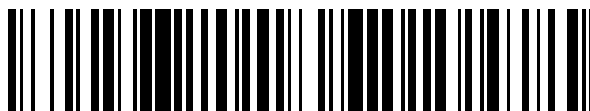


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 528 739**

51 Int. Cl.:

**C12N 1/15** (2006.01)

**C12N 1/19** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **22.12.2004 E 04806613 (8)**

97 Fecha y número de publicación de la concesión europea: **12.11.2014 EP 1696864**

54 Título: **Métodos para eliminar la manosilfosforilación de glucanos en la producción de glucoproteínas**

30 Prioridad:

**24.12.2003 US 532461 P**

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

**12.02.2015**

73 Titular/es:

**GLYCOFI, INC. (100.0%)  
21 LAFAYETTE STREET, SUITE 200  
LEBANON, NEW HAMPSHIRE 03766, US**

72 Inventor/es:

**BOBROWICZ, PIOTR;  
STADHEIM, TERRANCE y  
WILDT, STEFAN**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

**ES 2 528 739 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

Métodos para eliminar la manosilfosforilación de glucanos en la producción de glucoproteínas

5 **Campo de la invención**

La presente invención se refiere a la eliminación de la transferencia de manosilfosfato en glucanos de glucoproteínas, y se refiere además a eliminar genes responsables de la adición de restos de manosilfosfato en glucanos en *Pichia pastoris*.

10 En particular, la invención se refiere a modificar por ingeniería genética células hospedadoras de *P. pastoris* para producir glucanos sin restos de manosilfosfato.

15 **Antecedentes de la invención**

La capacidad para producir proteínas humanas recombinantes ha conducido a avances importantes en el cuidado de la salud humana y sigue siendo un área activa de descubrimiento de fármacos. Muchas proteínas terapéuticas requieren la adición cotraduccional de glucanos a restos de asparagina específicos (*N*-glucosilación) de la proteína para asegurar la actividad de función estructural apropiada y posterior estabilidad en suero humano. Para el uso terapéutico en seres humanos, las glucoproteínas requieren *N*-glucosilación de tipo humano. Las líneas celulares de mamífero (células de ovario de hámster Chino (CHO) así como células retinianas humanas) que pueden imitar el procesamiento de glucoproteína de tipo humano tiene varios inconvenientes incluyendo bajos títulos de proteínas, largos tiempo de fermentación, productos heterogéneos y problemas de contenimiento viral continuados. Por lo tanto, el uso de sistemas de expresión de levadura y hongos filamentosos que tienen procesamiento más económico, menos obstáculos de seguridad y que producen rendimientos de proteína heteróloga más robustos se han investigado exhaustivamente como células hospedadoras para productos terapéuticos humanos.

En levadura y hongos filamentosos, se producen glucoproteínas que tienen oligosacáridos que son diferentes de los de glucoproteínas derivadas de mamífero. Específicamente en levadura, los oligosacáridos de cadena externa son hipermanosilados que consisten en 30-150 restos de manosa (Kukuruzinska *et al.*, 1987, Annu. Rev. Biochem. 56: 915-944). Además, el manosilfosfato se transfiere con frecuencia tanto al núcleo como a las cadenas de azúcares externas de glucoproteínas producidas en levadura (Ballou, 1990, Method Enzymol. 185: 440-470). Es más importante que se ha mostrado que estos glucanos manosilfosforilados de glucoproteínas producidas en la levadura, *Saccharomyces cerevisiae*, inducen una respuesta inmunitaria en conejos (Rosenfeld y Ballou, 1974, JBC, 249: 2319-2321). Por lo tanto, la eliminación de la manosilfosforilación en levadura y hongos filamentosos es esencial para la producción de glucoproteínas terapéuticas no inmunogénicas.

En *S. cerevisiae* hay al menos dos genes que participan en la transferencia de manosilfosfato. Los dos genes, *MNN4* y *MNN6* se han clonado, y los análisis de los productos génicos sugieren que actúan en la transferencia del manosilfosfato (para una revisión véase Jigami y Odani, 1999, Biochim. Biophys. Acta, 1426: 333-345). *MNN6* codifica una proteína de membrana de tipo II homóloga de la familia Kre2p/Mnt1p de proteínas que se ha caracterizado como  $\alpha$ -1,2 manosiltransferasas del Golgi implicadas en la *O*-manosilación y *N*-glucosilación (Lussier *et al.*, 1997, JBC, 272: 15527-15531). El mutante  $\Delta mnn6$  no muestra un defecto en la manosilfosforilación de los glucanos principales *in vivo*, pero muestra una reducción de la actividad manosilfosfato transferasa *in vitro* (Wang *et al.*, 1997, JBC, 272: 18117-18124). *Mnn4p* también es una proteína de membrana de tipo n potencial que es 33 % idéntica a la Yjr061p de *S. cerevisiae* (Odani *et al.*, 1996, Glycobiology, 6: 805-810; Hunter y Plowman, 1997, Trends in Biochem. Sci., 22: 18-22). Los mutantes tanto de  $\Delta mnn6$  como de  $\Delta mnn4$  reducen la transferencia de manosilfosfato. Sin embargo, el mutante doble  $\Delta mnn6 \Delta mnn4$  no reduce adicionalmente esta actividad. Estas observaciones sugieren la presencia de manosiltransferasas adicionales que añaden manosilfosfato a los glucanos principales.

Por lo tanto, a pesar de la reducción de la manosilfosforilación en *S. cerevisiae* con la disrupción de *MNN4*, *MNN6* o ambos en combinación, no hay pruebas de que la eliminación completa de la actividad manosilfosfato transferasa sea posible. Otros genes que afectan a los niveles de manosilfosfato se han identificado en *S. cerevisiae*. Estos genes incluyen *PMR1*, *VRG4*, *MNN2* y *MNN5*. *PMR1* codifica una  $Ca^{2+}/Mn^{2+}$ -ATPasa localizada en el Golgi requerida para la función normal del aparato de Golgi (Antebi y Fink, 1992, Mol. Biol. Cell, 3: 633-654); *Vrg4p* está implicado en el transporte del azúcar nucleotídico en el Golgi (Dean *et al.*, 1997, JBC, 272: 31908-31914), y *Mnn2p* y *Mnn5p* son  $\alpha$ 1,2-manosiltransferasas responsables del inicio de la ramificación en la cadena externa de glucanos ligados a *N* (Rayner y Munro, 1998, JBC, 273: 23836-23843). Para las cuatro proteínas, la reducción en los grupos de manosilfosfato unidos a los glucanos ligados a *N* parece ser una consecuencia de la disfunción del Golgi o una reducción en el tamaño de los glucanos ligados a *N* en lugar de un defecto específico en la actividad de transferencia de los grupos de manosilfosfato. El documento US-A 2002/0137134 describe la provisión de células eucariotas inferiores, en particular *Pichia pastoris*, que tienen una ruta de glucosilación modificada genéticamente que imita el procesamiento de las glucoproteínas en seres humanos. Las modificaciones pueden implicar la eliminación de manosilfosfato transferasas (*MNN4*, *MNN6*, o genes que complementan a los mutantes lbd).

Las proteínas expresadas en la levadura metilotrópica, *Pichia pastoris*, contienen glucanos manosilfosforilados (Miele, *et al.*, 1997, Biotech. Appl Biochem., 2: 79-83). Miura *et al.*, presentaron la identificación del gen *PNO1* (Fosforil manosiación de Oligosacáridos ligados a N) que tras su disrupción confiere una atenuación de la manosilfosforilación en glucoproteínas (documento WO 01/88143; Miura *et al.*, 2004, Gene, 324: 129-137). El gen *PNO1* codifica una proteína implicada en la transferencia de manosilfosfato a glucanos en *P. pastoris*. Su función específica, sin embargo, se desconoce. Como se ha mencionado, el mutante  $\Delta pno1$  reduce pero no anula la manosilfosforilación en N-glucanos en relación con un cepa de *P. pastoris* que tiene Pnolp de tipo silvestre.

En la actualidad, no existen métodos para eliminar la manosilfosforilación en glucoproteínas producidas en hospedadores fúngicos. Una cantidad residual de manosilfosforilación en glucoproteínas aún puede ser inmunogénica y, por lo tanto, es indeseable como productos terapéuticos humanos.

Lo que se necesita, por lo tanto, es un sistema de expresión basado en levadura u hongos filamentosos que produzca glucoproteínas que estén esencialmente sin glucanos manosilfosforilados.

### Sumario de la invención

La presente divulgación proporciona un método para eliminar restos de manosilfosfato en glucanos de glucoproteínas en un hospedador fúngico de *P. pastoris*.

La presente invención proporciona un hospedador fúngico que produce de forma normal glucoproteínas manosilfosforiladas o una fracción de las mismas, modificándose dicho hospedador fúngico para producir glucoproteínas esencialmente sin restos de manosilfosfato. En particular, la presente invención proporciona un hospedador fúngico que comprende una disrupción, una delección o una mutación combinadas de los genes *MNN4B* y *PNO1* en el que el hospedador fúngico es *Pichia pastoris* y es capaz de producir menos del 1 % de glucoproteínas manosilfosforiladas de N-glucanos totales en el que el gen *MNN4B* en el que hay que provocar una disrupción, suprimir o mutar codifica un ARN mensajero que se corresponde con un ADNc que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en

- (a) SEC ID N°: 3;
- (b) una secuencia de ácido nucleico que es una variante degradada de SEC ID N°: 3;
- (c) una secuencia de ácido nucleico al menos 50 %, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 98 %, al menos 99 % o al menos 99,9 % idéntica a SEC ID N°: 3;
- (d) una secuencia de ácido nucleico que codifica un polipéptido al menos 50 %, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 98 %, al menos 99 % o al menos 99,9 % idéntica a SEC ID N°: 4.

Además, la presente invención proporciona un método para producir composiciones de glucoproteínas en el hospedador fúngico de la invención que comprende propagar dicho hospedador fúngico y aislar los productos de glucoproteínas.

La presente divulgación proporciona además composiciones de glucoproteínas que están esencialmente sin glucoproteínas manosilfosforiladas. Dichas composiciones de glucoproteínas comprenden N-glucanos complejos que pueden usarse para aplicaciones terapéuticas.

La presente divulgación también proporciona polinucleótidos aislados que comprenden o consisten en secuencias de ácido nucleico seleccionadas del grupo que consiste en las secuencias codificantes de los *MNN4A*, *MNN4B* y *MNN4C* de *P. pastoris*; secuencias de ácido nucleico que son variantes degradadas de estas secuencias; y secuencias y fragmentos de ácido nucleico relacionados. La divulgación también proporciona polipéptidos aislados que comprenden o consisten en secuencias polipeptídicas seleccionadas del grupo que consiste en secuencias codificadas por los *MNN4A*, *MNN4B*, *MNN4C* de *P. pastoris*; secuencias polipeptídicas relacionadas, fragmentos y fusiones. También se proporcionan anticuerpos que se unen específicamente con los polipéptidos aislados desvelados.

### Breve descripción de los dibujos

La **Figura 1** representa la secuencia de ácido nucleico y aminoácidos de *MNN4A* de *P. pastoris*.

La **Figura 2** representa la secuencia de ácido nucleico y aminoácidos de *MNN4B* de *P. pastoris*.

La **Figura 3** representa la secuencia de ácido nucleico y aminoácidos de *MNN4C* de *P. pastoris*.

La **Figura 4** ilustra la estrategia de anulación por PCR de fusión de *MNN4B* de *P. pastoris* usando un marcador de resistencia a fármacos.

La **Figura 5A** muestra un cromatograma líquido de alto rendimiento para el control experimental negativo usando H<sub>2</sub>O como la muestra. B. muestra un cromatograma líquido de alto rendimiento para la muestra que contiene glucanos ligados a N de K3 purificado del sobrenadante de *P. pastoris* **YSH-44**. Los glucanos con manosilfosfato se eluyen entre 20 y 30 minutos. C. muestra un cromatograma líquido de alto rendimiento para una muestra que

contiene glucanos ligados a *N* de K3 purificado del sobrenadante de *P. pastoris* **YSH-49** ( $\Delta pno1$ ). Los glucanos con manosilfosfato se eluyen entre 20 y 30 minutos. **D.** muestra un cromatograma líquido de alto rendimiento para una muestra que contiene glucanos ligados a *N* de K3 purificado del sobrenadante de *P. pastoris* **YAS-130** ( $\Delta pno1 \Delta mnn4B$ ). Obsérvese la ausencia de glucanos manosilfosforilados entre 20 y 30 minutos.

La **Figura 6A** muestra un cromatograma líquido de alto rendimiento para la muestra que contiene glucanos ligados a *N* de K3 purificado del sobrenadante de *P. pastoris* **YSH-1** ( $\Delta och1$ ). Los glucanos con manosilfosfato se eluyen entre 20 y 30 minutos. **B.** muestra un cromatograma líquido de alto rendimiento para una muestra que contiene glucanos ligados a *N* de K3 purificado del sobrenadante de *P. pastoris* **YAS-164** ( $\Delta och1 \Delta mnn4A \Delta pno1$ ). Los glucanos con manosilfosfato se eluyen entre 20 y 30 minutos. **C.** muestra un cromatograma líquido de alto rendimiento para una muestra que contiene glucanos ligados a *N* de K3 purificado del sobrenadante de *P. pastoris* **YAS-174** ( $\Delta och1 \Delta mnn4A \Delta pno1 \Delta mnn4B$ ). Obsérvese la ausencia de glucanos manosilfosforilados entre 20 y 30 minutos.

La **Figura 7A** muestra un cromatograma líquido de alto rendimiento para la muestra de control experimental negativo que contiene H<sub>2</sub>O. **B.** muestra un cromatograma líquido de alto rendimiento para la muestra que contiene glucanos ligados a *N* de eritropoyetina expresada a partir de pBK291 (His-EPO) producido en la cepa de *P. pastoris* **BK248**. **C.** muestra un cromatograma líquido de alto rendimiento para la muestra que contiene glucanos ligados a *N* de His-EPO producido en la cepa de *P. pastoris* **BK244**. **D.** muestra un cromatograma líquido de alto rendimiento para la muestra que contiene glucanos ligados a *N* de CD40 expresado de pJC33 (His-CD40) producido en la cepa de *P. pastoris* **YJC12**. **E.** muestra un cromatograma líquido de alto rendimiento para la **YAS252**. Nota: los glucanos con manosilfosfato se eluyen entre 20 y 30 minutos.

La **Figura 8** A. muestra un cromatograma líquido de alto rendimiento para la muestra que contiene glucanos ligados a *N* de invertasa expresada de pPB147 producida en la cepa de *P. pastoris* **YAS252**.

La **Figura 9** muestra un alineamiento de homólogos de *MNN4/PNO1* en *P. pastoris* (Pp), *S. cerevisiae* (Sc), *Neurospora crassa* (Nc), *Aspergillus nidulans* (An), *Candida albicans* (Ca) y *Pichia angusta* (*Hansenula polymorpha*) (Pa) usando Clustal W de DNASTar.

#### Descripción detallada de la invención

A no ser que se defina de otro modo en el presente documento, los términos científicos y técnicos usados en relación con la presente invención tendrán los significados que entienden habitualmente los expertos habituales en la materia. Además, a no ser que se requiera de otro modo por el contexto, los términos singulares incluirán el plural y los términos plurales incluirán el singular. En general, las nomenclaturas usadas en relación con, y las técnicas de bioquímica, enzimología, molecular y biología celular, microbiología, genética y química de proteínas y ácidos nucleicos e hibridación descritas en el presente documento son las que se conocen bien y se usan habitualmente en este campo. Los métodos y técnicas de la presente invención se realizan generalmente de acuerdo con métodos convencionales bien conocidos en este campo y como se describe en diversas referencias generales y más específicas que se citan y se analizan a lo largo de la presente memoria descriptiva a no ser que se indique de otro modo. Véase, por ejemplo, Sambrook *et al.* Molecular Cloning. A Laboratory Manual, 2<sup>a</sup> ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1989); Ausubel *et al.*, Current Protocols in Molecular Biology, Green Publishing Associates (1992, y Suplementos hasta 2002); Harlow y Lane, Antibodies: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1990); Taylor y Drickamer, Introduction to Glycobiology, Oxford Univ. Press (2003); Worthington Enzyme Manual, Worthington Biochemical Corp., Freehold, NJ; Handbook of Biochemistry: Section A Proteins, Vol I, CRC Press (1976); Handbook of Biochemistry: Section A Proteins, Vol II, CRC. Press (1976); Essential of Glycobiology, Cold Spring Harbor Laboratory Press (1999).

Se entenderá que los siguientes términos, a no ser que se indique de otro modo, tendrán los siguientes significados:

La expresión "polinucleótido" o "molécula de ácido nucleico" se refiere a una forma polimérica de nucleótidos de al menos 10 bases de longitud. La expresión incluye moléculas de ADN (por ejemplo, ADNc o ADN genómico sintético) y moléculas de ARN (por ejemplo, ARNm o ARN sintético), así como análogos de ADN o ARN que contienen análogos de nucleótidos no naturales, enlaces internucleosídicos no nativos, o ambos. El ácido nucleico puede estar en cualquier conformación topológica. Por ejemplo, el ácido nucleico puede ser monocatenario, bicatenario, tricatenario, cuádruple, parcialmente bicatenario, ramificado, en horquilla, circular o en conformación de candado.

A no ser que se indique de otro modo, un "ácido nucleico que comprende SEC ID N°: X" se refiere a un ácido nucleico, al menos una parte del cual tiene (i) la secuencia de SEC ID N°: X, o (ii) una secuencia complementaria de SEC ID N°: X. La elección entre los dos se dicta por el contexto. Por ejemplo, si el ácido nucleico se usa como una sonda, la elección entre las dos se dicta por el requisito de que la sonda sea complementaria de la diana deseada.

Un ácido nucleico o polinucleótido "aislado" o "sustancialmente puro" (por ejemplo, un ARN, ADN o un polímero mixto) es uno que está sustancialmente separado de otros componentes celulares que acompañan de forma natural al polinucleótido nativo en su célula hospedadora natural, por ejemplo, ribosomas, polimerasas y secuencias genómicas con las que se asocia de forma natural. La expresión abarca un ácido nucleico o polinucleótido que (1) se ha retirado de su ambiente de origen natural, (2) no está asociado con todo o una parte de un polinucleótido en el que el polinucleótido aislado se encuentra en la naturaleza, (3) está unido operativamente con un polinucleótido con

el que no está unido en la naturaleza o (4) no aparece en la naturaleza. La expresión “aislado” o “sustancialmente puro” también puede usarse en referencia a aislados de ADN recombinantes o clonados, análogos polinucleotídicos sintetizados químicamente o análogos polinucleotídicos que se sintetizan biológicamente por sistemas heterólogos.

5 Sin embargo, “aislado” no requiere necesariamente que el ácido nucleico o polinucleótido descrito de este modo se haya retirado físicamente en sí mismo de su ambiente nativo. Por ejemplo, una secuencia de ácido nucleico endógena en el genoma de un organismo se considera “aislada” en el presente documento si se coloca una secuencia heteróloga adyacente a la secuencia de ácido nucleico endógena, de modo que la expresión de esta secuencia de ácido nucleico endógena está alterada. En este contexto, una secuencia heteróloga es una secuencia  
10 que no está adyacente de forma natural a la secuencia de ácido nucleico endógena, independientemente de si la secuencia heteróloga es en sí misma endógena o no (se origina de la misma célula hospedadora o descendencia de la misma) o exógena (se origina de una célula hospedadora diferente o descendencia de la misma). Como ejemplo, una secuencia promotora puede sustituir (por ejemplo, por recombinación homóloga) al promotor nativo de un gen en el genoma de una célula hospedadora, de modo que este gen tenga un patrón de expresión alterado. Este gen se volvería “aislado” porque está separado de al menos algunas de las secuencias que lo flanquean de forma natural.  
15

Un ácido nucleico también se considera “aislado” si contiene cualquier modificación que no aparezca de forma natural en el ácido nucleico correspondiente en un genoma. Por ejemplo, una secuencia codificante endógena se considera “aislada” si contiene una inserción, delección o una mutación puntual introducida de forma artificial, por ejemplo, por la intervención humana. Un “ácido nucleico aislado” también incluye un ácido nucleico integrado en un cromosoma de la célula hospedadora en un sitio heterólogo y una construcción de ácido nucleico presente como un episoma. Además, un “ácido nucleico aislado” puede estar sustancialmente sin otro material celular, o sustancialmente sin medio de cultivo cuando se produce por técnicas recombinantes, o sustancialmente sin precursores químicos u otros productos químicos cuando se sintetiza químicamente.  
20

25 Como se usa en el presente documento, la frase “variante degradada” de una secuencia de ácido nucleico de referencia abarca secuencias de ácido nucleico que pueden traducirse, de acuerdo con el código genético convencional, para proporcionar una secuencia de aminoácidos idéntica a la traducida de la secuencia de ácido nucleico de referencia. La expresión “oligonucleótido degradado” o “cebador degradado” se usa para indicar un oligonucleótido capaz de hibridar con secuencias de ácido nucleico diana que no son necesariamente idénticas en secuencia pero que son homólogas entre sí dentro de uno o más segmentos particulares.  
30

La expresión “porcentaje de identidad de secuencia” o “idéntico” en el contexto de secuencias de ácido nucleico se refiere a los restos en las dos secuencias que son iguales cuando se alinean para máxima correspondencia. La longitud de la comparación de identidad de secuencia puede ser sobre un tramo de al menos aproximadamente nueve nucleótidos, habitualmente al menos aproximadamente 20 nucleótidos, más habitualmente al menos aproximadamente 24 nucleótidos, normalmente al menos aproximadamente 28 nucleótidos, más normalmente al menos aproximadamente 32 nucleótidos y preferentemente al menos aproximadamente 36 o más nucleótidos. Hay varios algoritmos diferentes conocidos en la técnica que pueden usarse para medir la identidad de secuencias de nucleótidos. Por ejemplo, pueden compararse secuencias polinucleotídicas usando FASTA, Gap o Bestfit, que son programas en el Paquete de Wisconsin Versión 10.0, Genetics Computer Group (GCG), Madison, Wisconsin. FASTA proporciona alineamientos y porcentaje de identidad de secuencia de las regiones del mejor solapamiento entre las secuencias de consulta y de búsqueda. Pearson, *Methods Enzymol.* 183: 63-98 (1990). Por ejemplo, el porcentaje de identidad de secuencia entre las secuencias de ácido nucleico puede determinarse usando FASTA con sus parámetros por defecto (un tamaño de palabra de 6 y el factor de NOPAM para la matriz de puntuación) o usando Gap con sus parámetros por defecto como se proporciona en GCG Versión 6.1, incorporado en el presente documento por referencia. Como alternativa, pueden compararse secuencias usando el programa informático, BLAST (Altschul *et al.*, *J. Mol. Biol.* 215: 403-410 (1990); Gish y States, *Nature Genet.* 3: 266-272 (1993); Madden *et al.*, *Meth. Enzymol.* 266: 131-141 (1996); Altschul *et al.*, *Nucleic Acids Res.* 25: 3389-3402 (1997); Zhang y Madden, *Genome Res.* 7: 649-656 (1997)), especialmente blastp o tblastn (Altschul *et al.*, *Nucleic Acids Res.* 25: 3389-3402 (1997)).  
35  
40  
45  
50

La expresión “homología sustancial” o “similitud sustancial”, cuando se hace referencia a un ácido nucleico o fragmento del mismo, indica que, cuando se alinean de forma óptima con inserciones o delecciones de nucleótidos apropiadas con otro ácido nucleico (o su cadena complementaria), hay identidad de secuencia de nucleótidos en al menos aproximadamente el 50 %, más preferentemente el 60 % de las bases nucleotídicas, habitualmente al menos aproximadamente el 70 %, más habitualmente al menos aproximadamente el 80 %, preferentemente, al menos aproximadamente el 90 % y más preferentemente al menos aproximadamente el 95 %, 96 %, 97 %, 98 % o 99 % de las bases nucleotídicas, como se mide por cualquier algoritmo bien conocido de identidad de secuencia, tales como FASTA, BLAST o Gap, como se analizado anteriormente.  
55  
60

65 Como alternativa, existe homología o similitud sustancial cuando un ácido nucleico o fragmento del mismo hibrida con otro ácido nucleico, con una cadena de otro ácido nucleico o con la cadena complementaria del mismo, en condiciones de hibridación rigurosas. Las “condiciones de hibridación rigurosas” y “condiciones de lavado rigurosas” en el contexto de los experimentos de hibridación de ácido nucleico dependen de varios parámetros físicos diferentes. La hibridación de ácido nucleico se verá afectada por condiciones tales como la concentración salina, la

temperatura, los disolventes, la composición básica de las especies que hibridan, la longitud de las regiones complementarias y el número de desapareamientos de bases nucleotídicas entre los ácidos nucleicos que hibridan, como se apreciará fácilmente por los expertos en la materia. Un experto habitual en la materia conoce cómo variar estos parámetros para conseguir una rigurosidad particular de hibridación.

En general, la "hibridación rigurosa" se realiza a aproximadamente 25 °C por debajo del punto de fusión térmico ( $T_m$ ) para el híbrido de ADN específico en un conjunto de condiciones particular. El "lavado riguroso" se realiza a temperaturas aproximadamente 5 °C menores que la  $T_m$  para el híbrido de ADN específico en un conjunto de condiciones particular. La  $T_m$  es la temperatura a la que el 50 % de la secuencia diana hibrida con una sonda perfectamente coincidente. Véase Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1989), página 9.51.

Para fines del presente documento, las "condiciones rigurosas" se definen para hibridación en fase de solución como hibridación acuosa (es decir, sin formamida) en SSC 6X (en la que SSC 20X contiene NaCl 3,0 M y citrato sódico 0,3 M), SDS 1 % a 65 °C durante 8-12 horas, seguido de dos lavados en SSC 0,2X, SDS 0,1 % a 65 °C durante 20 minutos. Se apreciará por el experto en la materia que la hibridación a 65 °C sucederá a diferentes velocidades dependiendo de varios factores incluyendo la longitud y el porcentaje de identidad de las secuencias que se hibriden.

Los ácidos nucleicos (también denominados polinucleótidos) de la presente divulgación pueden incluir cadenas tanto con sentido como antisentido de ARN, ADNc, ADN genómico y formas sintéticas y polímeros mixtos de los anteriores. Pueden modificarse de forma química o bioquímica o pueden contener bases nucleotídicas no naturales o derivatizadas, como se apreciará fácilmente por los expertos en la materia. Dichas modificaciones incluyen, por ejemplo, marcadores, metilación, sustitución de uno o más de los nucleótidos de origen natural con un análogo, modificaciones internucleotídicas tales como enlaces sin carga (por ejemplo, metil fosfonatos, fosfotriésteres, fosforamidatos, carbamatos, etc.), enlaces con carga (por ejemplo, fosforotioatos, fosforoditioatos, etc.), restos pendientes (por ejemplo, polipéptidos), intercaladores (por ejemplo, acridina, psoraleno, etc.), quelantes, alquilantes y enlaces modificados (por ejemplo, ácidos nucleicos alfa anoméricos, etc.). También se incluyen moléculas sintéticas que imitan los polinucleótidos en su capacidad para unirse con una secuencia designada mediante enlaces de hidrógeno y otras interacciones químicas. Dichas moléculas se conocen en la técnica e incluyen, por ejemplo, en las que los enlaces peptídicos sustituyen enlaces fosfato en la cadena principal de la molécula. Otras modificaciones pueden incluir, por ejemplo, análogos en los que el anillo de ribosa contiene un resto de enlace u otra estructura tal como las modificaciones halladas en ácidos nucleicos "bloqueados".

El término "mutado" cuando se aplica a secuencias de ácido nucleico significa que los nucleótidos en una secuencia de ácido nucleico pueden insertarse, suprimirse o cambiarse en comparación con una secuencia de ácido nucleico de referencia. Puede realizarse una única disrupción en un locus (una mutación puntual) o pueden insertarse, suprimirse o cambiarse múltiples nucleótidos en un único locus. Además, pueden realizarse una o más alteraciones en cualquier variedad de loci dentro de una secuencia de ácido nucleico. Una secuencia de ácido nucleico puede mutarse por cualquier método conocido en la técnica incluyendo pero sin limitación técnicas de mutagénesis tales como "PCR propensa a errores" (un proceso para realizar PCR en condiciones en las que la fidelidad de copia de la ADN polimerasa es baja, de modo que se obtenga una alta tasa de mutaciones puntuales a lo largo de la longitud completa del producto de PCR; véase, por ejemplo Leung *et al.*, Technique, 1: 11-15 (1989) y Caldwell y Joyce, PCR Methods Applic. 2: 28-33 (1992)); y "mutagénesis dirigida a oligonucleótidos" (un proceso que permite la generación de mutaciones específicas en cualquier segmento de ADN clonado de interés; véase, por ejemplo, Reidhaar-Olson y Sauer, Science 241: 53-57 (1988)).

Se entiende que el término "vector" como se usa en el presente documento se refiere a una molécula de ácido nucleico capaz de transportar otro ácido nucleico con el que se ha unido. Un tipo de vector es un "plásmido", que se refiere a un bucle de ADN bicatenario circular en el que pueden ligarse segmentos de ADN adicionales. Otros vectores incluyen cósmidos, cromosomas artificiales bacterianos (BAC) y cromosomas artificiales de levadura (YAC). Otro tipo de vector es un vector viral, en el que pueden ligarse segmentos de ADN adicionales en el genoma viral (analizado en más detalle posteriormente). Ciertos vectores tienen capacidad de replicación autónoma en una célula hospedadora en la que se han introducido (por ejemplo, vectores que tienen un origen de replicación que actúa en la célula hospedadora). Otros vectores pueden integrarse en el genoma de una célula hospedadora tras su introducción en la célula hospedadora, y por lo tanto se replican junto con el genoma hospedador. Además, ciertos vectores preferidos son capaces de dirigir la expresión de genes con los que están unidos operativamente. Dichos vectores se denominan en el presente documento "vectores de expresión recombinante" (o simplemente, "vectores de expresión").

La expresión "secuencia marcadora" o "gen marcador" se refiere a la secuencia de ácido nucleico capaz de expresar una actividad que permite la selección positiva o negativa con respecto a la presencia o ausencia de la secuencia dentro de una célula hospedadora. Por ejemplo, el gen *URA5* de *P. pastoris* es un gen marcador porque su presencia puede seleccionarse por la capacidad de las células que contienen el gen para crecer en ausencia de uracilo. Su presencia también puede seleccionarse frente a la incapacidad de las células que contienen el gen para crecer en presencia de 5-FOA. Las secuencias marcadoras o genes no tienen que presentar necesariamente

capacidad de selección positiva y negativa. Los ejemplos no limitantes de secuencias o genes marcadores de *P. pastoris* incluyen *ADE1*, *ARG4*, *HIS4* y *URA3*.

