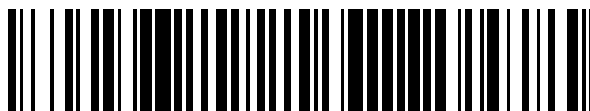


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 700 547**

51 Int. Cl.:

**C07K 14/435** (2006.01)

**C12N 15/00** (2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **02.03.2015 PCT/EP2015/054298**

87 Fecha y número de publicación internacional: **03.09.2015 WO15128507**

96 Fecha de presentación y número de la solicitud europea: **02.03.2015 E 15708175 (3)**

97 Fecha y número de publicación de la concesión europea: **07.11.2018 EP 3110834**

54 Título: **Variantes del pro-péptido del factor de apareamiento alfa**

30 Prioridad:

**28.02.2014 EP 14157172**

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

**18.02.2019**

73 Titular/es:

**NOVO NORDISK A/S (100.0%)**

**Novo Allé**

**2880 Bagsværd, DK**

72 Inventor/es:

**NØRGAARD, PER**

74 Agente/Representante:

**PONS ARIÑO, Ángel**

ES 2 700 547 T3

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

## DESCRIPCIÓN

Variantes del pro-péptido del factor de apareamiento alfa

Campo técnico

- 5 La presente invención se refiere a los campos técnicos de la expresión de proteínas y la química de proteínas donde un polipéptido se prepara mediante la expresión recombinante en levaduras.

Antecedentes de la invención

- 10 Las técnicas de la expresión de polipéptidos recombinantes permite la producción de grandes cantidades de polipéptidos convenientes que pueden usarse por ejemplo por su actividad biológica. Dichos polipéptidos se expresan frecuentemente como polipéptidos de fusión recombinantes en células huésped microbianas. El polipéptido de interés se une frecuentemente a un polipéptido pareja de fusión para aumentar el nivel de expresión, facilitar la secreción, aumentar la solubilidad, promover el plegamiento del polipéptido, para proteger el polipéptido contra la proteólisis no intencionada o para facilitar la purificación del polipéptido de interés.

- 15 Para asegurar la secreción de los polipéptidos expresados de manera recombinante a partir de levaduras, normalmente se fusiona un pre-pro péptido (frecuentemente denominado "líder") al extremo N-terminal del producto recombinante. La pre-secuencia asegura la translocación de la proteína de fusión al retículo endoplásmico (ER), que es el punto de partida de la vía secretora. La pro-secuencia asegura, además, el transporte desde el ER al aparato de Golgi, donde una proteasa endógena denominada Kex2p, frecuentemente se utiliza para escindir la pro-secuencia. El péptido recombinante procesado se secreta posteriormente al medio de cultivo, a partir del cual puede purificarse.

- 20 La pre-pro secuencia del Factor de apareamiento alfa se usa frecuentemente como líder para la secreción de polipéptidos recombinantes. Sin embargo, muchas otras secuencias pueden facilitar el proceso de secreción. La secuencia líder tiene una influencia importante, no solamente en las cantidades de péptido secretado, sino además en la calidad en términos de degradación y modificaciones postraduccionales como la O-glicosilación. Dado que la degradación y la O-glicosilación normalmente son eventos no deseados, es conveniente una secuencia líder que reduzca estas modificaciones al menor grado posible, y al mismo tiempo que maximice el rendimiento del polipéptido secretado.

- 30 El documento núm. EP 0121884 A2 describe la producción recombinante de insulina humana en *S. cerevisiae* mediante el uso del factor alfa de levadura.

El documento núm. EP 0324274A1 describe el uso de una secuencia líder truncada del factor alfa para mejorar la expresión y secreción de proteínas heterólogas en levadura.

- 35 Thim y otros (PNAS 83 (1986) 6766-6770) describen el uso del factor de apareamiento alfa para la secreción y el procesamiento de precursores de insulina en *Saccharomyces cerevisiae*.

El documento núm. WO95/34666 describe líderes sintéticos para producir polipéptidos secretados en *S. cerevisiae*.

- 40 Rakestraw y otros (Biotech. Bioeng. 103 (2009) 1192-1201) describen secuencias líder del factor de apareamiento alfa mutante que aumentan la secreción de anticuerpo de cadena sencilla y que aumentan los niveles de producción de IgG1 humana en *S. cerevisiae*.

- 45 Existe la necesidad de secuencias líder más específicas que aumenten el rendimiento del precursor del polipéptido, y que reduzcan la proporción del polipéptido recombinante que esté O-glicosilado. En particular existe la necesidad de secuencias líder para expresar péptidos de GLP-1 en levadura, donde dichos líderes aumenten el rendimiento de los péptidos de GLP-1 o los precursores de estos y disminuyan la O-glicosilación de los péptidos de GLP-1. Dichas secuencias líder más específicas pueden facilitar un mayor rendimiento del polipéptido recombinante así como disminuir la cantidad de impurezas O-glicosiladas.

Breve descripción de la invención

- 50 Un objeto de la presente invención es proporcionar células de levadura que tienen un nivel aumentado de la expresión de polipéptidos heterólogos. Otro objeto de la presente invención es proporcionar células de levadura que secretan el polipéptido recombinante con una cantidad reducida de variantes O-glicosiladas del polipéptido recombinante. En particular un objeto de la presente invención es proporcionar células de levadura que tienen tanto un nivel aumentado de expresión del polipéptido recombinante como una cantidad reducida de variantes O-glicosiladas. Un objeto de la presente invención es proporcionar un sistema de expresión mejorado para la expresión recombinante de péptidos de GLP-1 en células de levadura.

- 55 De acuerdo con un primer aspecto de la invención se proporciona un método para la expresión recombinante de un polipéptido que comprende un péptido de GLP-1 en levadura que comprende el cultivo de una cepa de levadura que comprende una secuencia de ADN que codifica una señal de procesamiento y secreción hacia el extremo corriente arriba del polipéptido, en donde dicha señal de procesamiento y secreción comprende una variante del pro-péptido

del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):



5 en donde

$X_{38}$  es F, L o V,

$X_{39}$  L, I, V o M,

$X_{40}$  es G o R,

$X_{41}$  es S, Y, L, I, V o M, y

10  $X_{42}$  es Y, W, L, V o M;

o en donde

$X_{38}$  es I o V,

$X_{39}$  L, I, V o M,

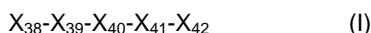
15  $X_{40}$  es A, G, Y, F, W, R, K, L, I, V o M,

$X_{41}$  es Y, F, o W, y

$X_{42}$  es L o I;

20 en donde dicho pro-péptido del factor de apareamiento  $\alpha$  son los residuos de aminoácidos 20-85 en el polipéptido: MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYLDLEGDFDVAVLFPFSNSTNNGLL FINTTASIAAKEEGVSLDKR (sec. con núm. de ident.:2) y dicha variante comprende de 1 a 15 sustituciones, deleciones o adiciones de aminoácidos respecto de dicho polipéptido.

En una modalidad del método para la expresión recombinante de un polipéptido en levadura la variante del pro-péptido del factor de apareamiento  $\alpha$  tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):

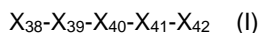


en donde  $X_{38}$  es V;  $X_{39}$  L, I, V o M;  $X_{40}$  es G o R;  $X_{41}$  es Y, y  $X_{42}$  es L.

30 En otra modalidad el polipéptido para la expresión recombinante es un péptido de GLP-1.

En otra modalidad el polipéptido para la expresión recombinante comprende GLP-1(7-37)[K34R], GLP-1(9-37)[K34R] o GLP-1(9-37)[K34R,G37K].

De acuerdo con un segundo aspecto de la invención se proporciona una variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):



en donde

$X_{38}$  es F, L, I o V;

$X_{39}$  L, I, V o M;

40  $X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;

$X_{41}$  es S, Y, F, W, L, I, V o M;

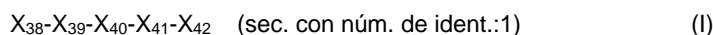
$X_{42}$  es Y, W, L, I, V, M o S;

con la condición de que  $X_{38}-X_{42}$  no sea VIGYS.

45 De acuerdo con un tercer aspecto de la invención se proporciona un precursor de GLP-1 que es un polipéptido de fusión que comprende:

Un pre-péptido,

50 Una variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):



en donde

$X_{38}$  es F, L, I o V;

55  $X_{39}$  L, I, V o M;

$X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;

$X_{41}$  es S, Y, F, W, L, I, V o M;

$X_{42}$  es Y, W, L, I, V, M o S;

con la condición de que  $X_{38}-X_{42}$  no sea VIGYS,

60

Opcionalmente un péptido de extensión, y

Un polipéptido de GLP-1.

5 De acuerdo con un cuarto aspecto de la invención se proporciona una secuencia de ADN que codifica la variante del pro-péptido del factor de apareamiento  $\alpha$  o el precursor de GLP-1.

De acuerdo con un quinto aspecto de la invención se proporciona un vector de expresión que comprende la secuencia de ADN que codifica la variante del pro-péptido del factor de apareamiento  $\alpha$  o el precursor de GLP-1.

10 De acuerdo con un sexto aspecto de la invención se proporciona una célula huésped que comprende el vector de expresión de acuerdo con la invención.

15 En una modalidad la célula huésped utilizada para la expresión tiene un gen de *pmt1* no funcional o no tiene un gen de *pmt1*. Sorprendentemente se ha encontrado que una célula huésped de este tipo disminuye la cantidad de polipéptido O-glicosilado independientemente de, y además de, la disminución obtenida a partir de la variante del pro-péptido del factor de apareamiento  $\alpha$  de la presente invención.

Breve descripción de los dibujos

20 La figura 1 muestra el plásmido de expresión mínimo como se usa en el ejemplo 1.  
La figura 2 muestra el precursor de GLP-1(7-37)[K34R] extendido en el extremo N-terminal que incluye el pre-pro-péptido del factor de apareamiento  $\alpha$  de tipo silvestre con indicación de la subsecuencia VIGYL del pro-péptido.

Descripción

25 De acuerdo con un primer aspecto de la invención se proporciona un método para la expresión recombinante de un polipéptido en levadura que comprende el cultivo de una cepa de levadura que comprende una secuencia de ADN que codifica una señal de procesamiento y secreción hacia el extremo corriente arriba del polipéptido, en donde dicha señal de procesamiento y secreción comprende una variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):

30 
$$X_{38}-X_{39}-X_{40}-X_{41}-X_{42} \quad (\text{sec. con núm. de ident.:1}) \quad (I)$$

en donde

$X_{38}$  es F, L, I o V;

$X_{39}$  L, I, V o M;

35  $X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;

$X_{41}$  es S, Y, F, W, L, I, V o M;

$X_{42}$  es Y, W, L, I, V, M o S;

con la condición de que  $X_{38}-X_{42}$  no sea VIGYS.

