementos integracionales pueden ser polinucleótidos no codificantes o codificantes. Por otro lado, el vector puede integrarse en el genoma de la célula hospedadora mediante recombinación no homóloga.

Para la replicación autónoma, el vector puede comprender adicionalmente un origen de replicación que permita al vector replicarse de manera autónoma en la célula hospedadora en cuestión. El origen de replicación puede ser cualquier replicador plasmídico que medie la replicación autónoma que funciona en una célula. La expresión "origen de replicación" o "replicador plasmídico" significa un polinucleótido que permite que un plásmido o vector se replique *in vivo*.

Son ejemplos de orígenes de replicación para su uso en una célula hospedadora de levadura el origen de replicación de 2 micrómetros, ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6.

Son ejemplos de orígenes de replicación útiles en una célula de hongo filamentoso AMA1 y ANS1 (Gems *et al.*, 1991, Gene 98: 61-67; Cullen *et al.*, 1987, Nucleic Acids Res. 15: 9163-9175; documento WO00/24883). El aislamiento del gen AMA1 y la construcción de plásmidos o vectores que comprenden el gen puede conseguirse de acuerdo con los métodos divulgados en el documento WO00/24883.

Más de una copia de un polinucleótido de la presente invención se puede insertar en una célula hospedadora para aumentar la producción de una proteína deseada. Se puede obtener un aumento en el número de copias del polinucleótido mediante la integración de al menos una copia adicional de la secuencia en el genoma de la célula hospedadora o mediante la inclusión de un gen marcador seleccionable amplificable con el polinucleótido donde las células que contienen copias amplificadas del gen marcador seleccionable y, por tanto, copias adicionales del polinucleótido, se pueden seleccionar mediante el cultivo de las células en presencia del agente seleccionable adecuado.

Los procedimientos utilizados para unir los elementos descritos anteriormente para construir los vectores recombinantes de la presente invención son conocidos para un experto en la materia (véase, por ejemplo, Sambrook *et al.*, 1989, Molecular Cloning, A Laboratory Manual, 2.ª edición, Cold Spring Harbor, Nueva York).

Realizaciones preferentes

1. Una célula hospedadora fúngica que tiene:

   a. una proteína Gsh1 modificada que tiene un nivel reducido de actividad en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
   b. un nivel de actividad o nivel reducido de expresión de la proteína Gsh1 en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
   c. un gen *GSH1* modificado que tiene un nivel reducido de actividad en relación con la actividad de un gen

GSH1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
d. un nivel reducido de expresión del gen *GSH1* en relación con el nivel de expresión de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica
en donde

(i) la proteína heteróloga es una proteína vegetal o animal, y/o
(ii) el cultivo es a una escala de al menos 1 litro, y/o
(iii) el método comprende además purificar la proteína heteróloga, y/o
(iv) el método comprende además formular la proteína heteróloga con un vehículo o diluyente terapéuticamente aceptable para producir así un producto terapéutico adecuado para la administración a un ser humano o un animal.

2. La célula hospedadora fúngica de la realización 1, en donde el nivel modificado es relativo al nivel de una célula hospedadora fúngica de referencia, tal como:

a. una célula hospedadora fúngica en la que la proteína Gsh1 comprende o consiste en la SEQ ID NO: 2,
b. S288C de *S. cerevisiae* o
c. DXY1 de *S. cerevisiae*.

3. El hospedador fúngico de cualquiera de las realizaciones 1 o 2 que tiene una modificación en:

e. la proteína Not4, y/o
f. el nivel de actividad o nivel de expresión de la proteína Not4, y/o
g. el gen *NOT4* y/o
h. el nivel de expresión del gen *NOT4*.

4. La célula hospedadora fúngica de la realización 3, en donde el nivel modificado es un nivel reducido.
5. La célula hospedadora fúngica de la realización 3, en donde el nivel modificado es un nivel aumentado.
6. La célula hospedadora fúngica de la realización 3 o 4, en donde el nivel modificado es relativo al nivel de una célula hospedadora fúngica de referencia, tal como:

a. una célula hospedadora fúngica en la que la proteína Not4 es una proteína Not4 de tipo silvestre,
b. una célula hospedadora fúngica en la que la proteína Not4 comprende o consiste en la SEQ ID NO: 6,
c. S288C de *S. cerevisiae* o
d. DXY1 de *S. cerevisiae*.

7. La célula hospedadora fúngica de acuerdo con cualquier realización anterior, que comprende una secuencia de nucleótidos que codifica una proteína deseada tal como una proteína heteróloga.
8. La célula hospedadora fúngica de acuerdo con la realización 7 en la que la proteína deseada se selecciona de albúmina, un anticuerpo monoclonal, un etopósido, una proteína sérica (tal como un factor de coagulación de la sangre), antiestasina, un péptido anticoagulante de garrapatas, transferrina, lactoferrina, endostatina, angiostatina, colágenos, inmunoglobulinas o moléculas basadas en inmunoglobulinas o fragmentos de las mismas (por ejemplo, un Small Modular ImmunoPharmaceutical™ ("SMIP") o dAb, fragmentos Fab', F(ab')2, scAb, scFv o fragmento scFv), una proteína con dominio Kunitz (tal como las descritas en el documento WO03/066824, con o sin fusiones de albúmina), interferones, interleucinas, IL-10, IL-11, IL-2, especies y subespecies de interferón α (alfa), especies y subespecies de interferón β (beta), especies y subespecies de interferón γ (gamma), leptina, CNTF, $CNTF_{Ax15}$, antagonista del receptor de IL-1, eritropoyetina (EPO) e imitadores de EPO, trombopoyetina (TPO) e imitadores de TPO, prosaptida, cianovirina-N, 5 hélices, péptido T20, péptido T1249, gp41 del VIH, gp120 del VIH, urocinasa, prourocinasa, tPA, hirudina, factor de crecimiento procedente de plaquetas, hormona paratiroidea, proinsulina, insulina, glucagón, péptidos similares al glucagón tales como exendina-4, GLP-1 o GLP-2, factor de crecimiento de tipo insulina, calcitonina, hormona del crecimiento, factor de crecimiento transformante β (beta), factor de necrosis tumoral, G-CSF, GM-CSF, M-CSF, FGF, factores de coagulación en formas tanto pre como activas, que incluyen, pero sin limitación, plasminógeno, fibrinógeno, trombina, pretrombina, protrombina, factor von Willebrand, antitripsina α_1, activadores de plasminógeno, factor VII, factor VIII, factor IX, factor X y factor XIII, factor de crecimiento nervioso, LACI, factor de crecimiento de células endoteliales procedente de plaquetas (PD-ECGF), glucosa oxidasa, colinesterasa sérica, aprotinina, proteína precursora de amiloide, inhibidor de intertripsina α, antitrombina III, especies de apo-lipoproteínas, proteína C, proteína S, un metabolito, un antibiótico o una variante o fragmento de cualquiera de los anteriores.
9. La célula hospedadora fúngica de acuerdo con la realización 7 u 8 en la que la proteína deseada comprende o consiste en una albúmina, variante, fragmento y/o fusión de la misma.
10. La célula hospedadora fúngica de acuerdo con la realización 9 en la que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 70 % de identidad con la SEQ ID NO: 10.
11. La célula hospedadora fúngica de acuerdo con la realización 9 en la que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 75, 80, 85, 90, 91, 92, 93, 95, 96, 97, 98 o 99 % de identidad con la SEQ ID NO: 10.

12. La célula hospedadora fúngica de acuerdo con la realización 11 en la que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 70 % de identidad con la SEQ ID NO: 10, preferentemente al menos un 75, 80, 85, 90, 91, 92, 93, 95, 96, 97, 98 o 99 % de identidad con la SEQ ID NO: 10 y comprende A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W o Y en una posición correspondiente a K573 de la SEQ ID NO: 10.

13. La célula hospedadora fúngica de acuerdo con la realización 12 en la que la albúmina o variante, fragmento y/o fusión de la misma comprende una P, H, W o Y en una posición correspondiente a K573 de la SEQ ID NO: 10.

14. La célula hospedadora fúngica de acuerdo con la realización 13 en la que la variante de albúmina, fragmento y/o fusión de la misma tiene al menos un 98 % de identidad con la SEQ ID NO: 10 y comprende una P en una posición correspondiente a K573 de la SEQ ID NO: 10.

15. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 11 a 14 en la que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 70 % de identidad con la SEQ ID NO: 10, preferentemente al menos un 75, 80, 85, 90, 91, 92, 93, 95, 96, 97, 98 o 99 % de identidad con la SEQ ID NO: 10 y comprende A, C, D, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y en una posición correspondiente a E492 de la SEQ ID NO: 10.

16. La célula hospedadora fúngica de acuerdo con la realización 15 en la que la albúmina o variante, fragmento y/o fusión de la misma comprende una G, D, F, H, M o R en una posición correspondiente a E492 de la SEQ ID NO: 10.

17. La célula hospedadora fúngica de acuerdo con la realización 16 en la que la variante de albúmina, fragmento y/o fusión de la misma tiene al menos un 98 % de identidad con la SEQ ID NO: 10 y comprende una G o D en una posición correspondiente a E492 de la SEQ ID NO: 10.

18. La célula hospedadora fúngica de acuerdo con la realización 17 en la que la variante de albúmina, fragmento y/o fusión de la misma tiene al menos un 98 % de identidad con la SEQ ID NO: 10 y comprende una G en una posición correspondiente a E492 de la SEQ ID NO: 10.

19. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 11 a 18 en la que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 70 % de identidad con la SEQ ID NO: 10, preferentemente al menos un 75, 80, 85, 90, 91, 92, 93, 95, 96, 97, 98 o 99 % de identidad con la SEQ ID NO: 10 y comprende A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W o Y en una posición correspondiente a K574 de la SEQ ID NO: 10.

20. La célula hospedadora fúngica de acuerdo con la realización 19 en la que la albúmina o variante, fragmento y/o fusión de la misma comprende una H, G, D, F, N, S o Y en una posición correspondiente a K574 de la SEQ ID NO: 10.

21. La célula hospedadora fúngica de acuerdo con la realización 21 en la que la variante de albúmina, fragmento y/o fusión de la misma tiene al menos un 98 % de identidad con la SEQ ID NO: 10 y comprende una D, F, G o H en una posición correspondiente a K574 de la SEQ ID NO: 10.

22. La célula hospedadora fúngica de acuerdo con la realización 21 en la que la variante de albúmina, fragmento y/o fusión de la misma tiene al menos un 98 % de identidad con la SEQ ID NO: 10 y comprende una H en una posición correspondiente a K574 de la SEQ ID NO: 10.

23. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 11 a 22 en la que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 70 % de identidad con la SEQ ID NO: 10, preferentemente al menos un 75, 80, 85, 90, 91, 92, 93, 95, 96, 97, 98 o 99 % de identidad con la SEQ ID NO: 10 y comprende A, C, D, E, F, G, H, I, K, L, M, N, P, R, S, T, V, W o Y en una posición correspondiente a Q580 de la SEQ ID NO: 10.

24. La célula hospedadora fúngica de acuerdo con la realización 23 en la que la albúmina o variante, fragmento y/o fusión de la misma comprende una K o R en una posición correspondiente a Q580 de la SEQ ID NO: 10.

25. La célula hospedadora fúngica de acuerdo con la realización 24 en la que la variante de albúmina, fragmento y/o fusión de la misma tiene al menos un 98 % de identidad con la SEQ ID NO: 10 y comprende una K en una posición correspondiente a K573 de la SEQ ID NO: 10.

26. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 11 a 24 en la que la fusión de albúmina comprende la variante de albúmina de SEQ ID NO: 45.

27. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 11 a 24 en la que la variante de albúmina comprende o consiste en la SEQ ID NO: 45.

28. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 9 a 27 en la que la fusión comprende un compañero de fusión que no es albúmina o una variante o un fragmento o fusión de la misma.

29. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 10 a 28 en la que la fusión comprende un compañero de fusión seleccionado de anticuerpo monoclonal, un etopósido, una proteína sérica (tal como un factor de coagulación de la sangre), antiestasina, un péptido anticoagulante de garrapatas, transferrina, lactoferrina, endostatina, angiostatina, colágenos, inmunoglobulinas o moléculas basadas en inmunoglobulinas o fragmentos de las mismas (por ejemplo, un Small Modular ImmunoPharmaceutical™ ("SMIP") o dAb, fragmentos Fab', F(ab')2, scAb, scFv o fragmento scFv), una proteína con dominio Kunitz (tal como las descritas en el documento WO03/066824, interferones, interleucinas, IL-10, IL-11, IL-2, especies y subespecies de interferón α (alfa), especies y subespecies de interferón β (beta), especies y subespecies de interferón γ (gamma), leptina, CNTF, $CNTF_{Ax15}$, antagonista del receptor de IL-1, eritropoyetina (EPO) e imitadores de EPO, trombopoyetina (TPO) e imitadores de TPO, prosaptida, cianovirina-N, 5 hélices, péptido T20, péptido T1249, gp41 del VIH, gp120 del VIH, urocinasa, prourocinasa, tPA, hirudina, factor de crecimiento procedente de plaquetas, hormona paratiroidea, proinsulina, insulina, glucagón, péptidos similares al glucagón

tales como exendina-4, GLP-1 o GLP-2, factor de crecimiento de tipo insulina, calcitonina, hormona del crecimiento, factor de crecimiento transformante β (beta), factor de necrosis tumoral, G-CSF, GM-CSF, M-CSF, FGF, factores de coagulación en formas tanto pre como activas, que incluyen, pero sin limitación, plasminógeno, fibrinógeno, trombina, pretrombina, protrombina, factor von Willebrand, antitripsina α_1, activadores de plasminógeno, factor VII, factor VIII, factor IX, factor X y factor XIII, factor de crecimiento nervioso, LACI, factor de crecimiento de células endoteliales procedente de plaquetas (PD-ECGF), glucosa oxidasa, colinesterasa sérica, aprotinina, proteína precursora de amiloide, inhibidor de intertripsina α, antitrombina III, especies de apo-lipoproteínas, proteína C, proteína S, un metabolito, un antibiótico o una variante o fragmento de cualquiera de los anteriores.

30. La célula hospedadora fúngica de acuerdo con la realización 28 o 29 en la que el compañero de fusión comprende o consiste en una proteína similar al glucagón o un análogo de la misma.

31. La célula hospedadora fúngica de acuerdo con la realización 30 en la que el compañero de fusión comprende o consiste en la SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 51 o SEQ ID NO: 52.

32. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 7 a 31 en la que la proteína deseada comprende o consiste en la SEQ ID NO: 16.

33. La célula hospedadora fúngica de acuerdo con cualquier realización anterior en la que el nivel de actividad y/o el nivel reducido de expresión de la proteína Gsh1 es relativo al nivel de actividad o al nivel de expresión de la proteína Gsh1 de una célula hospedadora fúngica precursora tal como una célula hospedadora fúngica de tipo silvestre.

34. La célula hospedadora fúngica de acuerdo con la realización 33, en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce a no más del 99 % del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

35. La célula hospedadora fúngica de acuerdo con la realización 34, en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce a no más del 95 % del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

36. La célula hospedadora fúngica de acuerdo con la realización 35, en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce a no más del 90 % del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

37. La célula hospedadora fúngica de acuerdo con la realización 36, en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce a no más del 80 % del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

38. La célula hospedadora fúngica de acuerdo con la realización 37, en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce a no más del 70 % del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

39. La célula hospedadora fúngica de acuerdo con la realización 38, en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce a no más del 60 % del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

40. La célula hospedadora fúngica de acuerdo con la realización 39 en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce a no más del 50 % del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

41. La célula hospedadora fúngica de acuerdo con la realización 40, en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce a no más del 40 % del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

42. La célula hospedadora fúngica de acuerdo con la realización 41, en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce a no más del 30 %, preferentemente no más del 20 %, del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

43. La célula hospedadora fúngica de acuerdo con la realización 42, en la que el nivel de actividad y/o el nivel de expresión de la proteína Gsh1 se reduce sustancialmente al 0 % del nivel de actividad o del nivel de expresión de la proteína Gsh1 de la célula hospedadora fúngica precursora.

44. La célula hospedadora fúngica de acuerdo con cualquier realización anterior, en la que la célula hospedadora carece de un gen *GSH1* funcional o una proteína Gsh1 funcional.

45. La célula hospedadora fúngica de acuerdo con cualquier realización anterior, en la que la célula hospedadora carece de un gen *GSH1* o proteína Gsh1.

46. La célula hospedadora fúngica de acuerdo con cualquier realización anterior en la que la proteína Gsh1 comprende una mutación en la posición correspondiente a una posición seleccionada de 47, 48, 49, 50, 51, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 409, 451, 452, 453, 454 y 455 de la SEQ ID NO: 2.

47. La célula hospedadora fúngica de acuerdo con la realización 46 en la que la posición corresponde a R125 de la SEQ ID NO: 2.

48. La célula hospedadora fúngica de acuerdo con la realización 46 o 47 en la que la posición corresponde a H409 de la SEQ ID NO: 2.

49. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 46 a 48 en la que la posición corresponde a P453 de la SEQ ID NO: 2.

50. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 46 a 49 en la que la mutación es una sustitución, preferentemente por un aminoácido no conservado.

51. La célula hospedadora fúngica de acuerdo con la realización 50 en la que la mutación en una posición correspondiente a la posición 125 de la SEQ ID NO: 2 es una sustitución por A, C, D, E, F, G, H, I, L, M, N, P, Q,

S, T, V, W o Y, preferentemente por C, D, E o G, más preferentemente por G.

52. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 46 a 51 en la que la mutación en una posición correspondiente a la posición 125 de la SEQ ID NO: 2 es de un aminoácido con carga positiva por un aminoácido alifático, aromático, hidrófobo, pequeño, minúsculo, polar o con carga negativa, preferentemente un aminoácido con carga negativa por un aminoácido minúsculo.

53. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 46 a 52 en la que la mutación en una posición correspondiente a la posición D49 de la SEQ ID NO: 2 es una sustitución por A, C, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y, preferentemente por H, K o R, más preferentemente por K o R.

54. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 46 a 53 en la que la mutación en una posición correspondiente a la posición 49 de la SEQ ID NO: 2 es de un aminoácido con carga negativa por un aminoácido con carga positiva.

55. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 46 a 54 en la que la mutación en una posición correspondiente a la posición H409 de la SEQ ID NO: 2 es una sustitución por A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W o Y, preferentemente por A, I, L, V, M, F, W, Y, más preferentemente por A, I, L, V, más preferentemente por L.

56. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 46 a 55 en la que la mutación en una posición correspondiente a la posición H409 de la SEQ ID NO: 2 es de un aminoácido aromático a un aminoácido alifático.

57. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 46 a 56 en la que la mutación en una posición correspondiente a la posición P453 de la SEQ ID NO: 2 es una sustitución por A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W o Y, preferentemente por A, I, L, V, M, F, W, Y, más preferentemente por A, I, L, V, más preferentemente por L.

58. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 46 a 57 en la que la mutación en una posición correspondiente a la posición P453 de la SEQ ID NO: 2 es de un aminoácido aromático a un aminoácido alifático.

59. La célula hospedadora fúngica de acuerdo con cualquier realización anterior en la que la proteína Gsh1 comprende o consiste en la SEQ ID NO: 4.

60. La célula hospedadora fúngica de acuerdo con cualquier realización anterior que comprende un gen *GSH1* modificado, por ejemplo, un polinucleótido que codifica la SEQ ID NO: 4.

61. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 3 a 60 en la que el nivel de actividad o el nivel de expresión modificado de la proteína Not4 está relacionado con el nivel de actividad o el nivel de expresión de la proteína Not4 de una célula hospedadora fúngica precursora tal como una célula hospedadora fúngica de tipo silvestre.

62. La célula hospedadora fúngica de acuerdo con la realización 61, en la que el nivel de actividad y/o el nivel de expresión de la proteína Not4 se reduce a no más del 90 % del nivel de actividad y/o del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica precursora.

63. La célula hospedadora fúngica de acuerdo con la realización 62, en la que el nivel de actividad y/o el nivel de expresión de la proteína Not4 se reduce a no más del 80 % del nivel de actividad y/o del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica precursora.

64. La célula hospedadora fúngica de acuerdo con la realización 63, en la que el nivel de actividad y/o el nivel de expresión de la proteína Not4 se reduce a no más del 70 % del nivel de actividad y/o del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica precursora.

65. La célula hospedadora fúngica de acuerdo con la realización 64, en la que el nivel de actividad y/o el nivel de expresión de la proteína Not4 se reduce a no más del 60 % del nivel de actividad y/o del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica precursora.

66. La célula hospedadora fúngica de acuerdo con la realización 65, en la que el nivel de actividad y/o el nivel de expresión de la proteína Not4 se reduce a no más del 50 % del nivel de actividad y/o del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica precursora.

67. La célula hospedadora fúngica de acuerdo con la realización 66, en la que el nivel de actividad y/o el nivel de expresión de la proteína Not4 se reduce a no más del 40 % del nivel de actividad y/o del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica precursora.

68. La célula hospedadora fúngica de acuerdo con la realización 67, en la que el nivel de actividad y/o el nivel de expresión de la proteína Not4 se reduce a no más del 30 % del nivel de actividad y/o del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica precursora.

69. La célula hospedadora fúngica de acuerdo con la realización 68, en la que el nivel de actividad y/o el nivel de expresión de la proteína Not4 se reduce a no más del 20 % del nivel de actividad y/o del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica precursora.

70. La célula hospedadora fúngica de acuerdo con la realización 69, en la que el nivel de actividad y/o el nivel de expresión de la proteína Not4 se reduce sustancialmente al 0 % del nivel de actividad y/o del nivel de expresión de la proteína Not4 de la célula hospedadora fúngica precursora.

71. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 3 a 70, en la que la célula hospedadora carece de un gen *NOT4* funcional o una proteína Not4 funcional.

72. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 3 a 71, en la que la célula hospedadora carece de un gen *NOT4* o proteína Not4.

73. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 3 a 72, en la que el gen *NOT4* o la proteína Not4 se muta para alterar la interacción de la proteína Not4 con una proteína Not1.

74. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 3 a 73 en la que la proteína Not4 comprende una mutación en la posición correspondiente a una posición seleccionada de 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469 o 470 de la SEQ ID NO: 6.

75. La célula hospedadora fúngica de acuerdo con la realización 74 en la que la posición se selecciona de una posición correspondiente a 429, 430, 434 o 437 de la SEQ ID NO: 6.

76. La célula hospedadora fúngica de acuerdo con la realización 74 o 75 en la que la posición se selecciona de una posición correspondiente a 463, 464 o 466 de la SEQ ID NO: 6.

77. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 74 a 76 en la que la posición se selecciona de una posición correspondiente a 442, 445, 447 o 452 de la SEQ ID NO: 6.

78. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 74 a 77 en la que la mutación es una sustitución, preferentemente por un aminoácido no conservado.

79. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 74 a 78 en la que la mutación en una posición correspondiente a la posición 429 de la SEQ ID NO: 6 es una sustitución de A, C, D, E, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y, preferentemente por G, A, V, L o I, más preferentemente por I, L o V, más preferentemente por I.

80. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 74 a 79, en la que la mutación en una posición correspondiente a la posición 429 de la SEQ ID NO: 6 es una sustitución de un aminoácido aromático por un aminoácido alifático.

81. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 74 a 80 en la que la proteína Not4 comprende o consiste en la SEQ ID NO: 8.

82. La célula hospedadora fúngica de acuerdo con cualquier realización anterior que comprende un gen *NOT4* modificado, por ejemplo, un polinucleótido que codifica la SEQ ID NO: 8.

83. La célula hospedadora fúngica de acuerdo con cualquier realización anterior en la que la célula hospedadora carece de un gen *NOT4* o una proteína Not4.

84. La célula hospedadora fúngica de acuerdo con cualquier realización anterior, en la que se sobreexpresa una o más de las siguientes chaperonas: *AHA1, CCT2, CCT3, CCT4, CCT5, CCT6, CCT7, CCT8, CNS1, CPR3, CPR6, ERO1, EUG1, FMO1, HCH1, HSP10, HSP12, HSP104, HSP26, HSP30, HSP42, HSP60, HSP78, HSP82, JEM1, MDJ1, MDJ2, MPD1, MPD2, PDI1, PFD1, ABC1, APJ1, ATP11, ATP12, BTT1, CDC37, CPR7, HSC82, KAR2, LHS1, MGE1, MRS11, NOB1, ECM10, SSA1, SSA2, SSA3, SSA4, SSC1, SSE2, SIL1, SLS1, ORM1, ORM2, PER1, PTC2, PSE1, UBI4* y *HAC1* o una *HAC1* truncada sin intrones, *TIM9, PAM18, TCP1* o una variante de las mismas.

85. La célula hospedadora fúngica de acuerdo con cualquier realización anterior en la que *KEX2,* o una variante o fragmento del mismo, se expresa o se sobreexpresa.

86. La célula hospedadora fúngica de acuerdo con la realización 84 u 85 en la que *PDI1* o una variante del mismo se sobreexpresa o *ERO1* o una variante del mismo se sobreexpresa.

87. La célula hospedadora fúngica de acuerdo con la realización 84 u 85 en la que *PDI1* o una variante del mismo se sobreexpresa y *ERO1* o una variante del mismo se sobreexpresa.

88. La célula hospedadora fúngica de acuerdo con la realización 84 u 85 en la que *PDI1* o una variante del mismo se sobreexpresa y *KEX2* o una variante del mismo se expresa o se sobreexpresa.

89. La célula hospedadora fúngica de acuerdo con la realización 84 u 85 en la que *ERO1* o una variante del mismo se sobreexpresa y *KEX2* o una variante del mismo se expresa o se sobreexpresa.

90. La célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 84 a 89 en la que *PDI1* o una variante del mismo se sobreexpresa y *ERO1* o una variante del mismo se sobreexpresa y *KEX2* o una variante del mismo se expresa o se sobreexpresa.

91. La célula hospedadora fúngica de acuerdo con cualquier realización anterior en la que el hospedador fúngico es una levadura o un hongo filamentoso.

92. La célula hospedadora fúngica, de acuerdo con cualquier realización anterior, en la que la célula hospedadora es un *Saccharomyces, Candida, Hansenula, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces* o *Yarrowia.*

93. La célula hospedadora fúngica de acuerdo con la realización 92 en la que el *Saccharomyces* es un *Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis* o *Saccharomyces oviformis,* preferentemente *Saccharomyces cerevisiae.*

94. Un cultivo de células hospedadoras fúngicas que contienen una secuencia polinucleotídica que codifica una proteína deseada, tal como una proteína heteróloga, caracterizada por que las células hospedadoras fúngicas tienen un nivel de actividad y/o un nivel de expresión reducido de la proteína Gsh1.

95. El cultivo de células hospedadoras fúngicas de acuerdo con la realización 94, que tiene un nivel de actividad y/o un nivel de expresión reducido de la proteína Not4.

96. El cultivo de células hospedadoras fúngicas de la realización 94 o 95 en el que las células hospedadoras son como se definen en cualquiera de las realizaciones 1 a 93.

97. Un método para producir una proteína deseada, tal como una proteína heteróloga, de una célula hospedadora fúngica que comprende:

a. proporcionar una célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 1 a 92 o un

cultivo de acuerdo con cualquiera de las realizaciones 94 a 96,
b. cultivar la célula hospedadora fúngica o el cultivo para producir la proteína deseada,
c. opcionalmente recuperar la proteína deseada,
d. opcionalmente purificar la proteína deseada,
e. opcionalmente formular la proteína deseada con un vehículo o diluyente terapéuticamente aceptable para producir así un producto terapéutico adecuado para la administración a un ser humano o un animal, y
f. opcionalmente proporcionar la proteína deseada en forma de dosificación unitaria.

98. Un método para aumentar el rendimiento de una proteína deseada (tal como una proteína heteróloga) que comprende:

a. proporcionar una célula hospedadora fúngica (tal como una levadura o un hongo filamentoso) que tenga:

1. una proteína Gsh1 modificada que tiene un nivel reducido de actividad en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
2. un nivel reducido de actividad de la proteína Gsh1 en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
3. un gen *GSH1* modificado que tiene un nivel reducido de actividad en relación con la actividad de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
4. un nivel reducido de expresión del gen *GSH1* en relación con el nivel de expresión de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica,
tal como una célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 1 a 92,
y

b. cultivar la célula hospedadora para producir la proteína deseada, y
c. opcionalmente recuperar la proteína deseada,
d. opcionalmente purificar la proteína deseada,
e. opcionalmente formular la proteína deseada con un vehículo o diluyente terapéuticamente aceptable para producir así un producto terapéutico adecuado para la administración a un ser humano o un animal, y
f. opcionalmente proporcionar la proteína deseada en forma de dosificación unitaria.

99. El método de acuerdo con la realización 97 o 98 en el que el cultivo se realiza en presencia de glutatión.
100. El método de acuerdo con la realización 99 en el que el glutatión está presente en al menos 0,05 mM.
101. El método de acuerdo con la realización 100 en el que el glutatión está presente de aproximadamente 0,05, 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450 mM a aproximadamente 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1, 2, 3, 4, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150, 200, 250, 300, 350, 400, 450, 500 mM.
102. El método de acuerdo con la realización 101 en el que el glutatión está presente de aproximadamente 1 a aproximadamente 10 mM.
103. El método de acuerdo con la realización 102 en el que el glutatión está presente de aproximadamente 3 a aproximadamente 7 mM.
104. El método de acuerdo con la realización 103 en el que el glutatión está presente en aproximadamente 5 mM.
105. El método de acuerdo con cualquiera de las realizaciones 97 a 104 en el que el rendimiento de la proteína deseada es al menos un 2 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica que tiene una proteína Gsh1 de tipo silvestre, tal como la SEQ ID NO: 2.
106. El método de acuerdo con la realización 105 en el que el rendimiento es al menos un 3, 4, 5, 6, 7, 8, 9, 10, 12,5, 15, 17,5, 20, 22,5, 25, 27,5, 30, 35, 40, 45 o al menos un 50 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia.
107. El método de acuerdo con la realización 105 o 106 en el que el rendimiento de la proteína deseada es al menos un 2 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica que tiene una proteína Gsh1 de SEQ ID NO: 2.
108. El método de acuerdo con la realización 107 en el que el rendimiento es al menos un 3, 4, 5, 6, 7, 8, 9, 10, 12,5, 15, 17,5, 20, 22,5, 25, 27,5, 30, 35, 40, 45 o al menos un 50 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia.
109. El método de acuerdo con cualquiera de las realizaciones 97 a 108, en el que la célula hospedadora fúngica tiene una modificación en:

1. la proteína Not4, y/o
2. el nivel de actividad (preferentemente reducido) de la proteína Not4, y/o
3. el gen *NOT4* y/o
4. el nivel de expresión (preferentemente reducido) del gen *NOT4*

tal como una célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 3 a 93.
110. El método de acuerdo con la realización 109 en el que el rendimiento de la proteína deseada es al menos

un 2 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica que tiene una proteína Not4 de tipo silvestre, tal como la SEQ ID NO: 6.