5 Las secuencias de control de la expresión “unidas operativamente” se refieren a un enlace en el que la secuencia de control de la expresión es contigua con el gen de interés para controlar el gen de interés, así como las secuencias de interés de la expresión que actúan en *trans* o a una distancia para controlar el gen de interés.

10 La expresión “secuencia de control de la expresión” como se usa en el presente documento se refiere a secuencias polinucleotídicas que son necesarias para afectar a la expresión de secuencias codificantes con las que están unidas operativamente. Las secuencias de control de la expresión son secuencias que controlan la transcripción, acontecimientos postranscripcionales y traducción de secuencias de ácido nucleico. Las secuencias de control de la expresión incluyen secuencias de inicio de la transcripción, terminación, promotoras y potenciadoras apropiadas; las señales de procesamiento de ARN eficaces tales como señales de corte y empalme y poliadenilación; secuencias que estabilizan el ARNm citoplasmático; secuencias que potencian la eficacia de traducción (por ejemplo, sitios de unión a ribosoma); secuencias que potencian la estabilidad proteica; y cuando se desee secuencias que potencian la secreción de proteínas. La naturaleza de dichas secuencias de control difiere dependiendo del organismo hospedador; en procariontes, dichas secuencias de control generalmente incluyen un promotor, sitio de unión a ribosoma y secuencia de terminación de la transcripción. Se entiende que la expresión “secuencias de control” incluye, al menos, todos los componentes cuya presencia es esencial para la expresión, y también puede incluir componentes adicionales cuya presencia es ventajosa, por ejemplo, secuencias líder y secuencias compañeras de fusión.

25 Se entiende que la expresión “célula hospedadora recombinante” (o simplemente “célula hospedadora”), como se usa en el presente documento, se refiere a una célula en la que se ha introducido un vector recombinante. Debería entenderse que se pretende que dichas expresiones se refieran no solamente a la célula objeto particular sino también a la descendencia de dicha célula. Debido a que pueden producirse ciertas modificaciones en generaciones sucesivas debido a mutación o influencias ambientales, dicha descendencia puede, de hecho, no ser idéntica a la célula parental, pero aún se incluyen dentro del alcance de la expresión “célula hospedadora” como se usa en el presente documento. Una célula hospedadora recombinante puede ser una célula aislada o línea celular que ha crecido en cultivo o puede ser una célula que reside en un tejido u organismo vivo.

35 El término “péptido” como se usa en el presente documento se refiere a un polipéptido corto, por ejemplo, uno que es normalmente de menos de aproximadamente 50 aminoácidos de longitud y más normalmente de menos de aproximadamente 30 aminoácidos de longitud. El término como se usa en el presente documento abarca análogos y miméticos que imitan la función estructural y por lo tanto biológica.

40 El término “polipéptido” abarca proteínas tanto de origen natural como de origen no natural, y fragmentos, mutantes, derivados y análogos de los mismos. Un polipéptido puede ser monomérico o polimérico. Además, un polipéptido puede comprender varios dominios diferentes cada uno de los cuales tiene una o más actividades distintas.

45 La expresión “proteína aislada” o “polipéptido aislado” es una proteína o un polipéptido que en virtud de su origen o fuente de derivación (1) no está asociado con componentes asociados de forma natural que lo acompañan en su estado nativo, (2) existe en una pureza no hallada en la naturaleza, pudiendo valorarse la pureza con respecto a la presencia de otro material celular (por ejemplo, está sin otras proteínas de la misma especie), (3) se expresa por una célula de diferente especie o (4) no aparece en la naturaleza (por ejemplo, es un fragmento de un polipéptido hallado en la naturaleza o incluye análogos de aminoácidos o derivados no hallados en la naturaleza o enlaces distintos de los enlaces peptídicos convencionales). Por lo tanto, un polipéptido que se sintetiza químicamente o se sintetiza en un sistema celular diferente de la célula de la que se origina de forma natural estará “aislado” de sus componentes asociados de forma natural. Puede hacerse que un polipéptido o proteína esté sustancialmente sin componentes asociados de forma natural por aislamiento, usando técnicas de purificación de proteínas bien conocidas en este campo. Como se ha definido de este modo, “aislado” no requiere necesariamente que la proteína, polipéptido, péptido u oligopéptido descrito de este modo se haya retirado físicamente de su ambiente nativo.

55 La expresión “fragmento polipeptídico” como se usa en el presente documento se refiere a un polipéptido que tiene una deleción, por ejemplo, una deleción amino terminal y/o carboxilo terminal en comparación con un polipéptido de longitud completa. En una realización preferida, el fragmento polipeptídico es una secuencia contigua en la que la secuencia de aminoácidos del fragmento es idéntica a las posiciones correspondientes en la secuencia de origen natural. Los fragmentos son normalmente de al menos 5, 6, 7, 8, 9 o 10 aminoácidos de longitud, preferentemente de al menos 12, 14, 16 o 18 aminoácidos de longitud, más preferentemente al menos 20 aminoácidos de longitud, más preferentemente al menos 25, 30, 35, 40 o 45 aminoácidos, y aún más preferentemente al menos 50 o 60 aminoácidos de longitud, y aún más preferentemente al menos 70 aminoácidos de longitud.

65 Un “derivado modificado” se refiere a polipéptidos o fragmentos de los mismos que son sustancialmente homólogos en secuencia estructural primaria pero que incluyen, por ejemplo, modificaciones químicas y bioquímicas *in vivo* o *in vitro* o que incorporan aminoácidos que no se encuentran en el polipéptido nativo. Dichas modificaciones incluyen, por ejemplo, acetilación, carboxilación, fosforilación, glucosilación, ubiquitinización, marcaje, por ejemplo, con

radionúclidos, y diversas modificaciones enzimáticas, como se apreciará fácilmente por los expertos en la materia. Se conoce bien en la técnica diversos métodos para marcar polipéptidos y de sustituyentes o marcadores útiles para dichos fines, e incluyen isótopos radiactivos tales como  $^{125}\text{I}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$  y  $^3\text{H}$ , ligandos que se unen con antiligandos marcados (por ejemplo, anticuerpos), fluoróforos, agentes quimioluminiscentes, enzimas y antiligandos que pueden actuar como miembros de un par de unión específico para un ligando marcado. La elección del marcador depende de la sensibilidad requerida, la facilidad de conjugación con el cebador, los requisitos de estabilidad y la instrumentación disponible. Se conocen bien en la técnica métodos para marcar polipéptidos. Véase, por ejemplo, Ausubel *et al.*, *Current Protocols in Molecular Biology*, Greene Publishing Associates (1992, y Suplementos hasta 2002).

La expresión "proteína de fusión" se refiere a un polipéptido que comprende un polipéptido o fragmento acoplado con secuencias de aminoácidos heterólogas. Las proteínas de fusión son útiles porque pueden construirse de modo que contengan dos o más elementos funcionales deseados de dos o más proteínas diferentes. Una proteína de fusión comprende al menos 10 aminoácidos contiguos de un polipéptido de interés, más preferentemente al menos 20 o 30 aminoácidos, aún más preferentemente al menos 40, 50 o 60 aminoácidos, aún más preferentemente al menos 75, 100 o 125 aminoácidos. Las fusiones que incluyen la totalidad de las proteínas de la presente invención tienen utilidad particular. El polipéptido heterólogo incluido dentro de la proteína de fusión de la presente invención es de al menos 6 aminoácidos de longitud, con frecuencia al menos 8 aminoácidos de longitud, y de forma útil al menos 15, 20 y 25 aminoácidos de longitud. Las fusiones que incluyen polipéptidos mayores, tales como una región Fc IgG e incluso proteínas completas, tales como la proteína verde fluorescente ("GFP"), proteínas que contienen cromóforos, tienen utilidad particular. Las proteínas de fusión pueden producirse de forma recombinante restringiendo una secuencia de ácido nucleico que codifica el polipéptido o fragmento del mismo en fase con una secuencia de ácido nucleico que codifica una proteína o péptido diferente y expresando después la proteína de fusión. Como alternativa, puede producirse una proteína de fusión de forma química reticulando el polipéptido o un fragmento del mismo con otra proteína.

La expresión "análogos no peptídicos" se refiere a un compuesto con propiedades que son análogas a las de un polipéptido de referencia. Un compuesto no peptídico también puede denominarse un "mimético peptídico" o un "peptidomimético". Véase, por ejemplo, Jones, *Amino Acid and Peptide Synthesis*, Oxford University Press (1992); Jung, *Combinatorial Peptide and Nonpeptide Libraries: A Handbook*, John Wiley (1997); Bodanszky *et al.*, *Peptide Chemistry-A Practical Textbook*, Springer Verlag (1993); *Synthetic Peptides: A Users Guide*, (Grant, ed., W. H. Freeman and Co., 1992); Evans *et al.*, *J. Med. Chem.* 30: 1229 (1987); Fauchere, *J. Adv. Drug Res.* 15: 29 (1986); Veber y Freidinger, *Trends Neurosci.*, 8: 392-396 (1985); y referencias citadas en cada una de las anteriores.

Dichos compuestos se desarrollan con frecuencia con la ayuda de modelos moleculares computarizados. Los miméticos peptídicos que son estructuralmente similares a péptidos útiles de la invención pueden usarse para producir un efecto equivalente y se prevé por lo tanto que sean parte de la invención.

Un "mutante polipeptídico" o una "muteína" se refieren a un polipéptido cuya secuencia contiene una inserción, duplicación, delección, reordenamiento o sustitución de uno o más aminoácidos en comparación con la secuencia de aminoácidos de una proteína nativa o de tipo silvestre. Una muteína puede tener una o más sustituciones puntuales de aminoácidos, en las que se ha cambiado un único aminoácido en una posición por otro aminoácido, una o más inserciones y/o delecciones, en las que se han insertado u suprimido uno o más aminoácidos, respectivamente, en la secuencia de la proteína de origen natural y/o truncamientos de la secuencia de aminoácidos en uno o ambos de los extremos carboxilo o amino terminales. Una muteína puede tener la misma actividad biológica pero preferentemente tiene una diferente en comparación con la proteína de origen natural.

Una muteína tiene al menos 50 % de homología de secuencia general con su homólogo de tipo silvestre. Se prefieren aún más muteínas que tengan al menos 70 %, 75 %, 80 %, 85 % o 90 % de homología de secuencia general con la proteína de tipo silvestre. En una realización aún más preferida, una muteína muestra al menos 95 % de identidad de secuencia, aún más preferentemente 98 %, aún más preferentemente 99 % y aún más preferentemente 99,9 % de identidad de secuencia general. La homología de secuencia puede medirse por cualquier algoritmo de análisis de secuencia habitual, tal como Gap o Bestfit.

Las sustituciones de aminoácidos pueden incluir las que: (1) reducen la susceptibilidad a proteólisis, (2) reducen la susceptibilidad a oxidación, (3) alteran la afinidad de unión para formar complejos de proteínas, (4) alteran la afinidad de unión o actividad enzimática, y (5) confieren o modifican otras propiedades fisicoquímicas o funcionales de dichos análogos.

Como se usa en el presente documento, los veinte aminoácidos convencionales y sus abreviaturas siguen el uso convencional. Véase *Immunology - A Synthesis* (Golub y Gren eds., Sinauer Associates, Sunderland, Mass., 2ª ed. 1991).

Los estereoisómeros (por ejemplo, aminoácidos D) de los veinte aminoácidos convencionales, aminoácidos no naturales tales como aminoácidos  $\alpha$  disustituidos, aminoácidos N-alquilo, y otros aminoácidos no convencionales también pueden ser componentes adecuados para polipéptidos de la presente invención. Los ejemplos de



aminoácidos no convencionales incluyen: 4-hidroxiprolina,  $\gamma$ -carboxiglutamato,  $\epsilon$ -*N,N,N*-trimetilisina,  $\epsilon$ -*N*-acetilisina, *O*-fosfoserina, *N*-acetilserina, *N*-formilmetionina, 3-metilhistidina, 5-hidroxilisina, *N*-metilarginina y otros aminoácidos e iminoácidos similares (por ejemplo, 4-hidroxiprolina). En la anotación polipeptídica usada en el presente documento, el extremo izquierdo corresponde al extremo amino terminal y el extremo derecho corresponde al extremo carboxilo terminal, de acuerdo con el uso convencional y la convención.

Una proteína tiene "homología" o es "homóloga" de una segunda proteína si la secuencia de ácido nucleico que codifica la proteína tiene una secuencia similar a la secuencia de ácido nucleico que codifica la segunda proteína. Como alternativa, una proteína tiene homología con una segunda proteína si las dos proteínas tienen secuencias de aminoácidos "similares". (Por lo tanto, se define que la expresión "proteínas homólogas" significa que las dos proteínas tienen secuencias de aminoácidos similares). En una realización preferida, una proteína homóloga es una que muestra al menos 65 % de homología de secuencia con la proteína de tipo silvestre, más preferentemente al menos 70 % de homología de secuencia. Se prefieren aún más proteínas homólogas que muestran al menos 75 %, 80 %, 85 % o 90 % de homología de secuencia con la proteína de tipo silvestre. En una realización aún más preferida, una proteína homóloga muestra al menos 95 %, 98 %, 99 % o 99,9 % de identidad de secuencia. Como se usa en el presente documento, se interpreta que la homología entre dos regiones de secuencias de aminoácidos (especialmente con respecto a similitudes estructurales predichas) implica similitud de función.

Cuando se usa "homólogo" en referencia a proteínas o péptidos, se reconoce que las posiciones de restos que no son idénticas con frecuencia difieren en sustituciones de aminoácidos conservativas. Una "sustitución de aminoácidos conservativa" es una en la que un resto de aminoácido se sustituye por otro resto de aminoácido que tiene una cadena lateral (grupo R) con propiedades químicas similares (por ejemplo, carga o hidrofobicidad). En general, una sustitución de aminoácidos conservativa no cambiará sustancialmente las propiedades funcionales de una proteína. En casos en los que dos o más secuencias de aminoácidos difieran entre sí por sustituciones conservativas, el porcentaje de identidad de secuencia o grado de homología puede ajustarse hacia arriba para corregir la naturaleza conservativa de la sustitución. Los expertos en la materia conocen bien medios para realizar este ajuste. Véase, por ejemplo, Pearson, 1994, *Methods Mol. Biol.* 24: 307-31 y 25: 365-89.

Los siguientes seis grupos contienen cada uno aminoácidos que son sustituciones conservativas entre sí: 1) Serina (S), Treonina (T); 2) Ácido Aspártico (D), Ácido Glutámico (E); 3) Asparagina (N), Glutamina (Q); 4) Arginina (R), Lisina (K); 5) Isoleucina (I), Leucina (L), Metionina (M), Alanina (A), Valina (V) y 6) Fenilalanina (F), Tirosina (Y), Triptófano (W).

La homología de secuencia para polipéptidos, que también se denomina porcentaje de identidad de secuencia, se mide normalmente usando software de análisis de secuencia. Véase, por ejemplo, el Paquete de Software de Análisis de Secuencia del Genetics Computer Group (GCG), Centro Biotecnológico de la Universidad de Wisconsin, 910 University Avenue, Madison, Wisconsin 53705. El software de análisis de proteínas empareja secuencias similares usando una medida de homología asignada a diversas sustituciones, deleciones y otras modificaciones, incluyendo sustituciones de aminoácidos conservativas. Por ejemplo, GCG contiene programas tales como "Gap" y "Bestfit" que pueden usarse con parámetros por defecto para determinar la homología de secuencia o identidad de secuencia entre polipéptidos estrechamente relacionados, tales como polipéptidos homólogos de diferentes especies de organismos o entre una proteína de tipo silvestre y una mutante de la misma. Véase, por ejemplo, GCG Versión 6.1.

Un algoritmo preferido cuando se compara una secuencia polipeptídica particular con una base de datos que contiene un gran número de secuencias de diferentes organismos es el programa informático BLAST (Altschul *et al.*, *J. Mol. Biol.* 215: 403-410 (1990); Gish y States, *Nature Genet.* 3: 266-272 (1993); Madden *et al.*, *Meth. Enzymol.* 266: 131-141 (1996); Altschul *et al.*, *Nucleic Acids Res.* 25: 3389-3402 (1997); Zhang y Madden, *Genome Res.* 7: 649-656 (1997)), especialmente blastp o tblastn (Altschul *et al.*, *Nucleic Acids Res.* 25: 3389-3402 (1997)).

Son parámetros preferidos para BLASTp:

Valor de expectativa: 10 (defecto); Filtro: segundos (defecto); Coste para abrir un hueco: 11 (defecto); Coste para extender un hueco: 1 (defecto); Alineamientos máximos: 100 (defecto); Tamaño de palabra: 11 (defecto); N° de descripciones: 100 (defecto); Matriz de Penalización: BLOWSUM62.

La longitud de las secuencias polipeptídicas comparadas con respecto a homología generalmente será de al menos aproximadamente 16 restos de aminoácidos, habitualmente al menos aproximadamente 20 restos, más habitualmente al menos aproximadamente 24 restos, normalmente al menos aproximadamente 28 restos y preferentemente más de aproximadamente 35 restos. Cuando se busca en una base de datos que contiene secuencias de un gran número de organismos diferentes, es preferible comparar secuencias de aminoácidos. La búsqueda en base de datos usando secuencias de aminoácidos puede medirse por algoritmos distintos de blastp conocidos en la técnica. Por ejemplo, las secuencias polipeptídicas pueden compararse usando FASTA, un programa en GCG Versión 6.1. FASTA proporciona alineamientos y el porcentaje de identidad de secuencia de las regiones del mejor solapamiento entre las secuencias de consulta y de búsqueda. Pearson, *Methods Enzymol.* 183: 63-98 (1990).

Por ejemplo, el porcentaje de identidad de secuencia entre las secuencias de aminoácidos puede determinarse usando FASTA con sus parámetros por defecto (un tamaño de palabra de 2 y la matriz de puntuación PAM250), como se proporciona en GCG Versión 6.1.

5 El término “región” como se usa en el presente documento se refiere a una parte físicamente contigua de la estructura primaria de una biomolécula. En el caso de proteínas, una región se define por una parte contigua de la secuencia de aminoácidos de esa proteína.

10 El término “dominio” como se usa en el presente documento se refiere a una estructura de una biomolécula que contribuye a una función conocida o sospechada de la biomolécula. Los dominios pueden ser coextensivos con regiones o partes de los mismos; los dominios también pueden incluir regiones distintas, no contiguas de una biomolécula. Los ejemplos de dominios proteicos incluyen, pero sin limitación, un dominio Ig, un dominio extracelular, un dominio transmembrana y un dominio citoplasmático.

15 Como se usa en el presente documento, el término “molécula” significa cualquier compuesto, incluyendo, pero sin limitación, una molécula pequeña, péptido, proteína, azúcar, nucleótido, ácido nucleico, lípido, etc., y dicho compuesto puede ser natural o sintético.

20 El término “eliminación” como se usa con respecto a manosilfosforilación se refiere a niveles de detección de glucanos manosilfosforilados que no indican restos de manosilfosfato detectables aparentes usando HPLC con el ajuste indicado.

25 A no ser que se defina de otro modo, todas las expresiones técnicas y científicas usadas en el presente documento tienen el mismo significado que se entiende habitualmente por un experto habitual en la materia a la que pertenece la presente invención. Se describen posteriormente métodos y materiales ejemplares.

En caso de conflicto, la presente memoria descriptiva, incluyendo definiciones, prevalecerá. Los materiales, métodos y ejemplos son solamente ilustrativos y no se pretende que sean limitantes.

30 Métodos para producir una cepa hospedadora fúngica que carece de manosilfosforilación en glucoproteínas

35 La presente divulgación proporciona métodos para eliminar la transferencia de manosilfosfato en glucanos de glucoproteínas en levadura o células hospedadoras fúngicas filamentosas que normalmente producen glucoproteínas que tienen manosilfosforilación. En una realización, la levadura o célula hospedadora fúngica filamentosas que normalmente produce glucoproteínas que tienen manosilfosforilación se modifica por ingeniería genética de modo que esté esencialmente sin manosilfosforilación en glucanos de glucoproteínas. En otra realización, los hospedadores fúngicos se modifican genéticamente para tener alterado, atenuado o mutado al menos un gen que codifica una proteína que participa en la manosilfosfato transferasa. Preferentemente, el método implica disrupción, atenuación o mutación de uno o más genes seleccionados de *MNN4A*, *MNN4B*, *MNN4C* y *PNO1*.

40 Usando genes conocidos que codifican manosilfosfato transferasas, se aislaron nuevos genes que codificaban manosilfosfato transferasa en *P. pastoris*. La secuencia génica de *MNN4* de *S. cerevisiae* (referencia de Genbank N° P36044) se sometió a blast frente al genoma de *P. pastoris* (Integrated Genomics, Chicago, IL). Esta búsqueda dio como resultado la identificación de tres ORF previamente desconocidas además del gen *PNO1*. Las tres ORF se designaron *MNN4A* (**SEC ID N°: 1**), *MNN4B* (**SEC ID N°: 3**) y *MNN4C* (**SEC ID N°: 5**). Estas ORF se amplificaron y posteriormente se secuenciaron y se muestran respectivamente en las **Figuras 1-3 (Ejemplo 1)**. Las secuencias de aminoácidos codificadas para *MNN4A* (**SEC ID N°: 2**), *MNN4B* (**SEC ID N°: 4**), *MNN4C* (**SEC ID N°: 6**) también se exponen en las **Figuras 1-3**.

50 Secuencias de ácido nucleico

55 En un aspecto, la presente divulgación proporciona una molécula de ácido nucleico que comprende o consiste en una secuencia que es una variante del gen *MNN4 A* de *P. pastoris* que tiene al menos 50 % de identidad con **SEC ID N°: 1**. La secuencia de ácido nucleico puede tener preferentemente al menos 65 %, 70 %, 75 % u 80 % de identidad con el gen de tipo silvestre. Aún más preferentemente, la secuencia de ácido nucleico puede tener 85 %, 90 %, 95 %, 98 %, 99 %, 99,9 % o aún mayor identidad con la **SEC ID N°: 1**. La presente divulgación también proporciona un polipéptido que comprende o consiste en una secuencia que es una variante del gen *MNN4A* de *P. pastoris* que tiene al menos 50 % de identidad con **SEC ID N°: 2**. La secuencia de aminoácidos puede tener preferentemente al menos 65 %, 70 %, 75 % u 80 % de identidad con el gen de tipo silvestre. Aún más preferentemente, la secuencia de aminoácidos puede tener 85 %, 90 %, 95 %, 98 %, 99 %, 99,9 % o aún mayor identidad con la **SEC ID N°: 2**.

65 En relación con la presente invención, el gen *MNN4B* de *P. pastoris* es particularmente útil en la eliminación de la transferencia de manosilfosfato en glucanos de glucoproteínas en una cepa de levadura. La presente divulgación proporciona una molécula de ácido nucleico que comprende o consiste en una secuencia que es una variante del

- gen *MNN4 B* de *P. pastoris* que tiene al menos 50 % de identidad con **SEC ID Nº: 3**. La secuencia de ácido nucleico puede tener preferentemente al menos 65 %, 70 %, 75 % u 80 % de identidad con el gen de tipo silvestre. Aún más preferentemente, la secuencia de ácido nucleico puede tener 85 %, 90 %, 95 %, 98 %, 99 %, 99,9 % o aún mayor identidad con la **SEC ID Nº: 3**. La presente divulgación también proporciona un polipéptido que comprende o
- 5 consiste en una secuencia que es una variante del gen *MNN4B* de *P. pastoris* que tiene al menos 50 % de identidad con **SEC ID Nº: 4**. La secuencia de aminoácidos puede tener preferentemente al menos 65 %, 70 %, 75 % u 80 % de identidad con el gen de tipo silvestre. Aún más preferentemente, la secuencia de aminoácidos puede tener 85 %, 90 %, 95 %, 98 %, 99 %, 99,9 % o aún mayor identidad con la **SEC ID Nº: 4**.
- 10 La presente divulgación proporciona una molécula de ácido nucleico que comprende o consiste en una secuencia que es una variante del gen *MNN4 C* de *P. pastoris* que tiene al menos 50 % de identidad con **SEC ID Nº: 5**. La secuencia de ácido nucleico puede tener preferentemente al menos 65 %, 70 %, 75 % u 80 % de identidad con el gen de tipo silvestre. Aún más preferentemente, la secuencia de ácido nucleico puede tener 85 %, 90 %, 95 %, 98 %, 99 %, 99,9 % o aún mayor identidad con la **SEC ID Nº: 5**. La presente divulgación también proporciona un
- 15 polipéptido que comprende o consiste en una secuencia que es una variante del gen *MNN4C* de *P. pastoris* que tiene al menos 50 % de identidad con **SEC ID Nº: 6**. La secuencia de aminoácidos puede tener preferentemente al menos 65 %, 70 %, 75 % u 80 % de identidad con el gen de tipo silvestre. Aún más preferentemente, la secuencia de aminoácidos puede tener 85 %, 90 %, 95 %, 98 %, 99 %, 99,9 % o aún mayor identidad con la **SEC ID Nº: 6**.
- 20 También se proporcionan vectores, incluyendo vectores de expresión y vectores de anulación que comprenden las moléculas de ácido nucleico anteriores. Un vector de anulación que comprende un *MNN4A*, *MNN4B* o *MNN4C* puede usarse para alterar el locus del gen *MNN4A*, *MNN4B* o *MNN4C*. Como alternativa, un vector de integración que comprende un marcador de resistencia a fármacos o un marcador auxotrófico se usa para alterar el locus del gen *MNN4*.
- 25 Combinación de anulaciones génicas de manosilfosforilación
- Cada uno de los tres genes de *P. pastoris* de nueva identificación, *MNN4A*, *MNN4B*, *MNN4C*, se altera usando la estrategia de solapamiento de PCR como se muestra en la **Figura 4** para determinar el efecto en la
- 30 manosilfosforilación. Los mutantes individuales,  $\Delta mnn4A$ ,  $\Delta mnn4B$  y  $\Delta mnn4C$  no mostraron una reducción significativa en la actividad de transferencia de manosilfosforilación en glucanos del dominio kringle 3 de la proteína plasminógeno humana (K3), mientras que el mutante  $\Delta pno1$  (YSH-49) presentó solamente una atenuación en la transferencia de manosilfosforilación reducida al 6 % (**Figura 5C**), pero no a los niveles descritos previamente en Miura *et al.* (documento WO 01/88143). Se ha postulado que diferentes glucoproteínas pueden presentar diversos
- 35 grados y tipos de glucosilación en la misma célula hospedadora (Montesino *et al.*, 1998, Prot. Expr. Purif. 14: 197-207). En una realización de la presente invención, se construyeron combinaciones de mutantes nulos, uno de los cuales, los mutantes dobles  $\Delta pno \Delta mnn4b$  en *P. pastoris* dio como resultado niveles indetectables de manosilfosforilación en glucanos de la proteína indicadora K3 (**Figura 5D**). De forma similar, otras glucoproteínas (por ejemplo, CD40 e invertasa) producidas a partir del mutante doble  $\Delta pno1 \Delta mnn4b$  en *P. pastoris* también dieron
- 40 como resultado falta de manosilfosforilación. El mutante doble, por lo tanto, produce diversas glucoproteínas de interés que no tienen manosilfosforilación en glucanos. En consecuencia, se proporciona un método para alterar una combinación de genes implicados en la transferencia de restos de manosilfosfato en glucanos de glucoproteínas en un hospedador (por ejemplo, *Pichia* sp.). Preferentemente, la combinación incluye la disrupción de *MNN4B* y *PNO1*.
- 45 En caso de que la disrupción del locus de *MNN4B* de *P. pastoris* solamente no confiera eliminación de manosilfosforilación en glucanos, se altera una combinación de genes de manosilfosforilación. La disrupción del locus *MNN4B* puede ser en combinación con al menos un segundo gen implicado en la transferencia de manosilfosfato, tal como *MNN4A*, *MNN4B*, *MNN4C* o *PNO1*. De acuerdo con la presente invención, el segundo gen es el gen *PNO1* de *P. pastoris* (referencia de Genbank Nº BD105434). Se contempla que un experto en la materia puede alterar o mutar cualquier gen implicado en la síntesis de oligosacáridos o un fragmento del mismo en
- 50 combinación con un *MNN4B* alterado o mutado, que daría como resultado la eliminación de la transferencia de manosilfosfato a glucanos en otros hospedadores fúngicos.
- En otra realización, el método posibilita la disrupción de un gen que codifica *MNN4B* (**SEC ID Nº: 3**) en un
- 55 hospedador (por ejemplo, *P. pastoris*) que ya tiene actividad manosilfosfato transferasa atenuada. Adicionalmente, se contempla que la eliminación de la transferencia de manosilfosfato a glucanos en otras especies de *Pichia* implica la disrupción o mutación de cualquier combinación de genes que tengan homología con *MNN4A*, *MNN4B*, *MNN4C* o *PNO1*.
- 60 En otro aspecto más de la invención cada uno de los tres genes de *P. pastoris* de nueva identificación, *MNN4A*, *MNN4B*, *MNN4C*, se alteró usando una estrategia de supresión de fusión como se describe en el **Ejemplo 3** para determinar si cualquier combinación de anulaciones de genes tiene un efecto en la manosilfosforilación de glucoproteínas expresadas en esta cepa mutante. Los mutantes  $\Delta mnn4A$ ,  $\Delta mnn4B$  y  $\Delta mnn4C$  individuales como con la estrategia de anulación de solapamiento de PCR (**Figura 4**) no mostraron una reducción de la actividad de
- 65 transferencia de manosilfosforilación en glucanos del dominio kringle 3 de proteína de plasminógeno humano (K3)

(datos no mostrados). Sin embargo, la proteína indicadora K3 expresada en un mutante nulo doble  $\Delta pno1 \Delta mnn4b$  (YAS174) está esencialmente sin ninguna manosilfosforilación (**Figura 6C**, compárese con **Figura 6A, B**). Obsérvese la ausencia de glucanos manosilfosforilados entre 20 y 30 minutos.