40 El término "secuencia líder" como se usa en la presente está destinado a referirse a una secuencia de aminoácidos que consiste en un pre-péptido (el péptido señal) y un pro-péptido. Ejemplos no limitantes de secuencias líder son por ejemplo la señal líder del factor alfa de *S. cerevisiae* y las secuencias líder sintéticas para levadura descritas en el documento WO95/34666.

45 "Pre-péptido" como se usa en la presente está destinado a referirse a un péptido señal que está presente como una secuencia *N-terminal* en la forma precursora de un polipéptido que se va a expresar. La función del péptido señal es facilitar la translocación del polipéptido hacia el retículo endoplásmico en la célula huésped. El péptido señal se escinde normalmente en el transcurso de este proceso. El péptido señal puede ser heterólogo u homólogo con respecto de la célula huésped que produce el polipéptido.

50 "Pro-péptido" como se usa en la presente está destinado a referirse a una secuencia peptídica cuya función es permitir que el polipéptido expresado se dirija del retículo endoplásmico al aparato de Golgi y después a una vesícula secretora para la secreción al medio de cultivo (es decir la exportación del polipéptido a través de la pared celular o al menos a través de la membrana celular hacia el espacio periplasmático de la célula de levadura).  
55 Ejemplos no limitantes de un pro-péptido son el pro-péptido del  $\alpha$ -factor de levadura (véanse los documentos US 4,546,082 y 4,870,008) y los pro-péptidos sintéticos descritos en los documentos US 5,395,922; 5,795,746; 5,162,498 y WO 98/32867. El pro-péptido contendrá preferentemente un sitio de procesamiento de la endopeptidasa en el extremo C-terminal, tal como una secuencia Lys-Arg o cualquier análogo funcional de esta.

60 El término "Factor de apareamiento  $\alpha$ " (MF $\alpha$ , MF $\alpha$  o MFalfa) como se usa en la presente está destinado a referirse a la prepro-secuencia de *Saccharomyces cerevisiae*, que comprende el pre-péptido del factor de apareamiento  $\alpha$

como los residuos de aminoácidos 1-19 y el pro-péptido del factor de apareamiento  $\alpha$  como los residuos de aminoácidos 20-85 en la estructura: MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYLDLEGDFDVAVL PFSNSTNNGLLFINTTIIASIAAKEEGVSLDKR (sec. con núm. de ident.:2).

El factor de apareamiento  $\alpha$  comprende la secuencia VIGYL (sec. con núm. de ident.:3) como los residuos de aminoácidos 38-42. Una variante del factor de apareamiento  $\alpha$  se ha usado para la expresión recombinante de polipéptidos que comprende la secuencia VIGYS (sec. con núm. de ident.:4) como los residuos de aminoácidos 38-42 en la secuencia del factor de apareamiento  $\alpha$ .

En el presente contexto los términos “polipéptido”, “proteína” y “péptido” pueden usarse indistintamente para designar un polipéptido. Debe entenderse que el término en particular utilizado no tiene limitación con relación al tamaño de la molécula (a menos que se indique directamente en el contexto particular).

Los residuos de aminoácidos se designan generalmente de acuerdo con una abreviatura de una única letra de acuerdo con la nomenclatura de la IUPAC, por ejemplo D significa ácido aspártico (Asp) y G significa glicina. Sin embargo, en algunos casos también se usa la correspondiente abreviatura de tres letras.

“Aminoácidos codificados genéticamente” como se usa en la presente está destinado a referirse al grupo que consiste en los siguientes aminoácidos: G, P, A, V, L, I, M, C, F, Y, W, H, K, R, Q, N, E, D, S, T así como cualquier modificación biológica de estos. Ejemplos no limitantes de tales modificaciones biológicas son por ejemplo amidación, glicosilación y formación de enlaces disulfuro.

“Análogos” como se usa en la presente está destinado a referirse a polipéptidos que se derivan del polipéptido de referencia mediante sustitución, delección y/o adición de uno o más residuos de aminoácidos a partir del polipéptido. Ejemplos no limitantes de un análogo de GLP-1(7-37) (sec. con núm. de ident.:5) son GLP-1(7-37)[K34R] (sec. con núm. de ident.: 6) donde el residuo 34 se ha sustituido por un residuo de arginina y GLP-1(9-37)[K34R] (sec. con núm. de ident.: 7) donde el residuo 34 se ha sustituido con un residuo de arginina y los residuos de aminoácidos 7-8 se han eliminado (mediante el uso de la numeración común de residuos de aminoácidos para los péptidos de GLP-1).

“Variante” como se usa en la presente con referencia a un polipéptido está destinado a referirse a una variante química del polipéptido que conserva sustancialmente la misma función principal que la proteína original. Por tanto una variante típicamente es una versión modificada de un polipéptido en donde se introducen pocas modificaciones según sea necesario para que el polipéptido modificado tenga alguna propiedad conveniente a la vez que conserva sustancialmente la misma función principal del polipéptido original. Ejemplos no limitantes de variantes polipeptídicas son por ejemplo polipéptidos extendidos, polipéptidos truncados, polipéptidos de fusión y análogos. Un ejemplo no limitante de una variante del pro-péptido del factor de apareamiento  $\alpha$  es L42S-Factor de apareamiento  $\alpha$ (20-85) (sec. con núm. de ident.:8). Un ejemplo no limitante de una variante de GLP-1(7-37) es GLP-1(7-37)[K34R].

En una modalidad, una variante de un polipéptido comprende de 1 a 2 sustituciones, delecciones o adiciones de aminoácidos en comparación con el polipéptido no modificado. En otra modalidad, una variante comprende de 1 a 5 sustituciones, delecciones o adiciones de aminoácidos en comparación con el polipéptido no modificado. En otra modalidad, una variante comprende de 1 a 15 sustituciones, delecciones o adiciones de aminoácidos respecto del polipéptido no modificado correspondiente.

De acuerdo con un segundo aspecto de la invención se proporciona una variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):

$$X_{38}-X_{39}-X_{40}-X_{41}-X_{42} \quad (\text{sec. con núm. de ident.:1}) \quad (I)$$

en donde

$X_{38}$  es F, L, I o V;

$X_{39}$  L, I, V o M;

$X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;

$X_{41}$  es S, Y, F, W, L, I, V o M;

$X_{42}$  es Y, W, L, I, V, M o S;

con la condición de que  $X_{38}-X_{42}$  no sea VIGYS. En una modalidad la variante del pro-péptido del factor de apareamiento  $\alpha$  no comprende VIGYS, VIDYS, VATYL, VIGYR, o AIGYL como  $X_{38}-X_{42}$ .

De acuerdo con un tercer aspecto de la invención se proporciona un precursor de GLP-1 que es un polipéptido de fusión que comprende:

Un pre-péptido,

Una variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):



en donde

- 5  $X_{38}$  es F, L, I o V;  
 $X_{39}$  L, I, V o M;  
 $X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;  
 $X_{41}$  es S, Y, F, W, L, I, V o M;  
 $X_{42}$  es Y, W, L, I, V, M o S;  
 10 con la condición de que  $X_{38}-X_{42}$  no sea VIGYS,

Opcionalmente un péptido de extensión, y

Un péptido de GLP-1.

- 15 El término "péptido de GLP-1", como se usa en la presente, está destinado para designar a GLP-1 (7-37), GLP-1 (7-36) amida así como análogos de estos, que puedan producirse mediante técnicas convencionales de ADN recombinante así como métodos sintéticos convencionales. Dichos péptidos de GLP-1 incluyen pero no se limitan al péptido 1 nativo similar al glucagón, por ejemplo los fragmentos peptídicos que comprenden GLP-1 (7-37) y  
 20 variantes funcionales de estos como las descritas en el documento WO 87/06941; los fragmentos peptídicos que comprenden GLP-1 (7-36) y derivados funcionales de estos como se describe en el documento WO 90/11296; los análogos de los péptidos activos de GLP-1 7-34, 7-35, 7-36, y 7-37 como se describe en el documento WO 91/11457; los fragmentos de GLP-1 truncados en el extremo N-terminal como se describe en el documento EP 0699686-A2; y los análogos y derivados de GLP-1 que incluyen un grupo imidazol N-terminal como se describe en el  
 25 documento EP 0708179-A2. Ejemplos no limitantes de un péptido de GLP-1 es GLP-1(7-37) y GLP-1(7-37)[K34R].

- El término "precursor de GLP-1" como se usa en la presente está destinado a referirse a un polipéptido que comprende un péptido extendido de GLP-1 donde la extensión sirve para facilitar la secreción, expresión o  
 30 recuperación del péptido de GLP-1. Ejemplos de precursores de GLP-1 pueden encontrarse en los documentos WO03/010186 y WO09/083549. Los precursores de GLP-1 están destinados a incluir péptidos de GLP-1 que tienen una pequeña extensión, por ejemplo 2-5 residuos de aminoácidos, así como péptidos de GLP-1 que tienen extensiones más largas que comprenden un pre-péptido y un pro-péptido.

- En una modalidad del método para la expresión recombinante de un polipéptido en levadura dicha secuencia de aminoácidos de la fórmula general (I) tiene una secuencia en donde

- 35  $X_{38}$  es F, L o V;  
 $X_{39}$  L, I, V o M;  
 $X_{40}$  es G o R;  
 $X_{41}$  es S, Y, L, I, V o M, y  
 40  $X_{42}$  es Y, W, L, V o M.

En otra modalidad dicha secuencia de aminoácidos de la fórmula general (I) tiene una secuencia en donde

- 45  $X_{38}$  es I o V;  
 $X_{39}$  L, I, V o M;  
 $X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;  
 $X_{41}$  es Y, F, o W, y  
 $X_{42}$  es L o I.

- 50 En otra modalidad dicha secuencia de aminoácidos de la fórmula general (I) tiene una secuencia en donde

- $X_{38}$  es V;  
 $X_{39}$  L, I, V o M;  
 $X_{40}$  es G o R;  
 $X_{41}$  es Y, y  
 55  $X_{42}$  es L.

En otra modalidad dicha secuencia de aminoácidos de la fórmula general (I) tiene una secuencia en donde  $X_{40}$  es R.  
 En otra modalidad dicha secuencia de aminoácidos de la fórmula general (I) tiene una secuencia en donde  $X_{40}$  es R y  $X_{42}$  es L.

- 60 Aún en otra modalidad dicha secuencia de aminoácidos de la fórmula general (I) tiene una secuencia en donde  $X_{38}$  es V,  $X_{39}$  es I,  $X_{40}$  es R,  $X_{41}$  es Y y  $X_{42}$  es L.

En otra modalidad el polipéptido que se expresa de manera recombinante es un péptido de GLP-1 o una variante de este, tal como GLP-1(7-37)[K34R] o GLP-1(9-37)[K34R].