111. El método de acuerdo con la realización 110 en el que el rendimiento es al menos un 3, 4, 5, 6, 7, 8, 9, 10, 12,5, 15, 17,5, 20, 22,5, 25, 27,5, 30, 35, 40, 45 o al menos un 50 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica que tiene una proteína Not4 de tipo silvestre, tal como la SEQ ID NO: 6.

112. El método de acuerdo con cualquiera de las realizaciones 109 a 111 en el que el rendimiento de la proteína deseada es al menos un 2 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica que tiene una proteína Gsh1 de SEQ ID NO: 2.

113. El método de acuerdo con la realización 112 en el que el rendimiento es al menos un 3, 4, 5, 6, 7, 8, 9, 10, 12,5, 15, 17,5, 20, 22,5, 25, 27,5, 30, 35, 40, 45 o al menos un 50 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica que tiene la proteína Gsh1 de SEQ ID NO: 2.

114. El método de acuerdo con cualquiera de las realizaciones 109 a 113 en el que el rendimiento de la proteína deseada es al menos un 2 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica que tiene la proteína Gsh1 de SEQ ID NO: 2 y la proteína Not4 de SEQ ID NO: 6.

115. El método de acuerdo con la realización 114 en el que el rendimiento de la proteína deseada es al menos un 3, 4, 5, 6, 7, 8, 9, 10, 12,5, 15, 17,5, 20, 22,5, 25, 27,5, 30, 35, 40, 45 o al menos un 50 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia tal como una célula hospedadora fúngica que tiene la proteína Gsh1 de SEQ ID NO: 2 y la proteína Not4 de SEQ ID NO: 6.

116. El método de acuerdo con cualquiera de las realizaciones 97 a 115 en el que la proteína deseada se selecciona de albúmina, un anticuerpo monoclonal, un etopósido, una proteína sérica (tal como un factor de coagulación de la sangre), antiestasina, un péptido anticoagulante de garrapatas, transferrina, lactoferrina, endostatina, angiostatina, colágenos, inmunoglobulinas o moléculas basadas en inmunoglobulinas o fragmentos de las mismas (por ejemplo, un Small Modular ImmunoPharmaceutical™ ("SMIP") o dAb, fragmentos Fab', F(ab')2, scAb, scFv o fragmento scFv), una proteína con dominio Kunitz (tal como las descritas en el documento WO03/066824, con o sin fusiones de albúmina), interferones, interleucinas, IL-10, IL-11, IL-2, especies y subespecies de interferón α (alfa), especies y subespecies de interferón β (beta), especies y subespecies de interferón γ (gamma), leptina, CNTF, $CNTF_{Ax15}$, antagonista del receptor de IL-1, eritropoyetina (EPO) e imitadores de EPO, trombopoyetina (TPO) e imitadores de TPO, prosaptida, cianovirina-N, 5 hélices, péptido T20, péptido T1249, gp41 del VIH, gp120 del VIH, urocinasa, prourocinasa, tPA, hirudina, factor de crecimiento procedente de plaquetas, hormona paratiroidea, proinsulina, insulina, glucagón, péptidos similares al glucagón tales como exendina-4, GLP-1 o GLP-2, factor de crecimiento de tipo insulina, calcitonina, hormona del crecimiento, factor de crecimiento transformante β (beta), factor de necrosis tumoral, G-CSF, GM-CSF, M-CSF, FGF, factores de coagulación en formas tanto pre como activas, que incluyen, pero sin limitación, plasminógeno, fibrinógeno, trombina, pretrombina, protrombina, factor von Willebrand, antitripsina α_1, activadores de plasminógeno, factor VII, factor VIII, factor IX, factor X y factor XIII, factor de crecimiento nervioso, LACI, factor de crecimiento de células endoteliales procedente de plaquetas (PD-ECGF), glucosa oxidasa, colinesterasa sérica, aprotinina, proteína precursora de amiloide, inhibidor de intertripsina α, antitrombina III, especies de apo-lipoproteínas, proteína C, proteína S, un metabolito, un antibiótico o una variante o fragmento de cualquiera de los anteriores.

117. El método de acuerdo con cualquiera de las realizaciones 97 a 116 en el que la proteína deseada comprende o consiste en una albúmina o variante, fragmento y/o fusión de la misma.

118. El método de acuerdo con la realización 117 en el que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 70 % de identidad con la SEQ ID NO: 10.

119. El método de acuerdo con la realización 118 en el que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 75, 80, 85, 90, 91, 92, 93, 95, 96, 97, 98 o 99 % de identidad con la SEQ ID NO: 10.

120. El método de acuerdo con la realización 119 en el que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 70 % de identidad con la SEQ ID NO: 10, preferentemente al menos un 75, 80, 85, 90, 91, 92, 93, 95, 96, 97, 98 o 99 % de identidad con la SEQ ID NO: 10 y comprende una A, C, D, E, F, G, H, I, L, M, N, P, Q, R, S, T, V, W o Y en una posición correspondiente a K573 de la SEQ ID NO: 10.

121. El método de acuerdo con la realización 120 en el que la albúmina o variante, fragmento y/o fusión de la misma comprende una P, H, W o Y en una posición correspondiente a K573 de la SEQ ID NO: 10.

122. El método de acuerdo con la realización 121 en el que la albúmina o variante, fragmento y/o fusión de la misma tiene al menos un 98 % de identidad con la SEQ ID NO: 10 y comprende una P en una posición correspondiente a K573 de la SEQ ID NO: 10.

123. El método de acuerdo con cualquiera de las realizaciones 117 a 122 en el que la fusión comprende un compañero de fusión que no es albúmina o una variante, fragmento y/o fusión de la misma.

124. El método de acuerdo con cualquiera de las realizaciones 118 a 123 en el que la fusión comprende un compañero de fusión seleccionado de anticuerpo monoclonal, un etopósido, una proteína sérica (tal como un factor de coagulación de la sangre), antiestasina, un péptido anticoagulante de garrapatas, transferrina, lactoferrina, endostatina, angiostatina, colágenos, inmunoglobulinas o moléculas basadas en inmunoglobulinas o fragmentos de las mismas (por ejemplo, un Small Modular ImmunoPharmaceutical™ ("SMIP") o dAb, fragmentos Fab', F(ab')2, scAb, scFv o fragmento scFv), una proteína con dominio Kunitz (tal como las descritas en el documento WO03/066824, interferones, interleucinas, IL-10, IL-11, IL-2, especies y subespecies de interferón α

(alfa), especies y subespecies de interferón β (beta), especies y subespecies de interferón γ (gamma), leptina, CNTF, $CNTF_{Ax15}$, antagonista del receptor de IL-1, eritropoyetina (EPO) e imitadores de EPO, trombopoyetina (TPO) e imitadores de TPO, prosaptida, cianovirina-N, 5 hélices, péptido T20, péptido T1249, gp41 del VIH, gp120 del VIH, urocinasa, prourocinasa, tPA, hirudina, factor de crecimiento procedente de plaquetas, hormona paratiroidea, proinsulina, insulina, glucagón, péptidos similares al glucagón tales como exendina-4, GLP-1 o GLP-2, factor de crecimiento de tipo insulina, calcitonina, hormona del crecimiento, factor de crecimiento transformante β (beta), factor de necrosis tumoral, G-CSF, GM-CSF, M-CSF, FGF, factores de coagulación en formas tanto pre como activas, que incluyen, pero sin limitación, plasminógeno, fibrinógeno, trombina, pretrombina, protrombina, factor von Willebrand, antitripsina α_1, activadores de plasminógeno, factor VII, factor VIII, factor IX, factor X y factor XIII, factor de crecimiento nervioso, LACI, factor de crecimiento de células endoteliales procedente de plaquetas (PD-ECGF), glucosa oxidasa, colinesterasa sérica, aprotinina, proteína precursora de amiloide, inhibidor de intertripsina α, antitrombina III, especies de apo-lipoproteínas, proteína C, proteína S, un metabolito, un antibiótico o una variante o fragmento de cualquiera de los anteriores.

125. El método de acuerdo con la realización 124 en el que el compañero de fusión comprende o consiste en una proteína similar al glucagón o un análogo de la misma.

126. El método de acuerdo con la realización 125 en el que el compañero de fusión comprende o consiste en la SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 51 o SEQ ID NO: 52.

127. El método de acuerdo con cualquiera de las realizaciones 116 a 126 en el que la proteína deseada comprende o consiste en la SEQ ID NO: 16.

128. El método de acuerdo con cualquiera de las realizaciones 97 a 127 en el que la célula hospedadora se cultiva a una escala de al menos 1 l.

129. El método de acuerdo con la realización 128 en el que la célula hospedadora se cultiva a una escala de al menos 2 l.

130. El método de acuerdo con la realización 129 en el que la célula hospedadora se cultiva a una escala de al menos 5 l.

131. El método de acuerdo con la realización 130 en el que la célula hospedadora se cultiva a una escala de al menos 10 l.

132. El método de acuerdo con la realización 131 en el que la célula hospedadora se cultiva a una escala de al menos 1000 l.

133. El método de acuerdo con la realización 132 en el que la célula hospedadora se cultiva a una escala de al menos 5000 l.

134. El método de acuerdo con cualquiera de las realizaciones 97 a 133 en el que la proteína deseada se secreta a partir de la célula hospedadora fúngica.

135. El método de acuerdo con la realización 134 en el que la proteína deseada resulta de una proteína inmadura que comprende un péptido señal.

136. El método de acuerdo con la realización 135 en el que el péptido señal comprende o consiste en la SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 41 o SEQ ID NO: 42 o un péptido señal que comprende el motivo pentapeptídico de SEQ ID NO: 20.

137. El método de acuerdo con la realización 136 en el que el péptido señal comprende o consiste en la SEQ ID NO: 19.

138. El método de acuerdo con la realización 136 en el que el péptido señal comprende o consiste en la SEQ ID NO: 24.

139. El método de acuerdo con la realización 136 en el que el péptido señal comprende o consiste en la SEQ ID NO: 42.

140. El método de acuerdo con cualquiera de las realizaciones 97 a 133 en el que la proteína deseada es intracelular.

141. Una proteína deseada (tal como una proteína heteróloga) producida por el método de acuerdo con cualquiera de las realizaciones 97 a 140.

142. La proteína deseada de acuerdo con la realización 141 para profilaxis, terapia o diagnóstico.

143. Una composición, tal como una composición farmacéutica, que comprende la proteína deseada de acuerdo con la realización 141 o 142 y un vehículo farmacéuticamente aceptable.

144. Un método de tratamiento que comprende administrar la proteína deseada de la realización 143 o 144 o la composición de la realización 143 a un paciente.

145. Un método para preparar una célula hospedadora fúngica de acuerdo con cualquiera de las realizaciones 1 a 93 o un cultivo de acuerdo con cualquiera de las realizaciones 94 a 96, comprendiendo el método modificar de manera genética una célula hospedadora fúngica (precursora) para reducir el nivel de expresión y/o el nivel de actividad de la proteína Gsh1.

146. El método de acuerdo con la realización 145, que comprende además modificar de manera genética una célula hospedadora fúngica (precursora) para reducir el nivel de expresión y/o el nivel de actividad de la proteína Not4.

147. Uso de un medio para reducir el nivel de expresión y/o el nivel de actividad de la proteína Gsh1 en una célula hospedadora fúngica para aumentar el rendimiento de una proteína deseada (tal como una proteína heteróloga) de la célula hospedadora fúngica, por ejemplo: mediante la mutación o eliminación del gen *GSH1*, dando lugar así a una proteína Gsh1 mutada o a una ausencia completa de la proteína Gsh1; mediante la eliminación o cambio del marco de lectura abierto del gen; mediante la mutación o cambio de las secuencias de control del gen *GSH1* tal como una secuencia promotora y/o una secuencia terminadora; mediante el bloqueo o

reducción de la transcripción del gen *GSH1*, por ejemplo, mediante la introducción de ARN de interferencia adecuado tal como ARNm antisentido, mediante la introducción, control o modificación de genes activadores transcripcionales adecuados o mediante la introducción de un agente que bloquea el nivel de actividad de la proteína Gsh1.

148. Uso de un medio para reducir el nivel de expresión y/o el nivel de actividad de la proteína Gsh1 y para reducir el nivel de actividad de la proteína Not4 en una célula hospedadora fúngica para aumentar el rendimiento de una proteína deseada (tal como una proteína heteróloga) de la célula hospedadora fúngica, por ejemplo: mediante la mutación o eliminación del gen *GSH1* y la mutación o eliminación del gen *NOT4*, dando así como resultado una proteína Gsh1 mutada o una ausencia completa de proteína Gsh1 y dando como resultado una proteína Not4 mutada o una ausencia completa de proteína Not4; mediante la eliminación o cambio del marco de lectura abierto del gen; mediante la mutación o cambio de las secuencias de control del gen *GSH1* y/o *NOT4* tal como una secuencia promotora y/o una secuencia terminadora; mediante el bloqueo o reducción de la transcripción del gen *GSH1* y/o *NOT4*, por ejemplo, mediante la introducción de ARN de interferencia adecuado tal como ARNm antisentido; mediante la introducción, control o modificación de genes activadores transcripcionales adecuados o mediante la introducción de un agente que bloquea el nivel de actividad de la proteína Gsh1 y/o la proteína Not4.

149. Una proteína Gsh1, que comprende al menos un 50 % de identidad con la SEQ ID NO: 2 y una mutación en una posición correspondiente a una o más posiciones seleccionadas de 47, 48, 49, 50, 51, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 409, 451, 452, 453, 454 y 455 de la SEQ ID NO: 2.

150. La proteína Gsh1 de acuerdo con la realización 149 en la que la posición corresponde a R125 de la SEQ ID NO: 2.

151. La proteína Gsh1 de acuerdo con la realización 149 o 150 en la que la posición corresponde a H409 de la SEQ ID NO: 2.

152. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 151 en la que la posición corresponde a P453 de la SEQ ID NO: 2.

153. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 152 en la que la mutación es una sustitución, preferentemente por un aminoácido no conservado.

154. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 150 a 153 en la que la mutación en una posición correspondiente a la posición R125 de la SEQ ID NO: 2 es una sustitución por A, C, D, E, F, G, H, I, L, M, N, P, Q, S, T, V, W o Y, preferentemente por C, D, E o G, más preferentemente por G.

155. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 154 en la que la mutación en una posición correspondiente a la posición R125 de la SEQ ID NO: 2 es una sustitución de un aminoácido con carga positiva por un aminoácido alifático, aromático, hidrófobo, pequeño, minúsculo, polar o con carga negativa, preferentemente un aminoácido con carga negativa por un aminoácido minúsculo.

156. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 155 en la que la mutación en una posición correspondiente a la posición D49 de la SEQ ID NO: 2 es una sustitución por A, C, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W o Y, preferentemente por H, K o R, más preferentemente por K o R.

157. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 156 en la que la mutación en una posición correspondiente a la posición 49 de la SEQ ID NO: 2 es de un aminoácido con carga negativa a un aminoácido con carga positiva.

158. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 157 en la que la mutación en una posición correspondiente a la posición H409 de la SEQ ID NO: 2 es una sustitución por A, C, D, E, F, G, I, K, L, M, N, P, Q, R, S, T, V, W o Y, preferentemente por A, I, L, V, M, F, W, Y, más preferentemente por A, I, L, V, más preferentemente por L.

159. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 158 en la que la mutación en una posición correspondiente a la posición H409 de la SEQ ID NO: 2 es de un aminoácido aromático a un aminoácido alifático.

160. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 159 en la que la mutación en una posición correspondiente a la posición P453 de la SEQ ID NO: 2 es una sustitución por A, C, D, E, F, G, H, I, K, L, M, N, Q, R, S, T, V, W o Y, preferentemente por A, I, L, V, M, F, W, Y, más preferentemente por A, I, L, V, más preferentemente por L.

161. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 160 en la que la mutación en una posición correspondiente a la posición P453 de la SEQ ID NO: 2 es de un aminoácido aromático a un aminoácido alifático.

162. La proteína Gsh1 de acuerdo con cualquiera de las realizaciones 149 a 161 que comprende o consiste en la SEQ ID NO: 4.

La invención se describe ahora en referencia a los siguientes ejemplos no limitantes:

Ejemplos

Ejemplo 1: Mutación del gen *GSH1 de Saccharomyces cerevisiae*

DP9 de *S. cerevisiae* tiene el genotipo cir° *MATa, leu2-3, leu2-112 ubc4 ura3 yap3::URA3 lys2 hsp150:LYS2* con *PDI1, URA3* y Ylplac211 integrados en el locus *PDI1* (Finnis *et al.,* 2010, Microbial Cell Factories 9: 87). Los inventores observaron que DP9 de *S. cerevisiae* (cuando se transformó con un plásmido codificante de albúmina)

fue capaz de producir albúmina humana recombinante con un rendimiento superior al de las cepas predecesoras, por ejemplo, DB1 de *S. cerevisiae*. La caracterización de DP9 de *S. cerevisiae* reveló un SNP en el gen *GSH1*. Para establecer si este SNP contribuyó a mejorar el rendimiento proteico de DP9 de *S. cerevisiae*, el SNP (A373G) se revirtió al tipo silvestre (es decir, A en la posición 373) como se describe a continuación. Por consiguiente, la proteína Gsh1 mutante (G125) también se revirtió al tipo silvestre (R125). El gen *GSH1* y la proteína Gsh1 de tipo silvestre se proporcionan en la SEQ ID NO: 1 y la SEQ ID NO: 2, respectivamente. El gen *GSH1* y la proteína Gsh1 mutante (es decir, que contienen el SNP) se proporcionan en la SEQ ID NO: 3 y la SEQ ID NO: 4, respectivamente.

El gen *GSH1* de *S. cerevisiae* está ubicado en el cromosoma X. El SNP (A373G) en el gen *GSH1* mutante se revirtió al tipo silvestre mediante el proceso de integración de un fragmento en el locus GSH1, que cambió la base 373G a A, revirtiendo así la proteína Gsh1 mutante (G en la posición 125) a la proteína Gsh1 de tipo silvestre (R en la posición 125).

Esto se logró amplificando primero, mediante PCR, un marcador de selección adecuado *(KanMX)* con cebadores de ADN que incluían secuencias de ADN en sus extremos 5', idénticas a las regiones cadena arriba del marco de lectura abierto de GSH1. Los cebadores de PCR fueron MBP267 y MBP268. *KanMX* confiere resistencia a la geneticina (G418).

Cebador MBP267:

5'-
ATACTATTGTAATTCAAAAAAAAAAAGCGAATCTTCCCATGCCTGTTGCTGCTCTTGAATGGCGACAGCCTATT
GCCCCAGTGTTCCCTCAACAACCTTGCGTACGCTGCAGGTCG-3' (SEQ ID NO: 29)

Cebador MBP268:

5'-
ACAGTTGTAGTCACGTGCGCGCCATGCTGACTAATGGCAGCCGTCGTTGGGCAGAAGAGAATTAGTATGGTA
CAGGATACGCTAATTGCGCTCCAACTACATCGATGAATTCGAGCTCG-3' (SEQ ID NO: 30)

Se realizó PCR para amplificar el gen *KanMX* del plásmido pDB5438 (Figura 2). Las condiciones fueron las siguientes: 100 ng del plásmido pDB5438, cada cebador 0,5 µM, dNTP 0,2 µM, desnaturalización inicial durante 30 segundos a 98 °C, a continuación, 35 ciclos a 98 °C durante 10 segundos, hibridación a 62 °C durante 30 segundos, extensión a 72 °C durante 2 minutos, seguido de una extensión final a 72 °C durante 4 minutos y enfriamiento a 4 °C, utilizando un termociclador Applied Biosystems 2720 y un kit de PCR de polimerasa de ADN de alta fidelidad NEB Q5 Hot Start (M0493S), volumen de reacción total de 50 µl, de acuerdo con las instrucciones del fabricante.

El producto, *5-GSH1 5'UTR-KanMX-GSH1* 5'UTR-3', se analizó mediante electroforesis en gel y se encontró que era del tamaño esperado, aproximadamente 1,7 kb. El producto de PCR amplificado se purificó. El producto purificado se utilizó para transformar una cepa de *S. cerevisiae* cepa que era de tipo silvestre para *GSH1* (es decir, SEQ ID NO: 1). La transformación se realizó utilizando un kit Sigma Yeast Transformation de acuerdo con las instrucciones del fabricante, excepto que después de la etapa donde se centrifuga la mezcla de transformación, el sedimento se volvió a suspender en 0,5 ml de medio YEPD y después se transfirió a un matraz de agitación que contenía 4,5 ml de YEPD. YEPD (g/l): 10 g de extracto técnico de levadura Bacto™, 20 g de peptona Bacto™, 20 g de glucosa.

El matraz de agitación se incubó durante 16 horas a 30 °C con agitación (200 rpm). El cultivo se centrifugó a 3.000 revoluciones por minuto (rpm) durante 5 minutos y se decantó el sobrenadante. A continuación, el sedimento se lavó en sorbitol 1 M y después se volvió a suspender en 0,5 ml de sorbitol 1 M. A continuación, se sembraron aproximadamente 100 µl en placas de agar G418 recién preparadas (concentración final de G418 de 300 µg/ml) y se incubaron boca abajo a 30 °C durante cuatro días. Las placas de agar G418 se prepararon de la siguiente manera: se disolvieron 0,17 g de base de nitrógeno de levadura (sin $(NH_4)2SO_4$ y sin aminoácidos), 0,1 g de ácido glutámico (sal monosódica, Sigma G-1626) y 0,069 g de polvo de CSM-Leu en 100 ml de H_2O (agua estéril para irrigación - apirógena, hipotónica) y se esterilizaron por filtración. A continuación, se añadieron 1,5 g de agar Bacto y la botella se calentó en un vaporizador durante una hora y después se enfrió a 55 °C en un baño de agua. Se añadieron y mezclaron 0,6 ml de geneticina (G418) 50 mg/ml y 4 ml de dextrosa estéril al 50 % (p/v). Se vertieron alícuotas de la mezcla en placas de Petri para que se solidificaran.

El ADN genómico se extrajo de transformantes resistentes a G418 y se utilizó como plantilla en una segunda PCR, utilizando los cebadores MBP288 y MBP289, para amplificar un fragmento de ORF de *GSH1* en 5'UTR *5'-GSH1 5'UTR-KanMX-GSH1* (SEQ ID NO: 31) que contiene la 5'UTR de GSH1 con el gen *KanMX* insertado, y la parte del ORF de *GSH1* en 5' (hasta la base 422)

Cebador MBP288: 5' — GATTTTATCGGTCAAAGG — 3' (SEQ ID NO: 32)
Cebador MBP289: 5' — CTATCTTGTCTCGCATATTC — 3' (SEQ ID NO: 33)

Los materiales de PCR, el método y las condiciones fueron los descritos anteriormente, excepto que se utilizó 1 µl

de ADN genómico, la temperatura de hibridación fue de 55 °C y el tiempo de extensión fue de 3 minutos en cada ciclo y finalmente de 7 minutos. El producto se analizó mediante electroforesis en gel y se encontró que tenía el tamaño esperado, aproximadamente 2,9 kb. Los productos de PCR amplificados se purificaron con un kit de purificación de PCR QIAquick de QIAGEN de acuerdo con las instrucciones del fabricante. El producto purificado se utilizó para transformar DP9 [pDB2305] (Figura 3) utilizando el método de transformación descrito anteriormente. DP9 de *S. cerevisiae* es una cepa que contiene el SNP de *GSH1* (A373G que da como resultado R125G). pDB2305 es un plásmido para la expresión de albúmina humana (Figura 3). El crecimiento y la selección en placas de agar G418 fueron como se describe anteriormente. El ADN genómico se extrajo de colonias resistentes y se utilizó PCR para amplificar un fragmento de 546 pb que cubría la región desde 100 pb cadena arriba del ORF hasta la base 446 del ORF, utilizando los cebadores MBP290 y MBP292. Se utilizaron el mismo kit de PCR y las mismas condiciones, excepto que las etapas de ciclado se cambiaron a una desnaturalización inicial de 98 °C durante 1 minuto, seguida de 35 ciclos a 98 °C durante 10 segundos, hibridación a 61 °C durante 20 segundos y extensión a 72 °C durante 2,5 minutos y después una extensión final a 72 °C durante 7 minutos.

Cebador MBP290: 5'— TCTCGAGTCACTTGTAGAAG — 3' (SEQ ID NO: 34)
Cebador MBP292: 5'— GAGCCCACATGCAAGTT — 3' (SEQ ID NO: 35)

Los productos se limpiaron con el kit de purificación de PCR Qiagen QIAquick 96. Se utilizó un kit de secuenciación cíclica Life Technologies BigDye Terminator v3.1 para la secuenciación de los productos de acuerdo con las instrucciones del fabricante, utilizando volúmenes de reacción totales de 50 µl, con -50 ng de los productos limpios como molde y 4 µl del cebador MBP291 1 µM. Las condiciones fueron las siguientes: Desnaturalización inicial a 96 °C durante 1 minuto después 25 ciclos con desnaturalización a 96 °C durante 10 segundos, hibridación a 50 °C durante 5 segundos, elongación a 60 °C durante 4 minutos y finalmente enfriamiento a 4 °C. Las reacciones de secuenciación se precipitaron y volvieron a suspender en HiDi (Applied Biosystems) y se analizaron en un analizador genético Applied Biosystems 3130xl.
Cebador MBP291: 5'— GTAGGGTGGTTTAGAGTATC — 3' (SEQ ID NO: 36)

El análisis de secuenciación mostró que un transformante tenía A de tipo silvestre en la posición 373 (R125), esta cepa se denominó PRG13 [pDB2305]. Dos transformantes todavía tenían la G en la posición 373 (G125), estas cepas se denominaron PSG10 [pDB2305] y PSG11 [pDB2305]. Estos tres transformantes se cultivaron en una placa de microvaloración (MTP, del inglés "microtiter plate") de 48 pocillos, que contenía 0,5 ml de BMMD (0,17 % (p/v) de base nitrogenada de levadura sin aminoácido y sulfato de amonio (Difco), sulfato de amonio 37,8 mM, ácido cítrico 36 mM, ortofosfato de hidrógeno disódico 126 mM pH 6,5, glucosa al 2 % (p/v), ajustado a pH 6,5 con NaOH) en cada pocillo. La MTP se incubó a 30 °C en cámara húmeda con agitación (200 rpm) durante 72 horas. A continuación, se transfirieron 50 µl de cultivo celular de cada pocillo a tres pocillos en una nueva MTP de 48 pocillos que contenía 0,45 ml de BMMD en cada pocillo. La nueva MTP se incubó a 30 °C en cámara húmeda con agitación (200 rpm) durante 96 horas.

El sobrenadante se aisló mediante centrifugación y la productividad de albúmina recombinante se determinó mediante análisis GP-HPLC con un sistema HPLC LC2010 (Shimadzu) equipado con detección UV bajo el control del programa informático del servidor cliente Shimadzu VP7.3. Se realizaron inyecciones de 75 µl en una columna TSK G3000SWXL de 7,8 mm de d.i. × 300 mm de longitud (Tosoh Bioscience), con una precolumna TSK SW de 6,0 mm de d.i. × 40 mm de longitud (Tosoh Bioscience). Las muestras se cromatografiaron en fosfato de sodio 25 mM, sulfato de sodio 100 mM, azida sódica al 0,05 % (p/v), pH 7,0 a 1 $ml \cdot min^{-1}$, con un tiempo de ejecución de 20 minutos. Las muestras se cuantificaron por detección UV a 280 nm, por área de máximos, respecto a un patrón de albúmina humana recombinante de concentración conocida (10 mg/ml) y se corrigieron por sus coeficientes de extinción relativos.

Como se muestra en la tabla 2, la presencia del SNP dio como resultado un aumento del 21 % en el rendimiento promedio de la albúmina.

Tabla 2: Productividad de la albúmina en PRG13 [pDB2305] en comparación con PSG10 [pDB2305] y PSG11 [pDB2305]

	GSH1 TS (PRG13 [pDB2305])	GSH1 con SNP A373G (PSG10 [pDB2305] y PSG11 [pDB2305])
Albúmina (rendimiento relativo)	100 % ± 3,91 %	121 % ± 5,58
Valor de p (prueba de la t): 8,89E-05		

El trabajo se repitió en otra cepa de *S. cerevisiae*. Brevemente, se realizó el mismo procedimiento de transformación en la cepa BXP10 [pDB2244] de *S. cerevisiae* (Figura 4), para revertir el SNP de GSH1 a tipo silvestre. pDB2244 es un plásmido que expresa albúmina humana. Las cepas se cultivaron en matraz de agitación, se incubaron aproximadamente 10 ml de medio de cultivo en un matraz de 50 ml, con agitación a 200 rpm a 30 °C. El rendimiento de la albúmina de un transformante que todavía contenía el SNP, es decir, BSG3 [pDB2244], se comparó con el rendimiento de la albúmina de un transformante con el SNP convertido a tipo silvestre, es decir, BRG5 [pDB2244]. BXP10 tiene el genotipo *MATa, leu2-3, leu2-122, can1, pra1, ubc4, ura3, yap3::URA3, lys2, hsp150::LYS2* y

pmt1::URA3.

Como se muestra en la tabla 3, la presencia del SNP dio como resultado un aumento del 68 % en el rendimiento de la albúmina (tres repeticiones para cada cepa).