#### 5 Sistema de expresión de glucoproteínas heterólogas

Usando técnicas establecidas para expresar glucoproteínas heterólogas en levadura y hongos filamentosos, se expresa un gen que codifica una glucoproteína terapéutica. Un sistema de expresión de proteínas recombinantes fúngico puede incluir normalmente promotores tales como AOX1, AOX2, u otros promotores inducibles, terminadores de la transcripción tales como CYC, marcadores seleccionables tales como URA3, URA5, G418, ADE1, ARG4, HIS4, Zeocina y señales de secreción tales como  $\alpha MF$  de *S. cerevisiae*. En una realización, este sistema de expresión se modifica para ser al menos un mutante de *mnn4B*. De acuerdo con la invención, las glucoproteínas se producen en *P. pastoris* que tiene una disrupción, delección o mutación combinada de los genes de *MNN4B* y *PNO1* como se define en las reivindicaciones. Pueden producirse glucoproteínas de interés por cualquier medio mediante el uso de los métodos desvelados en el presente documento. Puede proporcionarse producción de glucoproteínas por cualquier medio en una célula hospedadora, incluyendo acumulación en un compartimento intracelular o secreción de la célula a un sobrenadante de cultivo. Las células hospedadoras de la presente invención pueden propagarse o cultivarse por cualquier método conocido o contemplado en la técnica, incluyendo pero sin limitación tubos de cultivo, matraces, frascos rotatorios, matraces de agitación o fermentadores. Puede realizarse aislamiento y/o purificación de los productos de glucoproteínas por cualquier medio conocido o contemplado en la técnica tal como fraccionamiento, intercambio iónico, filtración en gel, cromatografía hidrófoba y cromatografía de afinidad. Se desvela un ejemplo de la producción y purificación de las glucoproteínas en el **Ejemplo 7**.

Las glucoproteínas expresadas sin glucanos manosilfosforilados usando los métodos descritos en el presente documento pueden incluir pero sin limitación: eritropoyetina, citocinas tales como interferón  $\alpha$ , interferón  $\beta$ , interferón  $\gamma$ , interferón  $\omega$ , TNF- $\alpha$ , CSF de granulocitos, GM-CSF, interleucinas tales como IL-1 $\alpha$ , factores de coagulación tales como factor VIII, factor IX, proteína C humana, antitrombina III y anticuerpos de trombotocina; IgG, IgA, IgD, IgE, IgM y fragmentos de los mismos, regiones Fc y Fab, cadena a del receptor de IgE soluble, uroquinasa, quimasa e inhibidor de tripsina de urea, proteína de unión a IGF, factor de crecimiento epidérmico, factor de liberación de hormonas del crecimiento, FSH, proteína de fusión de anexina V, angiostatina, factor de crecimiento endotelial vascular 2, factor inhibidor progenitor mielóide 1, osteoprotegerina, antitripsina  $\alpha$  1, DNasa II,  $\alpha$ -feto proteínas y glucocerebrosidasa.

#### 35 Producción de glucoproteínas complejas sin manosilfosforilación

En otro aspecto de la invención, la presente divulgación proporciona métodos para producir glucanos ligados a N complejos en hongos y levadura (por ejemplo, *P. pastoris*) que comprende eliminar la transferencia de manosilfosfato a glucanos en glucoproteínas. Dicho método proporciona una composición de glucoproteína que está esencialmente sin restos manosilfosfato en glucoproteínas. La célula hospedadora de la invención proporciona menos del 1 % de glucoproteínas manosilfosforiladas de N-glucanos totales. En una realización más preferida, la invención proporciona menos del 0,5 % de glucoproteínas manosilfosforiladas de N-glucanos totales.

En otro aspecto de la presente invención, las composiciones de glucoproteína están esencialmente sin restos de manosilfosfato en N-glucanos complejos. El método para producir dichos glucanos implica alterar los genes *PNO1* y *MNN4B* en una cepa hospedadora que expresa N-glucanos complejos (por ejemplo, *P. pastoris* YSH-44 que expresa la proteína indicadora K3) (Hamilton *et al.*, 2003, Science, 301: 1244-124). La cepa modificada por ingeniería genética que comprende alteraciones de *pno1 mnn4B*, designada YAS-130, carece de restos de manosilfosfato en glucanos de glucoproteínas (**Ejemplo 5**). Aunque una disrupción genética del gen *PNO1* en YSH-44 (designada YSH-49) reduce el % molar de glucanos que muestra manosilfosforilación (fracción ácida), aún permanecen restos de manosilfosfato (**Figura 5C**). El tratamiento de los glucanos de YSH-44 con hidrólisis ácida suave seguida de alcalina fosfatasa demuestra que la fracción ácida está comprendida por aproximadamente 5-15 % de glucanos totales. Esta cepa YSH-49 muestra una fracción ácida de aproximadamente el 6 %, que se compara favorablemente con la fracción ácida de aproximadamente el 9 % de la YSH-44 (**Figura 5B**).

Por el contrario, la **Figura 5D** muestra la eliminación de la transferencia de manosilfosfato a glucanos en *P. pastoris* YAS-130 ( $\Delta pno1 \Delta mnn4B$ ) en comparación con la **Figura 5A** de control ( $H_2O$ ), **Figura 5B** YSH-44 con aproximadamente 9 % de manosilfosforilación, y **Figura 5C** YSH-49 ( $\Delta pno1$ ) con aproximadamente 6 % de manosilfosforilación. En el presente documento se describe por primera vez una cepa de levadura modificada por ingeniería genética para estar esencialmente sin glucanos manosilfosforilados.

#### 60 Glucoproteínas terapéuticas producidas en levadura (*P. pastoris*)

Diferentes glucoproteínas pueden presentar diversos grados y tipos de glucosilación en la misma célula hospedadora (Montesino *et al.*, 1998). La presente invención proporciona métodos para producir diversas glucoproteínas en una cepa de levadura de *P. pastoris* recombinante que esencialmente carece de

manosilfosforilación. En consecuencia, el método implica modificar por ingeniería genética la expresión de una glucoproteína heteróloga en *P. pastoris*  $\Delta pno1 \Delta mnn4B$ . Como tal, la presente invención demuestra la eliminación de manosilfosforilación de glucanos en diversas glucoproteínas terapéuticas (**Figura 7A-E**, **Figura 8**).

5 Aunque la proteína indicadora K3 contiene un único sitio de glucosilación ligado a *N*, la proteína indicadora His-eritropoyetina (EPO) desvelada en el presente documento contiene tres sitios de glucosilación ligados a *N*, la proteína indicadora His-CD40 desvelada en el presente documento contiene dos sitios de glucosilación y la proteína His-invertasa desvelada en el presente documento contiene hasta 24 sitios de glucosilación. La eritropoyetina marcada con His (His-EPO) se expresa a partir de una cepa de *P. pastoris* que expresa manosilfosforilación en la **Figura 7B** y una cepa de *P. pastoris*  $\Delta pno1 \Delta mnn4B$  que carece de manosilfosforilación en la **Figura 7C**. Se expresa CD40 marcada con His (His-CD40) de la cepa de *P. pastoris* que expresa manosilfosforilación en la **Figura 7D** y cepa de *P. pastoris*  $\Delta pno1 \Delta mnn4b$  que carece de manosilfosforilación en la **Figura 7E**. Se expresa invertasa marcada con His de la cepa de *P. pastoris* que carece de manosilfosforilación en la **Figura 8**. La construcción de cepas para cada una de estas glucoproteínas se desvela en el **Ejemplo 6**.

#### 15 Identificación de homólogos de MNN4

En otro aspecto de la presente divulgación, se proporciona un método para identificar los homólogos de un gen *MNN4* en cualquier levadura preferentemente en *Pichia* sp. u hongos filamentosos. Un experto en la materia puede realizar una búsqueda de base de datos BLAST usando la secuencia de aminoácidos de *MNN4A*, *MNN4B*, *MNN4C* o *PNO1* (referencia de Genbank N° BD105434) frente al genoma de cualquier levadura, preferentemente *Pichia* y obtener los homólogos de cualquiera de estos genes. Con la identificación de los homólogos de *MNN4/PNO1* en levadura *Pichia*, un experto en la materia puede alterar o mutar posteriormente cualquier combinación de estos genes homólogos. Se muestra un alineamiento en la **Figura 9** de los homólogos de *MNN4/PNO1* en *P. pastoris*, *S. cerevisiae*, *Neurospora crassa*, *Aspergillus nidulans*, *Candida albicans* y *Pichia angusta* (*Hansenula polymorpha*). Tras explorar con respecto a la presencia de glucanos manosilfosforilados en proteínas expresadas de la cepa de *Pichia* (**Ejemplo 7**), un experto en la materia puede determinar el gen o la combinación de genes, que tras la disrupción confieren la expresión de glucoproteínas del hospedador de *Pichia* que están esencialmente sin manosilfosforilación.

Los genes alterados o genes que codifican proteínas que participan en la transferencia de manosilfosfato a glucanos de glucoproteínas son preferentemente de una cepa de levadura que pertenece al género *Pichia*. Las levaduras que pertenecen al género *Pichia* de acuerdo con la presente invención incluyen, pero sin limitación: *Pichia pastoris*, *Pichia finlandica*, *Pichia trehalophila*, *Pichia koclamae*, *Pichia membranaefaciens*, *Pichia methanolica*, *Pichia minuta* (*Ogataea minuta*, *Pichia lindneri*), *Pichia opuntiae*, *Pichia thermotolerans*, *Pichi salictaria*, *Pichia guercum*, *Pichia pijperi*, *Pichia stiptis*, y *Pichia angusta* (*Hansenula polymorpha*). *Pichia pastoris* se usa preferentemente entre estos. Otras levaduras y hongos filamentosos incluyen *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe*, *Saccharomyces* sp. *Hansenula polymorpha*, *Kluyveromyces* sp., *Candida* sp., *Candida albicans*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Trichoderma reesei*, *Chrysosporium lucknowense*, *Fusarium* sp., *Fusarium gramineum*, *Fusarium venenatum* y *Neurospora crassa*.

Los siguientes son ejemplos que ilustran las composiciones y métodos de la presente invención. Estos ejemplos no deberían interpretarse como limitantes, los ejemplos se incluyen solamente para fines de ilustración.

#### 45 **Ejemplo 1**

##### **Identificación y secuenciación de MNN4A, MNN4B, MNN4C en P. pastoris (Figuras 1-3)**

La secuencia proteica de *MNN4* de *Saccharomyces cerevisiae* (referencia de Genbank N° P36044) se sometió a blast frente a una secuencia genómica de *Pichia pastoris* (Integrated Genomics, Chicago, IL) para fases abiertas de lectura que codifican proteínas con homología. Esta búsqueda identificó tres ORF con regiones de homología con *MNN4p*. Estas ORF se designaron *MNN4A*, *MNN4B* y *MNN4C*. Cada uno de estos tres genes se secuenció posteriormente. Se descubrió que el gen *MNN4A* contenía una fase abierta de lectura que contenía 2580 bases de nucleótidos que codificaban 860 aminoácidos (**Figura 1**). Se descubrió que el gen *MNN4B* contenía una fase abierta de lectura que contenía 1956 bases de nucleótidos que codificaban 652 aminoácidos (**Figura 2**) y se descubrió que el gen *MNN4C* contenía una fase abierta de lectura que contenía 2289 bases de nucleótidos que codificaban 763 aminoácidos (**Figura 3**).

#### 60 **Ejemplo 2**

##### **Construcción de las cepas de P. pastoris: YSH-44 y YSH-1**

*P. pastoris* YSH-44 e YSH-1 se obtuvieron por ingeniería genética a partir de BK64-1, un mutante de delección de  $\Delta och1$  que secretaba K3, una proteína indicadora con un único sitio de glucosilación ligado a *N* (Choi *et al.*, 2003, PNAS, 100: 5022-5027; Hamilton *et al.*, 2003, Science, 301: 1244-1246). YSH-1 expresa glucoproteínas que tienen predominantemente *N*-glucanos GlcNAcMan<sub>5</sub>GlcNAc<sub>2</sub> e YSH-44 expresa glucoproteína que tiene

predominantemente *N*-glucanos GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>.

#### Delección del gen *PNO1* para la cepa YSH-44

5 El alelo de delección de *pnol* (*pnol::Hyg<sup>R</sup>*) en YSH-44 se generó por el método de solapamiento de PCR (Davidson *et al.*, 1999, Microbiol. 148: 2607-2615). Los cebadores PNK1 (5'-CATAGCCCCACTGCTAAGCC-AGAATTCTAATATG-3') (SEC ID Nº: 7) emparejado con PNK2 (5'-GCAGCGTACGAAGCTTCAGCTAGAATTGTTAAAGTGAATTATCAAG-TCT TTC-3') (SEC ID Nº: 8), PNK3 (5'-CAGATCCACTAGTGGCCTATGCAACAA-TATAGCACCTCTCAAATACAC GTTG-3') (SEC ID Nº: 9) emparejado con PNK4 (5'-TCITGAAGTAGATTTGGAGA-TTTTGGCGCTATG-3') (SEC ID Nº: 10) se usaron para amplificar las regiones flanqueantes 5' y 3' del gen *PNO1* del ADN genómico (NRRL-Y11430).  
 10 Los cebadores KAN1 (5'-AGCTGAAGCT-TCGTACGCTGC-3') (SEC ID Nº: 11) emparejado con KAN2 (5'-GCATAGGCCACTAGTGGATCTG-3') (SEC ID Nº: 12) se usaron para amplificar el marcador de resistencia a Hyg del vector pAG32 (Goldstein y McCusker, 1999, Yeast 14: 1541-1553). Después se usaron los cebadores PNK1 y PNK4 en una segunda reacción con los tres productos del primer ciclo de reacciones de PCR para generar un producto de solapamiento. El producto de PCR de fusión resultante se usó para transformar la cepa YSH-44, una cepa de *P. pastoris* obtenida por ingeniería genética que expresa predominantemente GlcNAc<sub>2</sub>Man<sub>3</sub>GlcNAc<sub>2</sub>. Los transformantes se seleccionaron en medio YPD (extracto de levadura 1 %, peptona 2 %, dextrosa 2 %) que contenía higromicina B 200 mg/ml. La integración apropiada del alelo de delección *pnol::Hyg<sup>R</sup>* se confirmó por PCR. Esta cepa  $\Delta$ *pnol1* se designó YSH-49.

#### 20 Ejemplo 3

##### Estrategia de anulación de *PNO1/MNN4B* en la cepa de *P. pastoris* YSH-49 (Figura 4)

25 Se consiguió una cepa de doble mutante YAS-130 ( $\Delta$ *pnol1*  $\Delta$ *mnn4b*) por solapamiento de PCR en YSH-49. Se usaron los cebadores TAS54 (TTCAACGAGTGACCAATGTAGA) (SEC ID Nº: 13) y TAS51 (CCAT-CCAGTGTCGAAAACGAGCTGGCGAACTTITCTGGGTCGAAG) (SEC ID Nº: 14) para amplificar el fragmento de ADN de 521 pb 5' del codón de partida predicho de ADN genómico de *Pichia pastoris* (NRRL-Y 11430). TAS51 contiene un saliente de 22 pb que es complementario del extremo 5' de un marcador de resistencia a fármacos. TAS49 (TGAAGACGTCCCCTTTGAACA) (SEC ID Nº: 15) y TAS52 (ACGAGGCAAGCTAAACAGATCTAGTTGTTTTTCTAT ATAAAAC) (SEC ID Nº: 16) se usaron para amplificar el fragmento de ADN de 503 pb 3' del codón de parada predicho. TAS52 también contiene un saliente de 22 pb que es complementario del extremo 3' del marcador de resistencia a fármacos. La PCR del marcador de resistencia a fármacos usó pAG29 (contiene el ORF de *pat*) como la fuente de ADN (Goldstein y McCusker, 1999). El marcador de resistencia a fármacos se amplificó usando los cebadores TAS53 (CTTCGACCCAGAAAAGTTCCGCCAGCTCG-TTTTCGACACTGGATGG) (SEC ID Nº: 17) y TAS50 (GTTTTATATAG-AAAAACAACACTAGATCTGTTTAGCTTGCCTCGT) (SEC ID Nº: 14). TAS53 tiene un saliente de 22 pb que es complementario de los 22 pb 5' del codón de inicio de *MNN4B* predicho. TAS50 tiene un saliente de 22 pb que es complementario de los 22 pb 3' del codón de parada de *MNN4B* predicho. El fragmento de *MNN4B* 5', fragmento de *MNN4B* 3' y el gen que confiere resistencia a un marcador seleccionable se combinaron en una relación equimolar y se usaron como un ADN molde con los cebadores TAS54 y TAS49 para la reacción de solapamiento de PCR.

##### Estrategia de anulación de *PNO1/MNN4B* en la cepa de *P. pastoris* YSH-1

45 YSH-1 se transformó por electroporación con pJN503b digerido con *Sfil* ( $\Delta$ *mnn4A*  $\Delta$ *pnol1::URA3*) para producir la cepa  $\Delta$ *och1*  $\Delta$ *mnn4A*  $\Delta$ *pnol1* YAS159. El marcador seleccionable de *URA3* se recuperó en esta cepa por contraselección con 5-FOA. La cepa resultante, YAS164 ( $\Delta$ *och1*  $\Delta$ *mnn4A*  $\Delta$ *pnol1*; *ura3*; *his4*; *ade1*; *arg4*), se transformó con pAS16 digerido con *Sfil* ( $\Delta$  *mnn4B::URA3*) dando lugar a la cepa  $\Delta$ *och1*  $\Delta$ *mnn4A*  $\Delta$ *pnol1*  $\Delta$ *mnn4B* YAS170. La cepa YAS170 se contraseleccionó posteriormente en 5-FOA para producir la cepa YAS174 ( $\Delta$ *och1*  $\Delta$ *mnn4A*  $\Delta$ *pnol1*  $\Delta$ *mnn4B*; *ura3*; *his4*; *ade1*; *arg4*). YAS174 representa por lo tanto una cepa de *Pichia pastoris* que es deficiente en la formación de cadena externa de manosa y desprovista de manosilfosfato en glucanos ligados a *N*.

#### 50 Ejemplo 4

##### Amplificación por PCR

55 Se usó un Mastercycler de Eppendorf para todas las reacciones de PCR. Las reacciones de PCR contenían ADN molde, dNTP 125 mM, 0,2 mM de cada uno de los cebadores directo e inverso, tampón de Taq polimerasa Ex (Takara Bio Inc.) y Taq polimerasa Ex. Los fragmentos de ADN 5' de la ORF de *MNN4B* predicha, 3' de la ORF de *MNN4B* predicha, y el marcador de resistencia a fármacos se amplificaron con 30 ciclos de 15 segundos a 97 °C, 15 segundos a 55 °C y 90 segundos a 72 °C con una etapa de desnaturalización inicial de 2 minutos a 97 °C y una etapa de extensión final de 7 minutos a 72 °C. Se separaron las muestras de PCR por electroforesis en gel de agarosa y las bandas de ADN se extrajeron y se purificaron usando un Kit de Extracción en Gel de Qiagen. Todas las purificaciones de ADN se eluyeron en Tris 10 mM, pH 8,0 excepto por la PCR final (solapamiento de los tres fragmentos) que se eluyeron en H<sub>2</sub>O desionizada.

65

**Ejemplo 5**Transformaciones de ADN, condiciones de cultivo para producción de glucanos complejos en *P. pastoris* para análisis de manosilfosforilación

Se preparó ADN para transformación añadiendo acetato de sodio a una concentración final de 0,3 M. Se añadió después etanol helado cien por cien a una concentración final de 70 % a la muestra de ADN. El ADN se sedimentó por centrifugación (12000 g x 10 min) y se lavó dos veces con etanol helado al 70 %. El ADN se secó y después se resuspendió en 50 ml de Tris 10 mM, pH 8,0. Se prepararon YSH-49 e YAS-130 expandiendo un cultivo de levadura en BMGY (glicerol mínimo tamponado: fosfato potásico 100 mM, pH 6,0; base nitrogenada de levadura 1,34 %; biotina  $4 \times 10^{-5}$  %; glicerol 1 %) a una D.O. de ~2-6. La levadura se hizo electrocompetente lavando 3 veces en sorbitol 1 M y resuspendiendo en ~1-2 ml de sorbitol 1 M. Se mezcló ADN (1-2 mg) con 100 ml de levadura competente y se incubó en hielo durante 10 minutos. La levadura se sometió después a electroporación con un BTX Electrocell Manipulator 600 usando los siguientes parámetros; 1,5 kV, 129 ohmios y 25 mF. Se añadió un mililitro de YPDS (extracto de levadura 1 %, peptona 2 %, dextrosa 2 %, sorbitol 1 M) a las células sometidas a electroporación. La levadura transformada se sembró posteriormente en placas de agar selectivo. Las células transformadas con construcciones de anulación que contenían el gen de resistencia *hph* se extendieron en placas de agar YPD Y+ (extracto de levadura 1 %, peptona 2 %, dextrosa 2 %, base nitrogenada de levadura sin aminoácidos 1,34 %) que contenían higromicina B 0,4 mg/ml. Las células transformadas con construcciones de anulación que contenían el gen de resistencia *pat* se extendieron en placas de agar con medio definido (base nitrogenada de levadura sin aminoácidos y  $\text{NH}_4\text{SO}_4$  1,34 %, dextrosa 2 %, L-prolina 0,1 %, biotina  $4 \times 10^{-5}$  %) que contenían glufosinato 0,6 mg/ml. Las colonias se pasaron en manchas a otra placa que contenía la misma selección farmacológica. El ADN se aisló de estas manchas y se analizó por PCR con respecto al reemplazo de la ORF de *MNN4B* de tipo silvestre con el marcador de resistencia a fármacos.

Se realizó exploración para anulaciones por amplificación por PCR (**Ejemplo 4**) de las partes tanto 5' como 3' de la construcción de anulación. Se usaron los cebadores TAS81 (TAGTCCAAGTACGA-AACGACACTATCG) (**SEC ID Nº: 19**) y TAS08 (AGGTGCGCACGTCAAGAC-TGTCAAGG) (**SEC ID Nº: 20**) para explorar la parte 5' de la construcción de anulación mientras que se usaron los cebadores TAS82 (ACGACGGTGAGTTCAAACAGTTTGGTT) (**SEC ID Nº: 21**) y TAS07 (TCGCTATACTGCTGTCGATTTCGATAC) (**SEC ID Nº: 22**) para explorar la parte 3' de la construcción de anulación. La observación de un producto de PCR en ambas exploraciones es indicativa de una anulación exitosa de la ORF *MNN4B* ya que los cebadores TAS08 y TAS07 se hibridan en los extremos 5' y 3' de la secuencia del marcador de resistencia a fármacos, respectivamente y TAS81 y TAS82 son complementarios de secuencias en el genoma que flanquea las regiones 5' y 3' de ADN usado en la construcción de anulación. Se exploraron noventa y seis transformantes dando cuatro resultados positivos como una anulación de *MNN4B*. Las cuatro cepas  $\Delta pno1 \Delta mnn4b$  expresaron la proteína indicadora K3 sin niveles detectables de manosiifosfato. Se muestra un ejemplo de esto en la **Figura 5D**.

**Ejemplo 6****Construcción de cepas para proteínas EPO, CD40 e invertasa marcadas con His Figuras 7.8**

Para eritropoyetina marcada con His (EPO), se amplificaron los primeros 166 aminoácidos de EPO de una biblioteca de ADNc de riñón humano (Clontech) e insertaron en el plásmido pPICZA 6 His C terminal (Invitrogen) en los sitios EcoRI y KpnI. Este plásmido (pBK291) se transformó en dos cepas de *P. pastoris*, dando como resultado la siguientes cepas que expresan EPO-6His: BK248 (*ura3, his4, ade1, arg4,  $\Delta och1::URA3$* ) y BK244 [YSH44 transformado con pBK116 y pBK284 que tiene las anulaciones *pno1mnn4b* (*pno1::Hyg<sup>R</sup>*) (*mnn4b::Kan<sup>R</sup>*) como se describe y se muestra en el **Ejemplo 2, Figura 4**. pBK116 resulta de un fragmento de ADN 3'UTR de AOX1 de 1551 pb aislado de NRRL11430 (ATCC) insertado en el plásmido de Invitrogen pPIC6A en el sitio AflIII y un fragmento de ADN 5'UTR de AOX1 de 1952 pb aislado de NRRL11430 insertado en el mismo plásmido pPIC6A en los sitios BglIII y BamHI con la retirada del fragmento de ADN PmeI/BamHI de 573 pb. Este pBK116 se digirió después con NotI y los fragmentos de NotI resultantes se transformaron en YSH44 para anular la proteína K3 indicadora. pBK284 resulta de un fragmento de ADN de 3196 pb que incluye el promotor de AOX1, ORF de AOX1 y secuencia de terminación de AOX1 aislada de NRRL11430 (ATCC) y clonada en el sitio de clonación múltiple del plásmido de Invitrogen pCR2.1-TOPO. Este plásmido se digirió después con MscI y BssHI para suprimir el gen de kanamicina. Esto dio como resultado pBK284 que se digirió con PmeI antes de la transformación en la cepa YSH44 transformada con pBK116 para integración en el locus del promotor de AOXI. Se muestra el análisis de glucanos de HPLC de EPO-6His en BK248 y BK244 en la **Figura 7B, C**. Para CD40 marcado con His, el ADN de CD40 humano se amplificó por PCR de phCD40/GemT (Pullen *et al.*, 1999, JBC, 274: 14246-14254) usando un cebador de EcoRI 5' y un cebador de His10-KpnI 3' para clonar en pPICZ aA dando como resultado pJC33. pJC33 se expresó en la cepa de *P. pastoris* **YJC12** (*ura3, his4, adel, arg4*) e **YAS252-2** (YAS-130 transformada con pBK116, pBK284 y pRCD465 (USSN 60/562424) que contiene galactosiltransferasa) dando como resultado **YAS252**. Se muestra el análisis de glucanos de HPLC de CD49-6His en YJC12 e YAS252 en la **Figura 7D, E**. Para invertasa marcada con His, se amplificó la secuencia de invertasa de longitud completa por PCR de ADN genómico de *Kluyveromyces lactis*, cepa CBS683, obtenida de Centraalbureau voor Schimmelcultures. La ORF de invertasa se amplificó usando cebadores 5' y 3' de extremos romos para inserción en el plásmido pPICZA (proporcionando el marcador 6His C terminal) en el

sitio PmII. Este pPB147 se transformó en la cepa de *P. pastoris* **YAS245-2** (YAS130 transformada con pBK116, pBK284 y pRCD465 (USSN 60/562424) que da como resultado **YAS253**). Se muestra el análisis de glucanos de HPLC de invertasa-6His en YAS253 en la **Figura 8**.

## 5 **Ejemplo 7**

### **Determinación de manosilfosforilación en *P. pastoris***

El alcance de la transferencia de manosilfosfato a glucanos ligados a *N* en las cepas mostradas en las **Figuras 5-8** se determinó secretando una proteína indicadora marcada con His (proteína kringle 3 en las Figuras 5,6; proteína eritropoyetina y proteína CD40 en la Figura 7 y proteína invertasa en la Figura 8) expresada bajo el control del promotor de AOX1 inducible por metanol. Brevemente, se inoculó un matraz de agitación que contenía BMGY con un cultivo de levadura nuevo (por ejemplo, YAS-130) y se cultivó hasta una D.O. de ~20. El cultivo se centrifugó y el sedimento celular se lavó con BMMY (metanol mínimo tamponado: igual que BMGY excepto metanol 0,5 % en lugar de glicerol 1 %). El sedimento celular se resuspendió en BMMY hasta un volumen 1/5 del cultivo de BMGY original y se colocó en un agitador durante 24 h. La proteína secretada se recogió por sedimentación de la biomasa por centrifugación y transfiriendo el medio de cultivo a un tubo nuevo. Las proteínas K3, EPO, CD40 e invertasa marcadas con His se purificaron después en una columna de afinidad de Ni y se digirieron con PNGasa (Choi *et al.*, 2003). Se separó el glucano de la proteína y después se marcó con 2-amino-benzamida (2-AB). El glucano marcado con 2-AB se liofilizó, se resuspendió en agua de uso para HPLC y se sometió a HPLC usando una columna GlycoSep C (Glyco, Novato, CA). Este análisis permite la separación de glucanos neutros y ácidos. Se determinó que estos glucanos estaban fosforilados a partir de experimentos con hidrólisis ácida suave que retira el grupo de manosa terminal, exponiendo el fosfato. Con tratamiento de fosfatasa alcalina posterior, el grupo de fosfato terminal puede escindirse, dejando un glucano neutro. Experimentos posteriores mostraron que los glucanos ligados a *N* fosforilados (glucanos ácidos) en todas las cepas migraban entre 20 y 30 minutos. Las condiciones de línea basal se evaluaron usando dH<sub>2</sub>O como un blanco. El porcentaje de fosforilación se calculó dividiendo las áreas de picos ácidos por la suma de los picos neutros y ácidos. Este análisis de HPLC se realizó en las condiciones posteriores.

### **Análisis de HPLC**

Las condiciones de HPLC son la siguientes: disolvente A (acetonitrilo), disolvente B (acetato de amonio 500 mM, 500 mM, pH 4,5) y disolvente C (agua). El caudal fue de 0.4 ml/minuto durante 50 minutos. Después de eluir de forma isocrática (A 20 %:C 80 %) durante 10 min se empleó un gradiente de disolvente lineal (A 20 %:B 0 %:C 80 % a A 20 %:B 50 %:C 30 %) durante 30 minutos para eluir los glucanos. La columna se equilibró con el disolvente (A 20 %:C 80 %) durante 20 min entre ciclos.