En una modalidad el polipéptido que se expresa de manera recombinante tiene una extensión N-terminal, es decir ubicada entre la variante del factor de apareamiento  $\alpha$  y el polipéptido que se va a producir. Esta extensión puede facilitar la expresión o secreción del polipéptido desde la célula huésped, o puede proteger parte del polipéptido de un procesamiento proteolítico no deseado en el extremo N-terminal. En otra modalidad la extensión del extremo N-terminal es un polipéptido que tiene de 2 a 10 residuos de aminoácidos o que tiene de aproximadamente 8 a aproximadamente 200 residuos de aminoácidos. Frecuentemente se usan extensiones más cortas en el extremo N-terminal cuando la extensión sirve para facilitar la expresión del polipéptido en una célula huésped, o cuando la extensión sirve para proteger un polipéptido de ser procesado proteolíticamente en el extremo N-terminal. En otra modalidad la extensión del extremo N-terminal se selecciona del grupo que consiste en EEK, EEA EK, HK, EEA HK, E(EA)2HK, E(EA)3HK, EEGHK, EHPK, EEGEPK, EEAHELK, EEAHEVK, EEAHEMK, EEAHEFK, EEAHEYK, EEAHEWK, EEAHEGNTTPK y EELDARLEALK. En otra modalidad la extensión del extremo N-terminal se selecciona del grupo que consiste en QPMYKR, GQPMYK, PGQPMY, KPGQPM, LKPGQP, QLKPGQ, LQLKPG, WLQLKP, HWLQLK, WHWLQL, AWHWLQ, EAWHWL, AEAWHW y EAEAWH.

Cuando el polipéptido expresado incluye una extensión N-terminal, se diseña para escindir esta extensión N-terminal mediante el uso de una proteasa, una peptidasa o mediante escisión química. Pueden usarse las proteasas como la tripsina, proteasa de *Acromobacter lyticus* y enterocinasa. La enzima proteolítica particular seleccionada para la escisión se determina frecuentemente por el polipéptido que se fabrica. Por tanto, el experto en la técnica frecuentemente seleccionará la enzima proteolítica sobre la base de la secuencia de los polipéptidos, específicamente la presencia de cualquier sitio de escisión interno primario o secundario, así como adaptar la extensión del extremo N-terminal para formar un buen sitio de escisión.

La eficiencia de escisión de una proteasa cuando se utiliza para escindir un polipéptido expresado con una extensión N-terminal puede determinarse mediante un simple ensayo de la siguiente manera: Una solución acuosa adecuada del polipéptido se incuba a un pH y temperatura que sean favorables para la proteasa y se extraen muestras de la mezcla de reacción en el tiempo. Tan pronto como se extraigan las muestras la actividad enzimática se inactiva. Después de recolectar todas las muestras a lo largo del periodo de tiempo de interés, la concentración del polipéptido correspondiente sin la extensión del extremo N-terminal se determina por ejemplo mediante análisis de HPLC. La representación de la concentración del polipéptido escindido en función del tiempo indicará el progreso de la reacción. La comparación de tales trazas de reacción para diferentes extensiones del extremo N-terminal del polipéptido permitirá una clasificación de las extensiones del extremo N-terminal de acuerdo con la capacidad de la proteasa para liberar el polipéptido sin extensión N-terminal.

La construcción de ácido nucleico que codifica el polipéptido puede ser adecuadamente de origen genómico, ADNc o sintético. Las alteraciones de la secuencia de aminoácidos se llevan a cabo mediante la modificación del código genético por técnicas bien conocidas.

La secuencia de ADN que codifica el polipéptido usualmente se inserta en un vector recombinante que puede ser cualquier vector, que puede someterse convenientemente a procedimientos de ADN recombinante, y la elección del vector dependerá frecuentemente de la célula huésped en la que se va a introducir. Por lo tanto, el vector puede ser un vector de replicación autónoma, es decir un vector, que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, por ejemplo un plásmido. Alternativamente, el vector puede ser uno que, cuando se introduce en una célula huésped, se integra en el genoma de la célula huésped y se replica junto con el(los) cromosoma(s) donde se ha integrado.

El vector es, preferentemente, un vector de expresión en el cual la secuencia de ADN que codifica al polipéptido se une operativamente a segmentos adicionales necesarios para la transcripción del ADN. El término, "unido operativamente" indica que los segmentos están dispuestos de manera que funcionan coordinadamente para sus propósitos previstos, por ejemplo la transcripción inicia en un promotor y procede a través de la secuencia de ADN que codifica el polipéptido hasta que termina dentro de un terminador.

Por lo tanto, los vectores de expresión para usar en la expresión del polipéptido comprenderán un promotor capaz de iniciar y dirigir la transcripción de un gen clonado o ADNc. El promotor puede ser cualquier secuencia de ADN, que muestre actividad transcripcional en la célula huésped de elección y puede derivarse de genes que codifican proteínas homólogas o heterólogas para la célula huésped.

Adicionalmente, los vectores de expresión para usar en la expresión del polipéptido comprenderán, además, una secuencia terminadora, una secuencia reconocida por una célula huésped para terminar la transcripción. La secuencia terminadora se une operativamente al extremo terminal 3' de la secuencia de ácido nucleico que codifica el polipéptido. Cualquier terminador que sea funcional en la célula huésped de elección puede usarse en la presente invención.

La expresión del polipéptido se diseña para dirigirse a la vía secretora para la expresión extracelular en el medio de cultivo. Los péptidos señal útiles para usar como el pre-péptido en líderes para la expresión en células huésped de levadura se obtienen por ejemplo de los genes para el factor alfa de *Saccharomyces cerevisiae* y la invertasa de *Saccharomyces cerevisiae*. Otros ejemplos de pre-péptidos útiles (péptidos señal) son el péptido señal de la proteasa aspártica 3 de levadura (Yps1) (Egel-Mitani y otros (1990) YEAST 6:127-137 y el documento US 5,726,038), la señal del factor alfa del gen *MF $\alpha$ 1* (Thorner (1981) en The Molecular Biology of the Yeast *Saccharomyces cerevisiae*, Strathern y otros, eds., pp 143-180, Cold Spring Harbor Laboratory, NY) y el documento US 4,870,008), el péptido señal de la amilasa salival de ratón (O. Hagenbuchle y otros, Nature 289, 1981, pp. 643-646), un péptido señal de la carboxipeptidasa modificada (L.A. Valls y otros, Cell 48, 1987, pp. 887-897) y el péptido señal de BAR1 de levadura (documento WO 87/02670).

Los procedimientos utilizados para ligar las secuencias de ADN que codifican el polipéptido, el promotor, el terminador y la secuencia señal de secreción, respectivamente, y para insertarlas en los vectores adecuados que contienen la información necesaria para la replicación, son bien conocidos por los expertos en la técnica (cf., por ejemplo, Sambrook y otros, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, Nueva York, 1989).

Muchas células de levadura contienen un plásmido endógeno, denominado el plásmido de 2 micras que contiene diversos elementos que aseguran su mantenimiento estable en la célula de levadura (Guerineau y otros, 1971, Biochem. Biophys. Res. Comm. 42(3):550-557). Todo o parte de este plásmido endógeno de 2 micras puede usarse junto con un gen recombinante, como un método para asegurar el mantenimiento estable de las secuencias necesarias para la expresión recombinante (Beggs J.D., 1978, Transformation of yeast by a replicating hybrid plasmid, Nature, 275:104-109). El presente inventor ha encontrado que cuando el plásmido endógeno de 2 micras está presente en la célula, el plásmido de expresión para la expresión recombinante solo necesita comprender el origen de replicación y la región STB. Los otros factores presentes en el plásmido endógeno de 2 micras pueden funcionar en trans. El origen de replicación y la región STB solo constituyen una pequeña parte del plásmido 2 $\mu$  endógeno.

En un aspecto la presente invención proporciona un plásmido de expresión que contiene solo el origen de replicación y la región STB del plásmido de 2 micras. Por tanto, este plásmido de expresión mínimo no contiene ninguna de la región FLP, la región 1 de repetición, la región REP1, D-proteína, la región 2 de repetición y la región REP2. En una modalidad este plásmido comprende un casete de expresión, una parte de *E.coli* que incluye un gen AmpR, una secuencia de *S. pombe* que codifica la triosa-fosfato-isomerasa como se describe en Russell, PR (1985, Transcription of the triose-phosphate gene of *Schizosaccharomyces pombe* initiates from a start point different from that in *Saccharomyces cerevisiae*, Gene, 40:125-130). En otra modalidad el plásmido mínimo no comprende un gen AmpR u otro gen de resistencia a antibióticos. Un gen de resistencia a antibióticos de este tipo es útil durante el trabajo de clonación por ejemplo en *E.coli* pero se prefiere eliminar el gen de resistencia a antibióticos en el plásmido usado para la expresión de proteínas recombinantes a escala industrial. El gen de resistencia a antibióticos puede volverse no funcional o eliminarse de la célula huésped mediante procedimientos bien conocidos, véase por ejemplo el documento WO 00/04172. El plásmido de expresión mínimo es útil para la expresión de un polipéptido en una levadura.

Los vectores de la presente invención preferentemente contienen uno o más marcadores de selección que permiten la selección fácil de las células transformadas. Un marcador de selección es un gen cuyo producto proporciona resistencia viral o a biocidas, resistencia a metales pesados, auxotrofías complementarias, y similares. Ejemplos de marcadores de selección bacterianos son los genes *dal* de *Bacillus subtilis* o *Bacillus licheniformis*, o los marcadores que confieren resistencia a antibióticos como resistencia a ampicilina, kanamicina, cloranfenicol o tetraciclina. Los marcadores de selección para usar en células auxótrofas de levadura incluyen *ADE2*, *HIS3*, *LEU2*, *LYS2*, *MET3*, *TRP1*, y *URA3*. Un marcador de selección preferido para levadura es el gen TPI de *Schizosaccharomyces pombe* (Russell (1985) Gene 40:125-130).

En el vector, la secuencia de polinucleótido se conecta operativamente a una secuencia promotora adecuada. El promotor puede ser cualquier secuencia de ácido nucleico que muestre actividad transcripcional en la célula huésped de elección, incluidos promotores mutantes, truncados, e híbridos, y puede obtenerse a partir de genes que codifican polipéptidos extracelulares o intracelulares ya sean homólogos o heterólogos con respecto a la célula huésped. Ejemplos de promotores útiles en células huésped de levadura son los promotores de *MF1*, *TPI1*,  $\alpha$ ADH2, *TDH3* o *PGK1* de *Saccharomyces cerevisiae*.

La construcción polinucleotídica de la invención típicamente también estará conectada operativamente a un terminador adecuado. En levadura un ejemplo de un terminador adecuado es el terminador de *TPI1* (Alber y otros (1982) J. Mol. Appl. Genet. 1:419-434), pero puede ser cualquier terminador endógeno de levadura.