Tabla 3: Productividad de la albúmina en BRG5 [pDB2244] y BSG3 [pDB2244]

	BRG5 [pDB2244] (*GSH1* TS)	BSG3 [pDB2244] *(GSH1* con SNP A373G
Albúmina (rendimiento relativo)	100 % ± 7,84	168 % ± 9,45
Valor de p (prueba de la t): 0,001		

Ejemplo 2: La eliminación del gen *GSH1* en *S. cerevisiae* mejoró la producción de proteína recombinante

El gen *GSH1* se eliminó en las mismas dos cepas de *S. cerevisiae* que se utilizaron en el ejemplo 1 para la reversión del SNP. La eliminación se logró mediante el reemplazo del gen *GSH1* con el marcador *KanMX.* Por consiguiente, las cepas resultantes no fueron capaces de producir ninguna proteína Gsh1. Para realizar la eliminación, el ADN requerido se ordenó como un inserto de plásmido flanqueado por sitios de restricción Acc65I. El fragmento de ADN estaba formado por *KanMX DNA* flanqueado en el extremo 5' por 300 pb inmediatamente cadena arriba del ORF de *GSH1* y en el extremo 3' por 300 pb inmediatamente cadena abajo del ORF de *GSH1* (SEQ ID NO: 43). El plásmido se digirió con Acc65I y el fragmento KanMX se purificó a partir de un gel de agarosa Trisacetato-EDTA (TAE) con el kit de extracción de gel QIAGEN. Este ADN se utilizó para transformar las cepas de levadura DP9 [pDB2305] y BXP10 [pDB2244], donde pDB2305 y pDB2244 son plásmidos que contienen un casete de expresión para seroalbúmina humana. Al final del procedimiento de transformación, las células de levadura se volvieron a suspender en medio YEPD y se cultivaron durante la noche a 30 °C y 200 rpm. Al día siguiente se sembraron en placas G418 (como se describe en el ejemplo 1) con L-glutatión añadido a 5 mM. Las colonias se sembraron en placas nuevas (igualmente con G418 y L-glutatión) y cuando las manchas crecieron, se extrajo el ADN genómico y se realizaron PCR utilizando cebadores que flanquearían las uniones si todo el fragmento de ADN se hubiera integrado de manera correcta. Se identificaron tres transformantes en los que se había eliminado GSH1 en BXP10 (los transformantes se denominaron BDG1, BDG2 y BDG10) y un transformante en el que se había eliminado GSH1 en DP9 (el transformante se denominó PDG14). Estas cepas, junto con la cepa descrita anteriormente con o sin la reversión del SNP, se cultivaron en placas de 48 pocillos en BMMD. Las cepas con la eliminación se cultivaron en pocillos complementados con L-glutatión 5 mM y las otras cepas se cultivaron con y sin L-glutatión. Los sobrenadantes se analizaron mediante GP-HPLC para cuantificar los niveles de expresión de la seroalbúmina humana.

Tabla 4: Productividad de la albúmina en cepas procedentes de BXP10: BRG5 [pDB2244] (*GSH1* TS), BSG3 [pDB2244] (SNP) y BDG1 [pDB2244], BDG2 [pDB2244] y BDG10 [pDB2244] *(ΔGSH1)*

	Sin L-glutatión			Con L-glutatión		
Cepa(s)	A	B	C	D	E	F
	BRG5 [pDB2244] (*GSH1* TS)	BSG3 [pDB2244] *(GSH1* con SNP)	BDG1 [pDB2244], BDG2 [pDB2244] y BDG10 [pDB2244] (todo *ΔGSH1)*	BRG5 [pDB2244] (*GSH1* TS)	BSG3 [pDB2244] *(GSH1* con SNP)	BDG1 [pDB2244], BDG2 [pDB2244] y BDG10 [pDB2244] (todo *ΔGSH1)*
Albúmina (rendimiento relativo)	100 % ± 1,47	132 % ± 5,70	(no sometido a ensayo)	100 % ± 6,67	127 % ± 6,84	137 % ± 9,74
(1 desviación típica) Eliminación del valor P (prueba de la t) (cultivado con L-glutatión) en comparación con TS (cultivado sin L-glutatión): 1,25E-08 (véase la tabla 4, columnas F y A) Eliminación del valor P (prueba de la t) en comparación con TS (ambos cultivados con L-glutatión): 8,90E-05 (véase la tabla 4, columnas F y D)						

Como se muestra en la tabla 4, la eliminación de *GSH1* en BXP10 [pDB2244], dio como resultado un aumento en el rendimiento de la albúmina en comparación con la cepa con GSH1 de tipo silvestre. Además, como también se ha mostrado en el ejemplo 1, la presencia del SNP en GSH1 aumentó el rendimiento de la albúmina en comparación con la cepa que contenía GSH1 de tipo silvestre.

Tabla 5: Productividad de la albúmina en cepas procedentes de DP9: PRG13 [pDB2305] (*GSH1* TS), PSG10 [pDB2305] *(GSH1* con SNP A373G) y PDG14 [pDB2305] *(ΔGSH1)*

	Sin L-glutatión			Con L-glutatión		
	A	B	C	D	E	F
	PRG13 [PDB2305] (*GSH1* TS)	PSG10 [PDB2305] *(GSH1* con SNP)	PDG14 [PDB2305] *(ΔGSH1)*	PRG13 [PDB2305] (*GSH1* TS)	PSG10 [PDB2305] *(GSH1* con SNP)	PDG14 [PDB2305] *(ΔGSH1)*

(continuación)

	Sin L-glutatión			Con L-glutatión		
	A	B	C	D	E	F
Albúmina (rendimiento relativo)	100 % ± 5,30 %	130 % ± 6,49 %	160 % ± 8,63 %	100 % ± 2,92 %	134 % ± 5,47 %	160 % ± 6,14 %
(1 desviación típica) Eliminación del valor P (prueba de la t) (cultivado con L-glutatión) en comparación con TS (cultivado sin L-glutatión): 8,45895E-09 (véase la tabla 5, columnas F y A) Eliminación del valor P (prueba de la t) en comparación con TS (ambos cultivados con L-glutatión): 1,02983E-05 (véase la tabla 5, columnas F y D)						

Como se muestra en la tabla 5, la eliminación de *GSH1* en DP9 [pDB2305], dio como resultado un aumento en el rendimiento de la albúmina en comparación con la cepa con GSH1 de tipo silvestre y en comparación con la cepa con el *GSH1* con el SNP.

Ejemplo 3: La mutación del gen *GSH1* mejora el rendimiento de fermentación de la albúmina a una escala de 10 l

Las cepas procedentes de BXP10 del ejemplo 1 se cultivaron a escala de 10 litros en fermentadores y la tasa de rendimiento de productividad celular específica (Y_{PXT}) se midió (Tabla 6). Las fermentaciones se llevaron a cabo de manera similar a la descrita para el ejemplo 5.

Tabla 6: Productividad de la albúmina en fermentaciones de 10 litros en cepas procedentes de BXP10

	BRG5 [pDB2244] (*GSH1* TS)	BSG3 [pDB2244] (GSH1 con SNP A373G)
Albúmina (rendimiento relativo, Y_{PXT}) (n = 6)	100	104,61

Se observó un aumento promedio del 4,6 % en la cepa que tenía GSH1 que contenía el SNP A373G (lo que dio lugar a R125G), esto fue estadísticamente significativo. Valor de p = 0,069

Ejemplo 4: El efecto positivo en el rendimiento de la proteína recombinante de *S. cerevisiae* proporcionado por una mutación en *GSH1* se ve reforzado por la presencia de una mutación en *NOT4*

Anteriormente (Documento PCT/US2016/068239, posteriormente publicado como documento WO 2017/112847), se identificó que la mutación de *NOT4* (también conocido como *MOT2)* da como resultado un mayor rendimiento de la proteína heteróloga producida en una cepa hospedadora.

Para determinar si combinar o no una mutación en *NOT4* con una mutación en GSH1 afecta aún más al rendimiento, se complementaron una, ambas o ninguna de las mutaciones en las células hospedadoras (mediante la transformación con un plásmido que contenía una versión de tipo silvestre de ninguno, uno o ambos genes) dando como resultado una serie de cepas que tenían fenotipos que se aproximan a:

(A) GSH1 que contiene el SNP, y *NOT4* que contiene el SNP,
(B) GSH1 que contiene el SNP y *NOT4* de tipo silvestre,
(C) GSH1 de tipo silvestre y *NOT4* que contiene el SNP, y
(D) GSH1 de tipo silvestre y *NOT4* de tipo silvestre.

Los rendimientos de la proteína heteróloga se midieron mediante análisis GP-HPLC de los sobrenadantes de los cultivos de las cepas resultantes.

Para generar plásmidos que contuvieran GSH1 y/o *NOT4* de tipo silvestre, se llevaron a cabo PCR con ADN genómico de la cepa de levadura DB1. Esta cepa tiene *NOT4* de tipo silvestre y GSH1 de tipo silvestre. Se utilizaron pares de cebadores (Tabla 7) para producir ADN que contenía los ORF de tipo silvestre y las secuencias cadena arriba y cadena abajo de *NOT4* y *GSH1.* Los cebadores se diseñaron para incluir sitios de enzimas de restricción para permitir la clonación en el plásmido YCP50 (Figura 5 y Rose *et al.*, 1987, *Gene*, 60, 237-243). Las PCR se realizaron de acuerdo con la tabla 8.

Tabla 7: Cebadores de PCR

Nombre del cebador	Secuencia (Bases subrayadas = sitio de restricción)	Sitio de restricción	SEQ ID NO:
MBP393	5' - TATTATGAATTCAAATGTTGAGCCCGAAGACG - 3'	*Eco*RI	37
MBP396	5' - TATTATAAGCTTAAATTAGCGAAGCAGGTTCC - 3'	*Hind*III	38
MBP397	5' - TATTATGCATGCCACGTATTCTTGTGCACACG - 3'	*Sph*I	39

(continuación)

Nombre del cebador	Secuencia (Bases subrayadas = sitio de restricción)	Sitio de restricción	SEQ ID NO:
MBP406	5' - TATTATGTCGACTACCACCTACACCAATAAGC - 3'	*Sal*I	40

Tabla 8: Condiciones de PCR

Gen diana	Cebadores	Temperatura de hibridación	Región genómica amplificada por PCR
NOT4	MBP393 + MBP396	64 °C	385 pb cadena arriba de ORF a 309 pb cadena abajo
GSH1	MBP397 + MBP406	62 °C	596 pb cadena arriba de ORF a 680 pb cadena abajo

Las condiciones de PCR fueron las siguientes: 0,1 µg de ADN genómico, 0,5 µM de cada cebador, dNTP 0,2 µM, desnaturalización inicial durante 30 segundos a 98 °C, a continuación, 35 ciclos a 98 °C durante 10 segundos, hibridación a temperatura superior durante 30 segundos, extensión a 72 °C durante 2 minutos, seguido de una extensión final a 72 °C durante 4 minutos y enfriamiento a 4 °C, utilizando un termociclador Applied Biosystems 2720 y un kit de PCR de polimerasa de ADN de alta fidelidad NEB Q5 Hot Start (M0493S), volumen de reacción total de 50 µl, de acuerdo con las instrucciones del fabricante. Los productos se analizaron mediante electroforesis en gel y se encontró que tenían los tamaños esperados, aproximadamente 2,5 kb para *NOT4* y aproximadamente 3,3 kb para GSH1. Los productos de PCR amplificados se purificaron con un kit de purificación de PCR QIAquick de QIAGEN de acuerdo con las instrucciones del fabricante.

Se realizaron digestiones con enzimas de restricción en los productos de PCR purificados y en el plásmido YCP50. Las enzimas de restricción y los tampones eran de New England Biolabs (NEB) y se siguieron los protocolos del fabricante. Los digestiones realizadas fueron: EcoRI + *Hind*III en el producto de PCR de *NOT4*, EcoRI + *Hind*III en YCP50, *SphI* + *Sal*I en el producto de PCR de GSH1 y *Sph*I y *Sal*I en YCP50. El tampón utilizado fue el tampón CutSmart y las enzimas de restricción fueron todas las enzimas NEB High Fidelity ("HF") (es decir, *Eco*RI-HF *etc.).* Los productos de digestión se incubaron durante aproximadamente 1,5 horas a 37 °C. Las digestiones de los productos de PCR se purificaron con el kit de purificación de PCR QIAquick de QIAGEN y el ADN de YCP50 se purificó a partir de un gel de agarosa TAE, con el kit de extracción de gel QIAquick de QIAGEN, de acuerdo con las instrucciones del fabricante.

Cada uno de los productos de PCR digeridos y purificados se unió al ADN de YCP50 que se había cortado con las mismas enzimas. Las uniones se realizaron utilizando el kit de unión NEB Quick (M2200S) de acuerdo con las instrucciones del fabricante y utilizando una relación molar de 3:1 para el producto de PCR:YCP50. A continuación, se utilizó un microlitro de cada reacción de unión para transformar células de *E. coli* competentes en NEB 5-α (eficacia elevada) (M2987I) de acuerdo con las instrucciones del fabricante y las células se sembraron en LB ampicilina. Se cultivaron durante la noche a partir de las colonias resultantes y se realizaron minipreparaciones de centrifugado QIAprep de QIAGEN. Se realizaron digestiones con enzimas de restricción para identificar plásmidos en los que la clonación parecía haber tenido éxito.

A continuación, el plásmido *YCP50-NOT4* se digirió con *SphI* + *Sal*I y el ADN purificado de un gel de agarosa TAE, con el kit de extracción de gel QIAquick de QIAGEN, de acuerdo con las instrucciones del fabricante. Este ADN se unió al producto de PCR de *GSH1* que había sido digerido con *Sph*I + *SalI* y después se purificó con el kit de purificación QIAquick PCR. De nuevo, la relación molar de ADN del inserto a la cadena principal fue de 3:1. Se transformaron células competentes en NEB 5-α (eficacia elevada) (como anteriormente) y se preparó ADN plasmídico (como anteriormente). Los plásmidos correctos se identificaron mediante digestiones de restricción.

Esto dio lugar a tres nuevos plásmidos: YCP50-*NOT4*, YCP50-*GSH1* e YCP50-*NOT4*-GSH1. Estos, e YCP50 solo, se utilizaron para transformar *ura3* de DYB7 [pDB2305] que contenía el plásmido pDB2305 para la expresión de la seroalbúmina humana recombinante. DYB7 se describe en Payne *et al*., (2008) Applied and Environmental Microbiology 74 (24) 7759-7766. Las transformaciones se realizaron utilizando un kit Sigma Yeast Transformation de acuerdo a las instrucciones del fabricante, excepto después de la etapa donde se centrifuga la mezcla de transformación, el sedimento se volvió a suspender en 200 µl de sorbitol 1 M y después se sembraron 100 µl en placas de agar BMMD. Las placas se incubaron a 30 °C hasta la aparición de colonias.

Se inocularon colonias individuales de cada transformación en pocillos de una placa de microvaloración de 48 pocillos, que contenía 0,5 ml de BMMD por pocillo. La placa se incubó a 30 °C y 200 rpm durante 2 días y después se mezcló un volumen igual de trehalosa al 40 % en cada pocillo y la placa se almacenó a -80 °C. En una fecha posterior, la placa se descongeló y se transfirieron 50 µl de cada pocillo a un pocillo de una placa nueva que contenía 450 µl de BMMD. Esta placa se incubó durante 2 días a 30 °C y 200 rpm y después se subcultivó en otra placa nueva (50 µl en 450 µl de BMMD). Esta placa se incubó durante 4 días. A continuación, la placa se centrifugó a 2000 rpm durante 5 minutos y se transfirieron 200 µl del sobrenadante de cada uno a frascos de HPLC. Las cantidades de seroalbúmina humana en los sobrenadantes se cuantificaron mediante GP-HPLC como se describe

en el ejemplo 1.

Los nuevos plásmidos, YCP50-NOT4, YCP50-*GSH1* y *YCP50-NOT4-GSH1*, se secuenciaron para confirmar que cada uno contenía los insertos correctos. La secuenciación se realizó como se describe en los ejemplos anteriores, excepto que en este caso se utilizaron de 150 a 300 ng de ADN plasmídico para cada reacción

Los rendimientos relativos de la albúmina, medido mediante análisis GP-HPLC de sobrenadantes de cultivo, se muestran en la tabla 9 (a continuación).

Tabla 9: Productividad de la albúmina en cepas que se aproximan a los fenotipos que contienen combinaciones de GSH1 (SNP o de tipo silvestre) y *NOT4* (SNP o de tipo silvestre)

	Plásmidos utilizados para la complementación			
	A	B	C	D
	YCP50	YCP50-*NOT4*	YCP50-*GSH1*	*YCP50-NOT4-GSH1*
Estado resultante de *GSH1*	SNP	SNP	TS*	TS*
Estado resultante de *NOT4*	SNP	TS*	SNP	TS*
Rendimiento de albúmina (relativo)	184 % ± 4,88	177 % ± 4,45	131 % ± 4,59	100 % ± 4,88
Valor de p (prueba de la t) en comparación con A	-	1,10E-03	4,20E-17	1,87E-20
*Nótese que debido al método de construcción de las cepas analizadas en la tabla 9, las cepas que contenían una copia TS de los genes que codifican GSH1 o *NOT4* también contenían las versiones con SNP (es decir, que codificaban Gsh1 R125G y Not4 F429I). Las versiones TS estaban en el plásmido y los genes que contenían el SNP estaban en el genoma. Debido a la complementación, el fenotipo resultante se aproxima al de TS				

Los rendimientos se muestran en relación con (D), la cepa en la que había complementación de ambos genes que contenían el SNP. Se vio que el *NOT4* con SNP (C), dio como resultado un aumento del 31 % en el rendimiento, el *GSH1* con SNP (B) dio como resultado un aumento del 77 % en el rendimiento y la presencia de ambos SNP (A), dio como resultado un aumento del 84 % en el rendimiento. Estos estudios de complementación demuestran que las mutaciones en *GSH1* y *NOT4* afectan de manera positiva el rendimiento de la albúmina, y cuando se combinan, el rendimiento aumenta aún más.

Se espera que dicho aumento en el rendimiento también se observe a mayor escala. Esto se debe a que, los ejemplos 1 y 2 demuestran que la mutación o eliminación de GSH1 afecta al rendimiento de la albúmina producida a partir de *S. cerevisiae* a pequeña escala, los ejemplos 3 y 5 demuestran que esto también se observa a una escala de 10 l y la solicitud anterior PCT/US2016/068239 demostró que la mutación o eliminación de *NOT4* afecta al rendimiento de la albúmina producida a partir de *S. cerevisiae* tanto a escala pequeña como de 10 l.

Ejemplo 5: La mutación del gen *GSH1* de *S. cerevisiae* mejoró el rendimiento de la albúmina a escala de 10 l

La productividad de las cepas de *S. cerevisiae DYB7 ura3* [pDB2305 / YCp50-*NOT4*], *DYB7 ura3* [pDB2305 / YCp50-*GSH1]* y *DYB7 ura3* [pDB2305 / *YCp50-NOT4-GSH1]* para la expresión de proteína recombinante se evaluaron mediante el crecimiento en un fermentador de 10 l (Wigley *et al.*, (2007) Genetic Engineering News. 27 (2): 40 - 42). La fermentación fue como se describe en el ejemplo 1 del documento WO97/33973 utilizando medio MW11 D, excepto que se utilizó el programa informático Wonderware Supervisory Control and Data Acquisition en lugar de MFCS, antes de su uso, el recipiente del fermentador también se sometió a un lavado con ácido cítrico, la reserva de oligoelementos comprendía $Na_2MoO_4.2H_2O$ en lugar de $Na_2MoO_4.5H_2O$, el pH se ajustó a pH 6,2 con solución de amoníaco, la introducción inicial de aire estéril en el recipiente fue de aproximadamente 1,0 vvm (es decir, 1,0 litro) en lugar de 0,5 vvm, durante la fermentación, el flujo de aire se incrementó en una etapa en lugar de dos para mantener un flujo de aire de aproximadamente 1,0 vvm, la tasa de crecimiento específico fue de aproximadamente 0,06 h^{-1} y la constante exponencial (K) se mantuvo en 0,06. "vvm" significa flujo de volumen de gas por unidad de volumen de líquido por minuto.

Los rendimientos relativos se muestran (gramos de producto por litro) en la tabla 10.

Tabla 10: Productividad de la albúmina a escala de 10 litros a partir de *DYB7* [pDB2305/*YCp50-NOT4]* y *DYB7* [pDB2305/*YCp50-NOT4-GSH1]*

Cepa	Estado de *GSH1*	Estado de *NOT4*	% g/l
DYB7 ura3 [pDB2305/YCp50-*NOT4-GSH1]*	TS	TS	100
DYB7 ura3 [pDB2305/YCp50-*NOT4*]	SNP A373G	TS	112

El rendimiento de la albúmina aumenta en una cepa que tiene GSH1 que contiene un SNP en comparación con una

cepa que tiene un fenotipo de GSH1 que se aproxima al tipo silvestre.

Ejemplo 6: La mutación o eliminación del gen GSH1 de *Saccharomyces cerevisiae* mejoró la expresión de una variante de albúmina y la eliminación de *GSH1* mejoró la expresión de una proteína de fusión de albúmina (albúmina-IL-1Ra) y scFv (vHvL)-FLAG

Las proteínas que se expresaron en este ejemplo fueron (a) una variante de albúmina (SEQ ID NO: 45) que contenía 4 mutaciones en relación con HSA TS (SEQ ID NO: 10), (b) IL-1Ra fusionado de manera genética con el extremo carboxiterminal de la seroalbúmina humana (SEQ ID NO: 47, "albúmina-IL-1Ra") y (c) el scFv, FITC8 (Evans *et al.*, 2010, Protein Expression and Purification 73:113-124, incluidas las referencias 16 y 17) con un marcador FLAG (DYKDDDDK, SEQ ID NO: 50) en su extremo carboxiterminal (SEQ ID NO: 49).

En la preparación para la expresión de albúmina-IL-1Ra, se cortó el plásmido pDB3936 (Figura 7) con las enzimas de restricción Acc65I y *Bam*HI y el plásmido pDB5912 (que contiene un casete de expresión de albúmina-IL-1Ra) se cortó con enzimas *NsiI* y *PvuI.* En la figura 8 se proporciona un mapa de plásmido para pDB5912, la secuencia de ADN que codifica la albúmina-IL-1Ra se muestra en la SEQ ID NO: 46. Las enzimas de restricción y los tampones eran de New England Biolabs. Ambos plásmidos digeridos se purificaron utilizando un kit de purificación Qiagen PCR siguiendo las instrucciones del fabricante.

Las 3 cepas, PDG14 [pDB2305], PSG11 [pDB2305] y PRG13 [pDB2305], se cultivaron en matraces de agitación en medio YEPD y se subcultivaron 3 veces para curarlos del plásmido (pDB2305). Las diluciones de los cultivos finales se sembraron en YEPD y después las colonias individuales de estas placas se sembraron en YEPD. Las manchas en YEPD se transfirieron aplacas con BMMD + L-glutatión 5 mM y se incubaron a 30 °C; una falta de crecimiento en BMMD + L-glutatión 5 mM identificó las células que se habían curado del plásmido. Cada una de las cepas de levadura curadas se transformó, utilizando el kit Sigma Yeast Transformation de acuerdo con las instrucciones del fabricante, con el plásmido pDB5862 (para la expresión de la variante de albúmina) (en la figura 9 se proporciona un mapa de plásmido para pDB5862, la secuencia de ADN que codifica la variante de albúmina se muestra en la SEQ ID NO: 44), pDB3029 (para la expresión de scFv (vHvL)-FLAG) (en la figura 9 se proporciona un mapa de plásmido para pDB3029, la secuencia de ADN que codifica scFv-FLAG se muestra en la SEQ ID NO: 48), o con los productos de digestión de restricción purificados de pDB3936 y pDB5912 (para la expresión de albúmina-IL-1RA del plásmido reparado en hueco pDB3936:GR:pDB5912). Las células se sembraron en BMMD y se incubaron durante 5 días a 30 °C. Se cultivaron seis transformantes de cada cepa en una MTP de 48 pocillos que contenía 0,5 ml de BMMD con L-glutatión 5 mM por pocillo. La placa se incubó durante 48 horas a 30 °C y 200 rpm en una cámara de humedad. A continuación, esta placa se subcultivó transfiriendo 50 µl de cada cultivo a 450 µl de BMMD con L-glutatión 5 mM en una nueva placa. Esta placa se incubó durante 96 horas.

El sobrenadante se aisló mediante centrifugación y la productividad de proteína recombinante (para la variante de albúmina, albúmina-IL-1Ra o ScFv) se determinó por GP-HPLC, como en el ejemplo 1.

Como se muestra en la tabla 11, la presencia del SNP (A373G) o la eliminación, dio como resultado un aumento en el rendimiento de la variante de albúmina. El rendimiento fue 15 % superior en la cepa que contenía el SNP en *GSH1* (PSG11) y 23 % superior en la cepa que contiene la eliminación de *GSH1* (PDG14) en comparación con la cepa con GSH1 de tipo silvestre (PRG13).

Tabla 11: Productividad de la variante de albúmina en PRG13 [pDB5862], PSG11 [pDB5862] y PDG14 [pDB5862]

	PRG13 [pDB5862]	PSG11 [pDB5862]	PPDG14 [pDB5862]
Variante de albúmina (rendimiento relativo)	100 % ± 19,74	115 % ± 18,63	123 % ± 20,34
Valor de p (prueba de la t) para SNP frente a *GSH1* TS: P = 0,052 Valor de p (prueba de la t) para eliminación frente a *GSH1* TS: P = 0,011			

Como se muestra en la tabla 12, la presencia de la eliminación dio como resultado un aumento en el rendimiento de la albúmina-IL-1Ra. El rendimiento fue un 14 % superior en la cepa que contenía la eliminación (PDG14) en comparación con la cepa con GSH1 de tipo silvestre (PRG13).

Tabla 12: Productividad de la albúmina-IL-1Ra en PRG13 [pDB3936:GR:pDB5912], PSG11 [pDB3936:GR:pDB5912] y PDG14 [pDB3936:GR:pDB5912]

	PRG13 [pDB3936:GR: pDB5912]	PSG11 [pDB3936:GR: pDB5912]	PPDG14 [pDB3936:GR: pDB5912]
Albúmina-IL-1Ra (rendimiento relativo)	100 % ± 20,69	88 % ± 15,75	114 ± 17,37
Valor de p (prueba de la t) para SNP frente a *GSH1* TS: P = 0,11 Valor de p (prueba de la t) para eliminación frente a *GSH1* TS: P = 0,049			

Como se muestra en la tabla 13, la presencia de la eliminación dio como resultado un aumento en el rendimiento de

ScFv-FLAG. El rendimiento 38 % superior en la cepa que contenía la eliminación de *GSH1* (PDG14) en comparación con la cepa con GSH1 de tipo silvestre (PRG13).