### **Texto de la lista de secuencias**

**SEC ID N°: 1** MNN4A (Figura 1)  
**SEC ID N°: 2** MNN4A AA (Figura 1)  
**SEC ID N°: 3** MNN4B (Figura 2)  
**SEC ID N°: 4** MNN4B AA (Figura 2)  
**SEC ID N°: 5** MNN4C (Figura 3)  
**SEC ID N°: 6** MNN4C AA (Figura 3)  
**SEC ID N°: 7** PNK1: CATAGCCCACTGCTAAGCCAGAATTCTAATATG

**SEC ID N°: 8** PNK2: GCAGCGTACGAAGCTTCAGCTAGAATTGTAAAGT-  
 GAATTATCAAGTCTTTC

**SEC ID N°: 9** PNK3: CAGATCCACTAGTGGCCTATGCAACAATATAG-  
 CACCTCTCAAATACACGTTG

**SEC ID N°: 10** PNK4: TCTTGAAGTAGATTTGGAGATTTTGCGCTATG  
**SEC ID N°: 11** KAN1: AGCTGAAGCTTCGTACGCTGC  
**SEC ID N°: 12** KAN2: GCATAGGCCACTAGTGGATCTG  
**SEC ID N°: 13** TAS54: TTCAACGAGTGACCAATGTAGA

**SEC ID N°: 14** TAS51: CCATCCAGTGTCGAAAACGAGCTGGC-  
 GAACTTTTCTGGGTCGAAG

**SEC ID N°: 15** TAS49: TGAAGACGTCCCCTTTGAACA



**SEC ID N°: 16 TAS52: ACGAGGCAAGCTAAACAGATCTAGTTGTTTTTTC-  
TATATAAAAC**

**SEC ID N°: 17 TAS53: CTTCGACCCAGAAAAGTTCGCCAGCTCGTTTTTC-  
GACACTGGATGG**

**SEC ID N°: 18 TAS50: GTTTTATATAGAAAAACAACACTAGATCTGTT-  
TAGCTTGCCTCGT**

**SEC ID N°: 19 TAS81: TAGTCCAAGTACGAAACGACACTATCG**

**SEC ID N°: 20 TAS08: AGCTGCGCACGTCAAGACTGTCAAGG**

5 **SEC ID N°: 21 TAS82: ACGACGGTGAGTTCAAACAGTTTGGTT**

**SEC ID N°: 22 TAS07: TCGCTATACTGCTGTCGATTCGATAC**

LISTADO DE SECUENCIAS

- 10 <110> GlycoFi. Inc.  
Terrance Stadheim  
Piotr, Bobrowicz  
Wildt, Stefan
- 15 <120> Métodos para eliminar la manosilfosforilación de glucanos en la producción de glucoproteínas
- <130> GFI-10
- <160> 29
- 20 <170> Patentin versión 3.3
- <210> 1
- <211> 2583
- 25 <212> ADN
- <213> *Pichia pastoris*
- <400> 1

atgaaagtat caaagcggtt gataccgagg agatctcgtc tcctcattat gatgatgcta 60  
 ctggttggtt accagctcgtt ggttttggtc ctaggattgg agagcgtctc tgaaggaaaa 120  
 tttagcaagct tgcttgactt gggcgattgg gatctagcta actcctcgtc atctatatcc 180  
 gatttcataa agctgaagct caaaggccaa aagacttacc acaaatttga tgaacatgct 240  
 ttgcgcccaa tggcaagaat tcaaagtaat gagaatggca agttggcgga ttacgagtct 300  
 acttcatcga agactgacgt aaccattcaa aatgltgaac ttggaagag attgagcgaa 360  
 gaagaataca ctacgaacc gcggataact ttgctgtgt atctgagcta cattcatcag 420  
 aggacttatg acaggtacgc gactagttac gtccttata acttgcgggt gccttttcg 480  
 tggctgact ggatagatct gacggcccta aatcaatact tggataaaac gaaaggctgc 540  
 gaggcagttt tccttagaga aagtgaggca actatgaagc ttaacaatat cactgttggtg 600  
 gactggcttg agggccttg cataactgat aatcacttc aaaattccgt aaactccaca 660  
 tatgcggaag agattaatag tcgggacatc ttgtcccta acttccatgt gtttggtat 720  
 tctgatgcta aagataatcc tcagcaaaaa atcttcaat ctaaactta tatcaactca 780  
 aagctgccgc tccaaaaaag ttgatattt ttaacagatg gaggtagtta cgcttgaca 840  
 gtcgaccgaa ctcaaaaataa aagaattcta aaatctggcc tgcttcaca ctttttca 900

aagaaaaaga aggaacacaa tctgcctcaa gaccaaaaaa ctttcacggtt tgaccccgta 960  
tacgaattca atagactgaa atctcaggtc aagccccgtc caatatcttc agaacctagt 1020  
attgattctg cttgaagga aaatgactac aagcftaaac tgaaagagtc gtcglttatt 1080  
titaattacg gaaggattct ttcgaactat gaagagcggc ttgagagtct aaatgacttc 1140  
gagaaatcgc actacgagtc cttagcttat tctccttgt tagaggcaag aaagttgccc 1200  
aagtattttg gcgaagttat attgaagaac ccacaagatg gtggaattca ttatgattac 1260  
agattcttca gcggactcat tgataaaact cagataaact attttgagga tgagactgaa 1320  
agaaagaaaa taatcatgcg tagacttctt cgaacttggc agtacttcac glatcacaat 1380  
aacattatca attggtatc gcacggttct ttactgtcat ggtattggga tggactttct 1440  
ttccatggg acaatgacat tgacgtacaa atgccataa tggagctgaa taactctgc 1500  
aaacagtcca acaattctct ggtcgtggag gatgittctc aagggittgg cagatactac 1560  
gttgattgca cgagcttctt ggcccagaga acgcgaggia atggtaacaa caacattgat 1620  
gcccgttita ttgatgtgc gtcgtgtctc ttactgaca ttacgggttt ggctttgact 1680  
ggatcaacaa tgcccaaaag atactccaat aagctgataa aacaaccgaa aaaatctacc 1740  
gactcaacag gatcgactcc tgagaacgga ctcaactagaa acttgaggca aaattgaaat 1800  
gcacaagttt acaactglag aaacgggtcat ttftaccaat actcggagct atctcctttg 1860  
aagttgtcga tagtagaagg tgcactcacc ctaataccca acgattttgt tactatattg 1920  
gaaactgagt accaaaggag aggtcttgaa aagaacacat atgcgaagta tctctacglt 1980  
ccagagcttc gactttggat gtcatacaat gacatctatg atatcttga aggtactaat 2040  
agtcatggcc gtcctttatc tgcaaagaca atggcgacta tcttctctcg gttaaactct 2100  
gacattaatc taaaaaagtt ttgcgcaat gatcatactt ttaagaacat ttattctact 2160  
ttcaacgtga cacgagtga cgaggaggaa ctgaagcatt tgatagtaaa ctatgaccaa 2220  
aataaacgga agtcggctga gtacaggcag ttcttggaat acttgcggtt tatgaatcca 2280  
atcagaaaag atctggtgac ttacgagagt aggttgaagg ctcttgatgg atacaatgag 2340  
gtcgaagaat tagaaaagaa gcaagagaat agggaaaaag aaagaaagga gaagaaggaa 2400

aaggaggaaa aagagaagaa ggaaaaggag gaaaaagaga agaaggaaaa ggaagaaaag 2460

gaaaagaagg aaaaggaaga aaaggagagg aaagagaagg aagaaaagga agaatatgaa 2520

gaagacgata atgagggcga acaaccaaca gaacaaaaga gccagcagga ggctaaagaa 2580

tag 2583

5 <210> 2  
 <211> 860  
 <212> PRT  
 <213> *Pichia pastoris*  
 <400> 2

Met Lys Val Ser Lys Arg Leu Ile Pro Arg Arg Ser Arg Leu Leu Ile  
 1 5 10 15

Met Met Met Leu Leu Val Val Tyr Gln Leu Val Val Leu Val Leu Gly  
 20 25 30

Leu Glu Ser Val Ser Glu Gly Lys Leu Ala Ser Leu Leu Asp Leu Gly  
 35 40 45

Asp Trp Asp Leu Ala Asn Ser Ser Leu Ser Ile Ser Asp Phe Ile Lys  
 50 55 60

Leu Lys Leu Lys Gly Gln Lys Thr Tyr His Lys Phe Asp Glu His Val  
 65 70 75 80

Phe Ala Ala Met Ala Arg Ile Gln Ser Asn Glu Asn Gly Lys Leu Ala  
 85 90 95

Asp Tyr Glu Ser Thr Ser Ser Lys Thr Asp Val Thr Ile Gln Asn Val  
 100 105 110

Glu Leu Trp Lys Arg Leu Ser Glu Glu Glu Tyr Thr Tyr Glu Pro Arg  
 115 120 125

Ile Thr Leu Ala Val Tyr Leu Ser Tyr Ile His Gln Arg Thr Tyr Asp  
 130 135 140

10

Arg Tyr Ala Thr Ser Tyr Ala Pro Tyr Asn Leu Arg Val Pro Phe Ser  
145                    150                    155                    160

Trp Ala Asp Trp Ile Asp Leu Thr Ala Leu Asn Gln Tyr Leu Asp Lys  
                  165                    170                    175

Thr Lys Gly Cys Glu Ala Val Phe Pro Arg Glu Ser Glu Ala Thr Met  
                  180                    185                    190

Lys Leu Asn Asn Ile Thr Val Val Asp Trp Leu Glu Gly Leu Cys Ile  
                  195                    200                    205

Thr Asp Lys Ser Leu Gln Asn Ser Val Asn Ser Thr Tyr Ala Glu Glu  
                  210                    215                    220

Ile Asn Ser Arg Asp Ile Leu Ser Pro Asn Phe His Val Phe Gly Tyr  
225                    230                    235                    240

Ser Asp Ala Lys Asp Asn Pro Gln Gln Lys Ile Phe Gln Ser Lys Ser  
                  245                    250                    255

Tyr Ile Asn Ser Lys Leu Pro Leu Pro Lys Ser Leu Ile Phe Leu Thr  
                  260                    265                    270

Asp Gly Gly Ser Tyr Ala Leu Thr Val Asp Arg Thr Gln Asn Lys Arg  
                  275                    280                    285

Ile Leu Lys Ser Gly Leu Leu Ser His Phe Phe Ser Lys Lys Lys Lys  
                  290                    295                    300

Glu His Asn Leu Pro Gln Asp Gln Lys Thr Phe Thr Phe Asp Pro Val  
305                    310                    315                    320

Tyr Glu Phe Asn Arg Leu Lys Ser Gln Val Lys Pro Arg Pro Ile Ser  
                  325                    330                    335

Ser Glu Pro Ser Ile Asp Ser Ala Leu Lys Glu Asn Asp Tyr Lys Leu  
                  340                    345                    350

Lys Leu Lys Glu Ser Ser Phe Ile Phe Asn Tyr Gly Arg Ile Leu Ser  
355 360 365

Asn Tyr Glu Glu Arg Leu Glu Ser Leu Asn Asp Phe Glu Lys Ser His  
370 375 380

Tyr Glu Ser Leu Ala Tyr Ser Ser Leu Leu Glu Ala Arg Lys Leu Pro  
385 390 395 400

Lys Tyr Phe Gly Glu Val Ile Leu Lys Asn Pro Gln Asp Gly Gly Ile  
405 410 415

His Tyr Asp Tyr Arg Phe Phe Ser Gly Leu Ile Asp Lys Thr Gln Ile  
420 425 430

Asn His Phe Glu Asp Glu Thr Glu Arg Lys Lys Ile Ile Met Arg Arg  
435 440 445

Leu Leu Arg Thr Trp Gln Tyr Phe Thr Tyr His Asn Asn Ile Ile Asn  
450 455 460

Trp Ile Ser His Gly Ser Leu Leu Ser Trp Tyr Trp Asp Gly Leu Ser  
465 470 475 480

Phe Pro Trp Asp Asn Asp Ile Asp Val Gln Met Pro Ile Met Glu Leu  
485 490 495

Asn Asn Phe Cys Lys Gln Phe Asn Asn Ser Leu Val Val Glu Asp Val  
500 505 510

Ser Gln Gly Phe Gly Arg Tyr Tyr Val Asp Cys Thr Ser Phe Leu Ala  
515 520 525

Gln Arg Thr Arg Gly Asn Gly Asn Asn Asn Ile Asp Ala Arg Phe Ile  
530 535 540

Asp Val Ser Ser Gly Leu Phe Ile Asp Ile Thr Gly Leu Ala Leu Thr

545                    550                    555                    560  
 Gly Ser Thr Met Pro Lys Arg Tyr Ser Asn Lys Leu Ile Lys Gln Pro  
                           565                    570                    575  
  
 Lys Lys Ser Thr Asp Ser Thr Gly Ser Thr Pro Glu Asn Gly Leu Thr  
                           580                    585                    590  
  
 Arg Asn Leu Arg Gln Asn Leu Asn Ala Gln Val Tyr Asn Cys Arg Asn  
                           595                    600                    605  
  
 Gly His Phe Tyr Gln Tyr Ser Glu Leu Ser Pro Leu Lys Leu Ser Ile  
                           610                    615                    620  
  
 Val Glu Gly Ala Leu Thr Leu Ile Pro Asn Asp Phe Val Thr Ile Leu  
                           625                    630                    635                    640  
  
 Glu Thr Glu Tyr Gln Arg Arg Gly Leu Glu Lys Asn Thr Tyr Ala Lys  
                           645                    650                    655  
  
 Tyr Leu Tyr Val Pro Glu Leu Arg Leu Trp Met Ser Tyr Asn Asp Ile  
                           660                    665                    670  
  
 Tyr Asp Ile Leu Gln Gly Thr Asn Ser His Gly Arg Pro Leu Ser Ala  
                           675                    680                    685  
  
 Lys Thr Met Ala Thr Ile Phe Pro Arg Leu Asn Ser Asp Ile Asn Leu  
                           690                    695                    700  
  
 Lys Lys Phe Leu Arg Asn Asp His Thr Phe Lys Asn Ile Tyr Ser Thr  
                           705                    710                    715                    720  
  
 Phe Asn Val Thr Arg Val His Glu Glu Glu Leu Lys His Leu Ile Val  
                           725                    730                    735  
  
 Asn Tyr Asp Gln Asn Lys Arg Lys Ser Ala Glu Tyr Arg Gln Phe Leu  
                           740                    745                    750

Glu Asn Leu Arg Phe Met Asn Pro Ile Arg Lys Asp Leu Val Thr Tyr  
 755 760 765

Glu Ser Arg Leu Lys Ala Leu Asp Gly Tyr Asn Glu Val Glu Glu Leu  
 770 775 780

Glu Lys Lys Gln Glu Asn Arg Glu Lys Glu Arg Lys Glu Lys Lys Glu  
 785 790 795 800

Lys Glu Glu Lys Glu Lys Lys Glu Lys Glu Glu Lys Glu Lys Lys Glu  
 805 810 815

Lys Glu Glu Lys Glu Lys Lys Glu Lys Glu Glu Lys Glu Arg Lys Glu  
 820 825 830

Lys Glu Glu Lys Glu Glu Tyr Glu Glu Asp Asp Asn Glu Gly Glu Gln  
 835 840 845

Pro Thr Glu Gln Lys Ser Gln Gln Glu Ala Lys Glu  
 850 855 860

<210> 3  
 <211> 1959  
 <212> ADN  
 <213> *Pichia pastoris*

5

<400> 3

atgtcaaag aaacgtcaaa gaactgttt ggttcgataa atacctcaa tacggtggag 60  
 tatgtcatgt atatgatgct actactgact gcgtatlttt tgaaccacct gttgcatagt 120  
 ttggataaca tcaatcattt ggttgagtct gatgttaatt atcaactact tcaaagggia 180  
 acaaataaag tcaagctttt tgatgaggaa gcagtcttgc cctttgctaa gaatcicaat 240  
 agaagaactg aacgcittga tccaagggtg cctgtagctg cataccttcg aagccttcaa 300  
 gatcagtatt cggagcttcc acaagggtacc gacctgaatg atattccgcc cctggagggtt 360  
 tcttccact gggatgactg gtttaagttg ggaattgcat caaccittg ggacgccttc 420  
 gacaattaca acaagagaca aggagaaaat gcaatttctt acgagcagct ccaagcaata 480  
 ctgttaatg atttgaaga ttttctccc tacaccgcac atattcttca cagtaacgtg 540

10



gaagtctaca aafacagaac gattcctcaa aagatcgtct atatgtcaaa caagggctat 600  
 ttgaaactct tggtaaccga aaaggaaaaa ctatccaatg agggctctcg gagcatttc 660  
 catcagaaac aaggtggact taacgaattc agtagtctca atctcataga ggaggtgat 720  
 gcgttggatg aaatctatga ttcaaaggg ttgcctgctt gggatcctcc ctccctgag 780  
 gaactlgatg ctcagatga agattcaag ttcaatgcc aagaagaact ggcaaaggta 840  
 gagcaaatca aagaaccaa gctggaagac atattctatc aggaaggact gcaacacggg 900  
 attcaaacat tgccttcaga tgcaagtgtt tattttctg tgaattacgt tgaaaacgac 960  
 cctggaitac agtcccatca cttacacttc ccattttca gtggaatggt cttaccaaga 1020  
 gaaatccatt ctcagtga tcacatgaat aaggcgttt tctgtttgc aagacagcac 1080  
 ggttatgttg ttggttct ttatggtaac ttaattggat ggtattaca tggaataac 1140  
 caccctggg attcggacat cgaigccata atgcccattg cggagatggc aagaatggct 1200  
 catcaccata acaacacact aataatagag aacccccacg atggatatgg aacctattia 1260  
 ctgactattt ctccttggtt cacgaagaag acaagaggtg gfaaccatat tgatggctgt 1320  
 ttgtggacg ttaagagggg tacctacatc gacctcagtg caatttcagc tatgcacgga 1380  
 atatatcctg actgggtag agatgggtg aaagaaaacc ctaagaatct ggctctggcc 1440  
 gacaagaacg gtaattgga ccttactaga gatattctcc cattgaggag aacaatattc 1500  
 gaaggttctc gatcctacac cgtaaagac attgaagata cctgcttag aaactatgga 1560  
 gataaagtac tgataaacac agaactggca gaccatgaat ggcatgatga ctggaaaatg 1620  
 tgggtacaaa aaaagaaata ctgcacttat gaggaattg aagattacct gagtgctcat 1680  
 ggagggggtg aatacgacga agatggagta ttgacctgg aaggagcttg tggattgaa 1740  
 gaagtccgac aagattggat cattaccctg gaaagtgtaa atctcatat gaaggaatgg 1800  
 gaagctatcc agaggaacga atcaaccaca gagtatactg ctaaggatct tctcgttac 1860  
 aggccagatt cctcaaaaa tctattggat ggagttcca atcatggaaa tggaaatggt 1920  
 ggtaagatag agcatgtcaa actggaacac aacgactag 1959

<210> 4  
 <211> 652  
 <212> PRT  
 <213> *Pichia pastoris*  
 <400> 4

5

Met Phe Lys Glu Thr Ser Lys Asn Leu Phe Gly Ser Ile Asn Thr Phe  
 1            5                    10                    15

Asn Thr Val Glu Tyr Val Met Tyr Met Met Leu Leu Leu Thr Ala Tyr  
           20                    25                    30

Phe Leu Asn His Leu Leu His Ser Leu Asp Asn Ile Asn His Leu Val  
           35                    40                    45

Glu Ser Asp Val Asn Tyr Gln Leu Leu Gln Arg Val Thr Asn Lys Val  
           50                    55                    60

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

Arg Arg Thr Glu Arg Phe Asp Pro Arg Leu Pro Val Ala Ala Tyr Leu  
           85                    90                    95

Arg Ser Leu Gln Asp Gln Tyr Ser Glu Leu Pro Gln Gly Thr Asp Leu  
           100                    105                    110

Asn Asp Ile Pro Pro Leu Glu Val Ser Phe His Trp Asp Asp Trp Leu  
           115                    120                    125

Ser Leu Gly Ile Ala Ser Thr Phe Trp Asp Ala Phe Asp Asn Tyr Asn  
           130                    135                    140

Lys Arg Gln Gly Glu Asn Ala Ile Ser Tyr Glu Gln Leu Gln Ala Ile  
 145                    150                    155                    160

Leu Val Asn Asp Leu Glu Asp Phe Ser Pro Tyr Thr Ala His Ile Leu  
           165                    170                    175

His Ser Asn Val Glu Val Tyr Lys Tyr Arg Thr Ile Pro Gln Lys Ile

180                      185                      190  
 Val Tyr Met Ser Asn Lys Gly Tyr Phe Glu Leu Leu Val Thr Glu Lys  
     195                      200                      205  
 Glu Lys Leu Ser Asn Glu Gly Leu Trp Ser Ile Phe His Gln Lys Gln  
     210                      215                      220  
 Gly Gly Leu Asn Glu Phe Ser Ser Leu Asn Leu Ile Glu Glu Val Asp  
     225                      230                      235                      240  
 Ala Leu Asp Glu Ile Tyr Asp Ser Lys Gly Leu Pro Ala Trp Asp Pro  
     245                      250                      255  
 Pro Phe Pro Glu Glu Leu Asp Ala Ser Asp Glu Asp Phe Lys Phe Asn  
     260                      265                      270  
 Ala Thr Glu Glu Leu Ala Lys Val Glu Gln Ile Lys Glu Pro Lys Leu  
     275                      280                      285  
 Glu Asp Ile Phe Tyr Gln Glu Gly Leu Gln His Gly Ile Gln Thr Leu  
     290                      295                      300  
 Pro Ser Asp Ala Ser Val Tyr Phe Pro Val Asn Tyr Val Glu Asn Asp  
     305                      310                      315                      320  
 Pro Gly Leu Gln Ser His His Leu His Phe Pro Phe Phe Ser Gly Met  
     325                      330                      335  
 Val Leu Pro Arg Glu Ile His Ser Ser Val His His Met Asn Lys Ala  
     340                      345                      350  
 Phe Phe Leu Phe Ala Arg Gln His Gly Tyr Val Val Trp Phe Phe Tyr  
     355                      360                      365  
 Gly Asn Leu Ile Gly Trp Tyr Tyr Asn Gly Asn Asn His Pro Trp Asp  
     370                      375                      380

Ser Asp Ile Asp Ala Ile Met Pro Met Ala Glu Met Ala Arg Met Ala  
 385 390 395 400

His His His Asn Asn Thr Leu Ile Ile Glu Asn Pro His Asp Gly Tyr  
 405 410 415

Gly Thr Tyr Leu Leu Thr Ile Ser Pro Trp Phe Thr Lys Lys Thr Arg  
 420 425 430

Gly Gly Asn His Ile Asp Gly Arg Phe Val Asp Val Lys Arg Gly Thr  
 435 440 445

Tyr Ile Asp Leu Ser Ala Ile Ser Ala Met His Gly Ile Tyr Pro Asp  
 450 455 460

Trp Val Arg Asp Gly Val Lys Glu Asn Pro Lys Asn Leu Ala Leu Ala  
 465 470 475 480

Asp Lys Asn Gly Asn Trp Tyr Leu Thr Arg Asp Ile Leu Pro Leu Arg  
 485 490 495

Arg Thr Ile Phe Glu Gly Ser Arg Ser Tyr Thr Val Lys Asp Ile Glu  
 500 505 510

Asp Thr Leu Leu Arg Asn Tyr Gly Asp Lys Val Leu Ile Asn Thr Glu  
 515 520 525

Leu Ala Asp His Glu Trp His Asp Asp Trp Lys Met Trp Val Gln Lys  
 530 535 540

Lys Lys Tyr Cys Thr Tyr Glu Glu Phe Glu Asp Tyr Leu Ser Ala His  
 545 550 555 560

Gly Gly Val Glu Tyr Asp Glu Asp Gly Val Leu Thr Leu Glu Gly Ala  
 565 570 575

Cys Gly Phe Glu Glu Val Arg Gln Asp Trp Ile Ile Thr Arg Glu Ser  
 580 585 590

Val Asn Leu His Met Lys Glu Trp Glu Ala Ile Gln Arg Asn Glu Ser  
 595 600 605

Thr Thr Glu Tyr Thr Ala Lys Asp Leu Pro Arg Tyr Arg Pro Asp Ser  
 610 615 620

Phe Lys Asn Leu Leu Asp Gly Val Ser Asn His Gly Asn Gly Asn Val  
 625 630 635 640

Gly Lys Ile Glu His Val Lys Leu Glu His Asn Asp  
 645 650

<210> 5  
 <211> 2292  
 <212> ADN  
 <213> *Pichia pastoris*  
 <400> 5

5

aigagtggca atcctttct gttctctct tcaaatttg acttttctgg ttggatcat 60  
 tatagatcca ctgataaaga tcacttagct ctgatgttc tggattatga caaaaatcac 120  
 ttctctcca gaaactcccc cagtttgaaa tctcgtatic acttttatcg acataaattg 180  
 accactagaa agcaaattgg acttttcagc ggcagactga agcttttgt gcttgctctc 240  
 ttgtgttga tcacatttc tgcaatccac attccaatcc cttctcttt ggatattcta 300  
 ggttcccatg tcaaatacct gcccttacga gagaaagtcg atccggaaga ggcactccat 360  
 ctgcacggac tggatctctc ggtagcagag ctacctttt tcaatgatga catgatgtct 420  
 gaatttaact acgatcctag actaccacc gctttgatt tgaagttagt gttagatcat 480  
 ataagtglc gtaatggaac gttgatgct aagttaagg tccccttaa ctggaaactt 540  
 tgggtggatt tgcattcaag gttagttcca tctaatagtt ggtataatcg atttogaata 600  
 ccctcaggtc gtttcgaaac atgcgatgaa ttaagaggt tttcggaaat cactaagaat 660  
 cactttgaa cagacctga taattgctt gatatcgagt atgatactoc ggaaggttat 720  
 ccaaagttca aagtttgca tgcggaagat aaagctctc cttatgaagc acgtatcatt 780  
 tatggtgctt cttacctta ccacgaagca cagaatccta aaaggtgat atttttagga 840

10

ttgggcaagt ccaatgagtc ttgatctta ccagttgagg caaatgacag ttccaactta 900  
 atgcaattca accacgaata tgcaagaagc ttaacgata aaccttcgt ttctctgag 960  
 gaactgtca agaaggficc acigaccttg aattgaata gtgataaggt gctaccaatc 1020  
 aatgaactgg acgitalcaa agacaccccg cgcttaatga atcacaacaa ccagggactg 1080  
 agcatagaca agagctcatt tcaatgggat ctggaaaggg aattacagtt gtiagaacat 1140  
 agaaccagtc aagftaatga cgtggaaggg ctgatgcgg gtatttattc aacaattcaa 1200  
 tgtgaaatgc gctctatga cgattttca aaatacttcc atgaatcaaa agtctctggt 1260  
 aaalatcttc cttctggaga gcactatgac tggcgatttt ttaatggttt ttacctttct 1320  
 cagcaggaga atctagctgt cctgcacagg ttaggaagag catggctacg ctttctctgt 1380  
 gctgctgggt tacatacatg gattgctcac gggacactgt tgggttgga ttggaatggt 1440  
 ctgattctgc cgtgggatca ggatcttgat gtcaaataga ctgtacaatc attgtatctg 1500  
 ttggaagga attcaacag ctctcttga actgatgta gtattgaaga tggctacagc 1560  
 tcagcattgg gacattacta tattgacgtt ggatcctcct tctttgttag ggataaacta 1620  
 aatggtaaca atgctataga tgcacgttc gttgatactg agaccgggtt gtatgtgat 1680  
 ataactgcat tggctttac agatcactta aaactaaaac tcactaccaa agagaaagt 1740  
 gagctacaga aggttatgga tccaatgta aaggaaaaat tgcagtggt caaaaataaa 1800  
 tattcaacgg ccacgctacc ggggtgata gaaacagata ggaataaagt atctgatgcg 1860  
 ctagagaagc aatticatga ttcaagtc gacaatttg tcaacaaaga gttgtttcac 1920  
 tgtcgaaata accatttcta caaatatgga gaggttggcc gattacggag cactatgttt 1980  
 gagggcgficc ctgccctat accattgaa ttgagtcca tactgaaacg agaatacct 2040  
 aaaggictaa cttgaagca ttctccaat ctttttggg atccagtga ccgattgtgg 2100  
 glaccagaaa agaagaaaaa aattagacac atagagtttt cacttacgaa ggaagttaca 2160  
 gaaagccaca agaagaact tgcacagatc catgggaacg aaacgggtat aacctccgac 2220  
 ttgcataatt ctctttcag aatagatccc tggctgtctc gatacaggaa aaaaatgact 2280  
 aggagccaat aa 2292

<210> 6  
 <211> 763  
 <212> PRT  
 <213> *Pichia pastoris*

<400> 6

Met Ser Gly Asn Pro Phe Leu Phe Ser Pro Ser Asn Phe Asp Phe Ser  
 1            5            10            15

Gly Leu Asp His Tyr Arg Ser Thr Asp Lys Asp His Leu Ala Leu Asp  
           20            25            30

Val Leu Asp Tyr Asp Lys Asn His Phe Phe Ser Arg Asn Ser Pro Ser  
           35            40            45

Leu Lys Ser Arg Ile His Phe Tyr Arg His Lys Leu Thr Thr Arg Lys  
           50            55            60

Gln Ile Gly Leu Phe Ser Gly Arg Leu Lys Leu Phe Val Leu Ala Leu  
 65            70            75            80

Phe Val Leu Ile Thr Phe Ser Ala Ile His Ile Pro Ile Pro Phe Ser  
           85            90            95

Leu Asp Ile Leu Gly Ser His Val Lys Tyr Leu Pro Leu Arg Glu Lys  
           100            105            110

Val Asp Pro Glu Glu Ala Leu His Leu His Gly Leu Asp Leu Ser Val  
           115            120            125

Ala Glu Leu Pro Phe Phe Asn Asp Asp Met Met Ser Glu Phe Asn Tyr  
           130            135            140

Asp Pro Arg Leu Pro Thr Ala Leu Ile Leu Lys Leu Val Leu Asp His  
 145            150            155            160

Ile Ser Val Arg Asn Gly Thr Phe Asp Ala Lys Phe Lys Val Pro Phe  
           165            170            175

Asn Trp Lys Leu Trp Val Asp Leu His Ser Arg Leu Val Pro Ser Asn  
180 185 190

Ser Trp Tyr Asn Arg Phe Arg Leu Pro Ser Gly Arg Phe Glu Thr Cys  
195 200 205

Asp Glu Phe Lys Arg Phe Phe Gly Ile Thr Lys Asn His Phe Gly Thr  
210 215 220

Asp Leu Asp Asn Cys Val Asp Ile Glu Tyr Asp Thr Pro Glu Gly Tyr  
225 230 235 240

Pro Lys Phe Lys Val Leu His Ala Glu Asp Lys Ala Leu Pro Tyr Glu  
245 250 255

Ala Arg Ile Ile Tyr Gly Ala Ser Tyr Leu Tyr His Glu Ala Gln Asn  
260 265 270

Pro Lys Arg Leu Ile Phe Leu Gly Leu Gly Lys Ser Asn Glu Ser Leu  
275 280 285

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

His Glu Tyr Ala Arg Ser Phe Asn Asp Gln Pro Phe Val Ser Leu Glu  
305 310 315 320

Glu Leu Val Lys Lys Val Ser Leu Thr Leu Asn Leu Asn Ser Asp Lys  
325 330 335

Val Leu Pro Ile Asn Glu Leu Asp Val Ile Lys Asp Thr Pro Arg Leu  
340 345 350

Met Asn His Asn Asn Gln Gly Leu Ser Ile Asp Lys Ser Ser Phe Gln  
355 360 365

Trp Asp Leu Glu Arg Glu Leu Gln Leu Leu Glu His Arg Thr Ser Gln  
370 375 380



Val Asn Asp Val Glu Gly Leu Asp Ala Gly Ile Tyr Ser Thr Ile Gln  
 385 390 395 400