Los procedimientos usados para ligar la secuencia de polinucleótido de la invención, el promotor y el terminador, respectivamente, y para insertarlas en vectores adecuados de levadura que contienen la información necesaria para la replicación en levadura, son bien conocidos por los expertos en la técnica. Se entenderá que el vector puede construirse ya sea primero mediante la preparación de una construcción de ADN que contiene toda la secuencia de



ADN que codifica al polipéptido que se va a expresar, y posteriormente la inserción de este fragmento en un vector de expresión adecuado, o mediante la inserción secuencial de fragmentos de ADN que contienen la información genética para los elementos individuales (tal como la variante del factor de apareamiento  $\alpha$  de la invención, el polipéptido que se va a expresar opcionalmente que incluye una extensión N-terminal) seguido por el ensamblaje de los elementos mediante ligadura, métodos de clonación sin interrupciones o mediante clonación directa en la célula de levadura mediante recombinación homóloga.

La presente invención se refiere, además, a células huésped recombinantes, que comprenden una secuencia de polinucleótido que codifica la variante del factor de apareamiento  $\alpha$  de la invención y el polipéptido que se va a expresar. Un vector que comprende una secuencia de polinucleótido de este tipo se introduce en la célula huésped de manera que el vector se mantiene como un integrante cromosómico o como un vector extracromosómico autorreplicativo.

"Célula huésped" como se usa en la presente está destinado a referirse a un microorganismo que se usa para la expresión de un polipéptido de interés. Una célula huésped abarca cualquier progenie de una célula parental que no es idéntica a la célula parental debido a mutaciones que ocurren durante la replicación.

Una célula huésped adecuada para la presente invención es una célula de levadura. "Levadura" como se usa en la presente incluye levadura ascosporógena (Endomycetales), levadura basidiosporógena, y levadura que pertenece a los hongos imperfectos (Blastomycetes). Las levaduras ascosporógenas se dividen en las familias Spermophthoraceae y Saccharomycetaceae. La última está compuesta de cuatro subfamilias, Schizosaccharomycoidae (por ejemplo, género *Schizosaccharomyces*), Nadsonioideae, Lipomycoidae, y Saccharomycoidae (por ejemplo, géneros *Pichia*, *Kluyveromyces* y *Saccharomyces*). Las levaduras basidiosporógenas incluyen los géneros *Leucosporidium*, *Rhodospodium*, *Sporidiobolus*, *Filobasidium* y *Filobasidiella*. Las levaduras que pertenecen a los hongos imperfectos se dividen en dos familias, Sporobolomycetaceae (por ejemplo, géneros *Sorbolomyces* y *Bullera*) y Cryptococcaceae (por ejemplo, género *Candida*). Dado que la clasificación de las levaduras puede cambiar en el futuro, para los propósitos de esta invención, la levadura se define como se describe en *Biology and Activities of Yeast* (Skinner, F.A., Passmore, S.M., y Davenport, R.R., eds, Soc. App. Bacteriol. Symposium Series núm. 9, 1980. La biología de levaduras y la manipulación de la genética de levaduras se conocen bien en la técnica (véase, por ejemplo, *Biochemistry and Genetics of Yeast*, Bacil, M., Horecker, B.J., y Stopani, A.O.M., editores, 2da edición, 1987; *The Yeasts*, Rose, A.H., y Harrison, J.S., editores, 2da edición, 1987; y *The Molecular Biology of the Yeast Saccharomyces*, Strathern y otros, editores, 1981).

La célula huésped de levadura utilizada en el proceso de la invención puede ser cualquier organismo de levadura adecuado que, durante el cultivo, produzca grandes cantidades del polipéptido que se va a expresar.

Ejemplos de organismos de levadura adecuados son las cepas seleccionadas de una célula de una especie de *Candida*, *Kluyveromyces*, *Saccharomyces*, *Schizosaccharomyces*, *Pichia*, *Hansenula* y *Yarrowia*. En una modalidad, la célula huésped de levadura se selecciona de una *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, *Schizosaccharomyces pombe*, *Saccharomyces uvarum*, *Pichia kluyveri*, *Yarrowia lipolytica*, *Candida utilis*, *Candida cacaio* y *Geotrichum fermentans*. Otras células huésped de levadura útiles son una *Kluyveromyces lactis*, *Kluyveromyces fragilis*, *Hansenula polymorpha*, *Pichia pastoris*, *Yarrowia lipolytica*, *Schizosaccharomyces pombe*, *Ustilgo maylis*, *Candida maltose*, *Pichia guilliermondii* y *Pichia methanoliol* (cf. Gleeson y otros, J. Gen. Microbiol. 132, 1986, pp. 3459-3465; documentos US 4,882,279 y US 4,879,231). La transformación de las células de levadura puede realizarse por ejemplo mediante formación de protoplastos seguido por una transformación de una manera conocida *per se*.

La célula huésped para expresar el polipéptido es preferentemente una célula libre de cualquier gen funcional de resistencia a antibióticos. Aunque tales genes de resistencia a antibióticos son útiles durante las etapas de clonación iniciales, por ejemplo en *E.coli*, los genes de resistencia a antibióticos pueden volverse no funcionales o eliminarse de la célula huésped mediante procedimientos bien conocidos, véase por ejemplo el documento WO 00/04172.

"Medio" como se usa en la presente está destinado a referirse a una solución líquida para cultivar la célula huésped, es decir respaldar el crecimiento y la formación de productos de la levadura. Un medio adecuado para levadura es por ejemplo YPD o como se describe en el documento WO2008/037735. El medio contiene al menos una fuente de carbono, una o numerosas fuentes de nitrógeno, sales esenciales que incluyen sales de potasio, sodio, magnesio, fosfato y sulfato, metales traza, vitaminas solubles en agua, y auxiliares del proceso que incluyen pero sin limitarse a agentes antiespuma, inhibidores de proteasas, estabilizadores, ligandos e inductores. Las fuentes de carbono típicas son por ejemplo mono- o disacáridos. Las fuentes de nitrógeno típicas son, por ejemplo amoníaco, urea, aminoácidos, extracto de levadura, licor de maíz y proteínas completamente o parcialmente hidrolizadas. Los metales traza típicos son por ejemplo Fe, Zn, Mn, Cu, Mo y  $H_3BO_3$ . Las vitaminas solubles en agua típicas son por ejemplo biotina, pantotenato, niacina, tiamina, ácido p-aminobenzoico, colina, piridoxina, ácido fólico, riboflavina y ácido ascórbico.

El término "fermentación" como se usa en la presente está destinado a referirse a un proceso aséptico usado para propagar microorganismos sumergidos en un medio líquido. La fermentación se lleva a cabo preferentemente en tanques asépticos, agitados con líneas de suministro para la adición de gases comprimidos, estériles que consisten en, pero sin limitarse a, aire, oxígeno y amoníaco. Un tanque de fermentación puede contener sensores y dispositivos para monitorizar el pH, la temperatura, la presión, la velocidad de agitación, el nivel de oxígeno disuelto, el contenido líquido, el nivel de espuma, las velocidades de adición de la alimentación y las velocidades de adición de ácido y base. Además, el tanque de fermentación puede estar equipado con dispositivos ópticos para monitorizar los niveles de densidad celular, las concentraciones de metabolitos y los productos independientemente de su forma físico-química.

El producto deseado que se produce durante la fermentación está presente como un material extracelular soluble o como un material intracelular ya sea en la forma de material soluble o como material insoluble que incluye un material agregado. Se prefiere que esté presente como un material extracelular soluble. Un proceso de fermentación típicamente se lleva a cabo en tanques con un volumen de trabajo en el intervalo de 100 ml a 200 000 L. Un proceso de fermentación puede funcionar como un proceso discontinuo, un proceso alimentado de manera discontinua, un proceso repetido alimentado de manera discontinua o un proceso continuo.

Los polipéptidos secretados, una proporción importante de los cuales estará presente en el medio en una forma procesada correctamente, pueden recuperarse del medio por procedimientos convencionales que incluyen la separación de las células de levadura del medio mediante centrifugación, filtración o atrapamiento del polipéptido mediante una matriz de intercambio iónico o mediante una matriz de absorción de fase inversa, precipitación de los componentes proteicos del sobrenadante o el filtrado por medio de una sal, por ejemplo sulfato de amonio, seguido de purificación mediante una variedad de procedimientos cromatográficos, por ejemplo cromatografía de intercambio iónico, cromatografía de afinidad, o similares.

Las variantes novedosas del factor de apareamiento  $\alpha$  de la presente invención también facilitan la reducción de la O-glicosilación del polipéptido durante la expresión en levadura. Como tal las variantes del factor de apareamiento  $\alpha$  de la presente invención pueden usarse en un método mejorado para producir polipéptidos tal como los péptidos de GLP-1 en levadura. La expresión del polipéptido en una célula de levadura que tiene una capacidad reducida de O-glicosilación puede mantener el rendimiento mejorado de dicho precursor mientras que al mismo tiempo reduce incluso adicionalmente la fracción de dicho polipéptido que se O-glicosila durante la expresión.

Las O-manosiltransferasas de proteínas (PMT) inician el ensamblaje de O-manosil glicanos, una modificación esencial de las proteínas en los hongos. Las PMT se conservan en hongos y la familia de las PMT se clasifica filogenéticamente en las subfamilias *PMT1*, *PMT2* y *PMT4*, que difieren en la especificidad del sustrato proteico. Las O-manosiltransferasas de proteínas *Pmt1p* y *Pmt2p* catalizan la O-glicosilación de residuos de serina y treonina en proteínas en el retículo endoplásmico de levaduras mediante la transferencia de un residuo manosil a partir del dolicol fosfato-D-manosa (Gentzsch y otros, FEBS Lett 1995, 18, pp128-130). En *Saccharomyces cerevisiae* así como en muchas otras levaduras la familia de las PMT es altamente redundante, y solo la delección simultánea de miembros de las subfamilias de *PMT1/PMT2* y *PMT4* es letal (Girrbach y Strahl, J. Biol. Chem. 2003, 278, pp12554-62). El documento US 5,714,377 describe que las células de levadura que tienen una reducida capacidad de O-glicosilación a partir de una modificación de *PMT1/PMT2* son todavía viables y muestran buenas características de crecimiento en condiciones de fermentación industrial.

Modalidades no limitantes

La invención se describe adicionalmente mediante las siguientes modalidades no limitantes:

1. Un método para la expresión recombinante de un polipéptido en levadura que comprende el cultivo de una cepa de levadura que comprende una secuencia de ADN que codifica una señal de procesamiento y secreción hacia el extremo corriente arriba del polipéptido, en donde dicha señal de procesamiento y secreción comprende una variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):

$$X_{38}-X_{39}-X_{40}-X_{41}-X_{42} \quad (\text{sec. con núm. de ident.:1}) \quad (\text{I})$$

en donde

$X_{38}$  es F, L, I o V;  
 $X_{39}$  L, I, V o M;  
 $X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;  
 $X_{41}$  es S, Y, F, W, L, I, V o M;  
 $X_{42}$  es Y, W, L, I, V, M o S;  
 con la condición de que  $X_{38}-X_{42}$  no sea VIGYS.