Tabla 13: Productividad de ScFv (vHvL)-FLAG en PRG13 [pDB3029], PSG11 [pDB3029] y PDG14 [pDB3029]

	PRG13 [pDB3029]	PSG11 [pDB3029]	PDG14 [pDB3029]
ScFv (rendimiento relativo)	100 % ± 19,40	108 % ± 19,77	138 % ± 22,32
Valor de p (prueba de la t) para SNP frente a *GSH1* TS: P = 0,33 Valor de p (prueba de la t) para eliminación frente a *GSH1* TS: P = 0,0009			

LISTADO DE SECUENCIAS

<110> Albumedix Ltd

<120> CEPAS DE EXPRESIÓN DE PROTEÍNAS MEJORADAS

<130> 13169-WO-PCT

<150> EP17176932.6
<151> 20/06/2017

<160> 52

<170> PatentIn versión 3.5

<210> 1
<211> 2037
<212> ADN
<213> *Saccharomyces cerevisiae*

<400> 1

atgggactct tagctttggg cacgcctttg cagtggtttg agtctaggac gtacaatgaa 60

cacataaggg atgaaggtat cgagcagttg ttgtatattt tccaagctgc tggtaaaaga 120

gacaatgacc ctctttttg gggagacgag cttgagtaca tggttgtaga ttttgatgat 180

aaggagagaa attctatgct cgacgtttgc catgacaaga tactcactga gcttaatatg 240

gaggattcgt ccctttgtga ggctaacgat gtgagttttc accctgagta tggccggtat 300

atgttagagg caacaccagc ttctccatat ttgaattacg tgggtagtta cgttgaggtt 360

aacatgcaaa aaagacgtgc cattgcagaa tataagctat ctgaatatgc gagacaagat 420

agtaaaaata acttgcatgt gggctccagg tctgtccctt tgacgctgac tgtcttcccg 480

aggatgggat gccccgactt tattaacatt aaggatccgt ggaatcataa aaatgccgct 540

tccaggtctc tgtttttacc cgatgaagtc attaacagac atgtcaggtt tcctaacttg 600

acagcatcca tcaggaccag gcgtggtgaa aaagtttgca tgaatgttcc catgtataaa 660

gatatagcta ctccagaaac ggatgactcc atctacgatc gagattggtt tttaccagaa 720

gacaaagagg cgaaactggc ttccaaaccg ggtttcattt atatggattc catgggtttt 780

ggcatgggct gttcgtgctt acaagtgacc tttcaggcac ccaatatcaa caaggcacgt 840

tacctgtacg atgcattagt gaattttgca cctataatgc tagccttctc tgccgctgcg 900

cctgctttta aaggttggct agccgaccaa gatgttcgtt ggaatgtgat atctggtgcg 960

gtggacgacc gtactccgaa ggaaagaggt gttgcgccat tactacccaa atacaacaag 1020

aacggatttg gaggcattgc caaagacgta caagataaag tccttgaaat accaaagtca 1080

agatatagtt cggttgatct tttcttgggt gggtcgaaat ttttcaatag gacttataac 1140

gacacaaatg tacctattaa tgaaaaagta ttaggacgac tactagagaa tgataaggcg 1200

ccactggact atgatcttgc taaacatttt gcgcatctct acataagaga tccagtatct 1260

acattcgaag aactgttgaa tcaggacaac aaaacgtctt caaatcactt tgaaaacatc 1320

caaagtacaa attggcagac attacgtttt aaaccccccа cacaacaagc aaccccggac 1380

aaaaaggatt ctcctggttg gagagtggaa ttcagaccat ttgaagtgca actattagat 1440

tttgagaacg ctgcgtattc cgtgctcata tacttgattg tcgatagcat tttgaccttt 1500

tccgataata ttaacgcata tattcatatg tccaaagtat gggaaaatat gaagatagcc 1560

catcacagag atgctatcct atttgaaaaa tttcattgga aaaaatcatt tcgcaacgac 1620

accgatgtgg aaactgaaga ttattctata agcgagattt tccataatcc agagaatggt 1680

atatttcctc aatttgttac gccaatccta tgccaaaaag ggtttgtaac caaagattgg 1740

aaagaattaa agcattcttc caaacacgag agactatact attatttaaa gctaatttct 1800

gatagagcaa gcggtgaatt gccaacaaca gcaaaattct ttagaaattt tgtactacaa 1860

catccagatt acaaacatga ttcaaaaatt tcaaagtcga tcaattatga tttgctttct 1920

acgtgtgata gacttaccca tttagacgat tcaaaaggtg aattgacatc ctttttagga 1980

gctgaaattg cagaatatgt aaaaaaaaat aagccttcaa tagaaagcaa atgttaa 2037

<210> 2
<211> 678
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 2

Met Gly Leu Leu Ala Leu Gly Thr Pro Leu Gln Trp Phe Glu Ser Arg
1 5 10 15

Thr Tyr Asn Glu His Ile Arg Asp Glu Gly Ile Glu Gln Leu Leu Tyr
20 25 30

Ile Phe Gln Ala Ala Gly Lys Arg Asp Asn Asp Pro Leu Phe Trp Gly
35 40 45

Asp Glu Leu Glu Tyr Met Val Val Asp Phe Asp Asp Lys Glu Arg Asn
50 55 60

Ser Met Leu Asp Val Cys His Asp Lys Ile Leu Thr Glu Leu Asn Met
65 70 75 80

Glu Asp Ser Ser Leu Cys Glu Ala Asn Asp Val Ser Phe His Pro Glu
85 90 95

Tyr Gly Arg Tyr Met Leu Glu Ala Thr Pro Ala Ser Pro Tyr Leu Asn
100 105 110

Tyr Val Gly Ser Tyr Val Glu Val Asn Met Gln Lys Arg Arg Ala Ile

```
        115                 120                 125

Ala Glu Tyr Lys Leu Ser Glu Tyr Ala Arg Gln Asp Ser Lys Asn Asn
    130                 135                 140

Leu His Val Gly Ser Arg Ser Val Pro Leu Thr Leu Thr Val Phe Pro
145                 150                 155                 160

Arg Met Gly Cys Pro Asp Phe Ile Asn Ile Lys Asp Pro Trp Asn His
                165                 170                 175

Lys Asn Ala Ala Ser Arg Ser Leu Phe Leu Pro Asp Glu Val Ile Asn
            180                 185                 190

Arg His Val Arg Phe Pro Asn Leu Thr Ala Ser Ile Arg Thr Arg Arg
        195                 200                 205

Gly Glu Lys Val Cys Met Asn Val Pro Met Tyr Lys Asp Ile Ala Thr
    210                 215                 220

Pro Glu Thr Asp Asp Ser Ile Tyr Asp Arg Asp Trp Phe Leu Pro Glu
225                 230                 235                 240

Asp Lys Glu Ala Lys Leu Ala Ser Lys Pro Gly Phe Ile Tyr Met Asp
                245                 250                 255

Ser Met Gly Phe Gly Met Gly Cys Ser Cys Leu Gln Val Thr Phe Gln
            260                 265                 270

Ala Pro Asn Ile Asn Lys Ala Arg Tyr Leu Tyr Asp Ala Leu Val Asn
        275                 280                 285

Phe Ala Pro Ile Met Leu Ala Phe Ser Ala Ala Ala Pro Ala Phe Lys
    290                 295                 300

Gly Trp Leu Ala Asp Gln Asp Val Arg Trp Asn Val Ile Ser Gly Ala
305                 310                 315                 320

Val Asp Asp Arg Thr Pro Lys Glu Arg Gly Val Ala Pro Leu Leu Pro
                325                 330                 335

Lys Tyr Asn Lys Asn Gly Phe Gly Gly Ile Ala Lys Asp Val Gln Asp
            340                 345                 350

Lys Val Leu Glu Ile Pro Lys Ser Arg Tyr Ser Ser Val Asp Leu Phe
        355                 360                 365
```

```
Leu Gly Gly Ser Lys Phe Phe Asn Arg Thr Tyr Asn Asp Thr Asn Val
    370                 375                 380

Pro Ile Asn Glu Lys Val Leu Gly Arg Leu Leu Glu Asn Asp Lys Ala
385                 390                 395                 400

Pro Leu Asp Tyr Asp Leu Ala Lys His Phe Ala His Leu Tyr Ile Arg
                405                 410                 415

Asp Pro Val Ser Thr Phe Glu Glu Leu Leu Asn Gln Asp Asn Lys Thr
            420                 425                 430

Ser Ser Asn His Phe Glu Asn Ile Gln Ser Thr Asn Trp Gln Thr Leu
        435                 440                 445

Arg Phe Lys Pro Pro Thr Gln Gln Ala Thr Pro Asp Lys Lys Asp Ser
    450                 455                 460

Pro Gly Trp Arg Val Glu Phe Arg Pro Phe Glu Val Gln Leu Leu Asp
465                 470                 475                 480

Phe Glu Asn Ala Ala Tyr Ser Val Leu Ile Tyr Leu Ile Val Asp Ser
                485                 490                 495

Ile Leu Thr Phe Ser Asp Asn Ile Asn Ala Tyr Ile His Met Ser Lys
            500                 505                 510

Val Trp Glu Asn Met Lys Ile Ala His His Arg Asp Ala Ile Leu Phe
        515                 520                 525

Glu Lys Phe His Trp Lys Lys Ser Phe Arg Asn Asp Thr Asp Val Glu
    530                 535                 540

Thr Glu Asp Tyr Ser Ile Ser Glu Ile Phe His Asn Pro Glu Asn Gly
545                 550                 555                 560

Ile Phe Pro Gln Phe Val Thr Pro Ile Leu Cys Gln Lys Gly Phe Val
                565                 570                 575

Thr Lys Asp Trp Lys Glu Leu Lys His Ser Ser Lys His Glu Arg Leu
            580                 585                 590

Tyr Tyr Tyr Leu Lys Leu Ile Ser Asp Arg Ala Ser Gly Glu Leu Pro
        595                 600                 605

Thr Thr Ala Lys Phe Phe Arg Asn Phe Val Leu Gln His Pro Asp Tyr
    610                 615                 620
```

```
Lys His Asp Ser Lys Ile Ser Lys Ser Ile Asn Tyr Asp Leu Leu Ser
625                 630                 635                 640

Thr Cys Asp Arg Leu Thr His Leu Asp Asp Ser Lys Gly Glu Leu Thr
                645                 650                 655

Ser Phe Leu Gly Ala Glu Ile Ala Glu Tyr Val Lys Lys Asn Lys Pro
            660                 665                 670

Ser Ile Glu Ser Lys Cys
        675
```

<210> 3
<211> 2037
<212> ADN
<213> Secuencia artificial

<220>
<223> GSH1 con mutación A373G

<400> 3

```
atgggactct tagctttggg cacgcctttg cagtggtttg agtctaggac gtacaatgaa       60
cacataaggg atgaaggtat cgagcagttg ttgtatattt tccaagctgc tggtaaaaga      120
gacaatgacc ctctttttg gggagacgag cttgagtaca tggttgtaga ttttgatgat      180
aaggagagaa attctatgct cgacgtttgc catgacaaga tactcactga gcttaatatg      240
gaggattcgt ccctttgtga ggctaacgat gtgagttttc accctgagta tggccggtat      300
atgttagagg caacaccagc ttctccatat ttgaattacg tgggtagtta cgttgaggtt      360
aacatgcaaa aaggacgtgc cattgcagaa tataagctat ctgaatatgc gagacaagat      420
agtaaaaata acttgcatgt gggctccagg tctgtccctt tgacgctgac tgtcttcccg      480
aggatgggat gccccgactt tattaacatt aaggatccgt ggaatcataa aaatgccgct      540
tccaggtctc tgtttttacc cgatgaagtc attaacagac atgtcaggtt tcctaacttg      600
acagcatcca tcaggaccag gcgtggtgaa aaagtttgca tgaatgttcc catgtataaa      660
gatatagcta ctccagaaac ggatgactcc atctacgatc gagattggtt tttaccagaa      720
gacaaagagg cgaaactggc ttccaaaccg ggtttcattt atatggattc catgggtttt      780
ggcatgggct gttcgtgctt acaagtgacc tttcaggcac ccaatatcaa caaggcacgt      840
tacctgtacg atgcattagt gaattttgca cctataatgc tagccttctc tgccgctgcg      900
cctgctttta aaggttggct agccgaccaa gatgttcgtt ggaatgtgat atctggtgcg      960
gtggacgacc gtactccgaa ggaaagaggt gttgcgccat tactacccaa atacaacaag     1020
aacggatttg gaggcattgc caaagacgta caagataaag tccttgaaat accaaagtca     1080
agatatagtt cggttgatct tttcttgggt gggtcgaaat ttttcaatag gacttataac     1140
```

gacacaaatg tacctattaa tgaaaaagta ttaggacgac tactagagaa tgataaggcg 1200

ccactggact atgatcttgc taaacatttt gcgcatctct acataagaga tccagtatct 1260

acattcgaag aactgttgaa tcaggacaac aaaacgtctt caaatcactt tgaaaacatc 1320

caaagtacaa attggcagac attacgtttt aaaccccccа cacaacaagc aaccccggac 1380

aaaaaggatt ctcctggttg gagagtggaa ttcagaccat ttgaagtgca actattagat 1440

tttgagaacg ctgcgtattc cgtgctcata tacttgattg tcgatagcat tttgaccttt 1500

tccgataata ttaacgcata tattcatatg tccaaagtat gggaaaatat gaagatagcc 1560

catcacagag atgctatcct atttgaaaaa tttcattgga aaaaatcatt tcgcaacgac 1620

accgatgtgg aaactgaaga ttattctata agcgagattt tccataatcc agagaatggt 1680

atatttcctc aatttgttac gccaatccta tgccaaaaag ggtttgtaac caaagattgg 1740

aaagaattaa agcattcttc caaacacgag agactatact attatttaaa gctaatttct 1800

gatagagcaa gcggtgaatt gccaacaaca gcaaaattct ttagaaattt tgtactacaa 1860

catccagatt acaaacatga ttcaaaaatt tcaaagtcga tcaattatga tttgctttct 1920

acgtgtgata gacttaccca tttagacgat tcaaaaggtg aattgacatc ctttttagga 1980

gctgaaattg cagaatatgt aaaaaaaaat aagccttcaa tagaaagcaa atgttaa 2037

<210> 4
<211> 678
<212> PRT
<213> Secuencia artificial

<220>
<223> Gsh1 con mutación Arg125Gly

<400> 4

Met Gly Leu Leu Ala Leu Gly Thr Pro Leu Gln Trp Phe Glu Ser Arg
1 5 10 15

Thr Tyr Asn Glu His Ile Arg Asp Glu Gly Ile Glu Gln Leu Leu Tyr
20 25 30

Ile Phe Gln Ala Ala Gly Lys Arg Asp Asn Asp Pro Leu Phe Trp Gly
35 40 45

Asp Glu Leu Glu Tyr Met Val Val Asp Phe Asp Asp Lys Glu Arg Asn
50 55 60

Ser Met Leu Asp Val Cys His Asp Lys Ile Leu Thr Glu Leu Asn Met
65 70 75 80

Glu Asp Ser Ser Leu Cys Glu Ala Asn Asp Val Ser Phe His Pro Glu

```
                85                  90                  95

Tyr Gly Arg Tyr Met Leu Glu Ala Thr Pro Ala Ser Pro Tyr Leu Asn
            100                 105                 110

Tyr Val Gly Ser Tyr Val Glu Val Asn Met Gln Lys Gly Arg Ala Ile
        115                 120                 125

Ala Glu Tyr Lys Leu Ser Glu Tyr Ala Arg Gln Asp Ser Lys Asn Asn
    130                 135                 140

Leu His Val Gly Ser Arg Ser Val Pro Leu Thr Leu Thr Val Phe Pro
145                 150                 155                 160

Arg Met Gly Cys Pro Asp Phe Ile Asn Ile Lys Asp Pro Trp Asn His
                165                 170                 175

Lys Asn Ala Ala Ser Arg Ser Leu Phe Leu Pro Asp Glu Val Ile Asn
            180                 185                 190

Arg His Val Arg Phe Pro Asn Leu Thr Ala Ser Ile Arg Thr Arg Arg
        195                 200                 205

Gly Glu Lys Val Cys Met Asn Val Pro Met Tyr Lys Asp Ile Ala Thr
    210                 215                 220

Pro Glu Thr Asp Asp Ser Ile Tyr Asp Arg Asp Trp Phe Leu Pro Glu
225                 230                 235                 240

Asp Lys Glu Ala Lys Leu Ala Ser Lys Pro Gly Phe Ile Tyr Met Asp
                245                 250                 255

Ser Met Gly Phe Gly Met Gly Cys Ser Cys Leu Gln Val Thr Phe Gln
            260                 265                 270

Ala Pro Asn Ile Asn Lys Ala Arg Tyr Leu Tyr Asp Ala Leu Val Asn
        275                 280                 285

Phe Ala Pro Ile Met Leu Ala Phe Ser Ala Ala Ala Pro Ala Phe Lys
    290                 295                 300

Gly Trp Leu Ala Asp Gln Asp Val Arg Trp Asn Val Ile Ser Gly Ala
305                 310                 315                 320

Val Asp Asp Arg Thr Pro Lys Glu Arg Gly Val Ala Pro Leu Leu Pro
                325                 330                 335
```

```
Lys Tyr Asn Lys Asn Gly Phe Gly Gly Ile Ala Lys Asp Val Gln Asp
            340                 345                 350

Lys Val Leu Glu Ile Pro Lys Ser Arg Tyr Ser Ser Val Asp Leu Phe
        355                 360                 365

Leu Gly Gly Ser Lys Phe Phe Asn Arg Thr Tyr Asn Asp Thr Asn Val
    370                 375                 380

Pro Ile Asn Glu Lys Val Leu Gly Arg Leu Leu Glu Asn Asp Lys Ala
385                 390                 395                 400

Pro Leu Asp Tyr Asp Leu Ala Lys His Phe Ala His Leu Tyr Ile Arg
                405                 410                 415

Asp Pro Val Ser Thr Phe Glu Glu Leu Leu Asn Gln Asp Asn Lys Thr
            420                 425                 430

Ser Ser Asn His Phe Glu Asn Ile Gln Ser Thr Asn Trp Gln Thr Leu
        435                 440                 445

Arg Phe Lys Pro Pro Thr Gln Gln Ala Thr Pro Asp Lys Lys Asp Ser
    450                 455                 460

Pro Gly Trp Arg Val Glu Phe Arg Pro Phe Glu Val Gln Leu Leu Asp
465                 470                 475                 480

Phe Glu Asn Ala Ala Tyr Ser Val Leu Ile Tyr Leu Ile Val Asp Ser
                485                 490                 495

Ile Leu Thr Phe Ser Asp Asn Ile Asn Ala Tyr Ile His Met Ser Lys
            500                 505                 510

Val Trp Glu Asn Met Lys Ile Ala His His Arg Asp Ala Ile Leu Phe
        515                 520                 525

Glu Lys Phe His Trp Lys Lys Ser Phe Arg Asn Asp Thr Asp Val Glu
    530                 535                 540

Thr Glu Asp Tyr Ser Ile Ser Glu Ile Phe His Asn Pro Glu Asn Gly
545                 550                 555                 560

Ile Phe Pro Gln Phe Val Thr Pro Ile Leu Cys Gln Lys Gly Phe Val
                565                 570                 575

Thr Lys Asp Trp Lys Glu Leu Lys His Ser Ser Lys His Glu Arg Leu
            580                 585                 590
```

Tyr Tyr Tyr Leu Lys Leu Ile Ser Asp Arg Ala Ser Gly Glu Leu Pro
595 600 605

Thr Thr Ala Lys Phe Phe Arg Asn Phe Val Leu Gln His Pro Asp Tyr
610 615 620

Lys His Asp Ser Lys Ile Ser Lys Ser Ile Asn Tyr Asp Leu Leu Ser
625 630 635 640

Thr Cys Asp Arg Leu Thr His Leu Asp Asp Ser Lys Gly Glu Leu Thr
645 650 655

Ser Phe Leu Gly Ala Glu Ile Ala Glu Tyr Val Lys Lys Asn Lys Pro
660 665 670

Ser Ile Glu Ser Lys Cys
675

<210> 5
<211> 1764
<212> ADN
<213> *Saccharomyces cerevisiae*

<400> 5

atgatgaatc cacacgttca agaaaatttg caagcaatcc acaacgcctt aagcaatttt 60

gatacgtcat ttttatcgga ggatgaagaa gattattgcc ctctttgtat tgaaccaatg 120

gatattactg ataaaaattt ttttccttgt ccctgtggtt atcaaatttg tcaattttgc 180

tacaataata tcagacaaaa tccagaatta aatggccgtt gcccagcatg tcgtcgtaaa 240

tatgatgacg agaacgtcag atacgtcaca ttatctccgg aggagttaaa aatggagaga 300

gccaagctcg ctaggaagga gaaagaaaga aagcatagag aaaaagaacg taaagagaat 360

gaatatacga ataggaaaca tttatctggt accagagtta tccaaaagaa tttagtgtac 420

gttgttggca tcaatcctcc tgttccatac gaggaagttg cgcccactct gaaatctgaa 480

aaatattttg gccaatatgg taagataaat aagattgtgg ttaatagaaa aacaccccat 540

tctaacaaca caaccagcga gcattatcac catcattcac caggatatgg cgtttacata 600

accttcggat ccaaggacga tgctgcaaga tgtatagctc aggtagacgg gacgtatatg 660

gatggccgcc tgatcaaagc tgcctacggt actactaaat actgttcttc ttatttaaga 720

ggattgccat gcccaaatcc caactgtatg tttttgcatg aacctggtga agaagctgat 780

tcttttaata aaagagaact ccacaataaa caacaagcgc aacagcaaag tggcggaact 840

gcattcacta gatctggaat acacaacaat atatctacca gtaccgctgg ttcaaatacc 900

aatttactaa gtgaaaattt cacaggcaca ccttcaccgg cggcgatgag ggctcagtta 960

catcatgaca gccatacaaa cgctggaaca ccggtattaa cacctgctcc ggtccctgca 1020

gggtcaaatc cttggggagt tactcaatca gcaacacctg taacctctat caatctctct 1080

aaaaacagca gctccataaa cttgccaaca ttaaatgatt ctctgggcca tcatactacc 1140

cccacaacag agaataccat cacaagtacg acaactacta ccaataccaa tgctacaagt 1200

cactcccatg gtagcaagaa gaagcaatct cttgctgcag aggaatacaa agatccttat 1260

gacgcactag ggaatgctgt tgacttttg gatgcaagac tacattctct atcaaattat 1320

cagaagcgcc ctatatctat caaatccaat attattgacg aagaaactta taaaaagtat 1380

ccgtctttgt tttcttggga caagattgag gcctcaaaga aaagtgacaa tacattagcc 1440

aacaaacttg tggagatcct ggctataaag ccaatagact acactgcttc tgtcgttcaa 1500

ttcttgcaga gtgtcaatgt tggtgtaaat gacaatatta caatcacaga taatacgaaa 1560

actcccaccc aaccaataag actgcaaacc gtctcacagc aaatccaacc accattaaac 1620

gtcagtaccc ctccaccggg tatctttggt ccacaacata aggttcctat tcagcagcaa 1680

caaatgggtg atacaagctc aagaaattcc tctgatttac taaatcaact aatcaacgga 1740

aggaaaatta tcgccggtaa ttaa 1764

<210> 6
<211> 587
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 6

```
Met Met Asn Pro His Val Gln Glu Asn Leu Gln Ala Ile His Asn Ala
1               5                  10                  15

Leu Ser Asn Phe Asp Thr Ser Phe Leu Ser Glu Asp Glu Glu Asp Tyr
            20                  25                  30

Cys Pro Leu Cys Ile Glu Pro Met Asp Ile Thr Asp Lys Asn Phe Phe
        35                  40                  45

Pro Cys Pro Cys Gly Tyr Gln Ile Cys Gln Phe Cys Tyr Asn Asn Ile
    50                  55                  60

Arg Gln Asn Pro Glu Leu Asn Gly Arg Cys Pro Ala Cys Arg Arg Lys
65                  70                  75                  80

Tyr Asp Asp Glu Asn Val Arg Tyr Val Thr Leu Ser Pro Glu Glu Leu
                85                  90                  95

Lys Met Glu Arg Ala Lys Leu Ala Arg Lys Glu Lys Glu Arg Lys His
            100                 105                 110
```

```
Arg Glu Lys Glu Arg Lys Glu Asn Glu Tyr Thr Asn Arg Lys His Leu
        115                 120                 125

Ser Gly Thr Arg Val Ile Gln Lys Asn Leu Val Tyr Val Val Gly Ile
    130                 135                 140

Asn Pro Pro Val Pro Tyr Glu Glu Val Ala Pro Thr Leu Lys Ser Glu
145                 150                 155                 160

Lys Tyr Phe Gly Gln Tyr Gly Lys Ile Asn Lys Ile Val Val Asn Arg
                165                 170                 175

Lys Thr Pro His Ser Asn Asn Thr Thr Ser Glu His Tyr His His His
            180                 185                 190

Ser Pro Gly Tyr Gly Val Tyr Ile Thr Phe Gly Ser Lys Asp Asp Ala
        195                 200                 205

Ala Arg Cys Ile Ala Gln Val Asp Gly Thr Tyr Met Asp Gly Arg Leu
    210                 215                 220

Ile Lys Ala Ala Tyr Gly Thr Thr Lys Tyr Cys Ser Ser Tyr Leu Arg
225                 230                 235                 240

Gly Leu Pro Cys Pro Asn Pro Asn Cys Met Phe Leu His Glu Pro Gly
                245                 250                 255

Glu Glu Ala Asp Ser Phe Asn Lys Arg Glu Leu His Asn Lys Gln Gln
            260                 265                 270

Ala Gln Gln Gln Ser Gly Gly Thr Ala Phe Thr Arg Ser Gly Ile His
        275                 280                 285

Asn Asn Ile Ser Thr Ser Thr Ala Gly Ser Asn Thr Asn Leu Leu Ser
    290                 295                 300

Glu Asn Phe Thr Gly Thr Pro Ser Pro Ala Ala Met Arg Ala Gln Leu
305                 310                 315                 320

His His Asp Ser His Thr Asn Ala Gly Thr Pro Val Leu Thr Pro Ala
                325                 330                 335

Pro Val Pro Ala Gly Ser Asn Pro Trp Gly Val Thr Gln Ser Ala Thr
            340                 345                 350

Pro Val Thr Ser Ile Asn Leu Ser Lys Asn Ser Ser Ser Ile Asn Leu
        355                 360                 365
```

```
Pro Thr Leu Asn Asp Ser Leu Gly His His Thr Thr Pro Thr Thr Glu
    370                 375                 380

Asn Thr Ile Thr Ser Thr Thr Thr Thr Thr Asn Thr Asn Ala Thr Ser
385                 390                 395                 400

His Ser His Gly Ser Lys Lys Lys Gln Ser Leu Ala Ala Glu Glu Tyr
                405                 410                 415

Lys Asp Pro Tyr Asp Ala Leu Gly Asn Ala Val Asp Phe Leu Asp Ala
            420                 425                 430

Arg Leu His Ser Leu Ser Asn Tyr Gln Lys Arg Pro Ile Ser Ile Lys
        435                 440                 445

Ser Asn Ile Ile Asp Glu Glu Thr Tyr Lys Lys Tyr Pro Ser Leu Phe
    450                 455                 460

Ser Trp Asp Lys Ile Glu Ala Ser Lys Lys Ser Asp Asn Thr Leu Ala
465                 470                 475                 480

Asn Lys Leu Val Glu Ile Leu Ala Ile Lys Pro Ile Asp Tyr Thr Ala
                485                 490                 495

Ser Val Val Gln Phe Leu Gln Ser Val Asn Val Gly Val Asn Asp Asn
            500                 505                 510

Ile Thr Ile Thr Asp Asn Thr Lys Thr Pro Thr Gln Pro Ile Arg Leu
        515                 520                 525

Gln Thr Val Ser Gln Gln Ile Gln Pro Pro Leu Asn Val Ser Thr Pro
    530                 535                 540

Pro Pro Gly Ile Phe Gly Pro Gln His Lys Val Pro Ile Gln Gln Gln
545                 550                 555                 560

Gln Met Gly Asp Thr Ser Ser Arg Asn Ser Ser Asp Leu Leu Asn Gln
                565                 570                 575

Leu Ile Asn Gly Arg Lys Ile Ile Ala Gly Asn
            580                 585
```

<210> 7
<211> 1764
<212> ADN
<213> Secuencia artificial

<220>
<223> NOT4 con mutación T1285A

<400> 7

atgatgaatc cacacgttca agaaaatttg caagcaatcc acaacgcctt aagcaatttt 60

gatacgtcat ttttatcgga ggatgaagaa gattattgcc ctctttgtat tgaaccaatg 120

gatattactg ataaaaattt ttttccttgt ccctgtggtt atcaaatttg tcaattttgc 180

tacaataata tcagacaaaa tccagaatta aatggccgtt gcccagcatg tcgtcgtaaa 240

tatgatgacg agaacgtcag atacgtcaca ttatctccgg aggagttaaa aatggagaga 300

gccaagctcg ctaggaagga gaaagaaaga aagcatagag aaaaagaacg taaagagaat 360

gaatatacga ataggaaaca tttatctggt accagagtta tccaaaagaa tttagtgtac 420

gttgttggca tcaatcctcc tgttccatac gaggaagttg cgcccactct gaaatctgaa 480

aaatattttg gccaatatgg taagataaat aagattgtgg ttaatagaaa aacaccccat 540

tctaacaaca caaccagcga gcattatcac catcattcac caggatatgg cgtttacata 600

accttcggat ccaaggacga tgctgcaaga tgtatagctc aggtagacgg gacgtatatg 660

gatggccgcc tgatcaaagc tgcctacggt actactaaat actgttcttc ttatttaaga 720

ggattgccat gcccaaatcc caactgtatg tttttgcatg aacctggtga agaagctgat 780

tcttttaata aaagagaact ccacaataaa caacaagcgc aacagcaaag tggcggaact 840

gcattcacta gatctggaat acacaacaat atatctacca gtaccgctgg ttcaaatacc 900

aatttactaa gtgaaaattt cacaggcaca ccttcaccgg cggcgatgag ggctcagtta 960

catcatgaca gccatacaaa cgctggaaca ccggtattaa cacctgctcc ggtccctgca 1020

gggtcaaatc cttggggagt tactcaatca gcaacacctg taacctctat caatctctct 1080

aaaaacagca gctccataaa cttgccaaca ttaaatgatt ctctgggcca tcatactacc 1140

cccacaacag agaataccat cacaagtacg acaactacta ccaataccaa tgctacaagt 1200

cactcccatg gtagcaagaa gaagcaatct cttgctgcag aggaatacaa agatccttat 1260

gacgcactag ggaatgctgt tgacattttg gatgcaagac tacattctct atcaaattat 1320

cagaagcgcc ctatatctat caaatccaat attattgacg aagaaactta taaaaagtat 1380

ccgtctttgt tttcttggga caagattgag gcctcaaaga aaagtgacaa tacattagcc 1440

aacaaacttg tggagatcct ggctataaag ccaatagact acactgcttc tgtcgttcaa 1500

ttcttgcaga gtgtcaatgt tggtgtaaat gacaatatta caatcacaga taatacgaaa 1560

actcccaccc aaccaataag actgcaaacc gtctcacagc aaatccaacc accattaaac 1620

gtcagtaccc ctccaccggg tatctttggt ccacaacata aggttcctat tcagcagcaa 1680

caaatgggtg atacaagctc aagaaattcc tctgatttac taaatcaact aatcaacgga 1740

aggaaaatta tcgccggtaa ttaa 1764

<210> 8
<211> 587
<212> PRT
<213> Secuencia artificial

<220>

<223> Not4 con mutación Phe429Ile

<400> 8

```
Met Met Asn Pro His Val Gln Glu Asn Leu Gln Ala Ile His Asn Ala
1               5                  10                  15

Leu Ser Asn Phe Asp Thr Ser Phe Leu Ser Glu Asp Glu Glu Asp Tyr
            20                  25                  30

Cys Pro Leu Cys Ile Glu Pro Met Asp Ile Thr Asp Lys Asn Phe Phe
        35                  40                  45

Pro Cys Pro Cys Gly Tyr Gln Ile Cys Gln Phe Cys Tyr Asn Asn Ile
    50                  55                  60

Arg Gln Asn Pro Glu Leu Asn Gly Arg Cys Pro Ala Cys Arg Arg Lys
65                  70                  75                  80

Tyr Asp Asp Glu Asn Val Arg Tyr Val Thr Leu Ser Pro Glu Glu Leu
                85                  90                  95

Lys Met Glu Arg Ala Lys Leu Ala Arg Lys Glu Lys Glu Arg Lys His
            100                 105                 110

Arg Glu Lys Glu Arg Lys Glu Asn Glu Tyr Thr Asn Arg Lys His Leu
        115                 120                 125

Ser Gly Thr Arg Val Ile Gln Lys Asn Leu Val Tyr Val Val Gly Ile
    130                 135                 140

Asn Pro Pro Val Pro Tyr Glu Glu Val Ala Pro Thr Leu Lys Ser Glu
145                 150                 155                 160

Lys Tyr Phe Gly Gln Tyr Gly Lys Ile Asn Lys Ile Val Val Asn Arg
                165                 170                 175

Lys Thr Pro His Ser Asn Asn Thr Thr Ser Glu His Tyr His His His
            180                 185                 190

Ser Pro Gly Tyr Gly Val Tyr Ile Thr Phe Gly Ser Lys Asp Asp Ala
        195                 200                 205
```

```
Ala Arg Cys Ile Ala Gln Val Asp Gly Thr Tyr Met Asp Gly Arg Leu
    210                 215                 220

Ile Lys Ala Ala Tyr Gly Thr Thr Lys Tyr Cys Ser Ser Tyr Leu Arg
225                 230                 235                 240

Gly Leu Pro Cys Pro Asn Pro Asn Cys Met Phe Leu His Glu Pro Gly
                245                 250                 255

Glu Glu Ala Asp Ser Phe Asn Lys Arg Glu Leu His Asn Lys Gln Gln
            260                 265                 270

Ala Gln Gln Gln Ser Gly Gly Thr Ala Phe Thr Arg Ser Gly Ile His
        275                 280                 285

Asn Asn Ile Ser Thr Ser Thr Ala Gly Ser Asn Thr Asn Leu Leu Ser
    290                 295                 300

Glu Asn Phe Thr Gly Thr Pro Ser Pro Ala Ala Met Arg Ala Gln Leu
305                 310                 315                 320

His His Asp Ser His Thr Asn Ala Gly Thr Pro Val Leu Thr Pro Ala
                325                 330                 335

Pro Val Pro Ala Gly Ser Asn Pro Trp Gly Val Thr Gln Ser Ala Thr
            340                 345                 350

Pro Val Thr Ser Ile Asn Leu Ser Lys Asn Ser Ser Ser Ile Asn Leu
        355                 360                 365

Pro Thr Leu Asn Asp Ser Leu Gly His His Thr Thr Pro Thr Thr Glu
    370                 375                 380

Asn Thr Ile Thr Ser Thr Thr Thr Thr Thr Asn Thr Asn Ala Thr Ser
385                 390                 395                 400

His Ser His Gly Ser Lys Lys Lys Gln Ser Leu Ala Ala Glu Glu Tyr
                405                 410                 415

Lys Asp Pro Tyr Asp Ala Leu Gly Asn Ala Val Asp Ile Leu Asp Ala
            420                 425                 430

Arg Leu His Ser Leu Ser Asn Tyr Gln Lys Arg Pro Ile Ser Ile Lys
        435                 440                 445

Ser Asn Ile Ile Asp Glu Glu Thr Tyr Lys Lys Tyr Pro Ser Leu Phe
    450                 455                 460
```

```
Ser Trp Asp Lys Ile Glu Ala Ser Lys Lys Ser Asp Asn Thr Leu Ala
465                 470                 475                 480

Asn Lys Leu Val Glu Ile Leu Ala Ile Lys Pro Ile Asp Tyr Thr Ala
                485                 490                 495

Ser Val Val Gln Phe Leu Gln Ser Val Asn Val Gly Val Asn Asp Asn
            500                 505                 510

Ile Thr Ile Thr Asp Asn Thr Lys Thr Pro Thr Gln Pro Ile Arg Leu
        515                 520                 525

Gln Thr Val Ser Gln Gln Ile Gln Pro Pro Leu Asn Val Ser Thr Pro
    530                 535                 540

Pro Pro Gly Ile Phe Gly Pro Gln His Lys Val Pro Ile Gln Gln Gln
545                 550                 555                 560

Gln Met Gly Asp Thr Ser Ser Arg Asn Ser Ser Asp Leu Leu Asn Gln
                565                 570                 575

Leu Ile Asn Gly Arg Lys Ile Ile Ala Gly Asn
            580                 585
```