Cys Glu Met Arg Ser Met Tyr Asp Phe Ser Lys Tyr Phe His Glu Ser  
 405 410 415

Lys Val Ser Gly Lys Tyr Leu Pro Ser Gly Glu His Tyr Asp Trp Arg  
 420 425 430

Phe Phe Asn Gly Phe Tyr Leu Ser Gln Gln Glu Asn Leu Ala Val Leu  
 435 440 445

His Arg Leu Gly Arg Ala Trp Leu Arg Phe Ser Arg Ala Ala Gly Leu  
 450 455 460

His Thr Trp Ile Ala His Gly Thr Leu Leu Gly Trp Tyr Trp Asn Gly  
 465 470 475 480

Leu Ile Leu Pro Trp Asp Gln Asp Leu Asp Val Gln Met Thr Val Gln  
 485 490 495

Ser Leu Tyr Leu Leu Gly Arg Asn Phe Asn Ser Ser Leu Val Thr Asp  
 500 505 510

Val Ser Ile Glu Asp Gly Tyr Ser Ser Ala Leu Gly His Tyr Tyr Ile  
 515 520 525

Asp Val Gly Ser Ser Phe Phe Val Arg Asp Lys Leu Asn Gly Asn Asn  
 530 535 540

Ala Ile Asp Ala Arg Phe Val Asp Thr Glu Thr Gly Leu Tyr Val Asp  
 545 550 555 560

Ile Thr Ala Leu Ala Phe Thr Asp His Leu Lys Leu Lys Leu Thr Thr  
 565 570 575

Lys Glu Lys Val Glu Leu Gln Lys Val Met Asp Pro Asn Val Lys Glu  
 580 585 590

Lys Leu Gln Trp Ile Lys Asn Lys Tyr Ser Thr Ala Thr Leu Pro Gly  
 595 600 605

Val Ile Glu Thr Asp Arg Asn Lys Val Ser Asp Ala Leu Glu Lys Gln  
 610 615 620

Phe His Asp Phe Lys Phe Asp Asn Phe Val Asn Lys Glu Leu Phe His  
 625 630 635 640

Cys Arg Asn Asn His Phe Tyr Lys Tyr Gly Glu Val Gly Arg Leu Arg  
 645 650 655

Ser Thr Met Phe Glu Gly Val Pro Ala Leu Ile Pro Phe Glu Phe Glu  
 660 665 670

Ser Ile Leu Lys Arg Glu Tyr Pro Lys Gly Leu Thr Leu Lys His Phe  
 675 680 685

Ser Asn His Phe Trp Asp Pro Val Asn Arg Leu Trp Val Pro Glu Lys  
 690 695 700

Lys Lys Lys Ile Arg His Ile Glu Phe Ser Leu Thr Lys Glu Val Thr  
 705 710 715 720

Glu Ser His Lys Lys Glu Leu Ala Gln Ile His Gly Asn Glu Thr Gly  
 725 730 735

Ile Thr Ser Asp Phe Ala Tyr Ser Pro Phe Arg Ile Asp Pro Trp Leu  
 740 745 750

Ser Arg Tyr Arg Lys Lys Met Thr Arg Ser Gln  
 755 760

5 <210> 7  
 <211> 33  
 <212> ADN  
 <213> Artificial  
 <220>  
 <223> Cebador oligonucleotídico degradado

10 <400> 7  
 catagcccac tgctaagcca gaattcta atg 33

15 <210> 8  
 <211> 51  
 <212> ADN  
 <213> Artificial

# ES 2 528 739 T3

<220>  
 <223> Cebador oligonucleotídico degradado

<400> 8  
 5 gcagcgtacg aagcttcagc tagaattgta aagtgaatta tcaagcttt c 51

<210> 9  
 <211> 52  
 <212> ADN  
 10 <213> Artificial

<220>  
 <223> Cebador oligonucleotídico degradado

<400> 9  
 15 cagatccact agtggcctat gcaacaatat agcacctctc aaatacacgt tg 52

<210> 10  
 <211> 32  
 20 <212> ADN  
 <213> Artificial

<220>  
 <223> Cebador oligonucleotídico degradado

<400> 10  
 25 tcttgaagta gatttgaga ttttgcgcta tg 32

<210> 11  
 <211> 21  
 30 <212> ADN  
 <213> Artificial

<220>  
 <223> Cebador oligonucleotídico degradado

<400> 11  
 35 agctgaagct tcgtacgctg c 21

<210> 12  
 <211> 22  
 40 <212> ADN  
 <213> Artificial

<220>  
 <223> Cebador oligonucleotídico degradado

<400> 12  
 50 gcataggcca ctagtgatc tg 22

<210> 13  
 <211> 22  
 <212> ADN  
 55 <213> Artificial

<220>  
 <223> Cebador oligonucleotídico degradado

<400> 13  
 60 ttcaacgagt gaccaatgta ga 22

<210> 14  
 <211> 45  
 <212> ADN  
 65 <213> Artificial

ES 2 528 739 T3

<220>  
 <223> Cebador oligonucleotídico degradado  
  
 <400> 14  
 5 ccatccagtg tcgaaaacga gctggcgaac tttctgggt cgaag 45  
  
 <210> 15  
 <211> 21  
 <212> ADN  
 10 <213> Artificial  
  
 <220>  
 <223> Cebador oligonucleotídico degradado  
  
 <400> 15  
 15 tgaagacgtc cccttgaac a 21  
  
 <210> 16  
 <211> 44  
 20 <212> ADN  
 <213> Artificial  
  
 <220>  
 <223> Cebador oligonucleotídico degradado  
 25  
 <400> 16  
 acgaggcaag ctaaacagat ctagtgtt tttctatata aaac 44  
  
 <210> 17  
 <211> 45  
 30 <212> ADN  
 <213> Artificial  
  
 <220>  
 <223> Cebador oligonucleotídico degradado  
 35  
 <400> 17  
 cttcgacca gaaaagtgc ccagctcgtt ttcgacctg gatgg 45  
  
 <210> 18  
 <211> 44  
 40 <212> ADN  
 <213> Artificial  
  
 <220>  
 <223> Cebador oligonucleotídico degradado  
 45  
 <400> 18  
 50 gttttatata gaaaaaaca ctagatctgt ttagcttgcc tcgt 44  
  
 <210> 19  
 <211> 27  
 <212> ADN  
 <213> Artificial  
 55  
 <220>  
 <223> Cebador oligonucleotídico degradado  
  
 <400> 19  
 60 tagtccaagt acgaaacgac actatcg 27  
  
 <210> 20  
 <211> 27  
 <212> ADN  
 65 <213> Artificial

ES 2 528 739 T3

<220>

<223> Cebador oligonucleotídico degradado

5 <400> 20  
acgacggtga gttcaaacag tttggt 27

<210> 21

<211> 26

<212> ADN

10 <213> Artificial

<220>

<223> Cebador oligonucleotídico degradado

15 <400> 21  
tcgctatact gctgtcgatt cgatac 26

<210> 22

<211> 777

20 <212> PRT

<213> *Pichia pastoris*

<400> 22

Met Thr Leu Arg Ser Ala Ile Lys Ala Arg Thr Ser Lys Gly Leu Ile  
1 5 10 15

Gly Ala Val Ile Ile Ala Ser Ile Ile Phe Phe Thr Thr Val Thr Phe  
20 25 30

Tyr Asp Glu Ser Lys Ile Val Gly Ile Ile Arg Val Ser Asp Thr Tyr  
35 40 45

Thr Gly His Ser Ala Val Ser Ser Thr Phe Asn Ala Ser Ser Val Val  
50 55 60

Ser Asp Asn Lys Ile Asn Gly Tyr Gly Leu Pro Leu Ile Asp Thr Glu  
65 70 75 80

Ser Asn Ser Arg Tyr Glu Asp Pro Asp Asp Ile Ser Ile Glu Asn Glu  
85 90 95

25

Leu Arg Tyr Arg Ile Ala Gln Ser Thr Lys Glu Glu Glu Asn Met Trp  
100 105 110

Lys Leu Asp Thr Thr Leu Thr Glu Ala Ser Leu Lys Ile Pro Asn Ile  
115 120 125

Gln Ser Phe Glu Leu Gln Pro Phe Lys Glu Arg Leu Asp Asn Ser Leu  
130 135 140

Tyr Asn Ser Lys Asn Ile Gly Asn Phe Tyr Phe Tyr Asp Pro Arg Leu  
145 150 155 160

Thr Phe Ser Val Tyr Leu Lys Tyr Ile Lys Asp Lys Leu Ala Ser Gly  
165 170 175

Ser Thr Thr Asn Leu Thr Ile Pro Phe Asn Trp Ala His Phe Arg Asp  
180 185 190

Leu Ser Ser Leu Asn Pro Tyr Leu Asp Ile Lys Gln Glu Asp Lys Val  
195 200 205

Ala Cys Asp Tyr Phe Tyr Glu Ser Ser Asn Lys Asp Lys Arg Lys Pro  
210 215 220

Thr Gly Asn Cys Ile Glu Phe Lys Asp Val Arg Asp Glu His Leu Ile  
225 230 235 240

Gln Tyr Gly Ile Ser Ser Lys Asp His Leu Pro Gly Pro Phe Ile Leu  
245 250 255

Lys Ser Leu Gly Ile Pro Met Gln His Thr Ala Lys Arg Leu Glu Ser  
260 265 270

Asn Leu Tyr Leu Leu Thr Gly Ala Pro Val Pro Leu Ser Leu Ser Phe  
275 280 285

Met Thr Lys Lys Gly Leu Tyr Gln Val Gly Val Asp Gln Thr Gly Lys

290                    295                    300  
 Leu Asp Pro Asn Ile Ala Arg Thr Glu Leu Trp Glu Phe Tyr Lys Asn  
 305                    310                    315                    320  
  
 Gly Lys Glu Asn Leu Gln Phe Asn Ala Gln Glu Glu Leu Ser His Leu  
                   325                    330                    335  
  
 Ile Glu Thr Val Pro Ser Ser Ser Asn Ser Ser Ser Gly Glu Gly Tyr  
                   340                    345                    350  
  
 Phe Thr Thr Glu Leu Lys Glu Asn Asn Phe Glu Leu Pro Leu Ser Lys  
                   355                    360                    365  
  
 Asn Asp Phe Thr Phe Asp Asp Ser Glu Val Glu Ser Leu Ile Lys Gly  
                   370                    375                    380  
  
 Leu Ser Glu Gln Asp Leu Asp Leu His Thr Gln Arg Tyr Lys Glu Ser  
 385                    390                    395                    400  
  
 Leu Gln Tyr Ser Phe Ala Thr Arg Glu Asn Asp Val Lys Lys Tyr Phe  
                   405                    410                    415  
  
 Tyr Glu Ala Arg Met Ile Ile Asn Thr Val Asn Lys Glu Gly Gly Ala  
                   420                    425                    430  
  
 His Tyr Asp Trp Arg Phe Phe Asn Gly Ala Met Asn His Glu Ser Ser  
                   435                    440                    445  
  
 Gly Phe Thr Glu Glu Glu Arg Gln Leu Arg Lys Arg Ser Val Leu His  
                   450                    455                    460  
  
 Arg Leu Leu Arg Asn Trp Leu Val Phe Asn Tyr Gln Gln Gly Ser Pro  
 465                    470                    475                    480  
  
 Thr Trp Leu Ala His Gly Thr Leu Leu Ser Trp Tyr Trp Asn Ser Leu  
                   485                    490                    495

Met Phe Pro Trp Asp Tyr Asp Ile Asp Val Gln Met Pro Ile Lys Ser  
500 505 510

Leu Asn Asn Leu Cys Ala Asn Phe Asn Gln Ser Leu Ile Ile Glu Asp  
515 520 525

Leu Thr Glu Gly Tyr Ser Ser Phe Phe Leu Asp Cys Gly Ser Ser Ile  
530 535 540

Thr His Arg Thr Lys Gly Lys Gly Leu Asn Phe Ile Asp Ala Arg Phe  
545 550 555 560

Ile Asn Val Glu Thr Gly Leu Tyr Ile Asp Ile Thr Gly Leu Ser Thr  
565 570 575

Ser Gln Ser Ala Arg Pro Pro Arg Phe Ser Asn Ala Ser Lys Lys Asp  
580 585 590

Pro Ile Tyr Asn Cys Arg Asn Asn His Phe Tyr Ser His Asn Asn Ile  
595 600 605

Ala Pro Leu Lys Tyr Thr Leu Met Glu Gly Val Pro Ser Phe Ile Pro  
610 615 620

Gln Gln Tyr Glu Glu Ile Leu Arg Glu Glu Tyr Thr Thr Gly Leu Thr  
625 630 635 640

Ser Lys His Tyr Asn Gly Asn Phe Phe Met Thr Gln Leu Asn Leu Trp  
645 650 655

Leu Glu Arg Asp Pro Met Leu Ala Leu Val Pro Ser Ser Lys Tyr Glu  
660 665 670

Ile Glu Gly Gly Gly Val Asp His Asn Lys Ile Ile Lys Ser Ile Leu  
675 680 685

Glu Leu Ser Asn Ile Lys Lys Leu Glu Leu Leu Asp Asp Asn Pro Asp  
690 695 700



Ile Leu Glu Glu Val Ile Arg Thr Tyr Glu Leu Thr Ser Ile His His  
 705                      710                      715                      720

Lys Glu Met Gln Tyr Leu Ser Ser Val Lys Pro Asp Gly Asp Arg Ser  
                     725                      730                      735

Met Gln Ser Asn Asp Ile Thr Ser Ser Tyr Gln Glu Phe Leu Ala Ser  
                     740                      745                      750

Leu Lys Lys Phe Gln Pro Leu Arg Lys Asp Leu Phe Gln Phe Glu Arg  
                     755                      760                      765

Ile Asp Leu Ser Lys His Arg Lys Gln  
 770                      775

<210> 23  
 <211> 1178  
 <212> PRT  
 <213> *Saccharomyces cerevisiae*

5

<400> 23

Met Leu Gln Arg Ile Ser Ser Lys Leu His Arg Arg Phe Leu Ser Gly  
 1                      5                      10                      15

Leu Leu Arg Val Lys His Tyr Pro Leu Arg Arg Ile Leu Leu Pro Leu  
                     20                      25                      30

Ile Leu Leu Gln Ile Ile Ile Ile Thr Phe Ile Trp Ser Asn Ser Pro  
                     35                      40                      45

Gln Arg Asn Gly Leu Gly Arg Asp Ala Asp Tyr Leu Leu Pro Asn Tyr  
                     50                      55                      60

Asn Glu Leu Asp Ser Asp Asp Asp Ser Trp Tyr Ser Ile Leu Thr Ser  
 65                      70                      75                      80

Ser Phe Lys Asn Asp Arg Lys Ile Gln Phe Ala Lys Thr Leu Tyr Glu  
                     85                      90                      95

10

Asn Leu Lys Phe Gly Thr Asn Pro Lys Trp Val Asn Glu Tyr Thr Leu  
100 105 110

Gln Asn Asp Leu Leu Ser Val Lys Met Gly Pro Arg Lys Gly Ser Lys  
115 120 125

Leu Glu Ser Val Asp Glu Leu Lys Phe Tyr Asp Phe Asp Pro Arg Leu  
130 135 140

Thr Trp Ser Val Val Leu Asn His Leu Gln Asn Asn Asp Ala Asp Gln  
145 150 155 160

Pro Glu Lys Leu Pro Phe Ser Trp Tyr Asp Trp Thr Thr Phe His Glu  
165 170 175

Leu Asn Lys Leu Ile Ser Ile Asp Lys Thr Val Leu Pro Cys Asn Phe  
180 185 190

Leu Phe Gln Ser Ala Phe Asp Lys Glu Ser Leu Glu Ala Ile Glu Thr  
195 200 205

Glu Leu Gly Glu Pro Leu Phe Leu Tyr Glu Arg Pro Lys Tyr Ala Gln  
210 215 220

Lys Leu Trp Tyr Lys Ala Ala Arg Asn Gln Asp Arg Ile Lys Asp Ser  
225 230 235 240

Lys Glu Leu Lys Lys His Cys Ser Lys Leu Phe Thr Pro Asp Gly His  
245 250 255

Gly Ser Pro Lys Gly Leu Arg Phe Asn Thr Gln Phe Gln Ile Lys Glu  
260 265 270

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

Tyr Ile Leu Thr Thr Gln Ser His Pro Leu Ser Ile Ser Ile Ile Glu  
290 295 300

Ser Asp Asn Ser Thr Tyr Gln Val Pro Leu Gln Thr Glu Lys Ser Lys  
 305                    310                    315                    320

Asn Leu Val Gln Ser Gly Leu Leu Gln Glu Tyr Ile Asn Asp Asn Ile  
 325                    330                    335

Asn Ser Thr Asn Lys Arg Lys Lys Asn Lys Gln Asp Val Glu Phe Asn  
 340                    345                    350

His Asn Arg Leu Phe Gln Glu Phe Val Asn Asn Asp Gln Val Asn Ser  
 355                    360                    365

Leu Tyr Lys Leu Glu Ile Glu Glu Thr Asp Lys Phe Thr Phe Asp Lys  
 370                    375                    380

Asp Leu Val Tyr Leu Ser Pro Ser Asp Phe Lys Phe Asp Ala Ser Lys  
 385                    390                    395                    400

Lys Ile Glu Glu Leu Glu Glu Gln Lys Lys Leu Tyr Pro Asp Lys Phe  
 405                    410                    415

Ser Ala His Asn Glu Asn Tyr Leu Asn Ser Leu Lys Asn Ser Val Lys  
 420                    425                    430

Thr Ser Pro Ala Leu Gln Arg Lys Phe Phe Tyr Glu Ala Gly Ala Val  
 435                    440                    445

Lys Gln Tyr Lys Gly Met Gly Phe His Arg Asp Lys Arg Phe Phe Asn  
 450                    455                    460

Val Asp Thr Leu Ile Asn Asp Lys Gln Glu Tyr Gln Ala Arg Leu Asn  
 465                    470                    475                    480

Ser Met Ile Arg Thr Phe Gln Lys Phe Thr Lys Ala Asn Gly Ile Ile  
 485                    490                    495

Ser Trp Leu Ser His Gly Thr Leu Tyr Gly Tyr Leu Tyr Asn Gly Met

500

505

510

Ala Phe Pro Trp Asp Asn Asp Phe Asp Leu Gln Met Pro Ile Lys His  
515 520 525

Leu Gln Leu Leu Ser Gln Tyr Phe Asn Gln Ser Leu Ile Leu Glu Asp  
530 535 540

Pro Arg Gln Gly Asn Gly Arg Tyr Phe Leu Asp Val Ser Asp Ser Leu  
545 550 555 560

Thr Val Arg Ile Asn Gly Asn Gly Lys Asn Asn Ile Asp Ala Arg Phe  
565 570 575

Ile Asp Val Asp Thr Gly Leu Tyr Ile Asp Ile Thr Gly Leu Ala Ser  
580 585 590

Thr Ser Ala Pro Ser Arg Asp Tyr Leu Asn Ser Tyr Ile Glu Glu Arg  
595 600 605

Leu Gln Glu Glu His Leu Asp Ile Asn Asn Ile Pro Glu Ser Asn Gly  
610 615 620

Glu Thr Ala Thr Leu Pro Asp Lys Val Asp Asp Gly Leu Val Asn Met  
625 630 635 640

Ala Thr Leu Asn Ile Thr Glu Leu Arg Asp Tyr Ile Thr Ser Asp Glu  
645 650 655

Asn Lys Asn His Lys Arg Val Pro Thr Asp Thr Asp Leu Lys Asp Leu  
660 665 670

Leu Lys Lys Glu Leu Glu Glu Leu Pro Lys Ser Lys Thr Ile Glu Asn  
675 680 685

Lys Leu Asn Pro Lys Gln Arg Tyr Phe Leu Asn Glu Lys Leu Lys Leu  
690 695 700

Tyr Asn Cys Arg Asn Asn His Phe Asn Ser Phe Glu Glu Leu Ser Pro  
 705                      710                      715                      720

Leu Ile Asn Thr Val Phe His Gly Val Pro Ala Leu Ile Pro His Arg  
                                  725                      730                      735

His Thr Tyr Cys Leu His Asn Glu Tyr His Val Pro Asp Arg Tyr Ala  
                                  740                      745                      750

Phe Asp Ala Tyr Lys Asn Thr Ala Tyr Leu Pro Glu Phe Arg Phe Trp  
                                  755                      760                      765

Phe Asp Tyr Asp Gly Leu Lys Lys Cys Ser Asn Ile Asn Ser Trp Tyr  
                                  770                      775                      780

Pro Asn Ile Pro Ser Ile Asn Ser Trp Asn Pro Asn Leu Leu Lys Glu  
 785                      790                      795                      800

Ile Ser Ser Thr Lys Phe Glu Ser Lys Leu Phe Asp Ser Asn Lys Val  
                                  805                      810                      815

Ser Glu Tyr Ser Phe Lys Asn Leu Ser Met Asp Asp Val Arg Leu Ile  
                                  820                      825                      830

Tyr Lys Asn Ile Pro Lys Ala Gly Phe Ile Glu Val Phe Thr Asn Leu  
                                  835                      840                      845

Tyr Asn Ser Phe Asn Val Thr Ala Tyr Arg Gln Lys Glu Leu Glu Ile  
                                  850                      855                      860

Gln Tyr Cys Gln Asn Leu Thr Phe Ile Glu Lys Lys Lys Leu Leu His  
 865                      870                      875                      880

Gln Leu Arg Ile Asn Val Ala Pro Lys Leu Ser Ser Pro Ala Lys Asp  
                                  885                      890                      895

Pro Phe Leu Phe Gly Tyr Glu Lys Ala Met Trp Lys Asp Leu Ser Lys  
                                  900                      905                      910

Ser Met Asn Gln Thr Thr Leu Asp Gln Val Thr Lys Ile Val His Glu  
 915                      920                      925

Glu Tyr Val Gly Lys Ile Ile Asp Leu Ser Glu Ser Leu Lys Tyr Arg  
 930                      935                      940

Asn Phe Ser Leu Phe Asn Ile Thr Phe Asp Glu Thr Gly Thr Thr Leu  
 945                      950                      955                      960

Asp Asp Asn Thr Glu Asp Tyr Thr Pro Ala Asn Thr Val Glu Val Asn  
 965                      970                      975

Pro Val Asp Phe Lys Ser Asn Leu Asn Phe Ser Ser Asn Ser Phe Leu  
 980                      985                      990

Asp Leu Asn Ser Tyr Gly Leu Asp Leu Phe Ala Pro Thr Leu Ser Asp  
 995                      1000                      1005

Val Asn Arg Lys Gly Ile Gln Met Phe Asp Lys Asp Pro Ile Ile  
 1010                      1015                      1020

Val Tyr Glu Asp Tyr Ala Tyr Ala Lys Leu Leu Glu Glu Arg Lys  
 1025                      1030                      1035

Arg Arg Glu Lys Lys Lys Lys Glu Glu Glu Glu Lys Lys Lys Lys  
 1040                      1045                      1050

Glu Glu Glu Glu Lys Lys Lys Lys Glu Glu Glu Glu Lys Lys Lys  
 1055                      1060                      1065

Lys Glu Glu Glu Glu Lys Lys Lys Lys Glu Glu Glu Glu Lys Lys  
 1070                      1075                      1080

Lys Lys Glu Glu Glu Glu Lys Lys Lys Gln Glu Glu Glu Glu Lys  
 1085                      1090                      1095

Lys Lys Lys Glu Glu Glu Glu Lys Lys Lys Gln Glu Glu Gly Glu  
 1100                      1105                      1110

Lys Met Lys Asn Glu Asp Glu Glu Asn Lys Lys Asn Glu Asp Glu  
 1115 1120 1125

Glu Lys Lys Lys Asn Glu Glu Glu Glu Lys Lys Lys Gln Glu Glu  
 1130 1135 1140

Lys Asn Lys Lys Asn Glu Asp Glu Glu Lys Lys Lys Gln Glu Glu  
 1145 1150 1155

Glu Glu Lys Lys Lys Asn Glu Glu Glu Lys Lys Lys Gln Glu  
 1160 1165 1170

Glu Gly His Ser Asn  
 1175

<210> 24  
 <211> 346  
 <212> PRT  
 <213> *Neurospora crassa*  
 <400> 24

5

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

Trp Arg Asp Arg Leu Leu Thr Leu Leu Met Ala Leu Thr Phe Val Leu  
 20 25 30

Ser Ser Leu Ala Ser Pro Leu Pro Ile Glu Gly Ala Val Val Lys Ala  
 35 40 45

Asn Asn Asn Asp Ala Val Ser Gln Pro Gln Ala Gln Ala Gln Ala Lys  
 50 55 60

Ala Glu Val Arg Gln Phe Ser Ala Pro Ala Gln Ala Gln Glu Ala Glu  
 65 70 75 80

Pro Ala Thr Ala Thr Thr Ser Asp Asp Thr Thr Asn Thr Asn Thr Asp  
 85 90 95

10

Asp Asp Asp Pro Leu Leu Pro Glu Arg Lys Tyr Phe His Glu Pro Gly  
100 105 110

Trp Thr Glu Glu Leu Ser His Tyr Asp Thr Arg Phe Phe Thr Ser Pro  
115 120 125

Val Pro Tyr Asp Pro His Leu Val His Leu Arg His Leu Ile Arg Ser  
130 135 140

Tyr Leu Leu Met Thr Ser Ser Arg Ser Leu Thr Thr Trp Leu Ala His  
145 150 155 160

Gly Thr Leu Leu Gly Trp Tyr Trp Asn Gly Ala Ile Met Pro Trp Asp  
165 170 175

Tyr Asp Leu Asp Val Gln Val Ser Asn Ile Thr Leu Gly Gln Met Ala  
180 185 190

Arg Asp Trp Asn Gln Thr Thr Phe Asp Tyr Val Tyr Thr Leu Ser Glu  
195 200 205

Glu Glu Glu Lys Glu Gly Leu Gly Lys Gln Gly Glu Val Thr Val Lys  
210 215 220

Lys Tyr Leu Leu Asp Val Asn Pro Tyr Trp Ala Gln Arg Thr Arg Leu  
225 230 235 240

Glu Gly Met Asn Val Ile Asp Ala Arg Trp Ile Asp Met Glu Asn Gly  
245 250 255

Met Tyr Val Asp Ile Thr Gly Leu Ser Glu Asp Arg Glu Glu Thr Gly  
260 265 270

Thr Arg Gln Gly Val Trp Ser Asp Lys Asn Tyr His Gly Tyr Gly Thr  
275 280 285

Arg Gln Ile Trp Pro Leu Arg Arg Thr Glu Phe Glu Gly Val Glu Ala



290 295 300

Trp Val Pro Trp Asp Val Glu Glu Ile Leu Lys Glu Glu Tyr Gly Val  
 305 310 315 320

Lys Ser Leu Thr Glu Glu Ser Phe Ala Gly His Gln Phe Asp His Gly  
 325 330 335

Arg Lys Gln Trp Val Lys Thr Glu Leu Ala  
 340 345

<210> 25  
 <211> 217  
 <212> PRT  
 <213> *Neurospora crassa*

5

<400> 25

Met Leu Leu Asn Trp Leu Leu Ile Val Thr Thr Leu Phe Pro Leu Ser  
 1 5 10 15

Thr Cys Ala Pro Ile Glu His Glu Asp Ala Ile Ala Ile Glu Gly Met  
 20 25 30

Ala Asp Gln Lys Asp Met Ser Gly Lys Ala Gly Asp Pro Pro Gln Lys  
 35 40 45

Tyr Phe His Glu Ser Thr Phe Ala Glu Lys Ala Leu Gly Tyr Gln Glu  
 50 55 60

Gln Lys Ala Ala Leu Lys Asn Leu Val Arg Thr Phe Leu Glu Thr Met  
 65 70 75 80

Arg Asp Leu Gly Ile Glu Thr Trp Leu Met His Gly Ser Leu Leu Gly  
 85 90 95

Trp Trp Trp Asn Lys Gln Ile Met Pro Trp Asp Ser Asp Ala Asp Val  
 100 105 110

Gln Val Thr Glu Ala Ser Met Tyr Phe Leu Ala Thr Tyr Tyr Asn Met

10

115                    120                    125  
 Ser Val Phe His Tyr Lys Thr Pro Arg Leu Pro Ala Gly Arg Asn Tyr  
 130                    135                    140  
 Met Leu Glu Val Asn Pro Asn Phe Ser Asn Gly Asp Gln Ser Asp Trp  
 145                    150                    155                    160  
 Leu Asn Val Ile Asp Ala Arg Trp Ile Asp Thr Glu Ser Gly Leu Phe  
 165                    170                    175  
 Ile Asp Ile Thr Thr Ala Arg Tyr Asn Leu Thr His Pro Ala Gly Glu  
 180                    185                    190  
 Gly Met Met Ser Cys Lys Asp Gly His Glu Phe Arg Val Thr Ile Ser  
 195                    200                    205  
 Thr Ser Val Lys Ser Ser Gly Gly Gly  
 210                    215

5 <210> 26  
 <211> 311  
 <212> PRT  
 <213> *Aspergillus niger*  
 <400> 26

Met His Lys Lys Ala Thr Leu Ala Leu Ala Ser Ala Ile Cys Ile Thr  
 1                    5                    10                    15  
 Ala Ala Thr Gly Leu Pro Gly Pro Val Leu Asp Ser Ala Pro Lys Ala  
 20                    25                    30  
 Ser Val His Gly Ser Val His Gly Ser Ile Leu Gly Thr Ala Ala Asp  
 35                    40                    45  
 Ile Asn Asp Pro Ser Tyr Leu Trp Thr Met Tyr Gly Leu Asn Thr Ser  
 50                    55                    60  
 Glu Glu Tyr Lys Tyr Phe Gln Glu Pro Gly Asn Asp Glu Ile His Ala

10

65                    70                    75                    80  
 His Tyr Asp Ser Arg Phe Phe Lys Asp Pro Val Pro Lys Glu His Arg  
                           85                    90                    95  
 Ser Gln Val Leu Thr His Ile Ile His Ser Tyr Phe Glu Phe Phe Asn  
                           100                    105                    110  
 Ser His Asn Leu Glu Thr Trp Leu Ala His Gly Thr Leu Leu Gly Trp  
                           115                    120                    125  
 Trp Trp Asn Gly Arg Ile Met Pro Trp Asp Trp Asp Ile Asp Thr Gln  
                           130                    135                    140  
 Val Ser Glu Ala Thr Leu Phe Arg Leu Ala Asp Glu Phe Asn Gly Thr  
                           145                    150                    155                    160  
 Val Ala Gln Tyr Asn Thr Thr Asn Pro Asp Thr Gln His Ser Tyr Leu  
                           165                    170                    175  
 Leu Asp Val Asn Pro Trp Ala Arg Gln Arg Asp Arg Gly Lys Gly Leu  
                           180                    185                    190  
 Asn Ile Ile Asp Ala Arg Trp Ile Asp Met Gln Thr Gly Leu Tyr Ile  
                           195                    200                    205  
 Asp Ile Thr Gly Leu Ser Lys Leu Asn Glu Glu Lys Pro Asn Glu Trp  
                           210                    215                    220  
 Gly Cys Lys Asn Asn His Asn Tyr Met Leu Ser Asp Ile Tyr Pro Leu  
                           225                    230                    235                    240  
 Arg Ala Ser Phe Phe Glu Gly Val Ala Ala Lys Val Pro Tyr Arg Tyr  
                           245                    250                    255  
 Glu Ser Val Leu Ile Asp Glu Tyr Gly Glu Lys Ala Leu Ser Glu Thr  
                           260                    265                    270