2. El método de acuerdo con la modalidad 1 en donde

- $X_{38}$  es F, L o V;  
 $X_{39}$  L, I, V o M;  
 $X_{40}$  es G o R;  
 $X_{41}$  es S, Y, L, I, V o M, y  
 $X_{42}$  es Y, W, L, V o M.
- 5
3. El método de acuerdo con la reivindicación 1 en donde  $X_{38}$  es I o V;
- 10
- $X_{39}$  L, I, V o M;  
 $X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;  
 $X_{41}$  es Y, F, o W, y  
 $X_{42}$  es L o I.
4. El método de acuerdo con cualquiera de las modalidades 1-3 en donde
- 15
- $X_{38}$  es V;  
 $X_{39}$  L, I, V o M;  
 $X_{40}$  es G o R;  
 $X_{41}$  es Y, y  
 $X_{42}$  es L.
- 20
5. El método de acuerdo con cualquiera de las modalidades 1-4, en donde cuatro de los residuos de aminoácidos en  $X_{38}$ - $X_{39}$ - $X_{40}$ - $X_{41}$ - $X_{42}$  son idénticos al residuo de aminoácido correspondiente en VIGYL o VIGYS.
6. El método de acuerdo con cualquiera de las modalidades 1-5, en donde al menos tres de los residuos de aminoácidos en  $X_{38}$ - $X_{39}$ - $X_{40}$ - $X_{41}$ - $X_{42}$  son idénticos al residuo de aminoácido correspondiente en VIGYL o VIGYS.
- 25
7. El método de acuerdo con cualquiera de las modalidades 1-6, en donde  $X_{41}$  es Y.
8. El método de acuerdo con cualquiera de las modalidades 1-6, en donde  $X_{41}$  es L.
- 30
9. El método de acuerdo con cualquiera de las modalidades 1-8, en donde  $X_{42}$  es L.
10. El método de acuerdo con cualquiera de las modalidades 1 y 5-8, en donde  $X_{42}$  es S.
11. El método de acuerdo con cualquiera de las modalidades 1 y 5-8, en donde  $X_{42}$  es Y, W, L, I, V o M.
- 35
12. El método de acuerdo con cualquiera de las modalidades 1-11, en donde  $X_{40}$  es R.
13. El método de acuerdo con la modalidad 12, en donde  $X_{38}$  es V,  $X_{39}$  es I,  $X_{41}$  es Y y  $X_{42}$  es L.
14. El método de acuerdo con la modalidad 12, en donde  $X_{38}$  es V,  $X_{39}$  es I,  $X_{41}$  es Y y  $X_{42}$  es S.
15. El método de acuerdo con cualquiera de las modalidades 1 y 12, en donde la fórmula (I) es VI- $X_{40}$ -YL o VI- $X_{40}$ -YS.
- 40
16. El método de acuerdo con la modalidad 15 en donde  $X_{40}$  es A, Y, F, W, R, K, L, I, V o M.
17. El método de acuerdo con cualquiera de las modalidades 1-16, en donde dicha variante del pro-péptido del factor de apareamiento  $\alpha$  tiene menos de 10 cambios de residuos de aminoácidos fuera de la secuencia  $X_{38}$ - $X_{42}$  en comparación con el pro-péptido del factor de apareamiento  $\alpha$  como se establece en la sec. con núm. de ident.:2 (residuos de aminoácidos 20-85).
- 45
18. El método de acuerdo con cualquiera de las modalidades 1-17, en donde dicha variante del pro-péptido del factor de apareamiento  $\alpha$  tiene menos de 5 cambios de residuos de aminoácidos fuera de la secuencia  $X_{38}$ - $X_{42}$  en comparación con el pro-péptido del factor de apareamiento  $\alpha$  como se establece en la sec. con núm. de ident.:2 (residuos de aminoácidos 20-85).
- 50
19. El método de acuerdo con cualquiera de las modalidades 1-18, en donde dicha variante del pro-péptido del factor de apareamiento  $\alpha$  tiene menos de 2 cambios de residuos de aminoácidos fuera de la secuencia  $X_{38}$ - $X_{42}$  en comparación con el pro-péptido del factor de apareamiento  $\alpha$  como se establece en la sec. con núm. de ident.:2 (residuos de aminoácidos 20-85).
- 55
20. El método de acuerdo con cualquiera de las modalidades 1-19, en donde dicha variante del pro-péptido del factor de apareamiento  $\alpha$  es parte de un prepro-péptido del factor de apareamiento  $\alpha$  que comprende un pre-péptido como la parte N-terminal fusionada a dicha variante del pro-péptido del factor de apareamiento  $\alpha$  como la parte C-terminal.

21. El método de acuerdo con la modalidad 20, en donde dicho pre-péptido es del péptido señal de la proteasa aspártica 3 de levadura (YAP3), de la señal del factor  $\alpha$  del gen MF $\alpha$ 1 de *S. cerevisiae* o una variante de estos.
- 5 22. El método de acuerdo con cualquiera de las modalidades 1-21, en donde dicha levadura porta al menos una modificación genética que reduce su capacidad de O-glicosilación.
23. El método de acuerdo con la modalidad 22, en donde dicha levadura porta al menos una modificación genética dentro de los genes de *PMT1* o *PMT2* que reduce su capacidad de O-glicosilación.
- 10 24. El método de acuerdo con cualquiera de las modalidades 22-23, en donde dicha levadura porta al menos una modificación genética que reduce su capacidad de O-glicosilación mediante la O-manosiltransferasa de proteínas 1 (*PMT1*) del polipéptido GLP-1(7-37)[K34R] cuando se expresa con el líder alfa en comparación con la levadura que porta los genes no modificados correspondientes.
- 15 25. El método de acuerdo con cualquiera de las modalidades 22-24, en donde dicha levadura porta al menos una modificación genética que reduce su capacidad de O-glicosilación mediante la O-manosiltransferasa de proteínas 2 (*PMT2*) del polipéptido GLP-1(7-37)[K34R] cuando se expresa con el líder alfa en comparación con la levadura que porta los genes no modificados correspondientes.
- 20 26. El método de acuerdo con cualquiera de las modalidades 22-25, en donde dicha capacidad de O-glicosilación se reduce en al menos un factor de 2.
27. El método de acuerdo con cualquiera de las modalidades 22-26, en donde dicha capacidad de O-glicosilación se reduce en al menos un factor de 4.
- 25 28. El método de acuerdo con cualquiera de las modalidades 22-27, en donde dicha al menos una modificación genética se ubica en la región codificante de *PMT1* o *PMT2*.
- 30 29. El método de acuerdo con cualquiera de las modalidades 22-27, en donde dicha al menos una modificación genética se ubica en las regiones responsables o implicadas en la expresión y/o regulación transcripcional de *PMT1* o *PMT2*.
30. El método de acuerdo con cualquiera de las modalidades 22-27, en donde se elimina el gen *PMT1* en dicha levadura.
- 35 31. El método de acuerdo con cualquiera de las modalidades 22-27, en donde se eliminan ambos genes *PMT1* y *PMT2* en dicha levadura.
- 40 32. El método de acuerdo con cualquiera de las modalidades 1-31, en donde dicho polipéptido comprende un péptido de GLP-1.
33. El método de acuerdo con la modalidad 32, en donde dicho polipéptido comprende GLP-1(9-37)[K34R] o GLP-1(9-37)[K34R].
- 45 34. El método de acuerdo con la modalidad 33, en donde dicho polipéptido es GLP-1(7-37)[K34R], GLP-1(9-37)[K34R] o GLP-1(9-37)[K34R,G37K] (sec. con núm. de ident.:44).
35. El método de acuerdo con cualquiera de las modalidades 32-34, en donde dicho polipéptido tiene una extensión N-terminal.
- 50 36. El método de acuerdo con la modalidad 35, en donde dicha extensión N-terminal se selecciona del grupo que consiste en EEK, EEAEK, HK, EEAHK, E(EA)2HK, E(EA)3HK, EEGHK, EHPK, EEGEPK, EEAEHLK, EEAEHVK, EEAEHMK, EEAEHFK, EEAEHYK, EEAEHWKEEGNTTPK y EELDARLEALK.
37. El método de acuerdo con la modalidad 35, en donde dicha extensión N-terminal se selecciona del grupo que consiste en DV, DVKPGQPLA, DVKPGQPEY, DVKPGQPLY, DVKPGQPLY, DVKPGQPLE, DVKPGQPMY y DVKPGQPMYDDDDK.
- 55 38. El método de acuerdo con la modalidad 35, en donde dicha extensión N-terminal se selecciona del grupo que consiste en QPMYKR, GQPMYK, PGQPMY, KPGQPM, LKPGQP, QKPGQ, LQLKPG, WLQLKP, HWLQLK, WHWLQL, AWHWLQ, EAWHWL, AEAWHW y EAEAWH.
- 60 39. Variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):
- 65  $X_{38}-X_{39}-X_{40}-X_{41}-X_{42}$  (sec. con núm. de ident.:1) (I)

en donde

$X_{38}$  es F, L, I o V;

$X_{39}$  L, I, V o M;

$X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;

5  $X_{41}$  es S, Y, F, W, L, I, V o M;

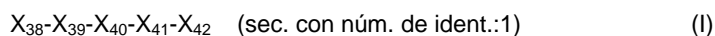
$X_{42}$  es Y, W, L, I, V, M o S;

con la condición de que  $X_{38}$ - $X_{42}$  no sea VIGYS.

10 40. La variante del pro-péptido del factor de apareamiento  $\alpha$  de acuerdo con la modalidad 39 en donde  $X_{38}$ - $X_{42}$  no es VIGYS, VIDYS, VATYL, VIGYR, o AIGYL.

41. Precursor de GLP-1 que es un polipéptido de fusión que comprende:  
Un pre-péptido,

15 Una variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):



20 en donde

$X_{38}$  es F, L, I o V;

$X_{39}$  L, I, V o M;

$X_{40}$  es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M;

$X_{41}$  es S, Y, F, W, L, I, V o M;

25  $X_{42}$  es Y, W, L, I, V, M o S;

con la condición de que  $X_{38}$ - $X_{42}$  no sea VIGYS,

Opcionalmente un péptido de extensión, y  
Un péptido de GLP-1.

30 42. El precursor de GLP-1 de acuerdo con la modalidad 41 en donde los péptidos comprendidos por dicho precursor de GLP-1 se fusionan de acuerdo con el orden en que son mencionados, es decir pre-péptido – variante del pro-péptido del factor de apareamiento  $\alpha$ , el péptido de extensión opcional y el péptido de GLP-1.

35 43. Secuencia de ADN que codifica el polipéptido de acuerdo con cualquiera de las modalidades 39-42.

44. Vector de expresión que comprende una secuencia de ADN de acuerdo con la modalidad 43.

40 45. Vector de expresión de acuerdo con la modalidad 44, en donde dicha secuencia de ADN que codifica un polipéptido que se va a expresar se une operativamente a un promotor corriente arriba y un terminador corriente abajo.