<210> 9
<211> 1755
<212> ADN
<213> *Homo sapiens*

<400> 9

gatgcacaca agagtgaggt tgctcatcgg tttaaagatt tgggagaaga aaatttcaaa 60

gccttggtgt tgattgcctt tgctcagtat cttcagcagt gtccatttga agatcatgta 120

aaattagtga atgaagtaac tgaatttgca aaaacatgtg ttgctgatga gtcagctgaa 180

aattgtgaca aatcacttca tacccttttt ggagacaaat tatgcacagt tgcaactctt 240

cgtgaaacct atggtgaaat ggctgactgc tgtgcaaaac aagaacctga gagaaatgaa 300

tgcttcttgc aacacaaaga tgacaaccca aacctccccc gattggtgag accagaggtt 360

gatgtgatgt gcactgcttt tcatgacaat gaagagacat ttttgaaaaa atacttatat 420

gaaattgcca gaagacatcc ttacttttat gccccggaac tccttttctt tgctaaaagg 480

tataaagctg cttttacaga atgttgccaa gctgctgata aagctgcctg cctgttgcca 540

aagctcgatg aacttcggga tgaagggaag gcttcgtctg ccaaacagag actcaagtgt 600

gccagtctcc aaaaatttgg agaaagagct ttcaaagcat gggcagtagc tcgcctgagc 660

cagagatttc ccaaagctga gtttgcagaa gtttccaagt tagtgacaga tcttaccaaa 720

gtccacacgg aatgctgcca tggagatctg cttgaatgtg ctgatgacag ggcggacctt 780

gccaagtata tctgtgaaaa tcaagattcg atctccagta aactgaagga atgctgtgaa 840

aaacctctgt tggaaaaatc ccactgcatt gccgaagtgg aaaatgatga gatgcctgct 900

gacttgcctt cattagctgc tgattttgtt gaaagtaagg atgtttgcaa aaactatgct 960

gaggcaaagg atgtcttcct gggcatgttt ttgtatgaat atgcaagaag gcatcctgat 1020

tactctgtcg tgctgctgct gagacttgcc aagacatatg aaaccactct agagaagtgc 1080

tgtgccgctg cagatcctca tgaatgctat gccaaagtgt tcgatgaatt taaacctctt 1140

gtggaagagc ctcagaattt aatcaaacaa aattgtgagc tttttgagca gcttggagag 1200

tacaaattcc agaatgcgct attagttcgt tacaccaaga aagtacccca agtgtcaact 1260

ccaactcttg tagaggtctc aagaaaccta ggaaaagtgg gcagcaaatg ttgtaaacat 1320

cctgaagcaa aaagaatgcc ctgtgcagaa gactatctat ccgtggtcct gaaccagtta 1380

tgtgtgttgc atgagaaaac gccagtaagt gacagagtca ccaaatgctg cacagaatcc 1440

ttggtgaaca ggcgaccatg cttttcagct ctggaagtcg atgaaacata cgttcccaaa 1500

gagtttaatg ctgaaacatt caccttccat gcagatatat gcacactttc tgagaaggag 1560

agacaaatca agaaacaaac tgcacttgtt gagctcgtga aacacaagcc caaggcaaca 1620

aaagagcaac tgaaagctgt tatggatgat ttcgcagctt ttgtagagaa gtgctgcaag 1680

gctgacgata aggagacctg ctttgccgag gagggtaaaa aacttgttgc tgcaagtcaa 1740

gctgccttag gctta 1755

<210> 10
<211> 585
<212> PRT
<213> *Homo sapiens*

<400> 10

```
Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1               5                  10                  15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
            20                  25                  30

Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
        35                  40                  45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
    50                  55                  60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65                  70                  75                  80
```

```
Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
                85                  90                  95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
            100                 105                 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
        115                 120                 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
    130                 135                 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145                 150                 155                 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
                165                 170                 175

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
            180                 185                 190

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
        195                 200                 205

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
    210                 215                 220

Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
225                 230                 235                 240

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
                245                 250                 255

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
            260                 265                 270

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
        275                 280                 285

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
    290                 295                 300

Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
305                 310                 315                 320

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
```

```
                325                 330                 335

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
            340                 345                 350

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
        355                 360                 365

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
    370                 375                 380

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
385                 390                 395                 400

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
                405                 410                 415

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
            420                 425                 430

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
        435                 440                 445

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
    450                 455                 460

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
465                 470                 475                 480

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
                485                 490                 495

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
            500                 505                 510

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
        515                 520                 525

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
    530                 535                 540

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
545                 550                 555                 560

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
                565                 570                 575
```

Ala Ala Ser Gln Ala Ala Leu Gly Leu
580 585

<210> 11
<211> 1827
<212> ADN
<213> *Homo sapiens*

<400> 11

atgaagtggg taagctttat ttcccttctt tttctcttta gctcggctta ttccaggagc 60

ttggataaaa gagatgcaca caagagtgag gttgctcatc ggtttaaaga tttgggagaa 120

gaaaatttca aagccttggt gttgattgcc tttgctcagt atcttcagca gtgtccattt 180

gaagatcatg taaaattagt gaatgaagta actgaatttg caaaaacatg tgttgctgat 240

gagtcagctg aaaattgtga caaatcactt catacccttt ttggagacaa attatgcaca 300

gttgcaactc ttcgtgaaac ctatggtgaa atggctgact gctgtgcaaa acaagaacct 360

gagagaaatg aatgcttctt gcaacacaaa gatgacaacc caaacctccc ccgattggtg 420

agaccagagg ttgatgtgat gtgcactgct tttcatgaca atgaagagac atttttgaaa 480

aaatacttat atgaaattgc cagaagacat ccttactttt atgccccgga actccttttc 540

tttgctaaaa ggtataaagc tgcttttaca gaatgttgcc aagctgctga taaagctgcc 600

tgcctgttgc caaagctcga tgaacttcgg gatgaaggga aggcttcgtc tgccaaacag 660

agactcaagt gtgccagtct ccaaaaattt ggagaaagag ctttcaaagc atgggcagta 720

gctcgcctga gccagagatt tcccaaagct gagtttgcag aagtttccaa gttagtgaca 780

gatcttacca aagtccacac ggaatgctgc catggagatc tgcttgaatg tgctgatgac 840

agggcggacc ttgccaagta tatctgtgaa aatcaagatt cgatctccag taaactgaag 900

gaatgctgtg aaaaacctct gttggaaaaa tcccactgca ttgccgaagt ggaaaatgat 960

gagatgcctg ctgacttgcc ttcattagct gctgattttg ttgaaagtaa ggatgtttgc 1020

aaaaactatg ctgaggcaaa ggatgtcttc ctgggcatgt ttttgtatga atatgcaaga 1080

aggcatcctg attactctgt cgtgctgctg ctgagacttg ccaagacata tgaaaccact 1140

ctagagaagt gctgtgccgc tgcagatcct catgaatgct atgccaaagt gttcgatgaa 1200

tttaaacctc ttgtggaaga gcctcagaat ttaatcaaac aaaattgtga gctttttgag 1260

cagcttggag agtacaaatt ccagaatgcg ctattagttc gttacaccaa gaaagtaccc 1320

caagtgtcaa ctccaactct tgtagaggtc tcaagaaacc taggaaaagt gggcagcaaa 1380

tgttgtaaac atcctgaagc aaaaagaatg ccctgtgcag aagactatct atccgtggtc 1440

ctgaaccagt tatgtgtgtt gcatgagaaa acgccagtaa gtgacagagt caccaaatgc 1500

tgcacagaat ccttggtgaa caggcgacca tgcttttcag ctctggaagt cgatgaaaca 1560

tacgttccca aagagtttaa tgctgaaaca ttcaccttcc atgcagatat atgcacactt 1620

tctgagaagg agagacaaat caagaaacaa actgcacttg ttgagctcgt gaaacacaag 1680

cccaaggcaa caaaagagca actgaaagct gttatggatg atttcgcagc ttttgtagag 1740

aagtgctgca aggctgacga taaggagacc tgctttgccg aggagggtaa aaaacttgtt 1800

gctgcaagtc aagctgcctt aggctta 1827

<210> 12
<211> 609
<212> PRT

<213> *Homo sapiens*

<400> 12

```
Met Lys Trp Val Thr Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1               5                  10                  15

Tyr Ser Arg Gly Val Phe Arg Arg Asp Ala His Lys Ser Glu Val Ala
            20                  25                  30

His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys Ala Leu Val Leu
        35                  40                  45

Ile Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe Glu Asp His Val
    50                  55                  60

Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr Cys Val Ala Asp
65                  70                  75                  80

Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr Leu Phe Gly Asp
                85                  90                  95

Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr Gly Glu Met Ala
            100                 105                 110

Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu Cys Phe Leu Gln
        115                 120                 125

His Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val Arg Pro Glu Val
    130                 135                 140

Asp Val Met Cys Thr Ala Phe His Asp Asn Glu Glu Thr Phe Leu Lys
145                 150                 155                 160

Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr Phe Tyr Ala Pro
                165                 170                 175

Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala Phe Thr Glu Cys
            180                 185                 190
```

```
Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro Lys Leu Asp Glu
        195                 200                 205

Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln Arg Leu Lys Cys
    210                 215                 220

Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys Ala Trp Ala Val
225                 230                 235                 240

Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe Ala Glu Val Ser
                245                 250                 255

Lys Leu Val Thr Asp Leu Thr Lys Val His Thr Glu Cys Cys His Gly
            260                 265                 270

Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu Ala Lys Tyr Ile
        275                 280                 285

Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys Glu Cys Cys Glu
    290                 295                 300

Lys Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu Val Glu Asn Asp
305                 310                 315                 320

Glu Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp Phe Val Glu Ser
                325                 330                 335

Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp Val Phe Leu Gly
            340                 345                 350

Met Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp Tyr Ser Val Val
        355                 360                 365

Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr Leu Glu Lys Cys
    370                 375                 380

Cys Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys Val Phe Asp Glu
385                 390                 395                 400

Phe Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile Lys Gln Asn Cys
                405                 410                 415

Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln Asn Ala Leu Leu
            420                 425                 430

Val Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr Pro Thr Leu Val
```

```
          435                 440                 445

Glu Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys Cys Cys Lys His
    450                 455                 460

Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr Leu Ser Val Val
465                 470                 475                 480

Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro Val Ser Asp Arg
                485                 490                 495

Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg Arg Pro Cys Phe
            500                 505                 510

Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys Glu Phe Asn Ala
        515                 520                 525

Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu Ser Glu Lys Glu
    530                 535                 540

Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu Val Lys His Lys
545                 550                 555                 560

Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met Asp Asp Phe Ala
                565                 570                 575

Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys Glu Thr Cys Phe
            580                 585                 590

Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln Ala Ala Leu Gly
        595                 600                 605

Leu
```

<210> 13
<211> 2108
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 13

```
Met Leu Ser Ala Thr Tyr Arg Asp Leu Asn Thr Ala Ser Asn Leu Glu
1               5                  10                  15

Thr Ser Lys Glu Lys Gln Ala Ala Gln Ile Val Ile Ala Gln Ile Ser
            20                  25                  30

Leu Leu Phe Thr Thr Leu Asn Asn Asp Asn Phe Glu Ser Val Glu Arg
```

```
        35                  40                  45

Glu Ile Arg His Ile Leu Asp Arg Ser Ser Val Asp Ile Tyr Ile Lys
    50                  55                  60

Val Trp Glu Arg Leu Leu Thr Leu Ser Ser Arg Asp Ile Leu Gln Ala
65                  70                  75                  80

Gly Lys Phe Leu Leu Gln Glu Asn Leu Leu His Arg Leu Leu Leu Glu
                85                  90                  95

Phe Ala Lys Asp Leu Pro Lys Lys Ser Thr Asp Leu Ile Glu Leu Leu
            100                 105                 110

Lys Glu Arg Thr Phe Asn Asn Gln Glu Phe Gln Lys Gln Thr Gly Ile
        115                 120                 125

Thr Leu Ser Leu Phe Ile Asp Leu Phe Asp Lys Ser Ala Asn Lys Asp
    130                 135                 140

Ile Ile Glu Ser Leu Asp Arg Ser Ser Gln Ile Asn Asp Phe Lys Thr
145                 150                 155                 160

Ile Lys Met Asn His Thr Asn Tyr Leu Arg Asn Phe Phe Leu Gln Thr
                165                 170                 175

Thr Pro Glu Thr Leu Glu Ser Asn Leu Arg Asp Leu Leu His Ser Leu
            180                 185                 190

Glu Gly Glu Ser Leu Asn Asp Leu Leu Ala Leu Leu Leu Ser Glu Ile
        195                 200                 205

Leu Ser Pro Gly Ser Gln Asn Leu Gln Asn Asp Pro Thr Arg Ser Trp
    210                 215                 220

Leu Thr Pro Pro Met Val Leu Asp Ala Thr Asn Arg Gly Asn Val Ile
225                 230                 235                 240

Ala Arg Ser Ile Ser Ser Leu Gln Ala Asn Gln Ile Asn Trp Asn Arg
                245                 250                 255

Val Phe Asn Leu Met Ser Thr Lys Tyr Phe Leu Ser Ala Pro Leu Met
            260                 265                 270

Pro Thr Thr Ala Ser Leu Ser Cys Leu Phe Ala Ala Leu His Asp Gly
        275                 280                 285
```

```
Pro Val Ile Asp Glu Phe Phe Ser Cys Asp Trp Lys Val Ile Phe Lys
    290                 295                 300

Leu Asp Leu Ala Ile Gln Leu His Lys Trp Ser Val Gln Asn Gly Cys
305                 310                 315                 320

Phe Asp Leu Leu Asn Ala Glu Gly Thr Arg Lys Val Ser Glu Thr Ile
                325                 330                 335

Pro Asn Thr Lys Gln Ser Leu Leu Tyr Leu Leu Ser Ile Ala Ser Leu
            340                 345                 350

Asn Leu Glu Leu Phe Leu Gln Arg Glu Glu Leu Ser Asp Gly Pro Met
        355                 360                 365

Leu Ala Tyr Phe Gln Glu Cys Phe Phe Glu Asp Phe Asn Tyr Ala Pro
    370                 375                 380

Glu Tyr Leu Ile Leu Ala Leu Val Lys Glu Met Lys Arg Phe Val Leu
385                 390                 395                 400

Leu Ile Glu Asn Arg Thr Val Ile Asp Glu Ile Leu Ile Thr Leu Leu
                405                 410                 415

Ile Gln Val His Asn Lys Ser Pro Ser Ser Phe Lys Asp Val Ile Ser
            420                 425                 430

Thr Ile Thr Asp Asp Ser Lys Ile Val Asp Ala Ala Lys Ile Ile Ile
        435                 440                 445

Asn Ser Asp Asp Ala Pro Ile Ala Asn Phe Leu Lys Ser Leu Leu Asp
    450                 455                 460

Thr Gly Arg Leu Asp Thr Val Ile Asn Lys Leu Pro Phe Asn Glu Ala
465                 470                 475                 480

Phe Lys Ile Leu Pro Cys Ala Arg Gln Ile Gly Trp Glu Gly Phe Asp
                485                 490                 495

Thr Phe Leu Lys Thr Lys Val Ser Pro Ser Asn Val Asp Val Val Leu
            500                 505                 510

Glu Ser Leu Glu Val Gln Thr Lys Met Thr Asp Thr Asn Thr Pro Phe
        515                 520                 525

Arg Ser Leu Lys Thr Phe Asp Leu Phe Ala Phe His Ser Leu Ile Glu
    530                 535                 540
```

```
Val Leu Asn Lys Cys Pro Leu Asp Val Leu Gln Leu Gln Arg Phe Glu
545                 550                 555                 560

Ser Leu Glu Phe Ser Leu Leu Ile Ala Phe Pro Arg Leu Ile Asn Phe
                565                 570                 575

Gly Phe Gly His Asp Glu Ala Ile Leu Ala Asn Gly Asp Ile Ala Gly
            580                 585                 590

Ile Asn Asn Asp Ile Glu Lys Glu Met Gln Asn Tyr Leu Gln Lys Met
        595                 600                 605

Tyr Ser Gly Glu Leu Ala Ile Lys Asp Val Ile Glu Leu Leu Arg Arg
    610                 615                 620

Leu Arg Asp Ser Asp Leu Pro Arg Asp Gln Glu Val Phe Thr Cys Ile
625                 630                 635                 640

Thr His Ala Val Ile Ala Glu Ser Thr Phe Phe Gln Asp Tyr Pro Leu
                645                 650                 655

Asp Ala Leu Ala Thr Thr Ser Val Leu Phe Gly Ser Met Ile Leu Phe
            660                 665                 670

Gln Leu Leu Arg Gly Phe Val Leu Asp Val Ala Phe Arg Ile Ile Met
        675                 680                 685

Arg Phe Ala Lys Glu Pro Pro Glu Ser Lys Met Phe Lys Phe Ala Val
    690                 695                 700

Gln Ala Ile Tyr Ala Phe Arg Ile Arg Leu Ala Glu Tyr Pro Gln Tyr
705                 710                 715                 720

Cys Lys Asp Leu Leu Arg Asp Val Pro Ala Leu Lys Ser Gln Ala Gln
                725                 730                 735

Val Tyr Gln Ser Ile Val Glu Ala Ala Thr Leu Ala Asn Ala Pro Lys
            740                 745                 750

Glu Arg Ser Arg Pro Val Gln Glu Met Ile Pro Leu Lys Phe Phe Ala
        755                 760                 765

Val Asp Glu Val Ser Cys Gln Ile Asn Gln Glu Gly Ala Pro Lys Asp
    770                 775                 780

Val Val Glu Lys Val Leu Phe Val Leu Asn Asn Val Thr Leu Ala Asn
785                 790                 795                 800
```

```
Leu Asn Asn Lys Val Asp Glu Leu Lys Lys Ser Leu Thr Pro Asn Tyr
                805                 810                 815

Phe Ser Trp Phe Ser Thr Tyr Leu Val Thr Gln Arg Ala Lys Thr Glu
            820                 825                 830

Pro Asn Tyr His Asp Leu Tyr Ser Lys Val Ile Val Ala Met Gly Ser
        835                 840                 845

Gly Leu Leu His Gln Phe Met Val Asn Val Thr Leu Arg Gln Leu Phe
    850                 855                 860

Val Leu Leu Ser Thr Lys Asp Glu Gln Ala Ile Asp Lys Lys His Leu
865                 870                 875                 880

Lys Asn Leu Ala Ser Trp Leu Gly Cys Ile Thr Leu Ala Leu Asn Lys
                885                 890                 895

Pro Ile Lys His Lys Asn Ile Ala Phe Arg Glu Met Leu Ile Glu Ala
            900                 905                 910

Tyr Lys Glu Asn Arg Leu Glu Ile Val Val Pro Phe Val Thr Lys Ile
        915                 920                 925

Leu Gln Arg Ala Ser Glu Ser Lys Ile Phe Lys Pro Pro Asn Pro Trp
    930                 935                 940

Thr Val Gly Ile Leu Lys Leu Leu Ile Glu Leu Asn Glu Lys Ala Asn
945                 950                 955                 960

Trp Lys Leu Ser Leu Thr Phe Glu Val Glu Val Leu Leu Lys Ser Phe
                965                 970                 975

Asn Leu Thr Thr Lys Ser Leu Lys Pro Ser Asn Phe Ile Asn Thr Pro
            980                 985                 990

Glu Val Ile Glu Thr Leu Ser Gly  Ala Leu Gly Ser Ile  Thr Leu Glu
        995                 1000                1005

Gln Gln  Gln Thr Glu Gln Gln  Arg Gln Ile Ile Leu  Met Gln Gln
    1010                1015                1020

His Gln  Gln Gln Met Leu Ile  Tyr Gln Gln Arg Gln  Gln Gln Gln
    1025                1030                1035

Gln Gln  Arg Gln Gln Gln Gln  Gln His His Ile Ser  Ala Asn Thr
```

```
    1040                1045                1050

Ile Ala  Asp Gln Gln Ala Ala  Phe Gly Gly Glu Gly  Ser Ile Ser
    1055                1060                1065

His Asp  Asn Pro Phe Asn Asn  Leu Leu Gly Ser Thr  Ile Phe Val
    1070                1075                1080

Thr His  Pro Asp Leu Lys Arg  Val Phe Gln Met Ala  Leu Ala Lys
    1085                1090                1095

Ser Val  Arg Glu Ile Leu Leu  Glu Val Val Glu Lys  Ser Ser Gly
    1100                1105                1110

Ile Ala  Val Val Thr Thr Thr  Lys Ile Ile Leu Lys  Asp Phe Ala
    1115                1120                1125

Thr Glu  Val Asp Glu Ser Lys  Leu Lys Thr Ala Ala  Ile Ile Met
    1130                1135                1140

Val Arg  His Leu Ala Gln Ser  Leu Ala Arg Ala Thr  Ser Ile Glu
    1145                1150                1155

Pro Leu  Lys Glu Gly Ile Arg  Ser Thr Met Gln Ser  Leu Ala Pro
    1160                1165                1170

Asn Leu  Met Ser Leu Ser Ser  Ser Pro Ala Glu Glu  Leu Asp Thr
    1175                1180                1185

Ala Ile  Asn Glu Asn Ile Gly  Ile Ala Leu Val Leu  Ile Glu Lys
    1190                1195                1200

Ala Ser  Met Asp Lys Ser Thr  Gln Asp Leu Ala Asp  Gln Leu Met
    1205                1210                1215

Gln Ala  Ile Ala Ile Arg Arg  Tyr His Lys Glu Arg  Arg Ala Asp
    1220                1225                1230

Gln Pro  Phe Ile Thr Gln Asn  Thr Asn Pro Tyr Ser  Leu Ser Leu
    1235                1240                1245

Pro Glu  Pro Leu Gly Leu Lys  Asn Thr Gly Val Thr  Pro Gln Gln
    1250                1255                1260

Phe Arg  Val Tyr Glu Glu Phe  Gly Lys Asn Ile Pro  Asn Leu Asp
    1265                1270                1275
```

```
Val Ile  Pro Phe Ala Gly Leu  Pro Ala His Ala Pro  Pro Met Thr
    1280                 1285                 1290

Gln Asn  Val Gly Leu Thr Gln  Pro Gln Gln Gln Gln  Ala Gln Met
    1295                 1300                 1305

Pro Thr  Gln Ile Leu Thr Ser  Glu Gln Ile Arg Ala  Gln Gln Gln
    1310                 1315                 1320

Gln Gln  Gln Leu Gln Lys Ser  Arg Leu Asn Gln Pro  Ser Gln Ser
    1325                 1330                 1335

Ala Gln  Pro Pro Gly Val Asn  Val Pro Asn Pro Gln  Gly Gly Ile
    1340                 1345                 1350

Ala Ala  Val Gln Ser Asp Leu  Glu Gln Asn Gln Arg  Val Leu Val
    1355                 1360                 1365

His Leu  Met Asp Ile Leu Val  Ser Gln Ile Lys Glu  Asn Ala Thr
    1370                 1375                 1380

Lys Asn  Asn Leu Ala Glu Leu  Gly Asp Gln Asn Gln  Ile Lys Thr
    1385                 1390                 1395

Ile Ile  Phe Gln Ile Leu Thr  Phe Ile Ala Lys Ser  Ala Gln Lys
    1400                 1405                 1410

Asp Gln  Leu Ala Leu Lys Val  Ser Gln Ala Val Val  Asn Ser Leu
    1415                 1420                 1425

Phe Ala  Thr Ser Glu Ser Pro  Leu Cys Arg Glu Val  Leu Ser Leu
    1430                 1435                 1440

Leu Leu  Glu Lys Leu Cys Ser  Leu Ser Leu Val Ala  Arg Lys Asp
    1445                 1450                 1455

Val Val  Trp Trp Leu Val Tyr  Ala Leu Asp Ser Arg  Lys Phe Asn
    1460                 1465                 1470

Val Pro  Val Ile Arg Ser Leu  Leu Glu Val Asn Leu  Ile Asp Ala
    1475                 1480                 1485

Thr Glu  Leu Asp Asn Val Leu  Val Thr Ala Met Lys  Asn Lys Met
    1490                 1495                 1500

Glu Asn  Ser Thr Glu Phe Ala  Met Lys Leu Ile Gln  Asn Thr Val
    1505                 1510                 1515
```

```
Leu Ser  Asp Asp Pro Ile Leu  Met Arg Met Asp Phe  Ile Lys Thr
    1520                1525                1530

Leu Glu  His Leu Ala Ser Ser  Glu Asp Glu Asn Val  Lys Lys Phe
    1535                1540                1545

Ile Lys  Glu Phe Glu Asp Thr  Lys Ile Met Pro Val  Arg Lys Gly
    1550                1555                1560

Thr Lys  Thr Thr Arg Thr Glu  Lys Leu Tyr Leu Val  Phe Thr Glu
    1565                1570                1575

Trp Val  Lys Leu Leu Gln Arg  Val Glu Asn Asn Asp  Val Ile Thr
    1580                1585                1590

Thr Val  Phe Ile Lys Gln Leu  Val Glu Lys Gly Val  Ile Ser Asp
    1595                1600                1605

Thr Asp  Asn Leu Leu Thr Phe  Val Lys Ser Ser Leu  Glu Leu Ser
    1610                1615                1620

Val Ser  Ser Phe Lys Glu Ser  Asp Pro Thr Asp Glu  Val Phe Ile
    1625                1630                1635

Ala Ile  Asp Ala Leu Gly Ser  Leu Ile Ile Lys Leu  Leu Ile Leu
    1640                1645                1650

Gln Gly  Phe Lys Asp Asp Thr  Arg Arg Asp Tyr Ile  Asn Ala Ile
    1655                1660                1665

Phe Ser  Val Ile Val Leu Val  Phe Ala Lys Asp His  Ser Gln Glu
    1670                1675                1680

Gly Thr  Thr Phe Asn Glu Arg  Pro Tyr Phe Arg Leu  Phe Ser Asn
    1685                1690                1695

Ile Leu  Tyr Glu Trp Ala Thr  Ile Arg Thr His Asn  Phe Val Arg
    1700                1705                1710

Ile Ser  Asp Ser Ser Thr Arg  Gln Glu Leu Ile Glu  Phe Asp Ser
    1715                1720                1725

Val Phe  Tyr Asn Thr Phe Ser  Gly Tyr Leu His Ala  Leu Gln Pro
    1730                1735                1740

Phe Ala  Phe Pro Gly Phe Ser  Phe Ala Trp Val Thr  Leu Leu Ser
    1745                1750                1755
```

```
His Arg  Met Leu Leu Pro Ile  Met Leu Arg Leu Pro  Asn Lys Ile
    1760                 1765                 1770

Gly Trp  Glu Lys Leu Met Leu  Leu Ile Ile Asp Leu  Phe Lys Phe
    1775                 1780                 1785

Leu Asp  Gln Tyr Thr Ser Lys  His Ala Val Ser Asp  Ala Val Ser
    1790                 1795                 1800

Val Val  Tyr Lys Gly Thr Leu  Arg Val Ile Leu Gly  Ile Ser Asn
    1805                 1810                 1815

Asp Met  Pro Ser Phe Leu Ile  Glu Asn His Tyr Glu  Leu Met Asn
    1820                 1825                 1830

Asn Leu  Pro Pro Thr Tyr Phe  Gln Leu Lys Asn Val  Ile Leu Ser
    1835                 1840                 1845

Ala Ile  Pro Lys Asn Met Thr  Val Pro Asn Pro Tyr  Asp Val Asp
    1850                 1855                 1860

Leu Asn  Met Glu Asp Ile Pro  Ala Cys Lys Glu Leu  Pro Glu Val
    1865                 1870                 1875

Phe Phe  Asp Pro Val Ile Asp  Leu His Ser Leu Lys  Lys Pro Val
    1880                 1885                 1890

Asp Asn  Tyr Leu Arg Ile Pro  Ser Asn Ser Leu Leu  Arg Thr Ile
    1895                 1900                 1905

Leu Ser  Ala Ile Tyr Lys Asp  Thr Tyr Asp Ile Lys  Lys Gly Val
    1910                 1915                 1920

Gly Tyr  Asp Phe Leu Ser Val  Asp Ser Lys Leu Ile  Arg Ala Ile
    1925                 1930                 1935

Val Leu  His Val Gly Ile Glu  Ala Gly Ile Glu Tyr  Lys Arg Thr
    1940                 1945                 1950

Ser Ser  Asn Ala Val Phe Asn  Thr Lys Ser Ser Tyr  Tyr Thr Leu
    1955                 1960                 1965

Leu Phe  Asn Leu Ile Gln Asn  Gly Ser Ile Glu Met  Lys Tyr Gln
    1970                 1975                 1980

Ile Ile  Leu Ser Ile Val Glu  Gln Leu Arg Tyr Pro  Asn Ile His
```