His Tyr Asn Asp Tyr Thr Trp Val Ser Lys Gln Glu Glu Trp Val Ser  
275 280 285

Asp Glu Ile Ile Ala Ala Glu Glu Lys Lys Lys Ala Lys Glu Gly Asp  
290 295 300

Lys Asp Gly Arg Gln Tyr Glu  
305 310

5 <210> 27  
<211> 261  
<212> PRT  
<213> *Aspergillus niger*  
  
<400> 27

Met Arg Leu Ser Thr Tyr Pro Ile Leu Phe Ala Phe Cys Gly Leu Ala  
1 5 10 15

Ser Val Arg Gly Glu Gly Glu Ile Thr Phe Glu Asp Val Arg Asp Lys  
20 25 30

Leu Pro Lys Thr Tyr Ser Gly Gln Gly Gly Glu Pro Gly Pro Lys Tyr  
35 40 45

Phe Lys Glu Ser Ser Phe Ala Glu Ser Val Leu Pro Glu Glu Glu Thr  
50 55 60

Leu Pro His Leu Ser Ala Leu Ile Gln Thr Tyr Leu Ser Thr Met Ala  
65 70 75 80

Asp Leu Gly Ala Glu Thr Trp Ile Met His Gly Ser Leu Leu Ala Trp  
85 90 95

Trp Trp Asn Gln Lys Ile Phe Pro Trp Asp Asn Asp Leu Asp Val Gln  
100 105 110

Ile Asn Glu Pro Thr Ile His Phe Leu Ala Asp Tyr Tyr Asn Met Thr  
115 120 125

10

Glu His His Phe Asp Leu Pro Asp Val Glu Gly Gly Arg Thr Tyr Leu  
130 135 140

Leu Glu Ile Asn Pro Asn Tyr Val Val Arg Ser Lys Leu Asp Lys Ala  
145 150 155 160

Asn Val Ile Asp Gly Arg Trp Ile Asp Thr Ser Ser Gly Leu Phe Ile  
165 170 175

Asp Ile Thr Ala Val Arg Ala Asp Asp Glu Arg Arg Ala Asn Gly Gln  
180 185 190

Pro Gly Ala Leu Met Cys Lys Asp Arg His Asn Phe Asp Glu Ser Glu  
195 200 205

Ile Tyr Pro Leu Arg Asn Ser Tyr Phe Glu Asp Val Pro Ala Lys Ile  
210 215 220

Pro Tyr Ala Tyr Thr Lys Leu Leu Gln Asp Glu Tyr Gly Ala Lys Ala  
225 230 235 240

Leu Thr Lys Thr Asn Tyr Gln Gly Cys Val Ile Leu Gln Glu Val Glu  
245 250 255

Phe Val Val Ser Thr  
260

<210> 28  
<211> 997  
<212> PRT  
<213> *Candida albicans*

5

<400> 28

Met Ser Asn Thr Ile Pro Gln Tyr Phe Ile Arg Ile Phe Asn Leu Ile  
1 5 10 15

Phe Ser Ala Arg Arg Lys Asn Phe Gln Leu Ala Leu Ile Ser Gly Leu  
20 25 30

10

Leu Phe Phe Gly Ser Phe Ala Ile Leu Ser Thr Thr Ser Tyr Ser Lys  
 35 40 45

Lys Phe Asn Tyr Phe Asp Asp Leu Ile Leu Lys Ile Tyr Asp Tyr Asn  
 50 55 60

Tyr Leu Thr Asn Asn Tyr Asn Ile Asp Tyr Leu Ala Lys Asn Asp Pro  
 65 70 75 80

Glu Ala Tyr Phe Asn Val Lys Val Gln Gln Ile Val Asp Glu Lys Lys  
 85 90 95

Gln His Asp Leu Glu Ser Lys Phe Trp Ser Leu Asp Thr Lys Ile Asn  
 100 105 110

Asp Asp Gln Ala Thr Leu Gln Ile Pro Ala Tyr Phe Thr Tyr Asn Lys  
 115 120 125

Pro Arg Asp Asn Lys Asn Leu Glu Asp Ser Glu Gln Ser Ser Lys Pro  
 130 135 140

Val Glu Lys Pro Leu Ile Gln Pro Phe Asp Pro Arg Phe Thr Leu Ala  
 145 150 155 160

Met Tyr Tyr Tyr Tyr Leu Asp Gln Gln Met Thr Thr Ala His His Asp  
 165 170 175

Ser Ser Ser Ser Ser Ser Gly Asn Ser Ile Thr Val Pro Phe Asn Trp  
 180 185 190

Tyr Asp Trp Val Asp Met Ser Val Leu Asn Lys Tyr Leu Leu Ala Pro  
 195 200 205

Asn Lys Asp Lys Pro Asp Cys Ser Ile Leu Asp Ala His Glu Asp Ala  
 210 215 220

Arg Lys Ile Glu Thr Glu Lys Lys Lys Met Glu Lys Leu Ala Lys Gln  
 225 230 235 240

Trp Asp Glu Asn Lys Arg Lys Ala Glu Glu Glu Lys Lys Lys Ala Glu  
245 250 255

Glu Asp Lys Lys Lys Glu Glu Glu Glu Lys Lys Lys Ala Glu Glu Glu  
260 265 270

Glu Lys Gln Arg His Glu Gln Glu Lys Gln Ala Leu Glu Glu Asp Lys  
275 280 285

Lys Lys Leu Glu Glu Glu Lys Lys Lys Ile Glu Glu Glu Lys Asn Lys  
290 295 300

Leu Gln Glu Gln Gln Gln Gln Gln Gln Glu Glu Lys Ala Asn Asp  
305 310 315 320

Gly Asn Gln Glu His Ser Lys Phe Val Lys Arg Asp Asp Glu Ile Lys  
325 330 335

Met Ser Thr Ser Gln Asp Lys Ser Asp Ser Asp Ala Asp Arg Ala Lys  
340 345 350

Ile Asp Met Thr Thr Phe Phe Asn Glu Ala Phe Glu Lys Leu Ser Asp  
355 360 365

Glu Asp Lys Ala Ser Val Ala Lys Asp Val Glu Asp Ala Val Lys Lys  
370 375 380

Ile Thr Gln Pro Ser Ser Trp Cys Val Pro Asn Ala Lys Leu Ser Ile  
385 390 395 400

Asp His Ser Asp Lys Gln Ile Val His Pro Gly Phe Asn Val Phe Lys  
405 410 415

Ser Pro Gly Arg Thr Thr Pro Gln Lys Ala Ile Ile Ala Gly Lys Ser  
420 425 430

Phe Leu Tyr Ser Tyr Ala Pro Pro Pro Ser Ser Ile Leu Phe Leu Thr  
435 440 445

Ser Glu Gly Ser Tyr Ser Val Asn Val Gln His Ser Ala Pro Leu Leu  
450 455 460

Arg Asn Gly Ile Pro Glu Ser Tyr Leu Ala Asn Asn Asn Phe Asp Val  
465 470 475 480

Ser Leu Asn Val Leu Gln Gln Leu His Lys Leu Lys Lys Asn His Lys  
485 490 495

Pro Asp Thr Ala Lys Val Ile Asn Asp Tyr Leu Leu His Ile Pro Lys  
500 505 510

Glu Ser Phe Lys Tyr Asp Pro Asp Ser Ile Ile Phe Asp Tyr Thr Lys  
515 520 525

Arg Leu Asp Lys Gly Glu Lys Leu Thr Ile Lys Glu Leu Lys Tyr Leu  
530 535 540

Gln Ser Leu Glu Tyr Ser Lys Asp Lys Val Ala His Gly Gly Pro Pro  
545 550 555 560

Lys Tyr Phe Ala Glu Ser Arg Leu Ile Gly Thr Thr Val Gly Asp His  
565 570 575

Tyr Asp Trp Arg Phe Phe Asn Gly Val Gln Phe Gly Thr Val Asp Gln  
580 585 590

Ser Leu Thr Leu His Arg Leu Ile Arg Thr Trp Leu Ser Phe Thr Arg  
595 600 605

Lys Ser Gly Ile Thr Thr Trp Ile Ala His Gly Ser Leu Leu Ser Trp  
610 615 620

Tyr Trp Asn Gly Met Ala Phe Pro Trp Asp Asn Asp Ile Asp Val Gln  
625 630 635 640

Val Pro Ile Met Asp Leu His Lys Leu Ser Leu Gln Phe Asn Gln Thr



645 650 655  
Ile Val Val Glu Asp Pro Glu Asp Gly Phe Gly Arg Tyr Phe Leu Asp  
660 665 670  
Ile Gly Ser Phe Ile Thr Leu Arg Glu Lys Gly Asn Gly Asn Asn Asn  
675 680 685  
Ile Asp Ala Arg Phe Ile Asp Ile Asp Thr Gly Leu Tyr Ile Asp Ile  
690 695 700  
Thr Ala Leu Ala Leu Ser Asn Ser Glu Thr Pro Lys Ser Asp Leu Ala  
705 710 715 720  
Glu Leu Pro Lys Asn Phe Glu Ile Lys Asp Asn Asn Tyr Lys Pro Ala  
725 730 735  
Asn Glu Leu Leu Gln Ile Tyr Asn Cys Arg Asn Asn His Phe Asn Ser  
740 745 750  
Tyr Asp Glu Leu Ser Pro Leu Met Lys Ser Ser Val Glu Gly Glu Ile  
755 760 765  
Gly Tyr Ile Pro Ser Arg Tyr Ser Thr Ile Leu Thr Arg Glu Tyr Arg  
770 775 780  
Ser Gly Leu Ser Ser Asn Ser His Gly Gly Tyr Ile Phe Ile Ala Lys  
785 790 795 800  
Leu Arg Leu Trp Val Lys Glu Glu Asp Leu Tyr Tyr Phe Ile Lys His  
805 810 815  
Arg Asp Gln Trp Thr Lys Tyr His Ser Phe Asn Thr Lys Leu Ser Gln  
820 825 830  
Asp Pro Ser Asn Thr Leu Leu Gln Asp Tyr Ser Tyr Leu Met Ser Glu  
835 840 845

Gln Glu Tyr.Glu Asn Leu Gln Tyr Ser Thr Asp Leu Glu His Asp Asn  
850 855 860

Pro Phe Lys Lys Thr Lys Lys Pro Leu Glu Leu Lys Asn Ser Glu Leu  
865 870 875 880

Glu Lys Leu Lys His Met Asn Glu Ser Glu Leu Leu Gln Phe Leu Asn  
885 890 895

Asn Asp Asp Ile Leu Ile Gln Phe Phe Asn Ala Lys Glu Phe Thr Ser  
900 905 910

Phe His Glu Ser Glu Ile Met Gln Leu Thr Phe Gly Lys Ser Thr Ala  
915 920 925

Lys Leu Met Ser Ser Ala Ile Asp Phe Pro Pro Ile Lys Tyr Glu Pro  
930 935 940

Tyr Leu Tyr Lys Leu Asn His Asp Leu Asp Thr Phe Glu Asn Lys Val  
945 950 955 960

Asp Arg Tyr Leu Ala Leu Gln Asp Ala Tyr Gln Gln Glu His Asn Asn  
965 970 975

Ser Pro Ser Gly Gly Ser Asp Asn Gly Phe Met Glu Ile Glu Glu Asp  
980 985 990

Leu Asp Phe Ala Phe  
995

<210> 29  
<211> 303  
<212> PRT  
<213> *Fichia angusta*

<400> 29

Met Leu Val Asn Ser Val Ser Pro Ser Lys Ser Glu Leu Asp Phe Asp  
1 5 10 15

5

10

Ser Leu Phe Thr Ala Glu Lys Asn Tyr Glu Phe Val Ile Pro Pro Asp  
 20 25 30

Arg Phe Asn Tyr Ser Tyr Asp Gln Ile Ile Glu Asn Tyr Glu Lys Arg  
 35 40 45

Ile Asp Glu Leu Asp Glu Lys Gln Leu Arg His Leu Gln Thr Leu Lys  
 50 55 60

Tyr Ser Arg Ser Ile Pro Ser Thr Lys Leu Lys Lys Ser Phe Arg Glu  
 65 70 75 80

Val Asn Ile Asn Trp Pro Ala Thr Tyr Asn Gly His Lys Val Thr Glu  
 85 90 95

Asn Gly Gly His Tyr Asp Phe Arg Phe Phe Asn Gly Phe Val Thr Glu  
 100 105 110

Ser Lys Leu Asn Glu Tyr Asp Asp Val Asn Glu Lys Arg Lys Ile Met  
 115 120 125

Leu His Arg Ile Ile His Thr Trp Leu Gln Phe Thr Tyr Lys Glu Gly  
 130 135 140

Ile Val Ser Phe Leu Ala His Gly Thr Leu Leu Ser Trp Tyr Trp Asn  
 145 150 155 160

Ala Leu Val Phe Glu Trp Asp Asn Asp Ile Asp Val Gln Met Pro Ile  
 165 170 175

Met Asp Phe Asp Arg Phe Cys Met Lys Tyr Asn Asn Ser Leu Ile Val  
 180 185 190

Glu Asp Val Gln His Gly Tyr Gly Lys Tyr Tyr Val Asp Cys Gly Pro  
 195 200 205

Tyr Pro Thr His Arg Thr Lys Gly Asn Gly Arg Asn Asn Ile Asp Ala  
 210 215 220

Arg Phe Ile Asp Val Asp Ser Gly Met Tyr Ile Asp Ile Thr Gly Leu  
225                    230                    235                    240

Ala Leu Thr Asp Thr Ile Lys Ile Pro Pro Arg Leu Glu Arg Leu Asp  
                  245                    250                    255

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

Pro Ala Glu Gln Thr Glu Gly Leu Ser Asp Pro Gly Ala Ser Arg Asn  
                  275                    280                    285

Val Lys Arg Ala Pro Val Lys Ser Asn Lys Gly Pro Glu Val Ser  
                  290                    295                    300

**REIVINDICACIONES**

- 5 1. Un hospedador fúngico que comprende una disrupción, una delección o una mutación combinadas de los genes *MNN4B* y *PNO1* en donde el hospedador fúngico es *Pichia pastoris* y es capaz de producir menos del 1 % de glucoproteínas manosilfosforiladas de los N-glucanos totales en donde el gen de *MNN4B* en el que hay que provocar una disrupción, suprimir o mutar codifica un ARN mensajero que se corresponde con un ADNc que comprende o consiste en una secuencia de ácido nucleico seleccionada del grupo que consiste en
- 10 (a) SEC ID N°: 3;  
(b) una secuencia de ácido nucleico que es una variante degradada de SEC ID N°: 3;  
(c) una secuencia de ácido nucleico al menos 50 %, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 98 %, al menos 99 % o al menos 99,9 % idéntica a SEC ID N°: 3;
- 15 (d) una secuencia de ácido nucleico que codifica un polipéptido que tiene la secuencia de aminoácidos de SEC ID N°: 4; y  
(e) una secuencia de ácido nucleico que codifica un polipéptido al menos 50 %, al menos 60 %, al menos 65 %, al menos 70 %, al menos 75 %, al menos 80 %, al menos 85 %, al menos 90 %, al menos 95 %, al menos 98 %, al menos 99 % o al menos 99,9 % idéntico a SEC ID N°: 4.
- 20 2. Un método para producir composiciones de glucoproteínas en un hospedador fúngico de la reivindicación 1 que comprende propagar dicho hospedador fúngico y aislar los productos de glucoproteínas.

Figura 1 (hoja 1)

1 ATG TTC AAA GAA ACG TCA AAG AAC TTG TTT GGT TCG ATA AAT  
 1▶ M F K E T S K N L F G S I N  
 43 ACC TTC AAT ACG GTG GAG TAT GTC ATG TAT ATG ATG CTA CTA  
 15▶ T F N T V E Y V M Y M M L L  
 85 CTG ACT GCG TAT TTT TTG AAC CAC CTG TTG CAT AGT TTG GAT  
 29▶ L T A Y F L N H L L H S L D  
 127 AAC ATC AAT CAT TTG GTT GAG TCT GAT GTT AAT TAT CAA CTA  
 43▶ N I N H L V E S D V N Y Q L  
 169 CTT CAA AGG GTA ACA AAT AAA GTC AAG CTT TTT GAT GAG GAA  
 57▶ L Q R V T N K V K L F D E E  
 211 GCA GTC TTG CCC TTT GCT AAG AAT CTC AAT AGA AGA ACT GAA  
 71▶ A V L P F A K N L N R R T E  
 253 CGC TTT GAT CCA AGG TTG CCT GTA GCT GCA TAC CTT CGA AGC  
 85▶ R F D P R L P V A A Y L R S  
 295 CTT CAA GAT CAG TAT TCG GAG CTT CCA CAA GGT ACC GAC CTG  
 99▶ L Q D Q Y S E L P Q G T D L  
 337 AAT GAT ATT CCG CCC CTG GAG GTT TCT TTC CAC TGG GAT GAC  
 113▶ N D I P P L E V S F H W D D  
 379 TGG TTA AGT TTG GGA ATT GCA TCA ACC TTT TGG GAC GCC TTC  
 127▶ W L S L G I A S T F W D A F  
 421 GAC AAT TAC AAC AAG AGA CAA GGA GAA AAT GCA ATT TCT TAC  
 141▶ D N Y N K R Q G E N A I S Y  
 463 GAG CAG CTC CAA GCA ATA CTT GPT AAT GAT TTG GAA GAT TTT  
 155▶ E Q L Q A I L V N D L E D F  
 505 TCT CCC TAC ACC GCA CAT ATT CTT CAC AGT AAC GTG GAA GTC  
 169▶ S P Y T A H I L H S N V E V  
 547 TAC AAA TAC AGA ACG ATT CCT CAA AAG ATC GTC TAT ATG TCA  
 183▶ Y K Y R T I P Q K I V Y M S  
 589 AAC AAG GGC TAT TTT GAA CTC TTG GTA ACC GAA AAG GAA AAA  
 197▶ N K G Y F E L L V T E K E K  
 631 CTA TCC AAT GAG GGT CTC TGG AGC ATT TTC CAT CAG AAA CAA  
 211▶ L S N E G L W S I F H Q K Q  
 673 GGT GGA CTT AAC GAA TTC AGT AGT CTC AAT CTC ATA GAG GAG  
 225▶ G G L N E F S S L N L I E E  
 715 GTT GAT GCG TTG GAT GAA ATC TAT GAT TCC AAA GGG TTG CCT  
 239▶ V D A L D E I Y D S K G L P  
 757 GCT TGG GAT CCT CCC TTC CCT GAG GAA CTT GAT GCT TCA GAT  
 253▶ A W D P P F P E E L D A S D  
 799 GAA GAT TTC AAG TTC AAT GCC ACA GAA GAA CTG GCA AAG GTA  
 267▶ E D F K F N A T E E L A K V

Figura 1 (hoja 2)

841 GAG CAA ATC AAA GAA CCA AAG CTG GAA GAC ATA TTC TAT CAG  
 281▶ E Q I K E P K L E D I F Y Q  
 883 GAA GGA CTG CAA CAC GGG ATT CAA ACA TTG CCT TCA GAT GCA  
 295▶ E G L Q H G I Q T L P S D A  
 925 AGT GTT TAT TTT CCT GTG AAT TAC GTT GAA AAC GAC CCT GGA  
 309▶ S V Y F P V N Y V E N D P G  
 967 TTA CAG TCC CAT CAC TTA CAC TTC CCA TTT TTC AGT GGA ATG  
 323▶ L Q S H H L H F P F F S G M  
 1009 GTC TTA CCA AGA GAA ATC CAT TCT TCA GTG CAT CAC ATG AAT  
 337▶ V L P R E I H S S V H H M N  
 1051 AAG GCG TTT TTC TTG TTT GCA AGA CAG CAC GGT TAT GTT GTT  
 351▶ K A F F L F A R Q H G Y V V  
 1093 TGG TTC TTT TAT GGT AAC TTA ATT GGA TGG TAT TAC AAT GCA  
 365▶ W F F Y G N L I G W Y Y N G  
 1135 AAT AAC CAC CCT TGG GAT TCG GAC ATC GAT GCC ATA ATG CCC  
 379▶ N N H P W D S D I D A I M P  
 1177 ATG GCG GAG ATG GCA AGA ATG GCT CAT CAC CAT AAC AAC ACA  
 393▶ M A E M A R M A H H H N N T  
 1219 CTA ATA ATA GAG AAC CCC CAC GAT GGA TAT GGA ACC TAT TTA  
 407▶ L I I E N P H D G Y G T Y L  
 1261 CTG ACT ATT TCT CCT TGG TTC ACG AAG AAG ACA AGA GGT GGT  
 421▶ L T I S P W F T K K T R G G  
 1303 AAC CAT ATT GAT GGT CGT TTT GTG GAC GPT AAG AGG GGT ACC  
 435▶ N H I D G R F V D V K R G T  
 1345 TAC ATC GAC CTC AGT GCA ATT TCA GCT ATG CAC GGA ATA TAT  
 449▶ Y I D L S A I S A M H G I Y  
 1387 CCT GAC TGG GTT AGA GAT GGT GTG AAA GAA AAC CCT AAG AAT  
 463▶ P D W V R D G V K E N P K N  
 1429 CTG GCT CTG GCC GAC AAG AAC GGT AAT TGG TAC CTT ACT AGA  
 477▶ L A L A D K N G N W Y L T R  
 1471 GAT ATT CTC CCA TTG AGG AGA ACA ATA TTC GAA GGT TCT CGA  
 491▶ D I L P L R R T I F E G S R  
 1513 TCC TAC ACC GTT AAA GAC ATT GAA GAT ACC CTG CTT AGA AAC  
 505▶ S Y T V K D I E D T L L R N  
 1555 TAT GGA GAT AAA GTA CTG ATA AAC ACA GAA CTG GCA GAC CAT  
 519▶ Y G D K V L I N T E L A D H  
 1597 GAA TGG CAT GAT GAC TGG AAA ATG TGG GTA CAA AAA AAG AAA  
 533▶ E W H D D W K M W V Q K K K  
 1639 TAC TGC ACT TAT GAG GAA TTT GAA GAT TAC CTG AGT GCT CAT  
 547▶ Y C T Y E E F E D Y L S A H

**Figura 1 (hoja 3)**

```

1681 GGA GGG GTT GAA TAC GAC GAA GAT GGA GTA TTG ACC TTG GAA
561▶ G G V E Y D E D G V L T L E
1723 GGA GCT TGT GGA TTT GAA GAA GTC CGA CAA GAT TGG ATC ATT
575▶ G A C G F E E V R Q D W I I
1765 ACC CGT GAA AGT GTA AAT CTT CAT ATG AAG GAA TGG GAA GCT
589▶ T R E S V N L H M K E W E A
1807 ATC CAG AGG AAC GAA TCA ACC ACA GAG TAT ACT GCT AAG GAT
603▶ I Q R N E S T T E Y T A K D
1849 CTT CCT CGT TAC AGG CCA GAT TCC TTC AAA AAT CTA TTG GAT
617▶ L P R Y R P D S F K N L L D
1891 GGA GTT TCC AAT CAT GGA AAT GGA AAT GTT GGT AAG ATA GAG
631▶ G V S N H G N G N V G K I E
1933 CAT GTC AAA CTT GAA CAC AAC GAC TAG
645▶ H V K L E H N D

```



**Figura 2 (hoja 1)**

1 ATG AGT GGC AAT CCT TTT CTG TTC TCT CCT TCA AAT TTT GAC  
 1▶ M S G N P F L F S P S N F D  
 43 TTT TCT GGT TTG GAT CAT TAT AGA TCC ACT GAT AAA GAT CAC  
 15▶ F S G L D H Y R S T D K D H  
 85 TTA GCT CTA GAT GTT CTC GAT TAT GAC AAA AAT CAC TTC TTC  
 29▶ L A L D V L D Y D K N H F F  
 127 TCC AGA AAC TCC CCC AGT TTG AAA TCT CGT ATT CAC TTT TAT  
 43▶ S R N S P S L K S R I H F Y  
 169 CGA CAT AAA TTG ACC ACT AGA AAG CAA ATT GGA CTT TTC AGC  
 57▶ R H K L T T R K Q I G L F S  
 211 GGC AGA CTG AAG CTT TTT GTG CTT GCT CTC TTT GTG TTG ATC  
 71▶ G R L K L F V L A L F V L I  
 253 ACA TTT TCT GCA ATC CAC ATT CCA ATC CCT TTC TCT TTG GAT  
 85▶ T F S A I H I P I P F S L D  
 295 ATT CTA GGT TCC CAT GTC AAA TAC CTG CCC TTA CGA GAG AAA  
 99▶ I L G S H V K Y L P L R E K  
 337 GTC GAT CCG GAA GAG GCA CTC CAT CTG CAC GGA CTG GAT CTC  
 113▶ V D P E E A L H L H G L D L  
 379 TCG GTA GCA GAG CTA CCT TTT TTC AAT GAT GAC ATG ATG TCT  
 127▶ S V A E L P F F N D D M M S  
 421 GAA TTT AAC TAC GAT CCT AGA CTA CCC ACC GCT TTG ATT TTG  
 141▶ E F N Y D P R L P T A L I L  
 463 AAG TTA GTG TTA GAT CAT ATA AGT GTG CGT AAT GGA ACG TTT  
 155▶ K L V L D H I S V R N G T F  
 505 GAT GCT AAG TTT AAG GTC CCC TTT AAC TGG AAA CTT TGG GTG  
 169▶ D A K F K V P F N W K L W V  
 547 GAT TTG CAT TCA AGG TTA GTT CCA TCT AAT AGT TGG TAT AAT  
 183▶ D L H S R L V P S N S W Y N  
 589 CGA TTT CGA TTA CCC TCA GGT CGT TTC GAA ACA TGC GAT GAA  
 197▶ R F R L P S G R F E T C D E  
 631 TTT AAG AGG TTT TTC GGA ATC ACT AAG AAT CAC TTT GGA ACA  
 211▶ F K R F F G I T K N H F G T  
 673 GAC CTT GAT AAT TGC GTT GAT ATC GAG TAT GAT ACT CCG GAA  
 225▶ D L D N C V D I E Y D T P E  
 715 GGT TAT CCA AAG TTC AAA GTT TTG CAT GCG GAA GAT AAA GCT  
 239▶ G Y P K F K V L H A E D K A  
 757 CTT CCT TAT GAA GCA CGT ATC ATT TAT GGT GCT TCT TAC CTT  
 253▶ L P Y E A R I I Y G A S Y L

**Figura 2 (hoja 2)**

799 TAC CAC GAA GCA CAG AAT CCT AAA AGG TTG ATA TTT TTA GGA  
 267▶ Y H E A Q N P K R L I F L G  
 841 TTG GGC AAG TCC AAT GAG TCT TTG ATC TTA CCA GTT GAG GCA  
 281▶ L G K S N E S L I L P V E A  
 883 AAT GAC AGT TCC AAC TTA ATG CAA TTC AAC CAC GAA TAT GCA  
 295▶ N D S S N L M Q F N H E Y A  
 925 AGA AGC TTT AAC GAT CAA CCT TTC GTT TCT CTT GAG GAA CTT  
 309▶ R S F N D Q P F V S L E E L  
 967 GTC AAG AAG GTT TCA CTG ACC TTG AAT TTG AAT AGT GAT AAG  
 323▶ V K K V S L T L N L N S D K  
 1009 GTG CTA CCA ATC AAT GAA CTG GAC GTT ATC AAA GAC ACC CCG  
 337▶ V L P I N E L D V I K D T P  
 1051 CGC TTA ATG AAT CAC AAC AAC CAG GGA CTG AGC ATA GAC AAG  
 351▶ R L M N H N N Q G L S I D K  
 1093 AGC TCA TTT CAA TGG GAT CTG GAA AGG GAA TTA CAG TTG TTA  
 365▶ S S F Q W D L E R E L Q L L  
 1135 GAA CAT AGA ACC AGT CAA GTT AAT GAC GTG GAA GGC CTT GAT  
 379▶ E H R T S Q V N D V E G L D  
 1177 GCG GGT ATT TAT TCA ACA ATT CAA TGT GAA ATG CGC TCT ATG  
 393▶ A G I Y S T I Q C E M R S M  
 1219 TAC GAT TTT TCA AAA TAC TTC CAT GAA TCA AAA GTC TCT GGT  
 407▶ Y D F S K Y F H E S K V S G  
 1261 AAA TAT CTT CCT TCT GGA GAG CAC TAT GAC TGG CGA TTT TTT  
 421▶ K Y L P S G E H Y D W R F F  
 1303 AAT GGT TTT TAC CTT TCT CAG CAG GAG AAT CTA GCT GTC CTG  
 435▶ N G F Y L S Q Q E N L A V L  
 1345 CAC AGG TTA GGA AGA GCA TGG CTA CGC TTT TCT CGT GCT GCT  
 449▶ H R L G R A W L R F S R A A  
 1387 GGT TTA CAT ACA TGG ATT GCT CAC GGG ACA CTG TTG GGT TGG  
 463▶ G L H T W I A H G T L L G W  
 1429 TAT TGG AAT GGT CTG ATT CTG CCG TGG GAT CAG GAT CTT GAT  
 477▶ Y W N G L I L P W D Q D L D  
 1471 GTT CAA ATG ACT GTA CAA TCA TTG TAT CTG TTG GGA AGG AAT  
 491▶ V Q M T V Q S L Y L L G R N  
 1513 TTC AAC AGC TCT CTT GTA ACT GAT GTT AGT ATT GAA GAT GGC  
 505▶ F N S S L V T D V S I E D G  
 1555 TAC AGC TCA GCA TTG GGA CAT TAC TAT ATT GAC GTT GGA TCC  
 519▶ Y S S A L G H Y Y I D V G S