46. Célula huésped que comprende el vector de expresión de acuerdo con cualquiera de las modalidades 44-45.

45 47. La célula huésped de acuerdo con la modalidad 46, que se selecciona del grupo que consiste en *Saccharomyces spp.*, *Pichia spp.*, *Hansenula spp.*, *Arxula spp.*, *Kluyveromyces spp.*, *Yarrowia spp.* y *Schizosaccharomyces spp.*

50 Todos los encabezamientos y sub-encabezamientos se usan en la presente solo por conveniencia y no deben interpretarse como limitantes de la invención de ninguna manera. El uso de cualquiera y todos los ejemplos, o lenguaje ilustrativo (por ejemplo, "tal como") proporcionado en la presente, está destinado meramente a ilustrar mejor la invención y no constituye una limitación sobre el alcance de la invención a menos que se indique de cualquier otra manera. Ningún lenguaje en la descripción debe interpretarse como que indica algún elemento no reivindicado como esencial para la práctica de la invención. Esta invención incluye todas las modificaciones y equivalentes de la materia mencionada en las reivindicaciones adjuntas a esta como permitidas por la ley aplicable.

## 55 EJEMPLOS

Ejemplos 1-45.

60 Se construyó un plásmido, que contenía el promotor de *TDH3*, el gen que codifica el pre-pro-péptido de MFalfa, el gen que codifica un GLP-1 extendido en el extremo N-terminal (DVKPGQPMYDDDDK-GLP-1(7-37)[K34R]) (sec. con núm. de ident.:45), una región mínima de 2 micras para el mantenimiento en levadura, y un marcador de selección, específicamente el gen *TPI* de *S.pombe* (véase la figura 1).

En este fondo del plásmido se introdujeron diversas mutaciones individuales en la región que codifica la posición 38-42 de la región de MFalfa-pro-péptido,  $X_{38}$ - $X_{42}$  = secuencia VIGYL (véase la figura 2). El experimento que usa la

secuencia VIGYL wt es el Ejemplo 1, c.f. tablas 1-5. Otros experimentos de referencia son los ejemplos 5-6, 10-11, 28-29, 37-38 y 44-45.

Estos plásmidos se introdujeron en una cepa de *Saccharomyces cerevisiae* que carece del gen *TPI1*, lo que permite la selección de transformantes que portan el plásmido en medio que contiene glucosa como única fuente de carbono. Los transformantes después de eso se cultivaron durante 3 días en matraces de agitación en 5 ml de medio relevante y los sobrenadantes del cultivo se analizaron mediante LCMS para determinar la concentración del péptido de GLP-1 secretado, y el grado de O-glicosilación del péptido. En las tablas 1-5 se muestra el rendimiento y el grado de O-glicosilación del tipo silvestre y de los mutantes de un único residuo de aminoácido de la secuencia X<sub>38</sub>-X<sub>42</sub> (se incluye un par de ejemplos de referencia sobre cada posición).

Tabla 1. Datos de la expresión de GLP-1 extendido en el extremo N-terminal bajo el control de mutantes del propéptido del factor de apareamiento  $\alpha$  que tienen una única mutación de aminoácidos en la posición 38 (X<sub>38</sub>). Las concentraciones de GLP-1 extendido en el extremo N-terminal e impureza O-glico se normalizan contra los datos de la misma expresión pero bajo el control del propéptido del factor de apareamiento  $\alpha$  que tiene la Val de tipo silvestre como X<sub>38</sub>.

Posición 38			
Ejemplo	Mutación	% de conc. respecto del peso	% de O-glico respecto del peso
1	Val	100	100
2	Ile	92	91
3	Phe	107	103
4	Leu	106	104
5	Asp	33	277
6	Pro	47	354

Tabla 2. Datos de la expresión de GLP-1 extendido en el extremo N-terminal bajo el control de mutantes del propéptido del factor de apareamiento  $\alpha$  que tienen una única mutación de aminoácidos en la posición 39 (X<sub>39</sub>). Las concentraciones de GLP-1 extendido en el extremo N-terminal e impureza O-glico se normalizan contra los datos de la misma expresión pero bajo el control del propéptido del factor de apareamiento  $\alpha$  que tiene la Ile de tipo silvestre como X<sub>39</sub>.

Posición 39			
Ejemplo	Mutación	% de conc. respecto del peso	% de O-glico respecto del peso
1	Ile	100	100
7	Leu	113	82
8	Val	108	76
9	Met	126	88
10	Glu	40	337
11	His	25	313

Tabla 3. Datos de la expresión de GLP-1 extendido en el extremo N-terminal bajo el control de mutantes del propéptido del factor de apareamiento  $\alpha$  que tienen una única mutación de aminoácidos en la posición 40 ( $X_{40}$ ). Las concentraciones de GLP-1 extendido en el extremo N-terminal y de impureza O-glico se normalizan contra los datos de la misma expresión pero bajo el control del propéptido del factor de apareamiento  $\alpha$  que tiene la Gly de tipo silvestre como  $X_{40}$ .

<b>Posición 40</b>			
<b>Ejemplo</b>	<b>Mutación</b>	<b>% de conc. respecto del peso</b>	<b>% de O-glico respecto del peso</b>
1	Gly	100	100
12	Ala	84	48
13	Ser	79	89
14	Thr	80	67
15	Asp	51	90
16	Glu	58	92
17	Gln	87	98
18	Tyr	60	49
19	Phe	50	23
20	Trp	53	18
21	Arg	106	14
22	Lys	93	36
23	His	74	83
24	Leu	42	13
25	Ile	18	1
26	Val	26	1
27	Met	60	36
28	Asn	81	112
29	Pro	46	253

Tabla 4. Datos de la expresión de GLP-1 extendido en el extremo N-terminal bajo el control de mutantes del propéptido del factor de apareamiento  $\alpha$  que tienen una única mutación de aminoácidos en la posición 41 ( $X_{41}$ ). Las concentraciones de GLP-1 extendido en el extremo N-terminal y la impureza O-glico se normalizan contra los datos de la misma expresión pero bajo el control del propéptido del factor de apareamiento  $\alpha$  que tiene la Tyr de tipo silvestre como  $X_{41}$ .

Posición 41			
Ejemplo	Mutación	% de conc. respecto del peso	% de O-glico respecto del peso
1	Tyr	100	100
30	Phe	69	74
31	Trp	79	73
32	Ser	107	182
33	Leu	129	104
34	Ile	111	171
35	Val	134	136
36	Met	131	151
37	Gly	50	366
38	Asp	27	590

Tabla 5. Datos de la expresión de GLP-1 extendido en el extremo N-terminal bajo el control de mutantes del propéptido del factor de apareamiento  $\alpha$  que tienen una única mutación de aminoácidos en la posición 42 ( $X_{42}$ ). Las concentraciones de GLP-1 extendido en el extremo N-terminal y la impureza O-glico se normalizan contra los datos de la misma expresión pero bajo el control del propéptido del factor de apareamiento  $\alpha$  que tiene la Leu de tipo silvestre como  $X_{42}$ .

Posición 42			
Ejemplo	Mutación	% de conc. respecto del peso	% de O-glico respecto del peso
1	Leu	100	100
39	Ile	98	64
40	Tyr	123	156
41	Trp	116	163
42	Val	103	111
43	Met	109	142
44	Asp	19	525
45	Pro	47	472



## Ejemplo 46.

Un plásmido similar a los de los ejemplos 1-45, que contiene el promotor de *TDH3*, el gen que codifica el pre-pro-péptido de MFalfa, el gen que codifica un GLP-1 extendido en el extremo N-terminal (DVKPGQPMYDDDDK-GLP-1(7-37)[K34R]), una región mínima de 2 micras para el mantenimiento en levadura, y un marcador de selección (Figura 1) se transformó en dos fondos de cepas diferentes: uno que contiene el gen *PMT1* (PMT +), el otro con eliminación del gen *PMT1* (PMT -). El gen de *PMT1* codifica una proteína manosiltransferasa implicada en la O-glicosilación.

El pro-péptido de MFalfa fue el de tipo silvestre (40G) o mutado (G40R). Las cepas se cultivaron en condiciones idénticas en cultivos continuos, con una tasa de dilución = 0.1, y pH 5.8. Las muestras se analizaron mediante HPLC para determinar la concentración del precursor de GLP-1 en el medio gastado y mediante LCMS para determinar la proporción del precursor de GLP-1 O-glicosilado.

Los resultados demuestran que los efectos de hecho son aditivos.

La mutación G40R redujo la O-glicosilación con más del 80 %. La delección de *PMT1* reduce la O-glicosilación en más del 80 %, en este caso se obtienen niveles de O-glicosilación cercanos al límite de detección. Además, la mutación G40R produce un aumento sorprendente en el rendimiento de GLP-1 extendido en el extremo N-terminal en comparación con la versión 40G de tipo silvestre, ambas en la cepa PMT+ así como en la cepa con *pmt1* eliminado (véase la tabla 6).

Estos resultados demuestran que la combinación de una mutación beneficiosa en el propéptido de MFalfa, por ejemplo G40R, y una cepa huésped con *PMT1* eliminado da como resultado una reducción aditiva en la O-glicosilación.

Tabla 6. Concentración máxima normalizada (rendimiento) del péptido DVKPGQPMYDDDDK-GLP-1(7-37)[K34R] en cultivos continuos cuando se usa un pro-péptido de MFalfa mutado en 40R en comparación con 40G de tipo silvestre correspondiente.

Cepa PMT+		Cepa PMT-	
40G	40R (respecto de 40G en la misma cepa)	40G	40R (respecto de 40G en la misma cepa)
100 %	133 %	100 %	169 %

## Ejemplo 47.

Para examinar el efecto de una de las mutaciones durante la expresión de otros péptidos de GLP-1, se construyeron plásmidos, que contenían el promotor de *TDH3*, el gen que codifica el pre-pro-péptido de MFalfa en la forma del tipo silvestre (40G) o una mutación de 40R, el gen que codifica GLP-1 extendido en el extremo N-terminal (DVKPGQPMYDDDDK-GLP-1(9-37)[K34R] o DVKPGQPMYDDDDK-GLP-1(9-37)[K34R,G37K]), una región mínima de 2 micras para el mantenimiento en levadura, y un marcador de selección, específicamente el gen de *TPI* de *S.pombe* (véase la figura 1).

Estos plásmidos se introdujeron en una cepa de *Saccharomyces cerevisiae* que carece del gen *TPI1*, lo que permite la selección de transformantes que portan el plásmido en medio que contiene glucosa como única fuente de carbono. Los transformantes después de eso se cultivaron durante 3 días en matraces de agitación en 5 ml de medio relevante y los sobrenadantes del cultivo se analizaron mediante LCMS para determinar la concentración del péptido de GLP-1 secretado, y el grado de O-glicosilación del péptido. La tabla 7 muestra los resultados que incluyen el rendimiento de péptido de GLP-1 y el grado de O-glicosilación del péptido de GLP-1 cuando se expresó del tipo silvestre (wt) y de la mutación de 40R.