1985 1990 1995

Thr Tyr Trp Phe Ser Phe Val Leu Met Asn Met Phe Lys Ser Asp
2000 2005 2010

Glu Trp Asn Asp Gln Lys Leu Glu Val Gln Glu Ile Ile Leu Arg
2015 2020 2025

Asn Phe Leu Lys Arg Ile Ile Val Asn Lys Pro His Thr Trp Gly
2030 2035 2040

Val Ser Val Phe Phe Thr Gln Leu Ile Asn Asn Asn Asp Ile Asn
2045 2050 2055

Leu Leu Asp Leu Pro Phe Val Gln Ser Val Pro Glu Ile Lys Leu
2060 2065 2070

Ile Leu Gln Gln Leu Val Lys Tyr Ser Lys Lys Tyr Thr Thr Ser
2075 2080 2085

Glu Gln Asp Asp Gln Ser Ala Thr Ile Asn Arg Arg Gln Thr Pro
2090 2095 2100

Leu Gln Ser Asn Ala
2105

<210> 14
<211> 31
<212> PRT
<213> Secuencia artificial

<220>
<223> fragmento de GLP-1

<400> 14

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
20 25 30

<210> 15
<211> 61
<212> PRT
<213> Secuencia artificial

<220>
<223> Repetición en tándem del fragmento GLP-1

<400> 15

```
His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1               5                  10                  15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg His Gly
            20                  25                  30

Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly Gln Ala
        35                  40                  45

Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly
    50                  55                  60
```

<210> 16
<211> 645
<212> PRT
<213> Secuencia artificial

<220>
<223> Albiglutida (repetición en tándem de GLP-1 fusionada con el extremo aminoterminal de la seroalbúmina humana)

<400> 16

```
His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1               5                  10                  15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg His Gly
            20                  25                  30

Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly Gln Ala
        35                  40                  45

Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Asp Ala His Lys
    50                  55                  60

Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu Glu Asn Phe Lys
65                  70                  75                  80

Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln Gln Cys Pro Phe
                85                  90                  95

Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu Phe Ala Lys Thr
            100                 105                 110

Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys Ser Leu His Thr
        115                 120                 125

Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu Arg Glu Thr Tyr
    130                 135                 140
```

```
Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro Glu Arg Asn Glu
145                 150                 155                 160

Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu Pro Arg Leu Val
                165                 170                 175

Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His Asp Asn Glu Glu
            180                 185                 190

Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg Arg His Pro Tyr
        195                 200                 205

Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg Tyr Lys Ala Ala
    210                 215                 220

Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala Cys Leu Leu Pro
225                 230                 235                 240

Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser Ser Ala Lys Gln
                245                 250                 255

Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu Arg Ala Phe Lys
            260                 265                 270

Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro Lys Ala Glu Phe
        275                 280                 285

Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys Val His Thr Glu
    290                 295                 300

Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp Arg Ala Asp Leu
305                 310                 315                 320

Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser Ser Lys Leu Lys
                325                 330                 335

Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His Cys Ile Ala Glu
            340                 345                 350

Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser Leu Ala Ala Asp
        355                 360                 365

Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala Glu Ala Lys Asp
    370                 375                 380

Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg Arg His Pro Asp
385                 390                 395                 400
```

```
Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr Tyr Glu Thr Thr
                405                 410                 415

Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu Cys Tyr Ala Lys
            420                 425                 430

Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro Gln Asn Leu Ile
        435                 440                 445

Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu Tyr Lys Phe Gln
    450                 455                 460

Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro Gln Val Ser Thr
465                 470                 475                 480

Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys Val Gly Ser Lys
                485                 490                 495

Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys Ala Glu Asp Tyr
            500                 505                 510

Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His Glu Lys Thr Pro
        515                 520                 525

Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser Leu Val Asn Arg
    530                 535                 540

Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr Tyr Val Pro Lys
545                 550                 555                 560

Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp Ile Cys Thr Leu
                565                 570                 575

Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala Leu Val Glu Leu
            580                 585                 590

Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu Lys Ala Val Met
        595                 600                 605

Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys Ala Asp Asp Lys
    610                 615                 620

Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val Ala Ala Ser Gln
625                 630                 635                 640

Ala Ala Leu Gly Leu
                645
```

<210> 17
<211> 24
<212> PRT

<213> Secuencia artificial

<220>
<223> Secuencia líder KEX

<400> 17

Met Lys Trp Val Ser Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Arg Ser Leu Asp Lys Arg
20

<210> 18
<211> 24
<212> PRT
<213> Secuencia artificial

<220>
<223> Secuencia líder KEX modificada

<400> 18

Met Lys Trp Val Ser Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Gly Ser Leu Asp Lys Arg
20

<210> 19
<211> 27
<212> PRT
<213> Secuencia artificial

<220>
<223> Factor α de apareamiento - Secuencia líder de fusión de la seroalbúmina humana

<400> 19

Met Lys Trp Val Ser Phe Ile Ser Leu Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Arg Ser Leu Asp Lys Arg Asp Ala His
20 25

<210> 20
<211> 5
<212> PRT
<213> Secuencia artificial

<220>
<223> Motivo pentapeptídico

<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> Xaa = Phe, Trp o Tyr

<220>

<221> MISC_FEATURE
<222> (2)..(2)
<223> Xaa = Ile, Leu, Val, Ala o Met

<220>
<221> MISC_FEATURE
<222> (3)..(3)
<223> Xaa = Leu, Val, Ala o Met

<220>
<221> MISC_FEATURE
<222> (4)..(4)
<223> Xaa = Ser o Thr

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> Xaa = Ile, Val, Ala o Met

<400> 20

Xaa Xaa Xaa Xaa Xaa
1 5

<210> 21
<211> 19
<212> PRT
<213> Secuencia artificial

<220>
<223> Secuencia líder

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> Xaa = Phe, Trp o Tyr

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> Xaa = Ile, Leu, Val, Ala o Met

<220>
<221> MISC_FEATURE
<222> (7)..(7)
<223> Xaa = Leu, Val, Ala o Met

<220>
<221> MISC_FEATURE
<222> (8)..(8)
<223> Xaa = Ser o Thr

<220>
<221> MISC_FEATURE
<222> (9)..(9)
<223> Xaa = Ile, Val, Ala o Met

<400> 21

Met Lys Trp Val Xaa Xaa Xaa Xaa Xaa Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Arg

<210> 22
<211> 24
<212> PRT
<213> Secuencia artificial

<220>
<223> Secuencia líder

<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> Xaa = Phe, Trp o Tyr

<220>
<221> MISC_FEATURE
<222> (6)..(6)
<223> Xaa = Ile, Leu, Val, Ala o Met

<220>
<221> MISC_FEATURE
<222> (7)..(7)
<223> Xaa = Leu, Val, Ala o Met

<220>
<221> MISC_FEATURE
<222> (8)..(8)
<223> Xaa = Ser o Thr

<220>
<221> MISC_FEATURE
<222> (9)..(9)
<223> Xaa = Ile, Val, Ala o Met

<220>
<221> PROPEP
<222> (20)..(24)

<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> Xaa = Gly o Ser

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> Xaa = Val o Leu

<220>
<221> MISC_FEATURE
<222> (22)..(22)
<223> Xaa = Phe o Asp

<220>
<221> MISC_FEATURE
<222> (23)..(23)
<223> Xaa = Arg o Lys

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> Xaa = Arg o Lys

<400> 22

Met Lys Trp Val Xaa Xaa Xaa Xaa Xaa Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Arg Xaa Xaa Xaa Xaa Xaa
20

<210> 23
<211> 24
<212> PRT
<213> Secuencia artificial

<220>
<223> Secuencia líder

<220>
<221> SEÑAL
<222> (1)..(19)
<223> Presecuencia

<220>
<221> PROPEP
<222> (20)..(24)

<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> Xaa = Gly o Ser

<220>
<221> MISC_FEATURE
<222> (21)..(21)
<223> Xaa = Val o Leu

<220>
<221> MISC_FEATURE
<222> (22)..(22)
<223> Xaa = Phe o Asp

<220>
<221> MISC_FEATURE
<222> (23)..(23)
<223> Xaa = Arg o Lys

<220>
<221> MISC_FEATURE
<222> (24)..(24)
<223> Xaa = Arg o Lys

<400> 23

Met Lys Trp Val Phe Ile Val Ser Ile Leu Phe Leu Phe Ser Ser Ala
1 5 10 15

Tyr Ser Arg Xaa Xaa Xaa Xaa Xaa
20

<210> 24
<211> 24
<212> PRT
<213> Secuencia artificial

<220>

<223> Secuencia líder

<400> 24

```
Met Lys Trp Val Phe Ile Val Ser Ile Leu Phe Leu Phe Ser Ser Ala
1               5                   10                  15

Tyr Ser Arg Ser Leu Asp Lys Arg
            20
```

<210> 25
<211> 522
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 25

```
Met Lys Phe Ser Ala Gly Ala Val Leu Ser Trp Ser Ser Leu Leu Leu
1               5                   10                  15

Ala Ser Ser Val Phe Ala Gln Gln Glu Ala Val Ala Pro Glu Asp Ser
            20                  25                  30

Ala Val Val Lys Leu Ala Thr Asp Ser Phe Asn Glu Tyr Ile Gln Ser
        35                  40                  45

His Asp Leu Val Leu Ala Glu Phe Phe Ala Pro Trp Cys Gly His Cys
    50                  55                  60

Lys Asn Met Ala Pro Glu Tyr Val Lys Ala Ala Glu Thr Leu Val Glu
65                  70                  75                  80

Lys Asn Ile Thr Leu Ala Gln Ile Asp Cys Thr Glu Asn Gln Asp Leu
                85                  90                  95
```

```
Cys Met Glu His Asn Ile Pro Gly Phe Pro Ser Leu Lys Ile Phe Lys
            100                 105                 110

Asn Ser Asp Val Asn Asn Ser Ile Asp Tyr Glu Gly Pro Arg Thr Ala
        115                 120                 125

Glu Ala Ile Val Gln Phe Met Ile Lys Gln Ser Gln Pro Ala Val Ala
    130                 135                 140

Val Val Ala Asp Leu Pro Ala Tyr Leu Ala Asn Glu Thr Phe Val Thr
145                 150                 155                 160

Pro Val Ile Val Gln Ser Gly Lys Ile Asp Ala Asp Phe Asn Ala Thr
                165                 170                 175

Phe Tyr Ser Met Ala Asn Lys His Phe Asn Asp Tyr Asp Phe Val Ser
            180                 185                 190

Ala Glu Asn Ala Asp Asp Asp Phe Lys Leu Ser Ile Tyr Leu Pro Ser
        195                 200                 205

Ala Met Asp Glu Pro Val Val Tyr Asn Gly Lys Lys Ala Asp Ile Ala
    210                 215                 220

Asp Ala Asp Val Phe Glu Lys Trp Leu Gln Val Glu Ala Leu Pro Tyr
225                 230                 235                 240

Phe Gly Glu Ile Asp Gly Ser Val Phe Ala Gln Tyr Val Glu Ser Gly
                245                 250                 255

Leu Pro Leu Gly Tyr Leu Phe Tyr Asn Asp Glu Glu Glu Leu Glu Glu
            260                 265                 270

Tyr Lys Pro Leu Phe Thr Glu Leu Ala Lys Lys Asn Arg Gly Leu Met
        275                 280                 285

Asn Phe Val Ser Ile Asp Ala Arg Lys Phe Gly Arg His Ala Gly Asn
    290                 295                 300

Leu Asn Met Lys Glu Gln Phe Pro Leu Phe Ala Ile His Asp Met Thr
305                 310                 315                 320

Glu Asp Leu Lys Tyr Gly Leu Pro Gln Leu Ser Glu Glu Ala Phe Asp
                325                 330                 335

Glu Leu Ser Asp Lys Ile Val Leu Glu Ser Lys Ala Ile Glu Ser Leu
            340                 345                 350
```

Val Lys Asp Phe Leu Lys Gly Asp Ala Ser Pro Ile Val Lys Ser Gln
355 360 365

Glu Ile Phe Glu Asn Gln Asp Ser Ser Val Phe Gln Leu Val Gly Lys
370 375 380

Asn His Asp Glu Ile Val Asn Asp Pro Lys Lys Asp Val Leu Val Leu
385 390 395 400

Tyr Tyr Ala Pro Trp Cys Gly His Cys Lys Arg Leu Ala Pro Thr Tyr
405 410 415

Gln Glu Leu Ala Asp Thr Tyr Ala Asn Ala Thr Ser Asp Val Leu Ile
420 425 430

Ala Lys Leu Asp His Thr Glu Asn Asp Val Arg Gly Val Val Ile Glu
435 440 445

Gly Tyr Pro Thr Ile Val Leu Tyr Pro Gly Gly Lys Lys Ser Glu Ser
450 455 460

Val Val Tyr Gln Gly Ser Arg Ser Leu Asp Ser Leu Phe Asp Phe Ile
465 470 475 480

Lys Glu Asn Gly His Phe Asp Val Asp Gly Lys Ala Leu Tyr Glu Glu
485 490 495

Ala Gln Glu Lys Ala Ala Glu Glu Ala Asp Ala Asp Ala Glu Leu Ala
500 505 510

Asp Glu Glu Asp Ala Ile His Asp Glu Leu
515 520

<210> 26
<211> 563
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 26

Met Arg Leu Arg Thr Ala Ile Ala Thr Leu Cys Leu Thr Ala Phe Thr
1 5 10 15

Ser Ala Thr Ser Asn Asn Ser Tyr Ile Ala Thr Asp Gln Thr Gln Asn
20 25 30

Ala Phe Asn Asp Thr His Phe Cys Lys Val Asp Arg Asn Asp His Val
35 40 45

```
Ser Pro Ser Cys Asn Val Thr Phe Asn Glu Leu Asn Ala Ile Asn Glu
    50                  55                  60

Asn Ile Arg Asp Asp Leu Ser Ala Leu Leu Lys Ser Asp Phe Phe Lys
65                  70                  75                  80

Tyr Phe Arg Leu Asp Leu Tyr Lys Gln Cys Ser Phe Trp Asp Ala Asn
                85                  90                  95

Asp Gly Leu Cys Leu Asn Arg Ala Cys Ser Val Asp Val Val Glu Asp
            100                 105                 110

Trp Asp Thr Leu Pro Glu Tyr Trp Gln Pro Glu Ile Leu Gly Ser Phe
        115                 120                 125

Asn Asn Asp Thr Met Lys Glu Ala Asp Asp Ser Asp Asp Glu Cys Lys
    130                 135                 140

Phe Leu Asp Gln Leu Cys Gln Thr Ser Lys Lys Pro Val Asp Ile Glu
145                 150                 155                 160

Asp Thr Ile Asn Tyr Cys Asp Val Asn Asp Phe Asn Gly Lys Asn Ala
                165                 170                 175

Val Leu Ile Asp Leu Thr Ala Asn Pro Glu Arg Phe Thr Gly Tyr Gly
            180                 185                 190

Gly Lys Gln Ala Gly Gln Ile Trp Ser Thr Ile Tyr Gln Asp Asn Cys
        195                 200                 205

Phe Thr Ile Gly Glu Thr Gly Glu Ser Leu Ala Lys Asp Ala Phe Tyr
    210                 215                 220

Arg Leu Val Ser Gly Phe His Ala Ser Ile Gly Thr His Leu Ser Lys
225                 230                 235                 240

Glu Tyr Leu Asn Thr Lys Thr Gly Lys Trp Glu Pro Asn Leu Asp Leu
                245                 250                 255

Phe Met Ala Arg Ile Gly Asn Phe Pro Asp Arg Val Thr Asn Met Tyr
            260                 265                 270

Phe Asn Tyr Ala Val Val Ala Lys Ala Leu Trp Lys Ile Gln Pro Tyr
        275                 280                 285

Leu Pro Glu Phe Ser Phe Cys Asp Leu Val Asn Lys Glu Ile Lys Asn
    290                 295                 300
```

```
Lys Met Asp Asn Val Ile Ser Gln Leu Asp Thr Lys Ile Phe Asn Glu
305                 310                 315                 320

Asp Leu Val Phe Ala Asn Asp Leu Ser Leu Thr Leu Lys Asp Glu Phe
                325                 330                 335

Arg Ser Arg Phe Lys Asn Val Thr Lys Ile Met Asp Cys Val Gln Cys
            340                 345                 350

Asp Arg Cys Arg Leu Trp Gly Lys Ile Gln Thr Thr Gly Tyr Ala Thr
        355                 360                 365

Ala Leu Lys Ile Leu Phe Glu Ile Asn Asp Ala Asp Glu Phe Thr Lys
    370                 375                 380

Gln His Ile Val Gly Lys Leu Thr Lys Tyr Glu Leu Ile Ala Leu Leu
385                 390                 395                 400

Gln Thr Phe Gly Arg Leu Ser Glu Ser Ile Glu Ser Val Asn Met Phe
                405                 410                 415

Glu Lys Met Tyr Gly Lys Arg Leu Asn Gly Ser Glu Asn Arg Leu Ser
            420                 425                 430

Ser Phe Phe Gln Asn Asn Phe Phe Asn Ile Leu Lys Glu Ala Gly Lys
        435                 440                 445

Ser Ile Arg Tyr Thr Ile Glu Asn Ile Asn Ser Thr Lys Glu Gly Lys
    450                 455                 460

Lys Lys Thr Asn Asn Ser Gln Ser His Val Phe Asp Asp Leu Lys Met
465                 470                 475                 480

Pro Lys Ala Glu Ile Val Pro Arg Pro Ser Asn Gly Thr Val Asn Lys
                485                 490                 495

Trp Lys Lys Ala Trp Asn Thr Glu Val Asn Asn Val Leu Glu Ala Phe
            500                 505                 510

Arg Phe Ile Tyr Arg Ser Tyr Leu Asp Leu Pro Arg Asn Ile Trp Glu
        515                 520                 525

Leu Ser Leu Met Lys Val Tyr Lys Phe Trp Asn Lys Phe Ile Gly Val
    530                 535                 540

Ala Asp Tyr Val Ser Glu Glu Thr Arg Glu Pro Ile Ser Tyr Lys Leu
```

545 550 555 560

Asp Ile Gln

<210> 27
<211> 2445
<212> ADN
<213> *Saccharomyces cerevisiae*

<400> 27

atgaaagtga ggaaatatat tactttatgc ttttggtggg ccttttcaac atccgctctt 60

gtatcatcac aacaaattcc attgaaggac catacgtcac gacagtattt tgctgtagaa 120

agcaatgaaa cattatcccg cttggaggaa atgcatccaa attggaaata tgaacatgat 180

gttcgagggc taccaaacca ttatgttttt tcaaaagagt tgctaaaatt gggcaaaaga 240

tcatcattag aagagttaca gggggataac aacgaccaca tattatctgt ccatgattta 300

ttcccgcgta acgacctatt taagagacta ccggtgcctg ctccaccaat ggactcaagc 360

ttgttaccgg taaaagaagc tgaggataaa ctcagcataa atgatccgct ttttgagagg 420

cagtggcact tggtcaatcc aagttttcct ggcagtgata taaatgttct tgatctgtgg 480

tacaataata ttacaggcgc aggggtcgtg gctgccattg ttgatgatgg ccttgactac 540

gaaaatgaag acttgaagga taatttttgc gctgaaggtt cttgggattt caacgacaat 600

accaatttac ctaaaccaag attatctgat gactaccatg gtacgagatg tgcaggtgaa 660

atagctgcca aaaaaggtaa caatttttgc ggtgtcgggg taggttacaa cgctaaaatc 720

tcaggcataa gaatcttatc cggtgatatc actacggaag atgaagctgc gtccttgatt 780

tatggtctag acgtaaacga tatatattca tgctcatggg gtcccgctga tgacggaaga 840

catttacaag gccctagtga cctggtgaaa aaggctttag taaaaggtgt tactgaggga 900

agagattcca aaggagcgat ttacgttttt gccagtggaa atggtggaac tcgtggtgat 960

aattgcaatt acgacggcta tactaattcc atatattcta ttactattgg ggctattgat 1020

cacaaagatc tacatcctcc ttattccgaa ggttgttccg ccgtcatggc agtcacgtat 1080

tcttcaggtt caggcgaata tattcattcg agtgatatca acggcagatg cagtaatagc 1140

cacggtggaa cgtctgcggc tgctccatta gctgccggtg tttacacttt gttactagaa 1200

gccaacccaa acctaacttg gagagacgta cagtatttat caatcttgtc tgcggtaggg 1260

ttagaaaaga acgctgacgg agattggaga gatagcgcca tggggaagaa atactctcat 1320

cgctatggct ttggtaaaat cgatgcccat aagttaattg aaatgtccaa gacctgggag 1380

aatgttaacg cacaaacctg gttttacctg ccaacattgt atgtttccca gtccacaaac 1440

tccacggaag agacattaga atccgtcata accatatcag aaaaaagtct tcaagatgct 1500

aacttcaaga gaattgagca cgtcacggta actgtagata ttgatacaga aattagggga 1560

actacgactg tcgatttaat atcaccagcg gggataattt caaaccttgg cgttgtaaga 1620

ccaagagatg tttcatcaga gggattcaaa gactggacat tcatgtctgt agcacattgg 1680

ggtgagaacg gcgtaggtga ttggaaaatc aaggttaaga caacagaaaa tggacacagg 1740

attgacttcc acagttggag gctgaagctc tttggggaat ccattgattc atctaaaaca 1800

gaaactttcg tctttggaaa cgataaagag gaggttgaac cagctgctac agaaagtacc 1860

gtatcacaat attctgccag ttcaacttct atttccatca gcgctacttc tacatcttct 1920

atctcaattg gtgtggaaac gtcggccatt ccccaaacga ctactgcgag taccgatcct 1980

gattctgatc caaacactcc taaaaaactt tcctctccta ggcaagccat gcattatttt 2040

ttaacaatat ttttgattgg cgccacattt ttggtgttat acttcatgtt ttttatgaaa 2100

tcaaggagaa ggatcagaag gtcaagagcg gaaacgtatg aattcgatat cattgataca 2160

gactctgagt acgattctac tttggacaat ggaacttccg gaattactga gcccgaagag 2220

gttgaggact tcgattttga tttgtccgat gaagaccatc ttgcaagttt gtcttcatca 2280

gaaaacggtg atgctgaaca tacaattgat agtgtactaa caaacgaaaa tccatttagt 2340

gaccctataa agcaaaagtt cccaaatgac gccaacgcag aatctgcttc caataaatta 2400

caagaattac agcctgatgt tcctccatct tccggacgat cgtga 2445

<210> 28
<211> 814
<212> PRT
<213> *Saccharomyces cerevisiae*

<400> 28

Met Lys Val Arg Lys Tyr Ile Thr Leu Cys Phe Trp Trp Ala Phe Ser
1 5 10 15

Thr Ser Ala Leu Val Ser Ser Gln Gln Ile Pro Leu Lys Asp His Thr
20 25 30

Ser Arg Gln Tyr Phe Ala Val Glu Ser Asn Glu Thr Leu Ser Arg Leu
35 40 45

Glu Glu Met His Pro Asn Trp Lys Tyr Glu His Asp Val Arg Gly Leu
50 55 60

Pro Asn His Tyr Val Phe Ser Lys Glu Leu Leu Lys Leu Gly Lys Arg
65 70 75 80

Ser Ser Leu Glu Glu Leu Gln Gly Asp Asn Asn Asp His Ile Leu Ser
85 90 95

```
Val His Asp Leu Phe Pro Arg Asn Asp Leu Phe Lys Arg Leu Pro Val
            100                 105                 110

Pro Ala Pro Pro Met Asp Ser Ser Leu Leu Pro Val Lys Glu Ala Glu
        115                 120                 125

Asp Lys Leu Ser Ile Asn Asp Pro Leu Phe Glu Arg Gln Trp His Leu
    130                 135                 140

Val Asn Pro Ser Phe Pro Gly Ser Asp Ile Asn Val Leu Asp Leu Trp
145                 150                 155                 160

Tyr Asn Asn Ile Thr Gly Ala Gly Val Val Ala Ala Ile Val Asp Asp
                165                 170                 175

Gly Leu Asp Tyr Glu Asn Glu Asp Leu Lys Asp Asn Phe Cys Ala Glu
            180                 185                 190

Gly Ser Trp Asp Phe Asn Asp Asn Thr Asn Leu Pro Lys Pro Arg Leu
        195                 200                 205

Ser Asp Asp Tyr His Gly Thr Arg Cys Ala Gly Glu Ile Ala Ala Lys
    210                 215                 220

Lys Gly Asn Asn Phe Cys Gly Val Gly Val Gly Tyr Asn Ala Lys Ile
225                 230                 235                 240

Ser Gly Ile Arg Ile Leu Ser Gly Asp Ile Thr Thr Glu Asp Glu Ala
                245                 250                 255

Ala Ser Leu Ile Tyr Gly Leu Asp Val Asn Asp Ile Tyr Ser Cys Ser
            260                 265                 270

Trp Gly Pro Ala Asp Asp Gly Arg His Leu Gln Gly Pro Ser Asp Leu
        275                 280                 285

Val Lys Lys Ala Leu Val Lys Gly Val Thr Glu Gly Arg Asp Ser Lys
    290                 295                 300

Gly Ala Ile Tyr Val Phe Ala Ser Gly Asn Gly Gly Thr Arg Gly Asp
305                 310                 315                 320

Asn Cys Asn Tyr Asp Gly Tyr Thr Asn Ser Ile Tyr Ser Ile Thr Ile
                325                 330                 335

Gly Ala Ile Asp His Lys Asp Leu His Pro Pro Tyr Ser Glu Gly Cys
            340                 345                 350
```

```
Ser Ala Val Met Ala Val Thr Tyr Ser Ser Gly Ser Gly Glu Tyr Ile
        355                 360                 365

His Ser Ser Asp Ile Asn Gly Arg Cys Ser Asn Ser His Gly Gly Thr
    370                 375                 380

Ser Ala Ala Ala Pro Leu Ala Ala Gly Val Tyr Thr Leu Leu Leu Glu
385                 390                 395                 400

Ala Asn Pro Asn Leu Thr Trp Arg Asp Val Gln Tyr Leu Ser Ile Leu
                405                 410                 415

Ser Ala Val Gly Leu Glu Lys Asn Ala Asp Gly Asp Trp Arg Asp Ser
            420                 425                 430

Ala Met Gly Lys Lys Tyr Ser His Arg Tyr Gly Phe Gly Lys Ile Asp
        435                 440                 445

Ala His Lys Leu Ile Glu Met Ser Lys Thr Trp Glu Asn Val Asn Ala
    450                 455                 460

Gln Thr Trp Phe Tyr Leu Pro Thr Leu Tyr Val Ser Gln Ser Thr Asn
465                 470                 475                 480

Ser Thr Glu Glu Thr Leu Glu Ser Val Ile Thr Ile Ser Glu Lys Ser
                485                 490                 495

Leu Gln Asp Ala Asn Phe Lys Arg Ile Glu His Val Thr Val Thr Val
            500                 505                 510

Asp Ile Asp Thr Glu Ile Arg Gly Thr Thr Thr Val Asp Leu Ile Ser
        515                 520                 525

Pro Ala Gly Ile Ile Ser Asn Leu Gly Val Val Arg Pro Arg Asp Val
    530                 535                 540

Ser Ser Glu Gly Phe Lys Asp Trp Thr Phe Met Ser Val Ala His Trp
545                 550                 555                 560

Gly Glu Asn Gly Val Gly Asp Trp Lys Ile Lys Val Lys Thr Thr Glu
                565                 570                 575

Asn Gly His Arg Ile Asp Phe His Ser Trp Arg Leu Lys Leu Phe Gly
            580                 585                 590

Glu Ser Ile Asp Ser Ser Lys Thr Glu Thr Phe Val Phe Gly Asn Asp
```

595 600 605

Lys Glu Glu Val Glu Pro Ala Ala Thr Glu Ser Thr Val Ser Gln Tyr
610 615 620

Ser Ala Ser Ser Thr Ser Ile Ser Ile Ser Ala Thr Ser Thr Ser Ser
625 630 635 640

Ile Ser Ile Gly Val Glu Thr Ser Ala Ile Pro Gln Thr Thr Thr Ala
645 650 655

Ser Thr Asp Pro Asp Ser Asp Pro Asn Thr Pro Lys Lys Leu Ser Ser
660 665 670

Pro Arg Gln Ala Met His Tyr Phe Leu Thr Ile Phe Leu Ile Gly Ala
675 680 685

Thr Phe Leu Val Leu Tyr Phe Met Phe Phe Met Lys Ser Arg Arg Arg
690 695 700

Ile Arg Arg Ser Arg Ala Glu Thr Tyr Glu Phe Asp Ile Ile Asp Thr
705 710 715 720

Asp Ser Glu Tyr Asp Ser Thr Leu Asp Asn Gly Thr Ser Gly Ile Thr
725 730 735

Glu Pro Glu Glu Val Glu Asp Phe Asp Phe Asp Leu Ser Asp Glu Asp
740 745 750

His Leu Ala Ser Leu Ser Ser Ser Glu Asn Gly Asp Ala Glu His Thr
755 760 765

Ile Asp Ser Val Leu Thr Asn Glu Asn Pro Phe Ser Asp Pro Ile Lys
770 775 780

Gln Lys Phe Pro Asn Asp Ala Asn Ala Glu Ser Ala Ser Asn Lys Leu
785 790 795 800

Gln Glu Leu Gln Pro Asp Val Pro Pro Ser Ser Gly Arg Ser
805 810

<210> 29
<211> 116
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP267

<400> 29

atactattgt aattcaaaaa aaaaaagcga atcttcccat gcctgttgct gctcttgaat 60

ggcgacagcc tattgcccca gtgttccctc aacaaccttg cgtacgctgc aggtcg 116

<210> 30
<211> 119
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP268

<400> 30

acagttgtag tcacgtgcgc gccatgctga ctaatggcag ccgtcgttgg gcagaagaga 60

attagtatgg tacaggatac gctaattgcg ctccaactac atcgatgaat tcgagctcg 119

<210> 31
<211> 2723
<212> ADN
<213> Secuencia artificial

<220>
<223> Fragmento de ORF de GSH1 en 5'UTR *5-GSH1* 5'UTR-KanMX-GSH1. Secuencia amplificada por MBP288 y MBP289 y utilizada para la reversión del SNP de GSH1.