## Figura 2 (hoja 3)

1597 TCC TTC TTT GTT AGG GAT AAA CTA AAT GGT AAC AAT GCT ATA  
 533▶ S F F V R D K L N G N N A I  
 1639 GAT GCA CGT TTC GTT GAT ACT GAG ACC GGG TTG TAT GTT GAT  
 547▶ D A R F V D T E T G L Y V D  
 1681 ATA ACT GCA TTG GCT TTT ACA GAT CAC TTA AAA CTA AAA CTC  
 561▶ I T A L A F T D H L K L K L  
 1723 ACT ACC AAA GAG AAA GTT GAG CTA CAG AAG GTT ATG GAT CCA  
 575▶ T T K E K V E L Q K V M D P  
 1765 AAT GTA AAG GAA AAA TTG CAG TGG ATC AAA AAT AAA TAT TCA  
 589▶ N V K E K L Q W I K N K Y S  
 1807 ACG GCC ACG CTA CCG GGT GTG ATA GAA ACA GAT AGG AAT AAA  
 603▶ T A T L P G V I E T D R N K  
 1849 GTA TCT GAT GCG CTA GAG AAG CAA TTT CAT GAT TTC AAG TTC  
 617▶ V S D A L E K Q F H D F K F  
 1891 GAC AAT TTT GTC AAC AAA GAG TTG TTT CAC TGT CGA AAT AAC  
 631▶ D N F V N K E L F H C R N N  
 1933 CAT TTC TAC AAA TAT GGA GAG GTT GGC CGA TTA CGG AGC ACT  
 645▶ H F Y K Y G E V G R L R S T  
 1975 ATG TTT GAG GGC GTT CCT GCC CTT ATA CCA TTT GAA TTT GAG  
 659▶ M F E G V P A L I P F E F E  
 2017 TCC ATA CTG AAA CGA GAA TAT CCT AAA GGT CTA ACT TTG AAG  
 673▶ S I L K R E Y P K G L T L K  
 2059 CAT TTC TCC AAT CAT TTT TGG GAT CCA GTG AAC CGA TTG TGG  
 687▶ H F S N H F W D P V N R L W  
 2101 GTA CCA GAA AAG AAG AAA AAA ATT AGA CAC ATA GAG TTT TCA  
 701▶ V P E K K K K I R H I E F S  
 2143 CTT ACG AAG GAA GTT ACA GAA AGC CAC AAG AAA GAA CTT GCA  
 715▶ L T K E V T E S H K K E L A  
 2185 CAG ATC CAT GGG AAC GAA ACG GGT ATA ACC TCC GAC TTC GCA  
 729▶ Q I H G N E T G I T S D F A  
 2227 TAT TCT CCT TTC AGA ATA GAT CCC TGG CTG TCT CGA TAC AGG  
 743▶ Y S P F R I D P W L S R Y R  
 2269 AAA AAA ATG ACT AGG AGC CAA TAA  
 757▶ K K M T R S Q

Figura 3 (hoja 1)

1 ATG AGT GGC AAT CCT TTT CTG TTC TCT CCT TCA AAT TTT GAC TTT TCT GGT TTG GAT CAT TAT AGA TCC  
 1 M S G N P F L F S P S N F D F S G L D H Y R S  
 70 ACT GAT AAA GAT CAC TTA GCT CTA GAT GTT CTC SAT TAT GAC AAA AAT CAC TTC TTC TCC AGA AAC TCC  
 24 T D K D H L A L D V L D Y D K N H F F S R N S  
 139 CCC ACT TTG AAA TCT CGT ATT CAC TTT TAT CGA CAT AAA TTG ACC ACT AGA AAG CAA ATT GGA CTT TTC  
 47 P S L K S R I H F Y R H K L T T R K Q I G L F  
 208 AGC GGC AGA CTG AAG CTT TTT GTG CTT GCT CTC TTT GTG TTG ATC ACA TTT TCT GCA ATC CAC ATT CCA  
 70 S G R L K L F V L A L F V L I T F S A I H I P  
 277 ATC CCT TTC TCT TTG GAT ATT CTA GGT TCC SAT GTC AAA TAC CTG CCC TTA CGA GAG AAA GTC GAT CCG  
 93 I P F S L D I L G S H V K Y L P L R E K V D P  
 346 GAA GAG GCA CTC CAT CTG CAC GGA CTG GAT CTC TCG GTA GCA GAG CTA CCT TTT AAT GAT GAC GAT  
 116 E E A L H L H G L D L S V A E L P F F N D D M  
 415 ATG TCT GAA TTT AAC TAC GAT CCT AGA CTA CCC ACC GCT TTG ATT TTG AAG TTA GTG TTA GAT CAT ATA  
 139 M S E F N Y D P R L P T A L I L K L V L D H I  
 484 AGT GTG CGT AAT GGA ACG TTT GAT GCT AAG TTT AAG GTC CCC TTT AAC TGG AAA CTT TGG GTG GAT TTG  
 162 S V R N G T F D A K F K V P F N W K L W V D L  
 553 CAT TCA AGG TTA GTT CCA TCT AAT AGT TGG TAT AAT CGA TTT CGA TTA CCC TCA GGT CGT TTC GAA ACA  
 185 H S R L V P S N S W Y N R F R L P S G R F E T  
 622 TGC GAT GAA TTT AAG AGG TTT TTC GGA ATC ACT AAG AAT CAC TTT GGA ACA GAC CTT GAT AAT TGC GTT  
 208 C D E F K R F F G I T K N H F G T D L D N C V  
 691 GAT ATC GAG TAT GAT ACT CCG GAA GGT TAT CCA AAG TTC AAA GTT TTG CAT GCG GAA GAT AAA GCT CTT  
 231 D I E Y D T P E G Y P K F K V L H A E D K A L  
 760 CCT TAT GAA GCA CGT ATC ATT TAT GGT GCT TCT TAC CTT TAC CAC GAA GCA CAG AAT CCT AAA AGG TTG  
 254 P Y E A R I I Y G A S Y L Y H E A Q N P K R L  
 829 ATA TTT TTA GGA TTG GGC AAG TCC AAT GAG TCT TTG ATC TTA CCA GTT GAG GCA AAT GAC AGT TCC AAC  
 277 I F L G L G K S N E S L I L P V E A N D S S N  
 898 TTA ATG CAA TTC AAC CAC GAA TAT GCA AGA AGC TTT AAC GAT CAA CCT TTC GTT TCT CTT GAG GAA CTT  
 300 L M Q F N H E Y A R S F N D Q P F V S L E E L  
 967 GTC AAG AAG GTT TCA CTG ACC TTG AAT TTG AAT AGT GAT AAG GTG CTA CCA ATC AAT GAA CTG GAC GTT  
 323 V K K V S L T L N L N S D K V L P I N E L D V  
 1036 ATC AAA GAC ACC CCG CGC TTA ATG AAT CAC AAC AAC GGA GCA CTG AGC ATA GAC AAG AGC TCA TTT CAA  
 346 I K D T P R L M N H N N Q G L S I D K S S F Q  
 1105 TGG GAT CTG GAA AGG GAA TTA CAG TTG TTA GAA CAT AGA ACC AGT CAA GTT AAT GAC GTG GAA GGC CTT  
 369 W D L E R E L Q L L E A H R T S Q V N D V E G L  
 1174 GAT GCG GGT ATT TAT TCA ACA ATT CAA TGT GAA ATG CCG TCT ATG TAC GAT TTT TCA AAA TAC TTC CAT  
 392 D A G I Y S T I Q C E M R S M Y D F S K Y F H  
 1243 GAA TCA AAA GTC TCT GGT AAA TAT CTT CCT TCT GGA GAG CAC TAT GAC TGG CGA TTT TTT AAT GGT TTT  
 415 E S K V S G K Y L P S G E H Y D W R F F N G F  
 1312 TAC CTT TCT CAG CAG GAG AAT CTA GCT GTC CTG CAC AGG TTA GGA AGA GCA TGG CTA CGC TTT TCT CGT  
 438 Y L S Q Q E N L A V L H R L G R A W L R F S R  
 1381 GCT GCT GGT TTA CAT ACA TGG ATT GCT CAC GGG ACA CTG TTG GGT TGG TAT TGG AAT GGT CTG ATT CTG  
 461 A A G L H T W I A H G T L L G W Y W N G L I L  
 1450 CCG TGG GAT CAG GAT CTT GAT GTT CAA ATG ACT GTA CAA TCA TTG TAT CTG TTG GGA AGG AAT TTC AAC  
 484 P W D Q D L D V Q M T V Q S L Y L L G R N F N  
 1519 AGC TCT CTT GTA ACT GAT GTT AGT ATT GAA GAT GGC TAC AGC TCA GCA TTG GGA CAT TAC TAT ATT GAC  
 507 S S L V T D V S I E D G Y S S A L G H Y Y I D  
 1588 GTT GGA TCC TCC TTC TTT GTC AGG GAT AAA CTA AAT GGT AAC AAT GCT ATA GAT GCA CGT TTC GTT GAT  
 530 V G S S F F V R D K L N G N N A I D A R F V D  
 1657 ACT GAG ACC GGG TTG TAT GTT GAT ATA ACT GCA TTG GCT TTT ACA GAT CAC TTA AAA CTA AAA CTC ACT  
 553 T E T G L Y V D I T A L A F T D H L K L K L T F  
 1726 ACC AAA GAG AAA GTT GAG CTA CAG AAG GTT ATG GAT CCA AAT GTA AAG GAA AAA TTG CAG TGG ATC AAA  
 576 T K E K V E L Q K V M D P N V K E K L Q W I K  
 1795 AAT AAA TAT TCA ACG GCC ACG CTA CCG GGT GTG ATA GAA ACA GAT AGG AAT AAA GTA TCT GAT GCG CTA  
 599 N K Y S T A T L P G V I E T D R N K V S D A L  
 1864 GAG AAG CAA TTT CAT GAT TTC AAG TTC GAC AAT TTT GTC AAC AAA GAG TTG TTT CAC TGT CGA AAT AAC  
 622 E K Q F H D F K F B N F V N K E L F H C R N N  
 1933 CAT TTC TAC AAA TAT GGA GAG GTT GGC CGA TTA CGG AGC ACT ATG TTT GAG GGC GTT CCT GCC CTT ATA  
 645 H F Y K Y G E V G R L R S T M F E G V P A L I  
 2002 CCA TTT GAA TTT GAG TCC ATA CTG AAA CGA GAA TAT CCT AAA GGT CTA ACT TTG AAG CAT TTC TCC AAT  
 668 P F E F E S I L K R E Y P K G L T L K H F S N  
 2071 CAT TTT TGG GAT CCA GTG AAC CGA TTG TGG GTA CCA AAG AAG AAA AAA ATT AGA CAC ATA GAG TTT  
 691 H F W D P V N R L W V P E K K K K I R H I E F

**Figura 3 (hoja 2)**

2140 TCA CTT ACG AAG GAA GTT ACA GAA AGC CAC AAG AAA GAA CTT GCA CAG ATC CAT GGG AAC GAA ACG GGT  
714▶ S L T K E V T E S H K K E L A Q I H G N E T G  
2209 ATA ACC TCC GAC TTC GCA TAT TCT CCT TTC AGA ATA GAT CCC TGG CTG TCT CGA TAC AGG AAA AAA ATG  
737▶ I T S D F A Y S P F R I D P W L S R Y R K K M  
2278 ACT AGG AGC CAA  
760▶ T R S Q

Figura 4

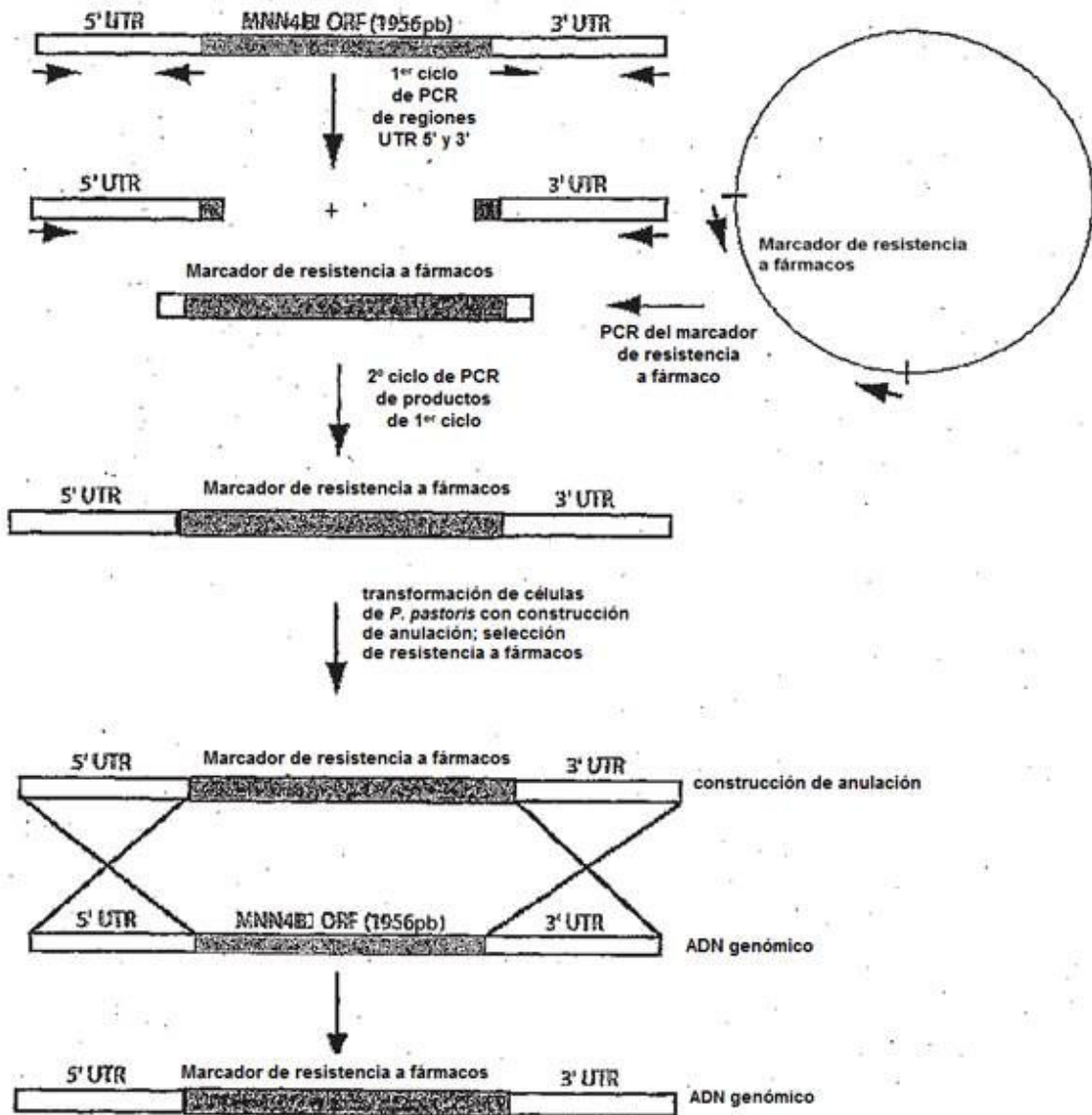
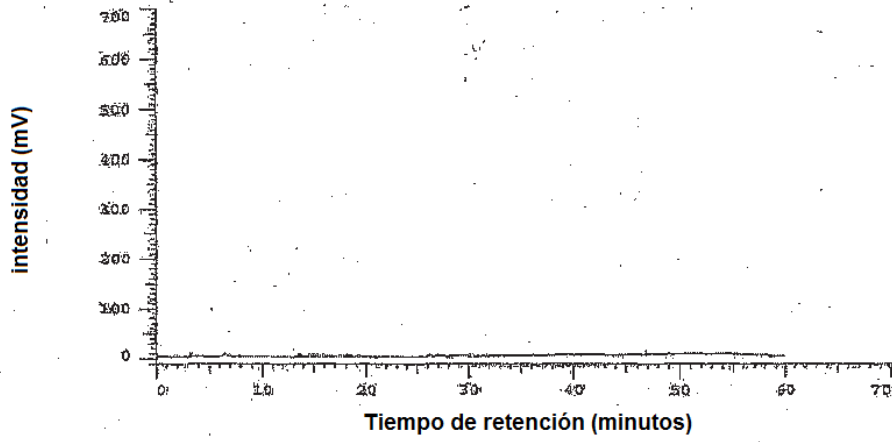
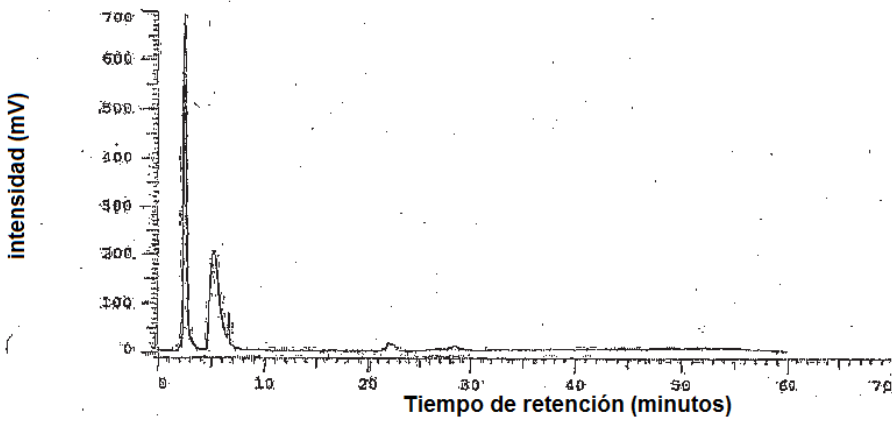


Figura 5 (hoja 1)

A.



B.



C.

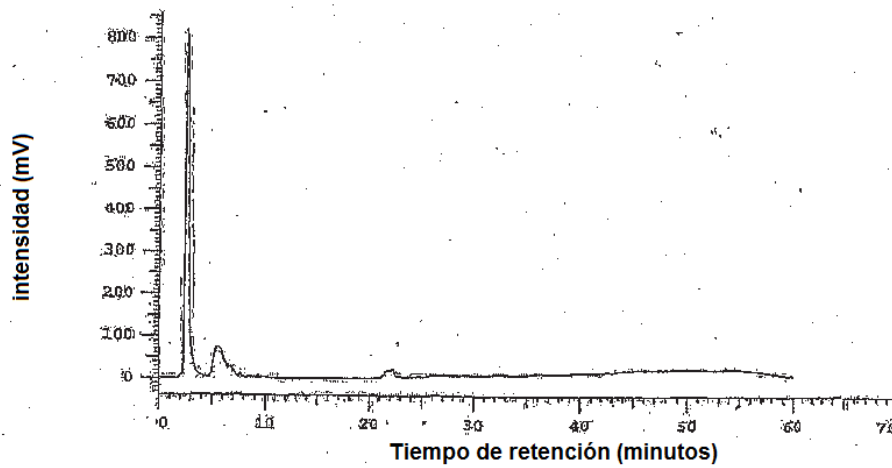


Figura 5 (hoja 2)

D.

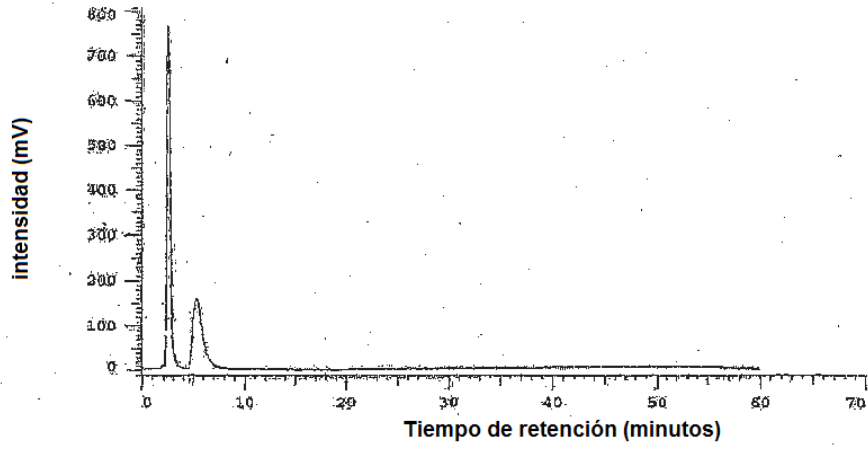




Figura 6

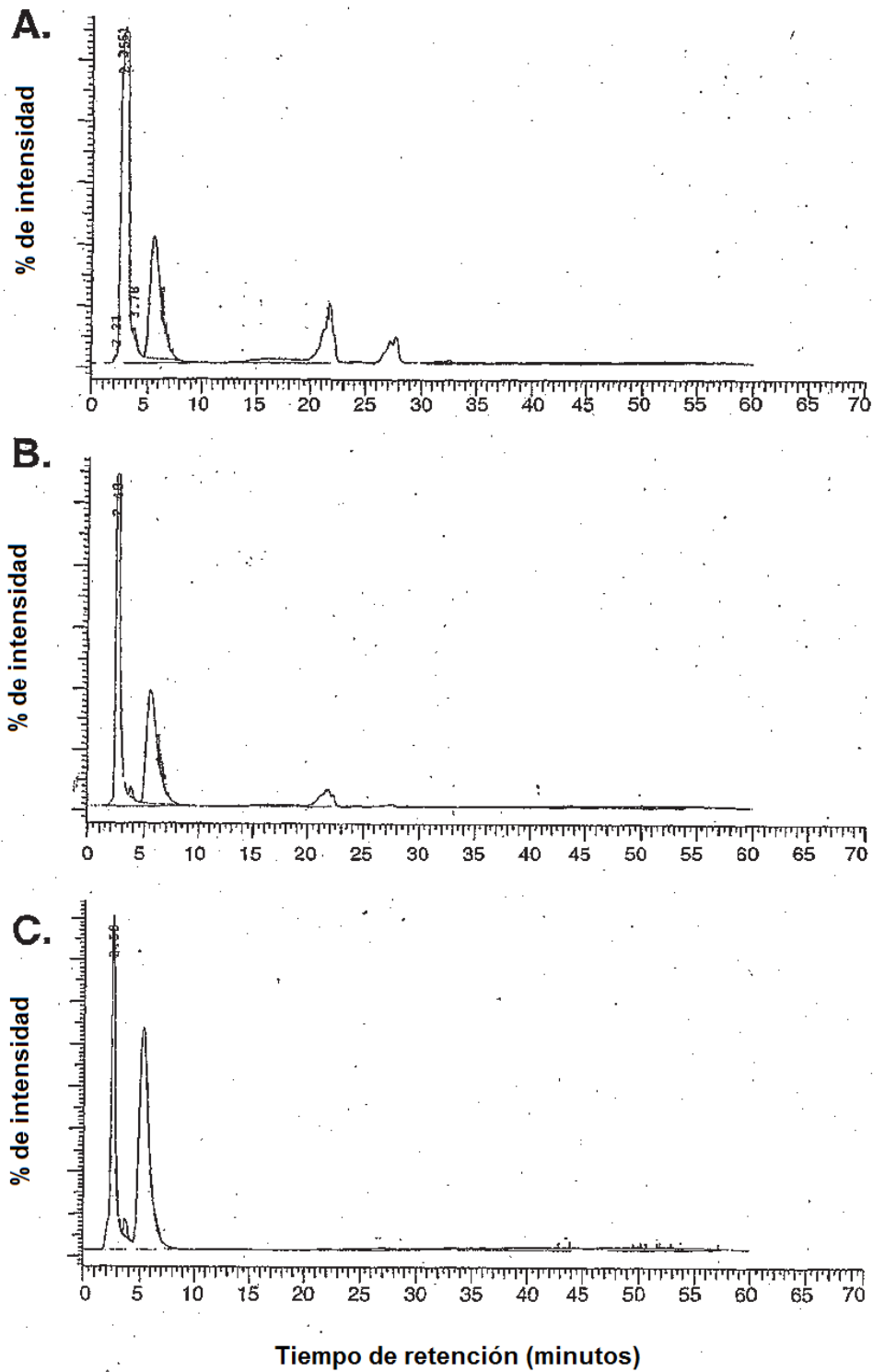


Figura 7 (hoja 1)

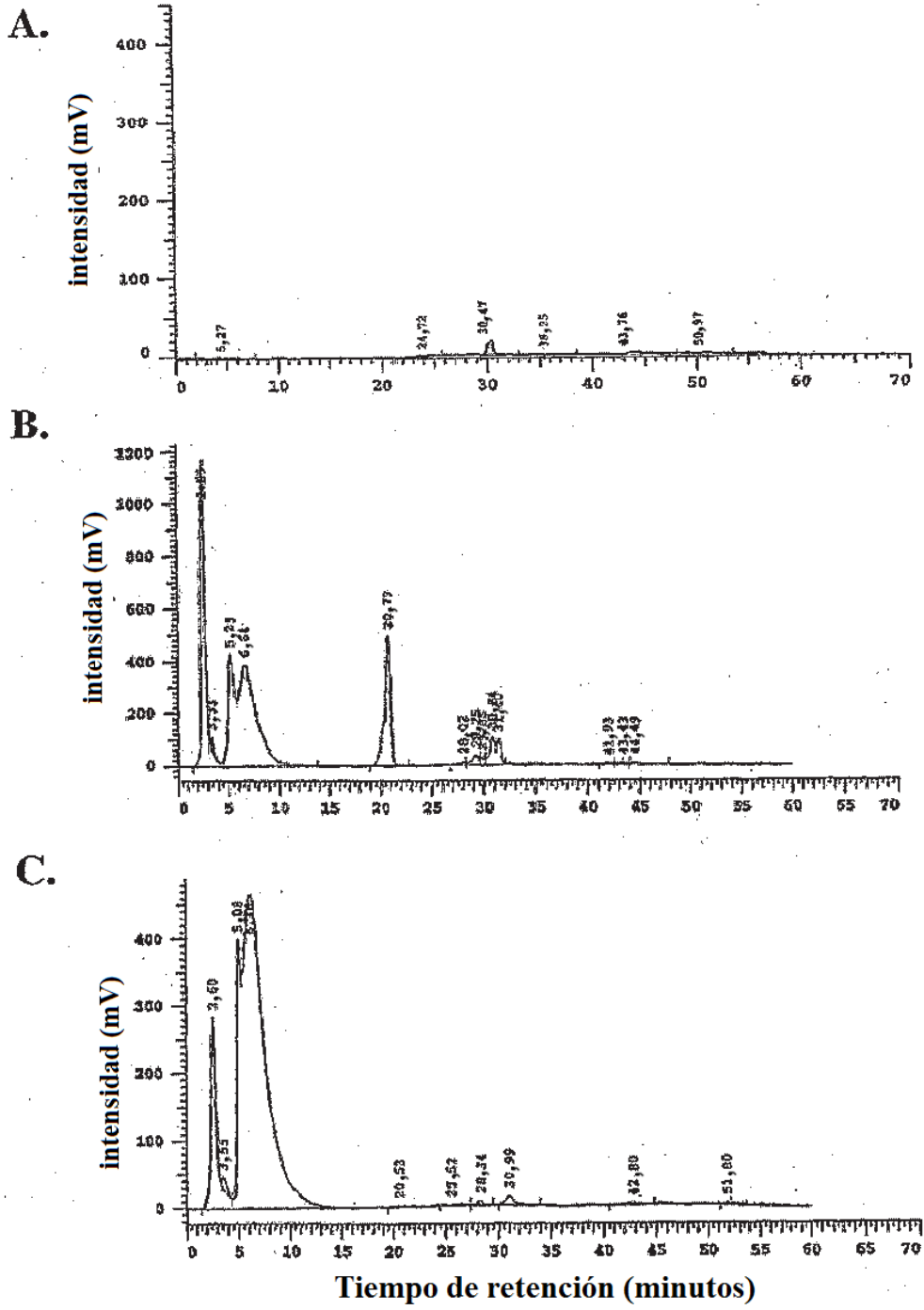
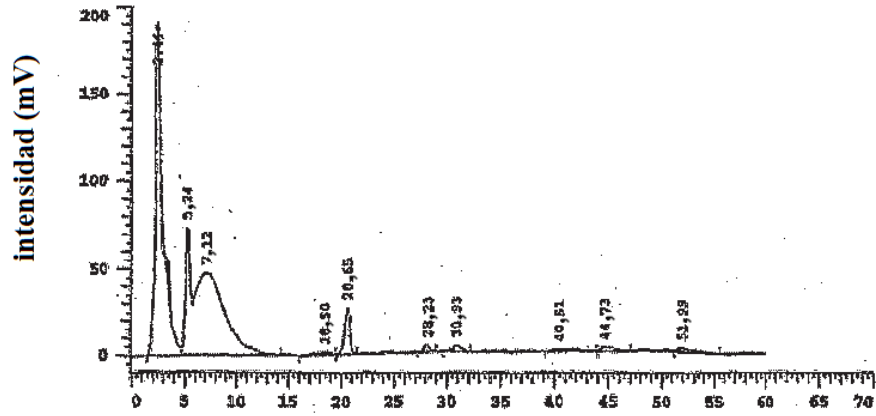
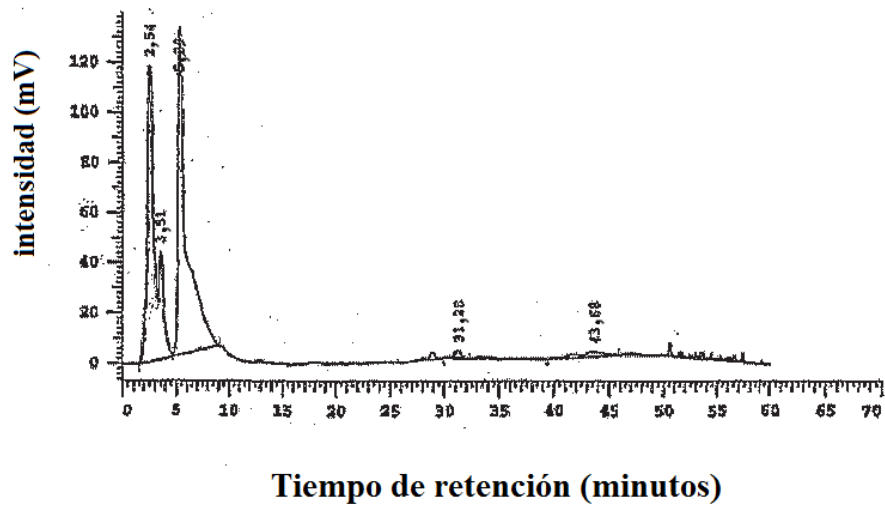


Figura 7 (hoja 2)

D.



E.