Tabla 7. Datos de la expresión del péptido de GLP-1 extendido en el extremo N-terminal específico bajo el control de mutantes del propéptido del factor de apareamiento  $\alpha$  con o sin una única sustitución de aminoácidos a Arg en la posición 40 ( $X_{40}$ ). Las concentraciones del péptido de GLP-1 extendido en el extremo N-terminal (rendimiento) y la impureza O-glicosilada se normalizan contra los datos de la expresión del mismo péptido de GLP-1 pero bajo el control del propéptido del factor de apareamiento  $\alpha$  que tiene la Gly de tipo silvestre como  $X_{40}$ .

Péptido de GLP-1	$X_{40}$ del pre-pro-péptido de MFalfa	Rendimiento (%) respecto del peso)	O-glicosilación (%) respecto del peso)
GLP-1(9-37)[K34R]	Gly (wt)	100	100
GLP-1(9-37)[K34R]	Arg	127	< 1 (BDL)
GLP-1(9-37)[K34R,G37K]	Gly (wt)	100	BDL
GLP-1(9-37)[K34R,G37K]	Arg	143	BDL

BDL: por debajo del límite de detección

Aunque determinadas características de la invención se han ilustrado y descrito en la presente, muchas modificaciones, sustituciones, cambios y equivalentes serán evidentes para un experto en la técnica. Por lo tanto, debe entenderse que las reivindicaciones adjuntas están destinadas a cubrir todas estas modalidades.

Listado de secuencias

<110> Novo Nordisk A/S

<120> Variantes del pro-péptido del factor de apareamiento alfa

<130> 130089WO01

<160> 44

<170> PatentIn versión 3.5

<210> 1

<211> 5

<212> PRT

<213> Secuencia artificial

<220>

<223> subsecuencia de variante del propéptido de MFalfa

<220>

<221> CARACTERÍSTICA\_MISC

<222> (1)..(1)

<223> Residuo de aminoácido que es F, L, I o V.

<220>

<221> CARACTERÍSTICA\_MISC

<222> (2)..(2)

<223> Residuo de aminoácido que es L, I, V o M.

<220>

<221> CARACTERÍSTICA\_MISC

<222> (3)..(3)

<223> Residuo de aminoácido que es A, G, S, E, Q, Y, F, W, R, K, H, L, I, V o M.

<220>

<221> CARACTERÍSTICA\_MISC

<222> (4)..(4)

<223> Residuo de aminoácido que es S, Y, F, W, L, I, V o M.

<220>

<221> CARACTERÍSTICA\_MISC

<222> (5)..(5)

<223> Residuo de aminoácido que es Y, W, L, I, V, M o S.

<400> 1

5

Xaa Xaa Xaa Xaa Xaa  
1 5

10

<210> 2

<211> 85

<212> PRT

<213> *Saccharomyces cerevisiae*

15

<400> 2

20

Met Arg Phe Pro Ser Ile Phe Thr Ala Val Leu Phe Ala Ala Ser Ser  
1 5 10 15

25

Ala Leu Ala Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln  
20 25 30

30

Ile Pro Ala Glu Ala Val Ile Gly Tyr Leu Asp Leu Glu Gly Asp Phe  
35 40 45

35

Asp Val Ala Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu  
50 55 60

40

Phe Ile Asn Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val  
65 70 75 80

45

Ser Leu Asp Lys Arg  
85

50

<210> 3

<211> 5

<212> PRT

<213> Secuencia artificial

55

<220>

<223> Secuencia VIGYL

<400> 3

60

Val Ile Gly Tyr Leu  
1 5

65

<210> 4

<211> 5

<212> PRT

<213> Secuencia artificial

<220>

<223> Secuencia VIGYS

5

<400> 4

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly  
1 5 10 15

10

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Lys Gly Arg Gly  
20 25 30

15

<210> 6

<211> 31

<212> PRT

<213> Secuencia artificial

20

<220>

<223> Análogo de GLP-1(7-37)

<400> 6

25

His Ala Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly  
1 5 10 15

30

Gln Ala Ala Lys Glu Phe Ile Ala Trp Leu Val Arg Gly Arg Gly  
20 25 30

<210> 7

35

<211> 29

<212> PRT

<213> Secuencia artificial

<220>

40

<223> Análogo de GLP-1(9-37)

<400> 7

Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly Gln Ala  
1 5 10 15

45

Ala Lys Glu Phe Ile Ala Trp Leu Val Arg Gly Arg Gly  
20 25

50

<210> 8

<211> 66

55

<212> PRT

<213> Secuencia artificial

<220>

<223> Variante del propéptido de MFalfa L42S-MFa(20-85)

60

<400> 8

5	Ala Pro Val Asn Thr Thr Thr Glu Asp Glu Thr Ala Gln Ile Pro Ala	1 5 10 15
10	Glu Ala Val Ile Gly Tyr Ser Asp Leu Glu Gly Asp Phe Asp Val Ala	20 25 30
15	Val Leu Pro Phe Ser Asn Ser Thr Asn Asn Gly Leu Leu Phe Ile Asn	35 40 45
20	Thr Thr Ile Ala Ser Ile Ala Ala Lys Glu Glu Gly Val Ser Leu Asp	50 55 60
25	Lys Arg	65
30	<210> 9 <211> 5 <212> PRT <213> Secuencia artificial	
35	<220> <223> Extensión 1  <400> 9	
40	Glu Glu Ala Glu Lys	1 5
45	<210> 10 <211> 5 <212> PRT <213> Secuencia artificial  <220> <223> Extensión 2  <400> 10	
50	Glu Glu Ala His Lys	1 5
55	<210> 11 <211> 7 <212> PRT  <213> Secuencia artificial	
60	<220> <223> Extensión 3  <400> 11	
65		

		<b>Glu Glu Ala Glu Ala His Lys</b>
5		<b>1 5</b>
	<210> 12	
	<211> 9	
10	<212> PRT	
	<213> Secuencia artificial	
	<220>	
	<223> Extensión 4	
15	<400> 12	
		<b>Glu Glu Ala Glu Ala Glu Ala His Lys</b>
20		<b>1 5</b>
	<210> 13	
	<211> 5	
	<212> PRT	
25	<213> Secuencia artificial	
	<220>	
	<223> Extensión 5	
30	<400> 13	
		<b>Glu Glu Gly His Lys</b>
35		<b>1 5</b>
	<210> 14	
	<211> 6	
	<212> PRT	
40	<213> Secuencia artificial	
	<220>	
	<223> Extensión 6	
45	<400> 14	
		<b>Glu Glu Gly Glu Pro Lys</b>
50		<b>1 5</b>
	<210> 15	
	<211> 7	
	<212> PRT	
	<213> Secuencia artificial	
55	<220>	
	<223> Extensión 7	
	<400> 15	
60		<b>Glu Glu Ala His Glu Leu Lys</b>
		<b>1 5</b>
65	<210> 16	
	<211> 7	

	<212> PRT	
	<213> Secuencia artificial	
	<220>	
5	<223> Extensión 8	
	<400> 16	
10		Glu Glu Ala His Glu Val Lys 1 5
	<210> 17	
	<211> 7	
15	<212> PRT	
	<213> Secuencia artificial	
	<220>	
	<223> Extensión 9	
20	<400> 17	
25		Glu Glu Ala His Glu Met Lys 1 5
	<210> 18	
30	<211> 7	
	<212> PRT	
	<213> Secuencia artificial	
	<220>	
35	<223> Extensión 10	
	<400> 18	
40		Glu Glu Ala His Glu Phe Lys 1 5
	<210> 19	
45	<211> 7	
	<212> PRT	
	<213> Secuencia artificial	
	<220>	
50	<223> Extensión 11	
	<400> 19	
55		Glu Glu Ala His Glu Tyr Lys 1 5
	<210> 20	
	<211> 15	
60	<212> PRT	
	<213> Secuencia artificial	
	<220>	
	<223> Extensión 12	
65	<400> 20	

	<b>Glu</b>	<b>Glu</b>	<b>Ala</b>	<b>His</b>	<b>Glu</b>	<b>Trp</b>	<b>Lys</b>	<b>Glu</b>	<b>Glu</b>	<b>Gly</b>	<b>Asn</b>	<b>Thr</b>	<b>Thr</b>	<b>Pro</b>	<b>Lys</b>
	<b>1</b>				<b>5</b>					<b>10</b>					<b>15</b>
5	<210> 21 <211> 11 <212> PRT <213> Secuencia artificial														
10	<220> <223> Extensión 13  <400> 21														
15		<b>Glu</b>	<b>Glu</b>	<b>Leu</b>	<b>Asp</b>	<b>Ala</b>	<b>Arg</b>	<b>Leu</b>	<b>Glu</b>	<b>Ala</b>	<b>Leu</b>	<b>Lys</b>			
		<b>1</b>				<b>5</b>					<b>10</b>				
20	<210> 22 <211> 6 <212> PRT <213> Secuencia artificial														
25	<220> <223> Extensión 14  <400> 22														
30	<b>Gln</b>	<b>Pro</b>	<b>Met</b>	<b>Tyr</b>	<b>Lys</b>	<b>Arg</b>									
	<b>1</b>				<b>5</b>										
35															
40	<210> 23 <211> 6 <212> PRT <213> Secuencia artificial														
45	<220> <223> Extensión 15  <400> 23														
50		<b>Gly</b>	<b>Gln</b>	<b>Pro</b>	<b>Met</b>	<b>Tyr</b>	<b>Lys</b>								
		<b>1</b>				<b>5</b>									
55	<210> 24 <211> 6 <212> PRT <213> Secuencia artificial														
60	<220> <223> Extensión 16  <400> 24														



5		Pro Gly Gln Pro Met Tyr
		1 5
	<210> 25	
	<211> 6	
10	<212> PRT	
	<213> Secuencia artificial	
	<220>	
	<223> Extensión 17	
15		
	<400> 25	
		Lys Pro Gly Gln Pro Met
20		1 5
	<210> 26	
	<211> 6	
25	<212> PRT	
	<213> Secuencia artificial	
	<220>	
	<223> Extensión 18	
30		
	<400> 26	
35		
	Leu Lys Pro Gly Gln Pro	
	1 5	
40	<210> 27	
	<211> 6	
	<212> PRT	
	<213> Secuencia artificial	
45	<220>	
	<223> Extensión 19	
	<400> 27	
50		Gln Leu Lys Pro Gly Gln
		1 5
	<210> 28	
55	<211> 6	
	<212> PRT	
	<213> Secuencia artificial	
	<220>	
60	<223> Extensión 20	
	<400> 28	

	Leu Gln Leu Lys Pro Gly
5	1 5
	<210> 29
	<211> 6
	<212> PRT
10	<213> Secuencia artificial
	<220>
	<223> Extensión 21
15	<400> 29
	Trp Leu Gln Leu Lys Pro
20	1 5
	<210> 30
	<211> 6
	<212> PRT
25	<213> Secuencia artificial
	<220>
	<223> Extensión 22
30	<400> 30
	His Trp Leu Gln Leu Lys
35	1 5
	<210> 31
	<211> 6
	<212> PRT
40	<213> Secuencia artificial
	<220>
	<223> Extensión 23
45	<400> 31
	Trp His Trp Leu Gln Leu
50	1 5
	<210> 32
	<211> 6
	<212> PRT
55	<213> Secuencia artificial
	<220>
	<223> Extensión 24
60	<400> 32
	Ala Trp His Trp Leu Gln
65	1 5
	<210> 33