<400> 31

gattttatcg gtcaaagggg aaatcaatgc gaaagacagt aatgatgaga gaaaaactct 60

ccgtaaccac caagtttggt tcagcgcgac gagattttta tcgattatcg agaaaaatac 120

ctgtatatct acatttctat gtcagtgata tatacttctt agataagtta tgccaccagt 180

gcatacgctt acgcacacac acgtattctt gtgcacacgc ctgttacttc ttgcagacat 240

cagacatact attgtaattc aaaaaaaaaa agcgaatctt cccatgcctg ttgctgctct 300

tgaatggcga cagcctattg ccccagtgtt ccctcaacaa ccttgcgtac gctgcaggtc 360

gacggatccc cgggttaatt aaggcgcgcc agatctgttt agcttgcctc gtccccgccg 420

ggtcacccgg ccagcgacat ggaggcccag aataccctcc ttgacagtct tgacgtgcgc 480

agctcagggg catgatgtga ctgtcgcccg tacatttagc ccatacatcc ccatgtataa 540

tcatttgcat ccatacattt tgatggccgc acggcgcgaa gcaaaaatta cggctcctcg 600

ctgcagacct gcgagcaggg aaacgctccc ctcacagacg cgttgaattg tccccacgcc 660

gcgcccctgt agagaaatat aaaaggttag gatttgccac tgaggttctt ctttcatata 720

cttcctttta aaatcttgct aggatacagt tctcacatca catccgaaca taaacaacca 780

tgggtaagga aaagactcac gtttcgaggc cgcgattaaa ttccaacatg gatgctgatt 840

tatatgggta taaatgggct cgcgataatg tcgggcaatc aggtgcgaca atctatcgat 900

tgtatgggaa gcccgatgcg ccagagttgt ttctgaaaca tggcaaaggt agcgttgcca 960

atgatgttac agatgagatg gtcagactaa actggctgac ggaatttatg cctcttccga 1020

ccatcaagca ttttatccgt actcctgatg atgcatggtt actcaccact gcgatccccg 1080

gcaaaacagc attccaggta ttagaagaat atcctgattc aggtgaaaat attgttgatg 1140

cgctggcagt gttcctgcgc cggttgcatt cgattcctgt ttgtaattgt ccttttaaca 1200

gcgatcgcgt atttcgtctc gctcaggcgc aatcacgaat gaataacggt ttggttgatg 1260

cgagtgattt tgatgacgag cgtaatggct ggcctgttga acaagtctgg aaagaaatgc 1320

ataagctttt gccattctca ccggattcag tcgtcactca tggtgatttc tcacttgata 1380

accttatttt tgacgagggg aaattaatag gttgtattga tgttggacga gtcggaatcg 1440

cagaccgata ccaggatctt gccatcctat ggaactgcct cggtgagttt tctccttcat 1500

tacagaaacg gctttttcaa aaatatggta ttgataatcc tgatatgaat aaattgcagt 1560

ttcatttgat gctcgatgag tttttctaat cagtactgac aataaaaaga ttcttgtttt 1620

caagaacttg tcatttgtat agtttttta tattgtagtt gttctatttt aatcaaatgt 1680

tagcgtgatt tatatttttt ttcgcctcga catcatctgc ccagatgcga agttaagtgc 1740

gcagaaagta atatcatgcg tcaatcgtat gtgaatgctg gtcgctatac tgctgtcgat 1800

tcgatactaa cgccgccatc cagtgtcgaa aacgagctcg aattcatcga tgtagttgga 1860

gcgcaattag cgtatcctgt accatactaa ttctcttctg cccaacgacg gctgccatta 1920

gtcagcatgg cgcgcacgtg actacaactg tggctggaaa ccttttcgtc ctccccggtt 1980

tttcagtgag ccgactctac tacaatgctt tttcattttt cactcagaaa aacctgcaat 2040

ttgccaaatt ggccatgctc tgtgcctccc ttgacaaagg acatcttccc tgtttataaa 2100

cggcggctta ccaaaagttg aagcttgttc ttgcctctta tgagtggagc aatcgattat 2160

attgaatcgt tgtgctggag tagttggatc tttccacgtg gtctcgagtc acttgtagaa 2220

gctgaaaatt gagcagattt agtatagggc tacattgtag ggtggtttag agtatcgaaa 2280

atatacatat agaagaataa aatgggactc ttagctttgg gcacgccttt gcagtggttt 2340

gagtctagga cgtacaatga acacataagg gatgaaggta tcgagcagtt gttgtatatt 2400

ttccaagctg ctggtaaaag agacaatgac cctctttttt ggggagacga gcttgagtac 2460

atggttgtag attttgatga taaggagaga aattctatgc tcgacgtttg ccatgacaag 2520

atactcactg agcttaatat ggaggattcg tccctttgtg aggctaacga tgtgagtttt 2580

caccctgagt atggccggta tatgttagag gcaacaccag cttctccata tttgaattac 2640

gtgggtagtt acgttgaggt taacatgcaa aaaagacgtg ccattgcaga atataagcta 2700

tctgaatatg cgagacaaga tag 2723

<210> 32
<211> 18
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP288

<400> 32
gattttatcg gtcaaagg 18

<210> 33
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP289

<400> 33
ctatcttgtc tcgcatattc 20

<210> 34
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP290

<400> 34
tctcgagtca cttgtagaag 20

<210> 35
<211> 17
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP292

<400> 35
gagcccacat gcaagtt 17

<210> 36
<211> 20
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP291

<400> 36
gtagggtggt ttagagtatc 20

<210> 37
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP393

<400> 37
tattatgaat tcaaatgttg agcccgaaga cg 32

<210> 38
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP396

<400> 38
tattataagc ttaaattagc gaagcaggtt cc 32

<210> 39
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP397

<400> 39
tattatgcat gccacgtatt cttgtgcaca cg 32

<210> 40
<211> 32
<212> ADN
<213> Secuencia artificial

<220>
<223> Cebador de PCR MBP406

<400> 40
tattatgtcg actaccacct acaccaataa gc 32

<210> 41
<211> 19
<212> PRT
<213> Secuencia artificial

<220>
<223> Péptido señal de invertasa modificada

<220>
<221> MISC
<222> (7)..(7)
<223> Xaa = Phe, Trp o Tyr

<220>
<221> MISC
<222> (8)..(8)
<223> Xaa = Ile, Leu, Val, Ala o Met

<220>
<221> MISC
<222> (9)..(9)
<223> Xaa = Leu, Val, Ala o Met

<220>
<221> MISC
<222> (10)..(10)
<223> Xaa = Ser o Thr

<220>
<221> MISC
<222> (11)..(11)
<223> Xaa = Ile, Val, Ala o Met

<400> 41

Met Leu Leu Gln Ala Phe Xaa Xaa Xaa Xaa Xaa Gly Phe Ala Ala Lys
1 5 10 15

Ile Ser Ala

<210> 42
<211> 19
<212> PRT
<213> Secuencia artificial

<220>
<223> Péptido señal de invertasa modificada

<400> 42

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

Ile Ser Ala

<210> 43
<211> 2118
<212> ADN
<213> Secuencia artificial

<220>
<223> KanMX flanqueado en el extremo 5' por 300 pb inmediatamente cadena arriba del ORF de GSH1 y en el extremo 3' por 300 pb inmediatamente cadena abajo del ORF de GSH1

<400> 43

ggtaccacaa tgctttttca tttttcactc agaaaaacct gcaatttgcc aaattggcca 60

tgctctgtgc ctcccttgac aaaggacatc ttccctgttt ataaacggcg gcttaccaaa 120

agttgaagct tgttcttgcc tcttatgagt ggagcaatcg attatattga atcgttgtgc 180

tggagtagtt ggatctttcc acgtggtctc gagtcacttg tagaagctga aaattgagca 240

gatttagtat agggctacat tgtagggtgg tttagagtat cgaaaatata catatagaag 300

aataaacgta cgctgcaggt cgacggatcc ccgggttaat taaggcgcgc cagatctgtt 360

tagcttgcct cgtccccgcc gggtcacccg gccagcgaca tggaggccca gaataccctc 420

cttgacagtc ttgacgtgcg cagctcaggg gcatgatgtg actgtcgccc gtacatttag 480

cccatacatc cccatgtata atcatttgca tccatacatt ttgatggccg cacggcgcga 540

agcaaaaatt acggctcctc gctgcagacc tgcgagcagg gaaacgctcc cctcacagac 600

gcgttgaatt gtccccacgc cgcgcccctg tagagaaata taaaaggtta ggatttgcca 660

ctgaggttct tctttcatat acttcctttt aaaatcttgc taggatacag ttctcacatc 720

acatccgaac ataaacaacc atgggtaagg aaaagactca cgtttcgagg ccgcgattaa 780

attccaacat ggatgctgat ttatatgggt ataaatgggc tcgcgataat gtcgggcaat 840

caggtgcgac aatctatcga ttgtatggga agcccgatgc gccagagttg tttctgaaac 900

atggcaaagg tagcgttgcc aatgatgtta cagatgagat ggtcagacta aactggctga 960

cggaatttat gcctcttccg accatcaagc attttatccg tactcctgat gatgcatggt 1020

tactcaccac tgcgatcccc ggcaaaacag cattccaggt attagaagaa tatcctgatt 1080

caggtgaaaa tattgttgat gcgctggcag tgttcctgcg ccggttgcat tcgattcctg 1140

tttgtaattg tccttttaac agcgatcgcg tatttcgtct cgctcaggcg caatcacgaa 1200

tgaataacgg tttggttgat gcgagtgatt ttgatgacga gcgtaatggc tggcctgttg 1260

aacaagtctg gaaagaaatg cataagcttt tgccattctc accggattca gtcgtcactc 1320

atggtgattt ctcacttgat aaccttattt ttgacgaggg gaaattaata ggttgtattg 1380

atgttggacg agtcggaatc gcagaccgat accaggatct tgccatccta tggaactgcc 1440

tcggtgagtt ttctccttca ttacagaaac ggctttttca aaaatatggt attgataatc 1500

ctgatatgaa taaattgcag tttcatttga tgctcgatga gtttttctaa tcagtactga 1560

caataaaaag attcttgttt tcaagaactt gtcatttgta tagttttttt atattgtagt 1620

tgttctattt taatcaaatg ttagcgtgat ttatattttt tttcgcctcg acatcatctg 1680

cccagatgcg aagttaagtg cgcagaaagt aatatcatgc gtcaatcgta tgtgaatgct 1740

ggtcgctata ctgctgtcga ttcgatacta acgccgccat ccagtgtcga aaacgagctc 1800

gaattcatcg atactccttt tacttcggtt gtgaaagaaa gttgacatta tcgatttggg 1860

tgacacggtg attgaaaaag caacgaccag tattatacct ctttttttta ttattcagtt 1920

tatatttttg caagtgatct taagcatttc tacacaaact tatgccaacg tgaccattta 1980

ttattttata tagcaaaaaa aaatgagggg ccttgcagaa caattgttgc gagtttctaa 2040

taacaagcac gtagaatatt ggccatttaa tttttctctt caatttatag aatggttgtg 2100

ttagtgacaa aaggtacc 2118

<210> 44
<211> 1755
<212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia de ácido nucleico que codifica una variante de albúmina con mutaciones E492G, K573P, K574H y Q580K en relación con la HSA ts

<400> 44

gacgctcaca agtctgaagt tgctcacaga ttcaaggact taggcgagga aaacttcaaa 60

gctttggttt tgattgcatt cgctcaatac ttacaacaat gcccattcga agaccacgtc 120

aaattggtta acgaagttac tgaattcgca aagacttgcg ttgcagacga atccgcagaa 180

aactgcgaca agtctttgca cacattgttc ggcgacaaat tgtgcacagt tgctacttta 240

agagaaacat acggcgaaat ggctgactgc tgcgctaaac aagagccaga aagaaacgaa 300

tgcttcttac aacacaagga cgacaaccca aacttaccaa gattggttag accagaagtt 360

gacgttatgt gcacagcatt ccacgacaac gaagaaactt tcttgaagaa gtacttgtac 420

gaaattgcta gaagacaccc atacttctac gctccagaat tgttgttctt cgctaaaaga 480

tacaaggctg cattcacaga atgctgccaa gcagctgaca aggctgcttg cttgttacca 540

aagttggacg aattgagaga cgaaggcaag gcttcttctg ctaagcaaag gttgaaatgc 600

gcttctttgc aaaagttcgg cgagagagca ttcaaggcat gggctgttgc tcgtttatct 660

caaagattcc caaaagcaga gttcgctgaa gtttccaagt tagttactga cttgacaaaa 720

gttcacactg aatgctgcca cggcgacttg ttagagtgcg ctgacgaccg tgctgactta 780

gccaaataca tttgcgaaaa ccaagactct atttcttcta agttaaagga gtgctgcgaa 840

aaaccgttgt tagagaaatc tcactgcatt gctgaagttg aaaacgacga aatgccagct 900

gacttgccat ctttagctgc tgacttcgtt gaatctaaag acgtttgcaa gaactacgca 960

gaagctaagg acgttttctt gggcatgttc ttatacgaat acgcaagaag acacccagac 1020

tactctgttg ttttgttgtt aagattggct aagacttacg aaactacatt agaaaagtgc 1080

tgcgctgccg cagacccaca cgaatgctac gctaaagttt tcgacgagtt caagccattg 1140

gttgaagaac cacaaaactt gattaagcaa aactgcgagt tattcgaaca attgggcgaa 1200

tacaaattcc aaaacgcctt gttagttaga tacactaaga aagttccaca agtttcaact 1260

ccaacattgg ttgaagtttc tcgtaactta ggcaaggttg gctctaagtg ctgcaaacac 1320

ccagaggcta agcgtatgcc atgcgctgaa gactacttgt ctgttgtttt gaaccagtta 1380

tgcgttttgc acgagaagac tccagtttct gaccgtgtta ctaagtgctg cacagaatct 1440

ttagttaaca gacgtccatg cttctcagct ttgggtgttg acgaaactta cgttccaaaa 1500

gaattcaacg ctgaaacttt cactttccac gctgacattt gcactttgtc tgaaaaggaa 1560

agacagatta agaaacaaac tgctttggtt gaattggtta agcacaagcc aaaggctact 1620

aaggaacaat tgaaggctgt tatggacgac ttcgctgctt tcgttgaaaa gtgctgcaaa 1680

gctgacgaca aggaaacttg cttcgctgag gaaggcccac atttggttgc agcttcaaaa 1740

gctgctttgg gcttg 1755

<210> 45
<211> 585
<212> PRT

<213> Secuencia artificial

<220>
<223> Variante de albúmina con mutaciones E492G, K573P, K574H y Q580K en relación con la HSA ts

<400> 45

```
Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1               5                  10                  15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
            20                  25                  30

Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
        35                  40                  45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
    50                  55                  60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65                  70                  75                  80

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
                85                  90                  95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
            100                 105                 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
        115                 120                 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
    130                 135                 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145                 150                 155                 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
                165                 170                 175

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
```

```
            180                 185                 190

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
        195                 200                 205

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
    210                 215                 220

Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
225                 230                 235                 240

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
                245                 250                 255

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
            260                 265                 270

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
        275                 280                 285

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
    290                 295                 300

Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
305                 310                 315                 320

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
                325                 330                 335

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
            340                 345                 350

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
        355                 360                 365

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
    370                 375                 380

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
385                 390                 395                 400

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
                405                 410                 415

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
            420                 425                 430
```

```
Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
        435                 440                 445

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
    450                 455                 460

Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
465                 470                 475                 480

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Gly Val Asp Glu Thr
                485                 490                 495

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
            500                 505                 510

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
        515                 520                 525

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
    530                 535                 540

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
545                 550                 555                 560

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Pro His Leu Val
                565                 570                 575

Ala Ala Ser Lys Ala Ala Leu Gly Leu
            580                 585
```

<210> 46
<211> 2253
<212> ADN
<213> Secuencia artificial

<220>
<223> IL-1 Ra fusionado con el extremo carboxiterminal de HSA

<400> 46

```
gacgctcaca agtccgaagt cgctcacaga ttcaaggact tgggtgaaga aaacttcaag       60

gctttggtct tgatcgcttt cgctcaatac ttgcaacaat gtccattcga agatcacgtc      120

aagttggtca acgaagttac cgaattcgct aagacttgtg ttgctgacga atccgcggaa      180

aactgtgaca agtccttgca caccttgttc ggtgataagt tgtgtactgt tgctaccttg      240

agagaaacct acggtgaaat ggctgactgt tgtgctaagc aagaaccaga aagaaacgaa      300

tgtttcttgc aacacaagga cgacaaccca aacttgccaa gattggttag accagaagtt      360

gacgtcatgt gtactgcttt ccacgacaac gaagaaacct tcttgaagaa gtacttgtac      420
```

gaaattgcta gaagacaccc atacttctac gctccagaat tgttgttctt cgctaagaga 480

tacaaggctg ctttcaccga atgttgtcaa gctgctgata aggctgcttg tttgttgcca 540

aagttggatg aattgagaga cgaaggtaag gcttcttccg ctaagcaaag attgaagtgt 600

gcttccttgc aaaagttcgg tgaaagagct ttcaaggctt gggctgtcgc tagattgtct 660

caaagattcc caaaggctga attcgctgaa gtttctaagt tggttactga cttgactaag 720

gttcacactg aatgttgtca cggtgacttg ttggaatgtg ctgatgacag agctgacttg 780

gctaagtaca tctgtgaaaa ccaagactct atctcttcca agttgaagga atgttgtgaa 840

aagccattgt tggaaaagtc tcactgtatt gctgaagttg aaaacgatga aatgccagct 900

gacttgccat ctttggctgc tgacttcgtt gaatctaagg acgtttgtaa gaactacgct 960

gaagctaagg acgtcttctt gggtatgttc ttgtacgaat acgctagaag acacccagac 1020

tactccgttg tcttgttgtt gagattggct aagacctacg aaactaccct cgagaagtgt 1080

tgtgctgctg ctgacccaca cgaatgttac gctaaggttt tcgatgaatt caagccattg 1140

gtcgaagaac cacaaaactt gatcaagcaa aactgtgaat tgttcgaaca attgggtgaa 1200

tacaagttcc aaaacgcttt gttggttaga tacactaaga aggtcccaca agtctccacc 1260

ccaactttgg ttgaagtctc tagaaacttg ggtaaggtcg gttctaagtg ttgtaagcac 1320

ccagaagcta agagaatgcc atgtgctgaa gattacttgt ccgtcgtttt gaaccaattg 1380

tgtgttttgc acgaaaagac cccagtctct gatagagtca ccaagtgttg tactgaatct 1440

ttggttaaca gaagaccatg tttctctgct ttggaagtcg acgaaactta cgttccaaag 1500

gaattcaacg ctgaaacttt caccttccac gctgatatct gtaccttgtc cgaaaaggaa 1560

agacaaatta agaagcaaac tgctttggtt gaattggtca agcacaagcc aaaggctact 1620

aaggaacaat tgaaggctgt catggatgat ttcgctgctt tcgttgaaaa gtgttgtaag 1680

gctgatgata aggaaacttg tttcgctgaa gaaggtaaga agttggtcgc tgcttcccaa 1740

gctgccttag gtttgggtgg ttctggtggt tccggtggtt ctggtggatc cggtggtcga 1800

ccctctggga gaaaatccag caagatgcaa gccttcagaa tctgggatgt taaccagaag 1860

accttctatc tgaggaacaa ccaactagtt gctggatact tgcaaggacc aaatgtcaat 1920

ttagaagaaa agatagatgt ggtacccatt gagcctcatg ctctgttctt gggaatccat 1980

ggagggaaga tgtgcctgtc ctgtgtcaag tctggtgatg agaccagact ccagctggag 2040

gcagttcaaa tcactgacct gagcgagaac agaaagcagg acaagcgctt cgccttcatc 2100

cgctcagaca gcggccccac caccagtttt gagtctgccg cctgccccgg ttggttcctc 2160

tgcacagcga tggaagctga ccagcccgtc agcctcacca atatgcctga cgaaggcgtc 2220

atggtcacca aattctactt ccaggaggac gag 2253

<210> 47
<211> 751
<212> PRT
<213> Secuencia artificial

<220>
<223> IL-1 Ra fusionado con el extremo carboxiterminal de HSA

<400> 47

```
Asp Ala His Lys Ser Glu Val Ala His Arg Phe Lys Asp Leu Gly Glu
1               5                   10                  15

Glu Asn Phe Lys Ala Leu Val Leu Ile Ala Phe Ala Gln Tyr Leu Gln
            20                  25                  30

Gln Cys Pro Phe Glu Asp His Val Lys Leu Val Asn Glu Val Thr Glu
        35                  40                  45

Phe Ala Lys Thr Cys Val Ala Asp Glu Ser Ala Glu Asn Cys Asp Lys
    50                  55                  60

Ser Leu His Thr Leu Phe Gly Asp Lys Leu Cys Thr Val Ala Thr Leu
65                  70                  75                  80

Arg Glu Thr Tyr Gly Glu Met Ala Asp Cys Cys Ala Lys Gln Glu Pro
                85                  90                  95

Glu Arg Asn Glu Cys Phe Leu Gln His Lys Asp Asp Asn Pro Asn Leu
            100                 105                 110

Pro Arg Leu Val Arg Pro Glu Val Asp Val Met Cys Thr Ala Phe His
        115                 120                 125

Asp Asn Glu Glu Thr Phe Leu Lys Lys Tyr Leu Tyr Glu Ile Ala Arg
    130                 135                 140

Arg His Pro Tyr Phe Tyr Ala Pro Glu Leu Leu Phe Phe Ala Lys Arg
145                 150                 155                 160

Tyr Lys Ala Ala Phe Thr Glu Cys Cys Gln Ala Ala Asp Lys Ala Ala
                165                 170                 175

Cys Leu Leu Pro Lys Leu Asp Glu Leu Arg Asp Glu Gly Lys Ala Ser
            180                 185                 190

Ser Ala Lys Gln Arg Leu Lys Cys Ala Ser Leu Gln Lys Phe Gly Glu
        195                 200                 205

Arg Ala Phe Lys Ala Trp Ala Val Ala Arg Leu Ser Gln Arg Phe Pro
```

```
    210                 215                 220

Lys Ala Glu Phe Ala Glu Val Ser Lys Leu Val Thr Asp Leu Thr Lys
225                 230                 235                 240

Val His Thr Glu Cys Cys His Gly Asp Leu Leu Glu Cys Ala Asp Asp
                245                 250                 255

Arg Ala Asp Leu Ala Lys Tyr Ile Cys Glu Asn Gln Asp Ser Ile Ser
            260                 265                 270

Ser Lys Leu Lys Glu Cys Cys Glu Lys Pro Leu Leu Glu Lys Ser His
        275                 280                 285

Cys Ile Ala Glu Val Glu Asn Asp Glu Met Pro Ala Asp Leu Pro Ser
    290                 295                 300

Leu Ala Ala Asp Phe Val Glu Ser Lys Asp Val Cys Lys Asn Tyr Ala
305                 310                 315                 320

Glu Ala Lys Asp Val Phe Leu Gly Met Phe Leu Tyr Glu Tyr Ala Arg
                325                 330                 335

Arg His Pro Asp Tyr Ser Val Val Leu Leu Leu Arg Leu Ala Lys Thr
            340                 345                 350

Tyr Glu Thr Thr Leu Glu Lys Cys Cys Ala Ala Ala Asp Pro His Glu
        355                 360                 365

Cys Tyr Ala Lys Val Phe Asp Glu Phe Lys Pro Leu Val Glu Glu Pro
    370                 375                 380

Gln Asn Leu Ile Lys Gln Asn Cys Glu Leu Phe Glu Gln Leu Gly Glu
385                 390                 395                 400

Tyr Lys Phe Gln Asn Ala Leu Leu Val Arg Tyr Thr Lys Lys Val Pro
                405                 410                 415

Gln Val Ser Thr Pro Thr Leu Val Glu Val Ser Arg Asn Leu Gly Lys
            420                 425                 430

Val Gly Ser Lys Cys Cys Lys His Pro Glu Ala Lys Arg Met Pro Cys
        435                 440                 445

Ala Glu Asp Tyr Leu Ser Val Val Leu Asn Gln Leu Cys Val Leu His
    450                 455                 460
```

```
Glu Lys Thr Pro Val Ser Asp Arg Val Thr Lys Cys Cys Thr Glu Ser
465                 470                 475                 480

Leu Val Asn Arg Arg Pro Cys Phe Ser Ala Leu Glu Val Asp Glu Thr
                485                 490                 495

Tyr Val Pro Lys Glu Phe Asn Ala Glu Thr Phe Thr Phe His Ala Asp
            500                 505                 510

Ile Cys Thr Leu Ser Glu Lys Glu Arg Gln Ile Lys Lys Gln Thr Ala
        515                 520                 525

Leu Val Glu Leu Val Lys His Lys Pro Lys Ala Thr Lys Glu Gln Leu
    530                 535                 540

Lys Ala Val Met Asp Asp Phe Ala Ala Phe Val Glu Lys Cys Cys Lys
545                 550                 555                 560

Ala Asp Asp Lys Glu Thr Cys Phe Ala Glu Glu Gly Lys Lys Leu Val
                565                 570                 575

Ala Ala Ser Gln Ala Ala Leu Gly Leu Gly Gly Ser Gly Gly Ser Gly
            580                 585                 590

Gly Ser Gly Gly Ser Gly Gly Arg Pro Ser Gly Arg Lys Ser Ser Lys
        595                 600                 605

Met Gln Ala Phe Arg Ile Trp Asp Val Asn Gln Lys Thr Phe Tyr Leu
    610                 615                 620

Arg Asn Asn Gln Leu Val Ala Gly Tyr Leu Gln Gly Pro Asn Val Asn
625                 630                 635                 640

Leu Glu Glu Lys Ile Asp Val Val Pro Ile Glu Pro His Ala Leu Phe
                645                 650                 655

Leu Gly Ile His Gly Gly Lys Met Cys Leu Ser Cys Val Lys Ser Gly
            660                 665                 670

Asp Glu Thr Arg Leu Gln Leu Glu Ala Val Gln Ile Thr Asp Leu Ser
        675                 680                 685

Glu Asn Arg Lys Gln Asp Lys Arg Phe Ala Phe Ile Arg Ser Asp Ser
    690                 695                 700

Gly Pro Thr Thr Ser Phe Glu Ser Ala Ala Cys Pro Gly Trp Phe Leu
705                 710                 715                 720
```

Cys Thr Ala Met Glu Ala Asp Gln Pro Val Ser Leu Thr Asn Met Pro
725 730 735

Asp Glu Gly Val Met Val Thr Lys Phe Tyr Phe Gln Glu Asp Glu
740 745 750

<210> 48
<211> 756
<212> ADN
<213> Secuencia artificial

<220>
<223> FLAG fusionada con el extremo carboxiterminal de scFv

<400> 48

gaagttcaat tgttggaatc tggtggtggt ttggttcaac ctggtggttc tttgagattg 60

tcttgtgctg cttctggttt actttttct aattattgga tgtcttgggt tagacaagct 120

ccaggtaaag gtttggaatg ggtttccggt atttcaggta atggtggtta tacttatttt 180

gctgattcag ttaaagatag atttactatt tctagagata attctaaaaa taccttatat 240

ttgcaaatga actctttgag agcagaagat actgctgttt attactgtgc aggtggtgac 300

ggttctggtt ggagtttttg gggtcaaggt actctagtta ccgtttcttc aggtggtggt 360

ggttctggtg gaggtggatc aggtggtgga ggatctcaat cagttttgac tcaaccacca 420

tctgcttcag gtactccagg tcaaagagtt accatttctt gtactggttc ttcttctaat 480

attggtgcag gttacgatgt tcattggtat caacaattgc caggtactgc tccaaaattg 540

ttgatttatg gtaacaacaa tagaccatct ggtgtcccag atagattttc tggttctaaa 600

tctggtactt ctgcttcttt ggctatttct ggtttaagat cagaagatga agctgattac 660

tactgtgctg cttgggatga ctctttgtct ggtagagttt tcggtggtgg tactaaattg 720

accgttttgg gtgattataa agatgatgac gataaa 756

<210> 49
<211> 252
<212> PRT
<213> Secuencia artificial

<220>
<223> FLAG fusionada con el extremo carboxiterminal de scFv

<400> 49

Glu Val Gln Leu Leu Glu Ser Gly Gly Gly Leu Val Gln Pro Gly Gly
1 5 10 15

Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser Asn Tyr
20 25 30

```
Trp Met Ser Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
        35                  40                  45

Ser Gly Ile Ser Gly Asn Gly Gly Tyr Thr Tyr Phe Ala Asp Ser Val
    50                  55                  60

Lys Asp Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Asn Thr Leu Tyr
65                  70                  75                  80

Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr Ala Val Tyr Tyr Cys
                85                  90                  95

Ala Gly Gly Asp Gly Ser Gly Trp Ser Phe Trp Gly Gln Gly Thr Leu
            100                 105                 110

Val Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly
        115                 120                 125

Gly Gly Gly Ser Gln Ser Val Leu Thr Gln Pro Pro Ser Ala Ser Gly
    130                 135                 140

Thr Pro Gly Gln Arg Val Thr Ile Ser Cys Thr Gly Ser Ser Ser Asn
145                 150                 155                 160

Ile Gly Ala Gly Tyr Asp Val His Trp Tyr Gln Gln Leu Pro Gly Thr
                165                 170                 175

Ala Pro Lys Leu Leu Ile Tyr Gly Asn Asn Asn Arg Pro Ser Gly Val
            180                 185                 190

Pro Asp Arg Phe Ser Gly Ser Lys Ser Gly Thr Ser Ala Ser Leu Ala
        195                 200                 205

Ile Ser Gly Leu Arg Ser Glu Asp Glu Ala Asp Tyr Tyr Cys Ala Ala
    210                 215                 220

Trp Asp Asp Ser Leu Ser Gly Arg Val Phe Gly Gly Gly Thr Lys Leu
225                 230                 235                 240

Thr Val Leu Gly Asp Tyr Lys Asp Asp Asp Asp Lys
                245                 250
```

<210> 50
<211> 8
<212> PRT
<213> Secuencia artificial

<220>
<223> FLAG

<400> 50

Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> 51
<211> 30
<212> PRT
<213> Secuencia artificial

<220>
<223> fragmento de GLP-1

<400> 51

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg
20 25 30

<210> 52
<211> 60
<212> PRT
<213> Secuencia artificial

<220>
<223> Repetición en tándem del fragmento GLP-1

<400> 52

His Gly Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly
1 5 10 15

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg His Gly
20 25 30

Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly Gln Ala
35 40 45

Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg
50 55 60

REIVINDICACIONES

1. Un método para producir una proteína heteróloga que comprende:

a. proporcionar una célula hospedadora fúngica que tiene:

1. una proteína Gsh1 modificada que tiene un nivel reducido de actividad en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
2. un nivel reducido de actividad de la proteína Gsh1 en relación con la actividad de una proteína Gsh1 de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
3. un gen *GSH1* modificado que tiene un nivel reducido de actividad en relación con la actividad de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y/o
4. un nivel reducido de expresión del gen *GSH1* en relación con el nivel de expresión de un gen *GSH1* de tipo silvestre de una célula hospedadora fúngica de referencia por lo demás idéntica, y

b. cultivar la célula hospedadora para producir la proteína heteróloga,
en donde

(i) la proteína heteróloga es una proteína vegetal o animal, y/o
(ii) el cultivo es a una escala de al menos 1 litro, y/o
(iii) el método comprende además purificar la proteína heteróloga, y/o
(iv) el método comprende además formular la proteína heteróloga con un vehículo o diluyente terapéuticamente aceptable para producir así un producto terapéutico adecuado para la administración a un ser humano o un animal.

2. El método de acuerdo con la reivindicación 1 para aumentar el rendimiento de la proteína heteróloga.

3. El método de acuerdo con la reivindicación 1 o 2 en el que la célula hospedadora fúngica es una levadura o un hongo filamentoso.

4. El método de acuerdo con cualquiera de las reivindicaciones 1 a 3 en el que la proteína Gsh1 comprende una mutación en la posición correspondiente a una posición seleccionada de 47, 48, 49, 50, 51, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 409, 451, 452, 453, 454 y 455 de la SEQ ID NO: 2.

5. El método de acuerdo con la reivindicación 4 en el que la mutación en una posición correspondiente a la posición 125 de la SEQ ID NO: 2 es una sustitución por C, D, E o G.

6. El método de acuerdo con cualquiera de las reivindicaciones 1 a 5 en el que la proteína Gsh1 comprende o consiste en la SEQ ID NO: 4.

7. El método de acuerdo con cualquiera de las reivindicaciones 1 a 3 en el que la célula hospedadora carece de un gen *GSH1* o una proteína Gsh1.

8. El método de acuerdo con cualquiera de las reivindicaciones 1 a 7 en el que el rendimiento de la proteína heteróloga es al menos un 2 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia.

9. El método de acuerdo con cualquiera de las reivindicaciones 1 a 8 en el que la célula hospedadora fúngica de referencia tiene la proteína Gsh1 de SEQ ID NO: 2.

10. El método de acuerdo con cualquiera de las reivindicaciones 1 a 9, en el que la célula hospedadora fúngica tiene una modificación en:

1. la proteína Not4, y/o
2. el nivel de actividad de la proteína Not4, y/o
3. el gen *NOT4* y/o
4. el nivel de expresión del gen *NOT4.*

11. El método de acuerdo con la reivindicación 10 en el que se reduce el nivel de actividad de la proteína Not4.

12. El método de la reivindicación 10 u 11 en el que se reduce el nivel de expresión del gen *NOT4*.

13. El método de acuerdo con cualquiera de las reivindicaciones 10 a 12 en el que la proteína Not4 comprende una mutación en la posición correspondiente a una posición seleccionada de 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469 o 470 de la SEQ ID NO: 6.

14. El método de acuerdo con la reivindicación 13 en el que la mutación en una posición correspondiente a la posición 429 de la SEQ ID NO: 6 es una sustitución por G, A, V, L o I.

15. El método de acuerdo con cualquiera de las reivindicaciones 10 a 13 en el que la proteína Not4 comprende o consiste en la SEQ ID NO: 8.

16. El método de acuerdo con cualquiera de las reivindicaciones 10 a 12 en el que la célula hospedadora carece de un gen *NOT4* o una proteína Not4.

17. El método de acuerdo con cualquiera de las reivindicaciones 10 a 16 en el que el rendimiento de la proteína heteróloga es al menos un 2 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia que tiene una proteína Not4 de tipo silvestre.

18. El método de acuerdo con la reivindicación 17 en donde la célula hospedadora fúngica de referencia tiene una proteína Not4 de SEQ ID NO: 6.

19. El método de acuerdo con cualquiera de las reivindicaciones 10 a 18 en el que el rendimiento de la proteína heteróloga es al menos un 2 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia que tiene la proteína Gsh1 de SEQ ID NO: 2.

20. El método de acuerdo con cualquiera de las reivindicaciones 10 a 19 en el que el rendimiento de la proteína heteróloga es al menos un 2 % superior al rendimiento a partir de una célula hospedadora fúngica de referencia que tiene la proteína Gsh1 de tipo silvestre y la proteína Not4 de tipo silvestre.

21. El método de acuerdo con la reivindicación 20 en el que la célula hospedadora fúngica de referencia tiene la proteína Gsh1 de SEQ ID NO: 2 y la proteína Not4 de SEQ ID NO: 6.

22. El método de acuerdo con cualquiera de las reivindicaciones 1 a 21 en el que la proteína heteróloga comprende o consiste en una albúmina o variante, fragmento y/o fusión de la misma.

23. El método de acuerdo con cualquiera de las reivindicaciones 1 a 22 en el que la célula hospedadora se cultiva a una escala de al menos 5 l.

24. El método de acuerdo con cualquiera de las reivindicaciones 1 a 23 que comprende además recuperar la proteína heteróloga.

25. El método de acuerdo con cualquiera de las reivindicaciones 1 a 24 que comprende además purificar la proteína heteróloga.

26. El método de acuerdo con cualquiera de las reivindicaciones 1 a 25 que comprende además formular la proteína heteróloga con un vehículo o diluyente terapéuticamente aceptable para producir así un producto terapéutico adecuado para la administración a un ser humano o un animal.

27. El método de acuerdo con cualquiera de las reivindicaciones 1 a 26 que comprende además proporcionar la proteína heteróloga en forma de dosificación unitaria.

FIG. 1

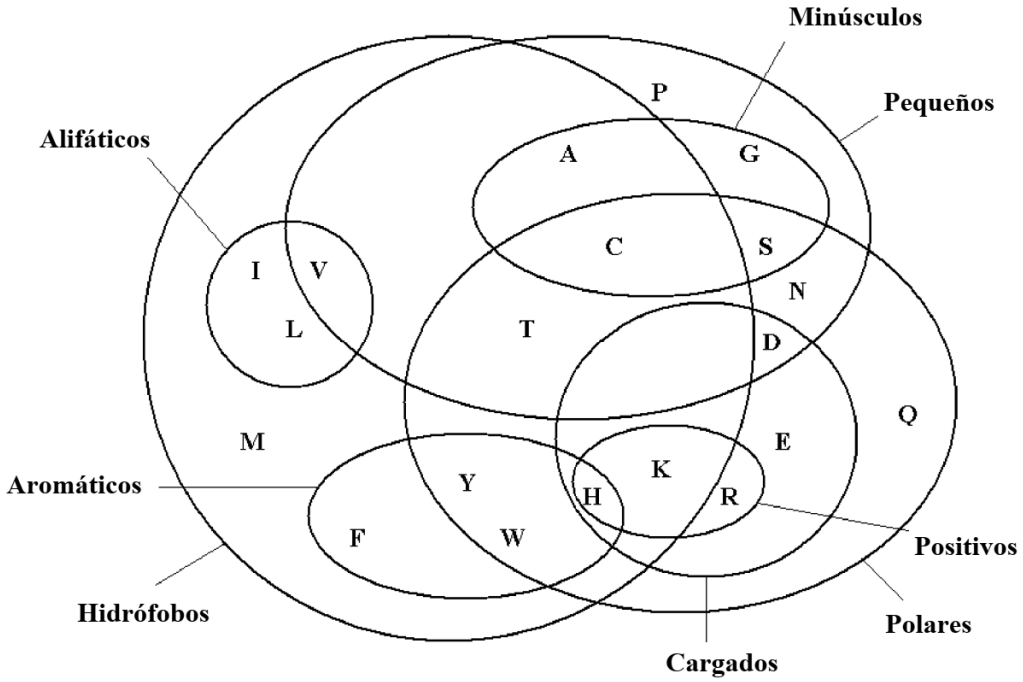

Minúsculos
Pequeños
Alifáticos
P
A
G
C
S
I
V
N
L
T
D
Q
M
E
K
Aromáticos
Y
H
R
F
W
Positivos
Hidrófobos
Polares
Cargados

FIG. 2.

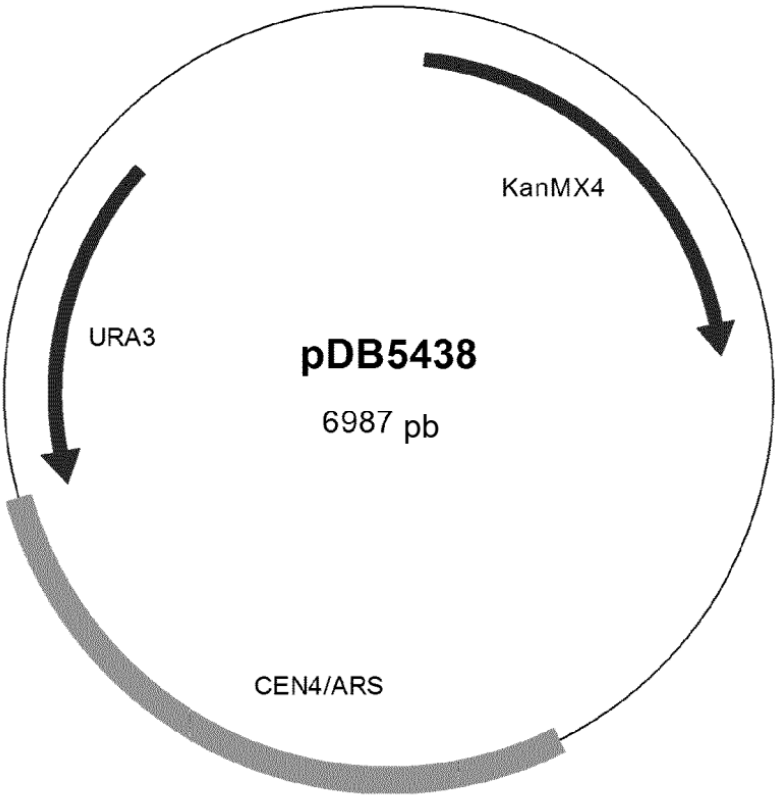


FIG. 3.

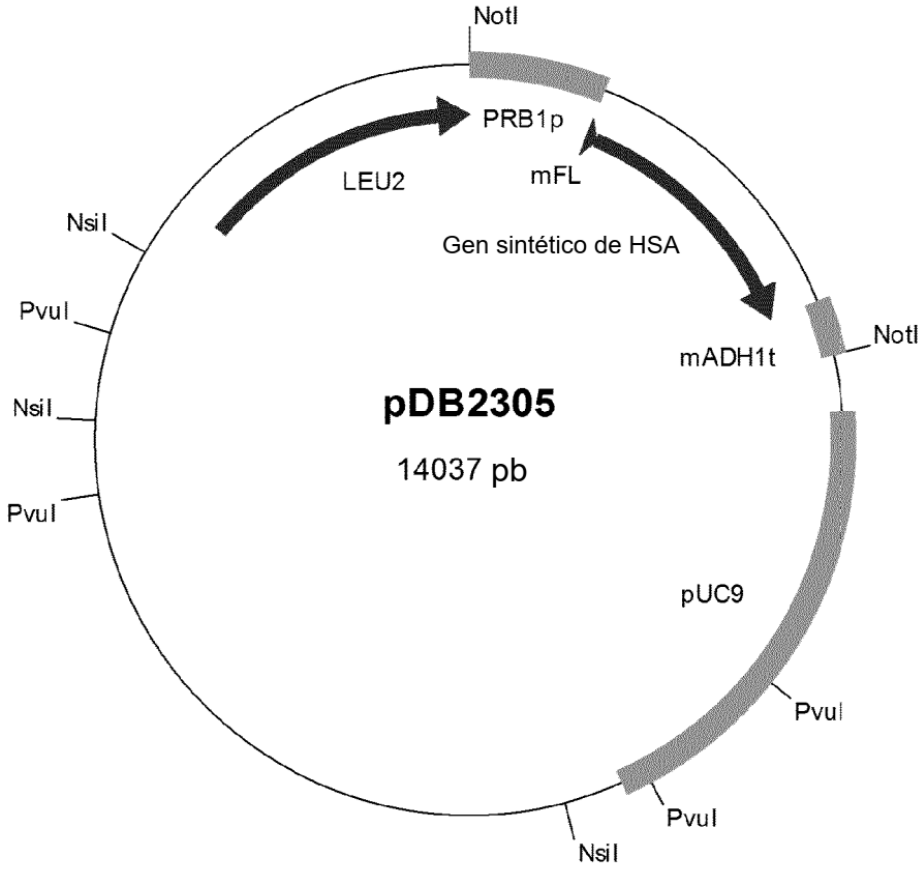

NotI
PRB1p
LEU2
mFL
Gen sintético de HSA
NsiI
PvuI
NsiI
PvuI
mADH1t
NotI
pDB2305
14037 pb
pUC9
PvuI
PvuI
NsiI

FIG. 4

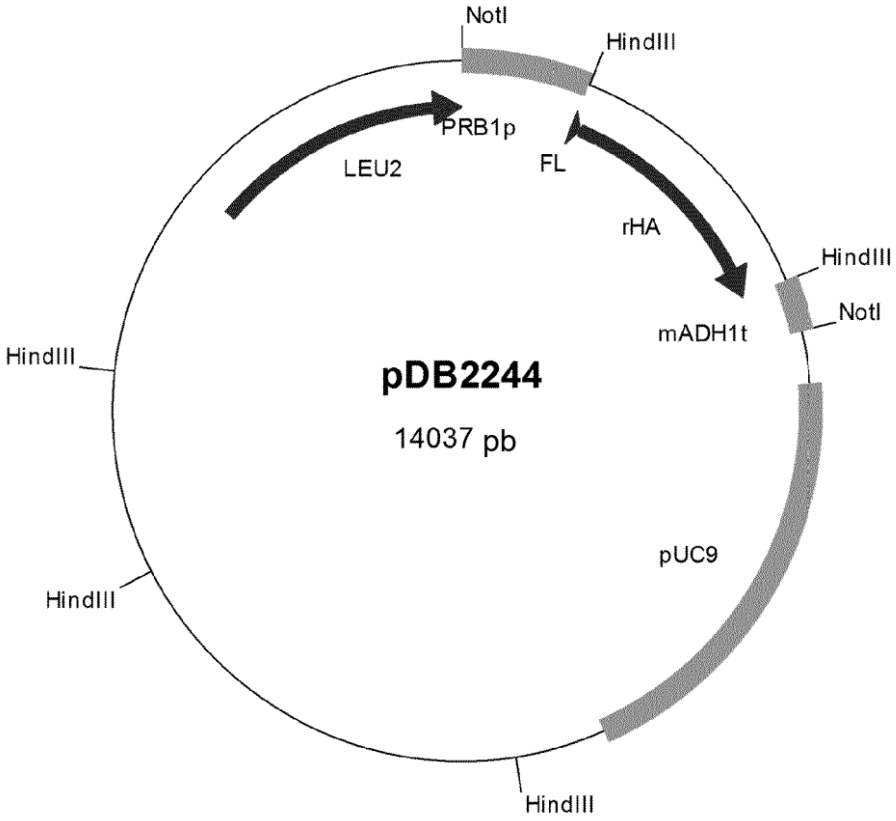

NotI
HindIII
PRB1p
LEU2
FL
rHA
HindIII
NotI
mADH1t
HindIII
pDB2244
14037 pb
pUC9
HindIII
HindIII

FIG. 5

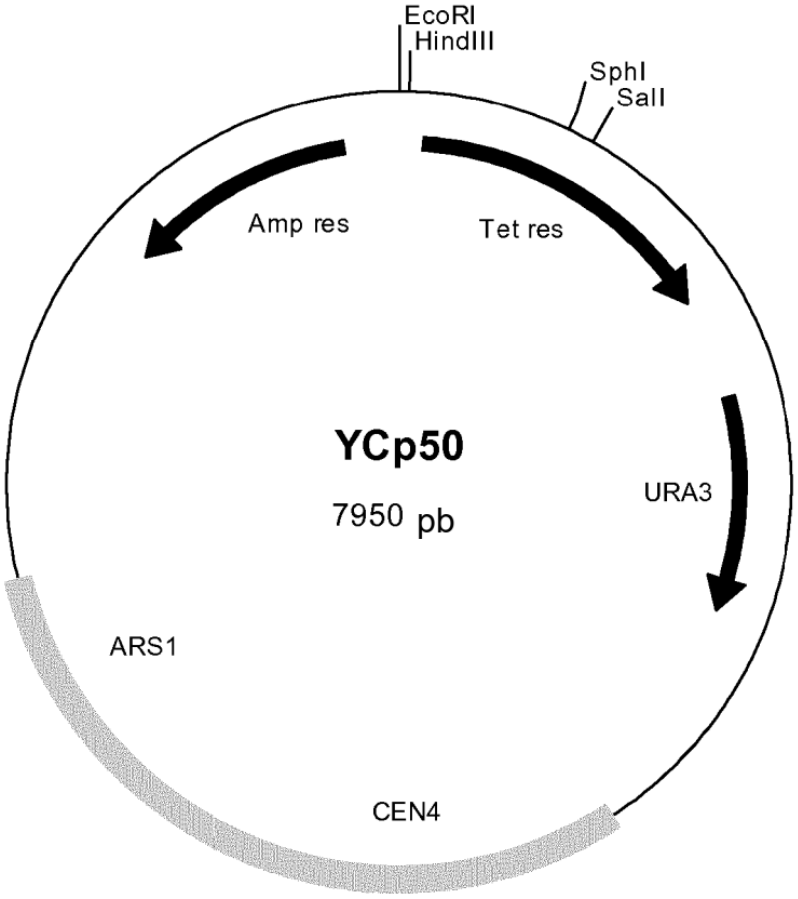

EcoRI
HindIII
SphI
SalI
Amp res
Tet res
YCp50
7950 pb
URA3
ARS1
CEN4

FIG. 6

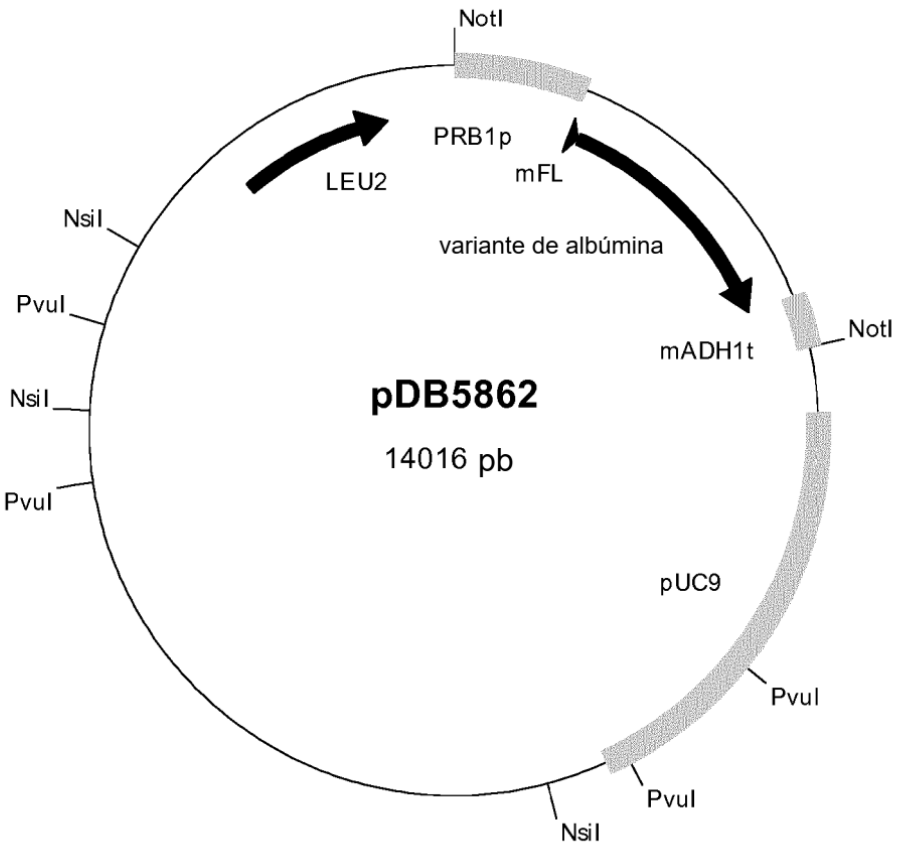

NotI
PRB1p
mFL
LEU2
variante de albúmina
NsiI
PvuI
NsiI
PvuI
mADH1t
NotI
pDB5862
14016 pb
pUC9
PvuI
PvuI
NsiI

FIG. 7

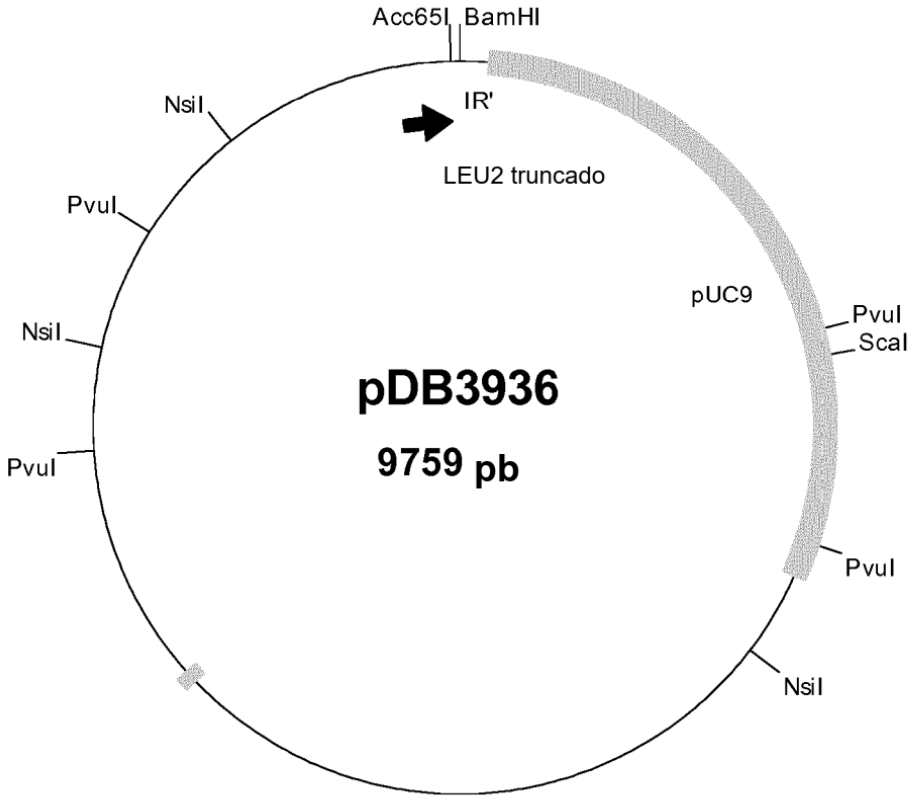

Acc65I
BamHI
IR'
LEU2 truncado
pUC9
PvuI
ScaI
PvuI
NsiI
PvuI
NsiI
PvuI
NsiI
pDB3936
9759 pb

FIG. 8

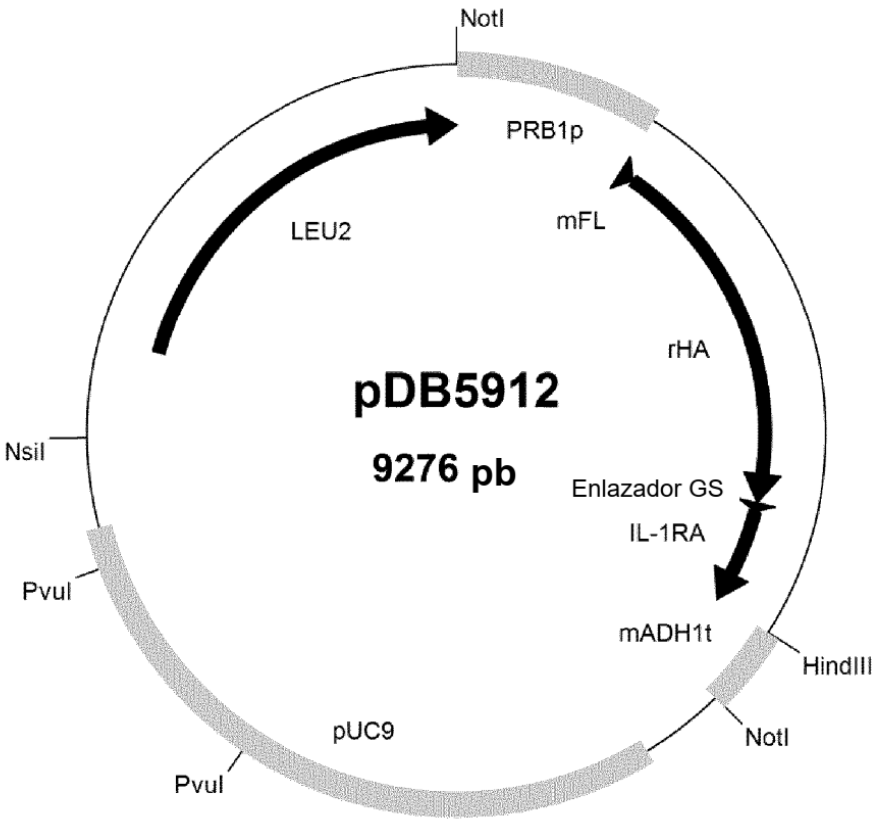
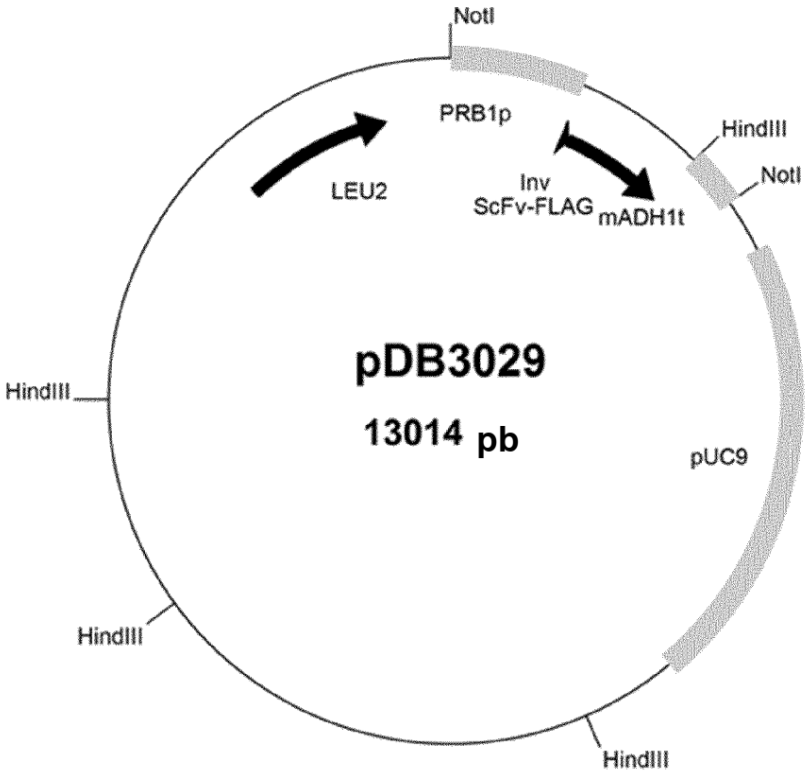

NotI
PRB1p
mFL
rHA
Enlazador GS
IL-1RA
mADH1t
HindIII
NotI
pUC9
PvuI
PvuI
NsiI
LEU2
pDB5912
9276 pb

FIG. 9


NotI
PRB1p
HindIII
NotI
LEU2
Inv
ScFv-FLAG
mADH1t
pDB3029
13014 pb
pUC9
HindIII
HindIII
HindIII