	<211> 6	
	<212> PRT	
	<213> Secuencia artificial	
5	<220>	
	<223> Extensión 25	
	<400> 33	
10	<210> 34	
		Glu Ala Trp His Trp Leu
		1 5
15	<211> 6	
	<212> PRT	
	<213> Secuencia artificial	
	<220>	
20	<223> Extensión 26	
	<400> 34	
25		Ala Glu Ala Trp His Trp
		1 5
	<210> 35	
	<211> 6	
30	<212> PRT	
	<213> Secuencia artificial	
	<220>	
	<223> Extensión 27	
35		
		Glu Ala Glu Ala Trp His
		1 5
40		
	<210> 36	
	<211> 9	
	<212> PRT	
45	<213> Secuencia artificial	
	<220>	
	<223> Extensión 28	
50	<400> 36	
		Asp Val Lys Pro Gly Gln Pro Leu Ala
		1 5
55	<210> 37	
	<211> 9	
	<212> PRT	
	<213> Secuencia artificial	
60	<220>	
	<223> Extensión 29	
	<400> 37	

5	Asp Val Lys Pro Gly Gln Pro Glu Tyr 1 5
10	<210> 38 <211> 9 <212> PRT <213> Secuencia artificial
15	<220> <223> Extensión 30 <400> 38
20	Asp Val Lys Pro Gly Glu Pro Leu Tyr 1 5
25	<210> 39 <211> 9 <212> PRT <213> Secuencia artificial
30	<220> <223> Extensión 31 <400> 39
35	Asp Val Lys Pro Gly Gln Pro Leu Tyr 1 5
40	<210> 40 <211> 9 <212> PRT <213> Secuencia artificial
45	<220> <223> Extensión 32 <400> 40
50	Asp Val Lys Pro Gly Gln Pro Leu Glu 1 5
55	<210> 41 <211> 9 <212> PRT <213> Secuencia artificial
60	<220> <223> Extensión 33 <400> 41
65	Asp Val Lys Pro Gly Gln Pro Met Tyr 1 5
	<210> 42 <211> 14

<212> PRT  
<213> Secuencia artificial

<220>  
<223> Extensión 34

<400> 42

10      Asp Val Lys Pro Gly Gln Pro Met Tyr Asp Asp Asp Asp Lys  
          1                         5                         10

<210> 43  
<211> 29  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> Análogo de GLP-1(9-37)

<400> 43

Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly Gln Ala  
 1 5 10 15

Ala Lys Glu Phe Ile Ala Trp Leu Val Arg Gly Arg Lys  
20 25

<210> 44  
<211> 45  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> Análogo extendido de GLP-1(7-37)

<400> 44

45      Asp Val Lys Pro Gly Gln Pro Met Tyr Asp Asp Asp Asp Lys His Ala  
          1               5                10                  15

50           Glu Gly Thr Phe Thr Ser Asp Val Ser Ser Tyr Leu Glu Gly Gln Ala  
               20                          25                          30

Ala Lys Glu Phe Ile Ala Trp Leu Val Arg Gly Arg Gly  
35 40 45

## REIVINDICACIONES

1. Un método para la expresión recombinante de un polipéptido que comprende un péptido de GLP-1 en levadura que comprende el cultivo de una cepa de levadura que comprende una secuencia de ADN que codifica una señal de procesamiento y secreción hacia el extremo corriente arriba del polipéptido, en donde dicha señal de procesamiento y secreción comprende una variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I):



en donde

$X_{38}$  es F, L o V,  
 $X_{39}$  L, I, V o M,  
 $X_{40}$  es G o R,  
 $X_{41}$  es S, Y, L, I, V o M, y  
 $X_{42}$  es Y, W, L, V o M;

o en donde

$X_{38}$  es I o V,  
 $X_{39}$  L, I, V o M,  
 $X_{40}$  es A, G, Y, F, W, R, K, L, I, V o M,  
 $X_{41}$  es Y, F, o W, y  
 $X_{42}$  es L o I;

en donde dicho pro-péptido del factor de apareamiento  $\alpha$  son los residuos de aminoácidos 20-85 en el polipéptido:

MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYLDLEGDFDVAVLPPFSNSTNNGLLFINTTIASIAAKE  
 EGVSLDKR (sec. con núm. de ident.:2) y dicha variante comprende de 1 a 15 sustituciones, delecciones o adiciones de aminoácidos respecto de dicho polipéptido.

2. El método de acuerdo con la reivindicación 1 en donde

$X_{38}$  es V;  
 $X_{39}$  L, I, V o M;  
 $X_{40}$  es G, R o K;  
 $X_{41}$  es Y, y  
 $X_{42}$  es L o I.

3. El método de acuerdo con la reivindicación 1 en donde

$X_{38}$  es V;  
 $X_{39}$  L, I, V o M;  
 $X_{40}$  es G o R;  
 $X_{41}$  es Y, y  
 $X_{42}$  es L.

4. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en donde tres de los residuos de aminoácidos en  $X_{38}-X_{39}-X_{40}-X_{41}-X_{42}$  son idénticos al residuo de aminoácido correspondiente en VIGYL.

5. El método de acuerdo con cualquiera de las reivindicaciones 1-4, en donde  $X_{40}$  es R.

6. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en donde  $X_{38}-X_{39}-X_{40}-X_{41}-X_{42}$  es VIRYL.

7. El método de acuerdo con cualquiera de las reivindicaciones 1-6, en donde dicha variante del pro-péptido del factor de apareamiento  $\alpha$  tiene menos de 10 cambios de residuos de aminoácidos fuera de la secuencia  $X_{38}-X_{42}$  en comparación con el pro-péptido del factor de apareamiento  $\alpha$  como se establece en la sec. con núm. de ident.:2 (residuos de aminoácidos 20-85).

8. El método de acuerdo con cualquiera de las reivindicaciones 1-7, en donde dicha levadura porta al menos una modificación genética que reduce su capacidad de O-glicosilación.

9. El método de acuerdo con la reivindicación 8, en donde el gen *PMT1* en dicha levadura se elimina.

10. El método de acuerdo con cualquiera de las reivindicaciones 1-8, en donde dicho polipéptido comprende o consiste en GLP-1(9-37)[K34R] o GLP-1(9-37)[K34R,G37K].

11. El método de acuerdo con la reivindicación 10, en donde dicho polipéptido comprende una extensión N-terminal.
12. Variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I) como se definió en cualquiera de las reivindicaciones anteriores, en donde dicho pro-péptido del factor de apareamiento  $\alpha$  son los residuos de aminoácidos 20-85 en el polipéptido:  
MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYLDLEGDFDVAVLFPFSNSTNNGLLFINTTIASIAAKE  
EGVSLDKR (sec. con núm. de ident.:2) y dicha variante comprende de 1 a 15 sustituciones, deleciones o adiciones de aminoácidos respecto de dicho polipéptido.
13. Precursor de GLP-1 que es un polipéptido de fusión que comprende:
  - Un pre-péptido,
  - Una variante del pro-péptido del factor de apareamiento  $\alpha$  que tiene al menos una sustitución en la secuencia VIGYL en las posiciones 38-42 para comprender la secuencia de aminoácidos de la fórmula general (I) como se definió en cualquiera de las reivindicaciones anteriores,
  - Opcionalmente un péptido de extensión, y
  - Un péptido de GLP-1,  
en donde dicho pro-péptido del factor de apareamiento  $\alpha$  son los residuos de aminoácidos 20-85 en el polipéptido:  
MRFPSIFTAVLFAASSALAAPVNTTTEDETAQIPAEAVIGYLDLEGDFDVAVLFPFSNSTNNGLLFINTTIASIAAKE  
EGVSLDKR (sec. con núm. de ident.:2) y dicha variante comprende de 1 a 15 sustituciones, deleciones o adiciones de aminoácidos respecto de dicho polipéptido.
14. Vector de expresión que comprende una secuencia de ADN que codifica el polipéptido de acuerdo con cualquiera de las reivindicaciones 12-13.
15. Célula huésped que comprende el vector de expresión de acuerdo con la reivindicación

Figura 1

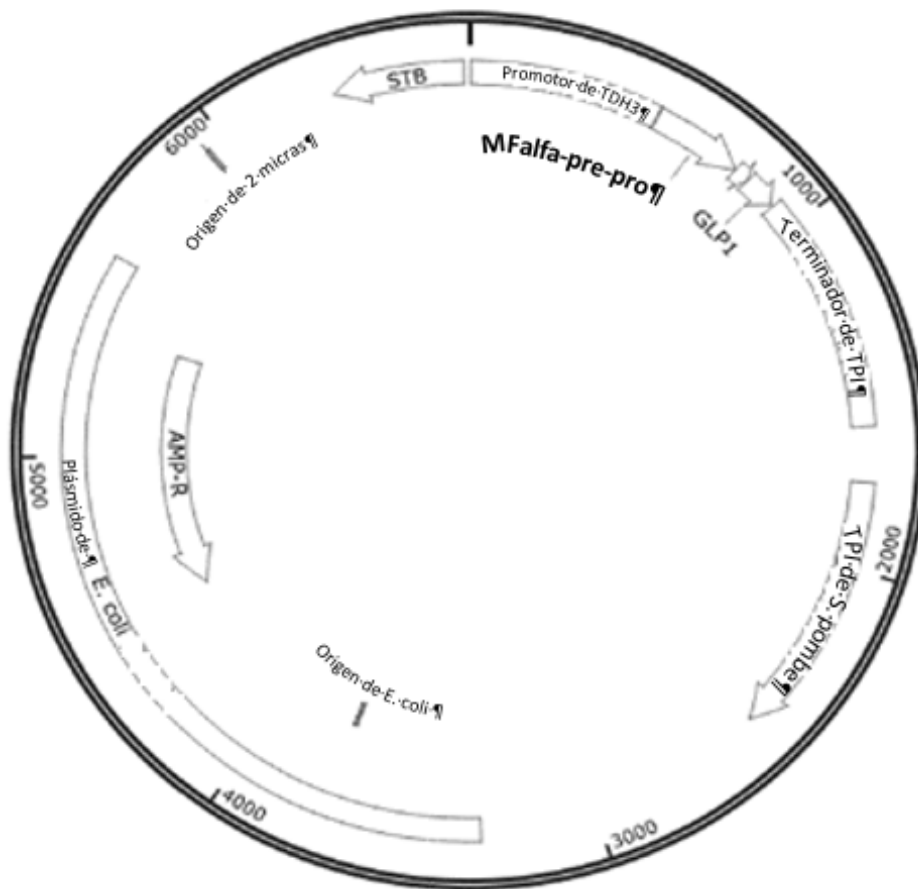




Figura 2