Tiempo de retención (minutos)

Figura 8

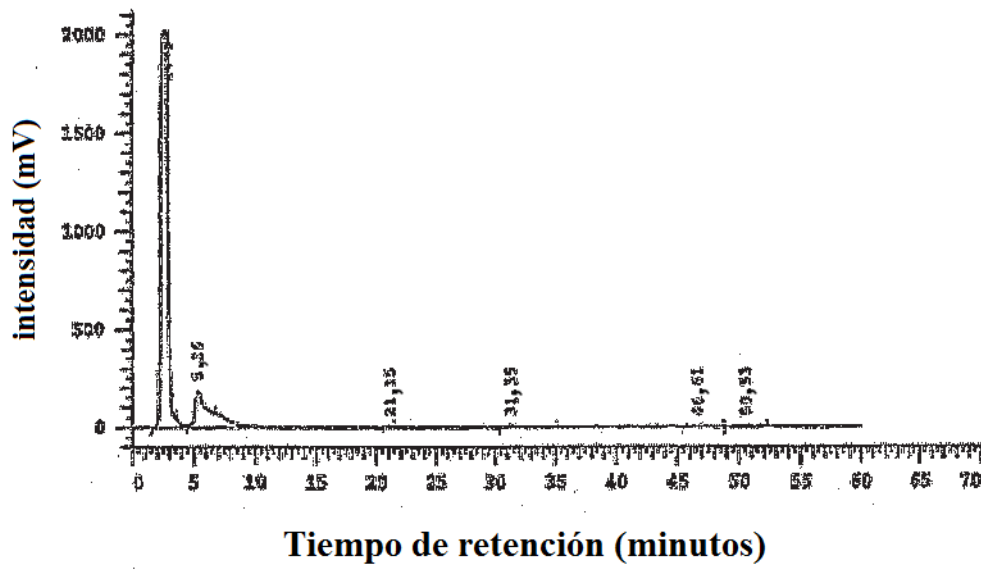


Figura 9 (hoja 1)

|    |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |         |         |     |     |     |     |     |         |         |         |         |         |         |        |
|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|---------|---------|-----|-----|-----|-----|-----|---------|---------|---------|---------|---------|---------|--------|
| 1  | --- | M   | K   | V   | S   | K   | R   | L   | I   | --- | --- | --- | --- | P   | R   | R   | S   | R   | L   | L   | PpMNN4A |         |     |     |     |     |     |         |         |         |         |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | PpMNN4B |     |     |     |     |     |         |         |         |         |         |         |        |
| 1  | --- | M   | S   | G   | N   | P   | F   | L   | F   | S   | P   | S   | N   | F   | D   | F   | S   | G   | L   | D   | H       | Y       | R   | S   | T   | D   | K   | PpMNN4C |         |         |         |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | M   | T   | L   | R   | S   | A   | I   | K   | A   | R   | T   | S       | K       | G   | L   | I   | G   | A   | PpPNO1  |         |         |         |         |         |        |
| 1  | M   | L   | Q   | R   | I   | S   | S   | K   | L   | H   | R   | R   | F   | L   | S   | G   | L   | L   | R   | V   | K       | H       | Y   | P   | L   | R   | R   | I       | L       | L       | ScMNN4  |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | NcMNN4A |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | NcMNN4B |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | AnMNN4A |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | AnMNN4B |         |         |        |
| 1  | --- | M   | S   | N   | T   | I   | P   | Q   | Y   | F   | I   | R   | I   | F   | N   | L   | I   | F   | S   | A   | R       | R       | K   | N   | F   | Q   | L   | CaMNN4  |         |         |         |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | PaMNN4  |         |         |        |
| 16 | I   | M   | M   | L   | L   | V   | V   | Y   | Q   | L   | V   | V   | L   | V   | L   | G   | L   | E   | S   | V   | S       | E       | G   | K   | L   | A   | S   | L       | L       | PpMNN4A |         |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | PpMNN4B |         |        |
| 27 | D   | H   | L   | A   | L   | D   | V   | L   | D   | Y   | D   | K   | N   | H   | F   | F   | S   | R   | N   | S   | P       | S       | L   | K   | S   | R   | I   | H       | F       | Y       | PpMNN4C |         |         |        |
| 19 | V   | I   | I   | A   | S   | I   | I   | F   | F   | T   | T   | V   | T   | F   | Y   | D   | E   | S   | K   | I   | V       | G       | I   | I   | R   | V   | S   | D       | T       | Y       | PpPNO1  |         |         |        |
| 31 | P   | L   | I   | L   | L   | Q   | I   | I   | I   | I   | T   | F   | I   | W   | S   | N   | S   | P   | Q   | R   | N       | G       | L   | G   | R   | D   | A   | D       | Y       | L       | ScMNN4  |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | NcMNN4A |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | NcMNN4B |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | AnMNN4A |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | AnMNN4B |         |        |
| 27 | A   | L   | I   | S   | G   | L   | L   | F   | F   | G   | S   | F   | A   | I   | L   | S   | T   | T   | S   | Y   | S       | K       | K   | F   | N   | Y   | F   | D       | D       | L       | CaMNN4  |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | ---     | PaMNN4  |        |
| 46 | D   | L   | G   | D   | W   | D   | L   | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | A   | N   | S       | S       | L   | S   | I   | S   | D   | F       | PpMNN4A |         |         |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | PpMNN4B |         |        |
| 57 | R   | H   | K   | L   | T   | T   | R   | K   | Q   | I   | G   | --- | --- | --- | --- | --- | --- | --- | L   | F   | S       | G       | R   | L   | K   | L   | F   | V       | PpMNN4C |         |         |         |         |        |
| 49 | T   | G   | H   | S   | A   | V   | S   | S   | T   | F   | N   | --- | --- | --- | --- | --- | --- | --- | A   | S   | S       | V       | V   | S   | D   | N   | K   | I       | PpPNO1  |         |         |         |         |        |
| 61 | L   | P   | N   | Y   | N   | E   | L   | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | D   | S   | D       | D       | S   | W   | Y   | S   | I   | ScMNN4  |         |         |         |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | NcMNN4A |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | NcMNN4B |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | AnMNN4A |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | AnMNN4B |         |        |
| 57 | I   | L   | K   | I   | Y   | D   | Y   | N   | Y   | L   | T   | N   | N   | Y   | N   | I   | D   | Y   | L   | A   | K       | N       | D   | P   | E   | A   | Y   | F       | N       | V       | CaMNN4  |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | ---     | PaMNN4  |        |
| 63 | I   | K   | L   | K   | L   | K   | G   | Q   | K   | T   | Y   | H   | K   | F   | D   | E   | H   | V   | F   | A   | A       | M       | A   | R   | I   | Q   | S   | N       | ---     | PpMNN4A |         |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | ---     | PpMNN4B |        |
| 78 | L   | A   | L   | F   | V   | L   | I   | T   | F   | S   | A   | I   | H   | I   | P   | I   | P   | F   | S   | L   | D       | I       | L   | G   | S   | H   | V   | K       | Y       | L       | PpMNN4C |         |         |        |
| 70 | N   | G   | Y   | G   | L   | P   | L   | I   | D   | T   | E   | S   | N   | S   | R   | Y   | E   | D   | P   | D   | D       | I       | S   | I   | E   | N   | E   | L       | R       | Y       | PpPNO1  |         |         |        |
| 78 | L   | T   | S   | S   | F   | K   | N   | D   | R   | K   | I   | Q   | F   | A   | K   | T   | L   | Y   | E   | N   | L       | K       | F   | G   | T   | N   | P   | K       | ---     | ScMNN4  |         |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | ---     | NcMNN4A |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | ---     | NcMNN4B |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | ---     | AnMNN4A |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | ---     | AnMNN4B |        |
| 87 | K   | V   | Q   | Q   | I   | V   | D   | E   | K   | K   | Q   | H   | D   | L   | E   | S   | K   | F   | W   | S   | L       | D       | T   | K   | I   | N   | D   | D       | Q       | A       | CaMNN4  |         |         |        |
| 1  | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | ---     | ---     | --- | --- | --- | --- | --- | ---     | ---     | ---     | ---     | ---     | ---     | PaMNN4 |

Figura 9 (hoja 2)

|     |    |   |         |
|-----|----|---|---------|
| 91  | -- | ENGKLADYESTSSKTDVTIQNVELWKRL              | PpMNN4A |
| 1   | -- | -----                                     | PpMNN4B |
| 108 |    | PLREKVDPEEALHLHGLDLSVAELPFFND             | PpMNN4C |
| 100 |    | RIAQSTKEEENMWKLDTTLTEASLKIPIQ             | PpPNO1  |
| 106 | -- | WVNEYTLQNDLLSVKMGPRKGSKLE/SVD             | ScMNN4  |
| 1   | -- | -----                                     | NcMNN4A |
| 1   | -- | -----                                     | NcMNN4B |
| 1   | -- | -----                                     | AnMNN4A |
| 1   | -- | -----                                     | AnMNN4B |
| 117 |    | TLQIPAYFTYNKPRDNKNLE DSEQSSKPE            | CaMNN4  |
| 1   | -- | -----                                     | PaMNN4  |
| 119 |    | SEEEYTYEPRITLAVYLSYIHQRTYDRYAT            | PpMNN4A |
| 1   | -- | -----MFKETSKN-----                        | PpMNN4B |
| 138 |    | MMSEFNYPDRRLPTALILKLVLDHISVRNGT           | PpMNN4C |
| 130 |    | SFELQPFKERLDMSLYNSKNIGNFYFYDPR            | PpPNO1  |
| 134 |    | ELKFYDFDPRRLTWSVVLNHLQNNDA DQPE           | ScMNN4  |
| 1   | -- | -----MWSSLTP-----                         | NcMNN4A |
| 1   | -- | -----                                     | NcMNN4B |
| 1   | -- | -----                                     | AnMNN4A |
| 1   | -- | -----                                     | AnMNN4B |
| 147 |    | KPLIQPFDRFTLAMYYYLDQQMTFAHHD              | CaMNN4  |
| 1   | -- | -----                                     | PaMNN4  |
| 149 |    | SYAP---YNLRV <b>PFSWADWIDLTA</b> LNQYLD   | PpMNN4A |
| 9   |    | -LFG---SINTFNTVEYVMYMM <b>LLLTAY</b>      | PpMNN4B |
| 168 |    | FDAK---FKV <b>PFNWKLWVDLHSRLVPSN</b>      | PpMNN4C |
| 160 |    | LTFS---VYLKYIKDKLASG <b>STTNLTIP</b>      | PpPNO1  |
| 163 |    | ---K <b>L PFSWYDWTTFHEL</b> NKLI S        | ScMNN4  |
| 8   |    | ---ARRQATTT <b>SWRDR</b> LLTL <b>LLMA</b> | NcMNN4A |
| 1   | -- | -----                                     | NcMNN4B |
| 1   | -- | ---MHKKATLALAS <b>AICITA</b>              | AnMNN4A |
| 1   | -- | -----                                     | AnMNN4B |
| 177 |    | SSSSSSSGNSITV <b>PFNWYDWVDM</b> SVLNKYLL  | CaMNN4  |
| 1   | -- | -----                                     | PaMNN4  |
| 176 |    | KTK-----GCEAVFPRE <b>SE</b> ATMKLN--      | PpMNN4A |
| 33  |    | F-----LNHLLHS <b>LD</b> NN--              | PpMNN4B |
| 193 |    | SWY-----NRFRLPSGRFET <b>CDEF</b> KRFFGI   | PpMNN4C |
| 185 |    | FNW-----AHFRDLSSLN <b>PYL</b> DIKQEDKVA   | PpPNO1  |
| 183 |    | IDKTVLPCNELFQSAFDK <b>ESLE</b> AIE TELGE  | ScMNN4  |
| 28  |    | -----LTFVLS <b>SL</b> ASP-----            | NcMNN4A |
| 1   | -- | -----                                     | NcMNN4B |
| 18  |    | -----                                     | AnMNN4A |
| 1   | -- | -----                                     | AnMNN4B |
| 207 |    | APNKDKPDCSILDAHEDARKIETE <b>KKMEK</b>     | CaMNN4  |
| 1   | -- | -----                                     | PaMNN4  |

Figura 9 (hoja 3)

|     |   |         |
|-----|---|---------|
| 197 | - - - - I T - - - - V V D W L E G L C I T D K S L Q N S V N     | PpMNN4A |
| 46  | - - - - - - - - - - - - H L V E S D V N Y Q - - - - - - - -     | PpMNN4B |
| 218 | T K N H F G - - - - - - T D L D N C V D I E Y D T P E G Y       | PpMNN4C |
| 210 | C D Y F Y E S S N K D K R K P T G N C I E F K D V R D E - -     | PpPNO1  |
| 213 | P L F L Y E - - - - - R P K Y A Q K L W Y K A A R N Q D R I K   | ScMNN4  |
| 39  | - - - - - - - - - - - - L P I E G A V - - - - - - - - - -       | NcMNN4A |
| 1   | -         | NcMNN4B |
| 18  | - - - - - - - - - - - - - - A T G - - - - - - - - - - - -       | AnMNN4A |
| 1   | -         | AnMNN4B |
| 237 | L A K Q W D E N K R K A E E E K K K A E E D K K K E E E E K     | CaMNN4  |
| 1   | - M L V N S V - -     | PaMNN4  |
| 219 | S T Y A E E I N S R D I L S P N F H - - - - - - - - - - - -     | PpMNN4A |
| 56  | - L L Q R V T N K V K L F D E E A V L P F - - - - - - - - - -   | PpMNN4B |
| 241 | P K F K V L H A E D K A L P Y E A R I I Y G A - - - - - - - -   | PpMNN4C |
| 238 | H L I Q Y G I S S K D H L P G P F I L K S - - - - - - - - - -   | PpPNO1  |
| 239 | D S K E L K K H C S K L F T P D G H G S P K G L R F N T Q F     | ScMNN4  |
| 46  | - - - - - - - - - - - - V K A N N N D A V S - - - - - - - - - - | NcMNN4A |
| 1   | -         | NcMNN4B |
| 21  | - - - - - - - - - - - - - - L P G P - - - - - - - - - - - -     | AnMNN4A |
| 1   | -         | AnMNN4B |
| 267 | K K A E E E E K Q R H E Q E K Q A L E E D K K K L E E E K K     | CaMNN4  |
| 7   | - - - - - - - - - - - - - - S P - - - - - - - - - - - - - -     | PaMNN4  |
| 237 | V F G Y S D A K D N P Q Q K I F Q - - - - - - - - - - - -       | PpMNN4A |
| 76  | - - - - - A K N L N R R T E R F D - - - - - - - - - - - -       | PpMNN4B |
| 264 | S Y L Y H E A Q N P K R L I F L G - - - - - - - - - - - -       | PpMNN4C |
| 259 | - L G I P M Q H T A K R L E S N L - - - - - - - - - - - -       | PpPNO1  |
| 269 | Q I K E L Y D K V R P E V Y Q L Q - - - - - - - - - - - -       | ScMNN4  |
| 56  | - - - - - - - - - - - - Q P Q A Q A Q - - - - - - - - - -       | NcMNN4A |
| 1   | - - - - - - - - - - - - - - M L L - - - - - - - - - - - -       | NcMNN4B |
| 25  | -         | AnMNN4A |
| 1   | -         | AnMNN4B |
| 297 | K I E E E K N K L Q E Q Q Q Q Q Q E E K A N D G N Q E H S       | CaMNN4  |
| 9   | -         | PaMNN4  |
| 254 | - S K S Y I N S K     | PpMNN4A |
| 89  | - P R L P V A A Y     | PpMNN4B |
| 281 | - L G K S N E S L     | PpMNN4C |
| 275 | - Y L L T G A P V     | PpPNO1  |
| 286 | - A R N Y I L T T     | ScMNN4  |
| 63  | - A K A E V R Q E     | NcMNN4A |
| 4   | - N W L L I V T T     | NcMNN4B |
| 25  | - V L D S A P K       | AnMNN4A |
| 1   | - M R L S T Y P I     | AnMNN4B |
| 327 | K F V K R D D E I K M S T S Q D K S D S D A D R A K I D M T     | CaMNN4  |
| 9   | - S K S - - - - -     | PaMNN4  |





Figura 9 (hoja 5)

|     |   |         |
|-----|---|---------|
| 310 | Q D Q K T F - - T F D P V Y E F N R L K S Q V K P - R P I S             | PpMNN4A |
| 171 | Y -           | PpMNN4B |
| 325 | K V S L T L - | PpMNN4C |
| 336 | L I E T V P - | PpPNO1  |
| 344 | K N K Q D V - - E F N H N R L F Q E F V N N D Q V N S L Y K             | ScMNN4  |
| 101 | -             | NcMNN4A |
| 35  | Q -           | NcMNN4B |
| 68  | -             | AnMNN4A |
| 32  | K L -         | AnMNN4B |
| 463 | L L R N G I P E S Y L A N N N F D V S L N V L Q Q L H K L K             | CaMNN4  |
| 31  | P D R -       | PaMNN4  |
| 337 | S E P S I D S A L K E N D Y K L K L K E S S F I F N Y G R I             | PpMNN4A |
| 172 | - - - - - T A H I L H S N V E V Y K Y R - - - - T I P Q K I V           | PpMNN4B |
| 338 | L P I N E L D V I K D T P R L M N H N N Q G L S I D K S S F             | PpMNN4C |
| 349 | G E G Y F T T E L K E N N F E L P L S K N D F T F D D S E V             | PpPNO1  |
| 372 | L E I E E T D K F T F D K D L V Y L S P S D F K F D A S K K             | ScMNN4  |
| 101 | -             | NcMNN4A |
| 36  | -             | NcMNN4B |
| 68  | -             | AnMNN4A |
| 34  | -             | AnMNN4B |
| 493 | K N H K P D T A K V I N D Y L L H I P K E S F K Y D P D S I             | CaMNN4  |
| 36  | -             | PaMNN4  |
| 367 | L S N Y E E R - - - - L E S L N D F E K S H Y E S L A Y S S             | PpMNN4A |
| 194 | Y M S N K G Y F E L - - - - L V T E K E K L S - - - - - - - -           | PpMNN4B |
| 368 | Q W D L E R E L Q L - L E H R T S Q V N D V E G L D A G I Y             | PpMNN4C |
| 379 | E S L I K G L S E Q - D L D L H T Q R Y K E S L Q Y S - - -             | PpPNO1  |
| 402 | I E E L E E Q K K L Y P D K F S A H N E N Y L N S L K N S V             | ScMNN4  |
| 101 | -             | NcMNN4A |
| 36  | -             | NcMNN4B |
| 68  | -             | AnMNN4A |
| 34  | -             | AnMNN4B |
| 523 | I F D Y T K R L D K - G E K L T I K E L K Y L Q S L E Y S K             | CaMNN4  |
| 42  | I E N Y E K R - - - - I D E L D E K Q L R H L Q T L K Y S R             | PaMNN4  |
| 393 | - - - - - L L E A R K L P K Y F G E V I L K N P - - - - -               | PpMNN4A |
| 213 | - - - - - - - - - - - N E G L W S I F H Q K Q G G L N E F S S L         | PpMNN4B |
| 397 | S T I Q C E M R S M Y D F S K Y F H E - - - - - - - - - - -             | PpMNN4C |
| 405 | - - - - - F A T R E N D V K K Y F Y E A R M I I N - - - - -             | PpPNO1  |
| 432 | - - - - - K T S P A L Q R K F F Y E A G A V K Q - - - - -               | ScMNN4  |
| 107 | -             | NcMNN4A |
| 47  | -             | NcMNN4B |
| 68  | -             | AnMNN4A |
| 46  | -             | AnMNN4B |
| 552 | D - - - - - K V A H G G P P K Y F A E S R L I G T - - - - -             | CaMNN4  |
| 68  | - - - - - S I P S T K L K K S F R E V N I N W P A T Y - - -             | PaMNN4  |





Figura 9 (hoja 8)

|     |   |                 |                             |         |
|-----|---|-----------------|-----------------------------|---------|
| 557 | L A L T G S T M   | - - - - - P K R | - - - - -                   | PpMNN4A |
| 455 | I S A M H   | - - - - -       | - - - - -                   | PpMNN4B |
| 564 | L A F T D H L K L K L T T K E K V E L Q K V M D P N V K E K | - - - - -       | - - - - -                   | PpMNN4C |
| 574 | L S T S Q   | - - - - -       | - - - - -                   | PpPNO1  |
| 590 | L A S T S A P   | - - - - - S R D | - - - - -                   | ScMNN4  |
| 264 | L S   | - - - - -       | - - - - -                   | NcMNN4A |
| 182 | A R Y N L   | - - - - -       | - - - - -                   | NcMNN4B |
| 213 | L S K L N   | - - - - -       | - - - - -                   | AnMNN4A |
| 181 | V R A D D   | - - - - -       | - - - - -                   | AnMNN4B |
| 707 | L A L S N S E T   | - - - - - P K   | - - - - -                   | CaMNN4  |
| 240 | L A L T D T   | - - - - -       | - - - - -                   | PaMNN4  |
| 568 | Y S N K L I K Q   | - - - - -       | - - - - -                   | PpMNN4A |
| 460 | - - - - - G I Y   | - - - - -       | - - - - -                   | PpMNN4B |
| 594 | L Q W I K N K Y   | - - - - -       | - - - - -                   | PpMNN4C |
| 579 | - - - - - S A R   | - - - - -       | - - - - -                   | PpPNO1  |
| 600 | Y L N S Y I E E R L Q E E H L D I N N I P E S N G E T A T L | - - - - -       | - - - - -                   | ScMNN4  |
| 266 | - - - - -   | - - - - -       | - - - - -                   | NcMNN4A |
| 187 | - - - - - T H   | - - - - -       | - - - - -                   | NcMNN4B |
| 218 | - - - - - E E K   | - - - - -       | - - - - -                   | AnMNN4A |
| 186 | - - - - - E R   | - - - - -       | - - - - -                   | AnMNN4B |
| 717 | - - S D L A E L   | - - - - -       | - - - - -                   | CaMNN4  |
| 246 | L - - - - I K I   | - - - - -       | - - - - -                   | PaMNN4  |
| 576 | P K K S T D   | - - - - -       | - - - - -                   | PpMNN4A |
| 463 | P D W V R D   | - - - - -       | - - - - -                   | PpMNN4B |
| 602 | S T A T L P   | - - - - -       | - - - - -                   | PpMNN4C |
| 582 | P P R F S N   | - - - - -       | - - - - -                   | PpPNO1  |
| 630 | P D K V D D G L V N M A T L N I T E L R D Y I T S D E N K N | - - - - -       | - - - - -                   | ScMNN4  |
| 266 | E D R E E T   | - - - - -       | - - - - -                   | NcMNN4A |
| 189 | P A G - - E   | - - - - -       | - - - - -                   | NcMNN4B |
| 221 | P N E W G   | - - - - -       | - - - - -                   | AnMNN4A |
| 188 | R A N G Q P   | - - - - -       | - - - - -                   | AnMNN4B |
| 723 | P K N   | - - - - -       | - - - - -                   | CaMNN4  |
| 249 | P   | - - - - -       | - - - - -                   | PaMNN4  |
| 582 | - - - - -   | - - - - -       | - - - - - S T G S T P E N G | PpMNN4A |
| 469 | - - - - -   | - - - - -       | - - - - - G V K E N P       | PpMNN4B |
| 608 | - - - - -   | - - - - -       | - - - - - G V I E T D       | PpMNN4C |
| 588 | - - - - -   | - - - - -       | - - - - - A S K K D P       | PpPNO1  |
| 660 | H K R V P T D T D L K D L L K K E L E E L P K S K T I E N K | - - - - -       | - - - - -                   | ScMNN4  |
| 272 | - - - - -   | - - - - -       | - - - - - G T R Q G         | NcMNN4A |
| 193 | - - - - -   | - - - - -       | - - - - - G M M S           | NcMNN4B |
| 226 | - - - - -   | - - - - -       | - - - - -                   | AnMNN4A |
| 194 | - - - - -   | - - - - -       | - - - - - G A L M           | AnMNN4B |
| 726 | - - - - -   | - - - - -       | - - - - - F E I K D N N     | CaMNN4  |
| 250 | - - - - -   | - - - - -       | - - - - -                   | PaMNN4  |

Figura 9 (hoja 9)

|     |  |         |
|-----|--|---------|
| 591 | L T R N L R Q N L - - - - - N A Q V <b>Y N C R N</b>                 | PpMNN4A |
| 475 | K N L - - - - - - - - - - - - - - - A L A D K N                      | PpMNN4B |
| 614 | R N K V S D A L E K Q F H D F K F D N F V N K E L F H C R N          | PpMNN4C |
| 594 | - I <b>Y N C R N</b>         | PpPNO1  |
| 690 | L N P K Q R Y F L N - - - - - - - - - - - E K L K L <b>Y N C R N</b> | ScMNN4  |
| 277 | - V W S D K N                | NcMNN4A |
| 197 | - C K D                    | NcMNN4B |
| 226 | - C K N                    | AnMNN4A |
| 198 | - C K D                    | AnMNN4B |
| 733 | Y K P A N - - - - - - - - - - - - - - - E L L Q I <b>Y N C R N</b>   | CaMNN4  |
| 250 | -                    | PaMNN4  |
| 609 | G H F Y Q Y S E L S P L K L S I V E G A L T L I P N D F V T          | PpMNN4A |
| 484 | G N W Y L T R D I L P L R R T I F E G S R S Y T V K D I E D          | PpMNN4B |
| 644 | N H F Y K Y G E V G R L R S T M F E G V P A L I P F E F E S          | PpMNN4C |
| 600 | N H F Y S H N N I A P L K Y T L M E G V P S F I P Q Q Y E E          | PpPNO1  |
| 710 | N H F N S F E E L S P L I N T V F H G V P A L I P H R R T Y          | ScMNN4  |
| 283 | Y H G Y G T R Q I W P L R R T E F E G V E A W V P W D V E E          | NcMNN4A |
| 200 | G H E F R V T I S T S V K S S - - - - - - - - - - - - - - -          | NcMNN4B |
| 229 | N H N Y M L S D I Y P L R A S F F E G V A A K V P Y R Y E S          | AnMNN4A |
| 201 | R H N F D E S E I Y P L R N S Y F E D V P A K I P Y A Y T K          | AnMNN4B |
| 748 | N H F N S Y D E L S P L M K S S V E G E I G Y L P S R Y S T          | CaMNN4  |
| 250 | -                    | PaMNN4  |
| 639 | I L E T E Y Q R R G - L E K N T Y A K Y L Y V P E L R L W M          | PpMNN4A |
| 514 | T L L R N Y G - D K V L I N T E L A D H E W H D D W K M W V          | PpMNN4B |
| 674 | I L K R E Y - - P K G L T L K H F S N H F W D P V N R L W V          | PpMNN4C |
| 630 | I L R E E Y - - T T G L T S K H Y N G N F F M T Q L N L W L          | PpPNO1  |
| 740 | C L H N E Y H V P D R Y A F D A Y K N T A Y L P E F R F W F          | ScMNN4  |
| 313 | I L K E E Y G - V K S L T E E S F A G H Q E F D H G R K Q W V        | NcMNN4A |
| 215 | - - - - - - - - - - - G G G  | NcMNN4B |
| 259 | V L I D E Y G - E K A L S E T H Y N D Y T W V S K Q E E W V          | AnMNN4A |
| 231 | L L Q D E Y - - G A K A L T K T N Y - - - - - - - - - - -            | AnMNN4B |
| 778 | I L T R E Y R S G - - L S S N S H G G Y I F I A K L R L W V          | CaMNN4  |
| 250 | -                    | PaMNN4  |
| 668 | S Y N D I Y D I L Q G T N S H G R P L S A K T M A T I F P R          | PpMNN4A |
| 543 | Q K K K - Y C T Y E E F      | PpMNN4B |
| 702 | P -                  | PpMNN4C |
| 658 | E R D P - M L A L V P S      | PpPNO1  |
| 770 | D Y D G - - - L K K C S N I N S W Y P N I P S I N S W N P N          | ScMNN4  |
| 342 | K T -                | NcMNN4A |
| 217 | -                    | NcMNN4B |
| 288 | S -                  | AnMNN4A |
| 247 | -                    | AnMNN4B |
| 806 | K E E D L Y Y F I K H R D Q W T K Y H S F N T K L S Q D P S          | CaMNN4  |
| 250 | -                    | PaMNN4  |





Figura 9 (hoja 12)

|      |   |         |
|------|---|---------|
| 785  | E K K Q E N R E K E R K E K K E K E E K E K K E K E E - - K   | PpMNN4A |
| 576  | A C G F E E V R Q D W I I T R E S V N L H M K E W E A I Q R   | PpMNN4B |
| 703  | E K K K K I R H I E F S L T K E V T E S H K K E L A Q I H G   | PpMNN4C |
| 699  | L D D N P D I L E E V I R T Y E L T S I H H K E M Q Y L S S   | PpPNO1  |
| 1037 | R K R R E K K K K E E E K K K E E E E K K K K E E E E K       | ScMNN4  |
| 346  |   | NcMNN4A |
| 217  |   | NcMNN4B |
| 289  | - - - - - D E I I A A E E K K - - - - - - - - - - - K         | AnMNN4A |
| 247  | Q G C V I L Q E V E F V V S T                                 | AnMNN4B |
| 917  | E I M Q L T F G K S T A K L M S S A I D F P P I K Y E P Y L   | CaMNN4  |
| 253  | - - - - - E R L D R Q R K A N N E Q G K S E D - - - A         | PaMNN4  |
| 813  | E K K E K E E K E K K E K E E K E R K E K E E K E E Y E E -   | PpMNN4A |
| 606  | N E S T T E Y T A K D L P R Y R P D S F K N L L D G V S N -   | PpMNN4B |
| 733  | N E T G I T S D F A Y S P - - - - - - - - - - - - - - -       | PpMNN4C |
| 729  | V K P D G D R S M Q S N D - - I T S S Y Q E F L A S L K K F   | PpPNO1  |
| 1067 | K K K E E E E K K K K E E E E K K K K E E E E K K K Q E E -   | ScMNN4  |
| 346  |   | NcMNN4A |
| 217  |   | NcMNN4B |
| 300  | A K - E G D -         | AnMNN4A |
| 261  |   | AnMNN4B |
| 947  | Y K L N H D L D T F E N K V D R Y L A L Q D A Y Q Q E H N -   | CaMNN4  |
| 272  | L P A E Q T E G L S D P G A S R N V K R A P V K - - - - -     | PaMNN4  |
| 842  | - D D N E G E Q P T E Q K S Q Q E A K E                       | PpMNN4A |
| 635  | - H G N G N V G K I E H V K L E H N D                         | PpMNN4B |
| 746  | - F R I D P W L S R Y R K K M T R S Q                         | PpMNN4C |
| 757  | Q P L R K D L F Q F E R I D L S K H R K Q                     | PpPNO1  |
| 1096 | - E E K K K K E E E E K K K K Q E E G E K M R N E D E E N K K | ScMNN4  |
| 346  |   | NcMNN4A |
| 217  |   | NcMNN4B |
| 305  | - - - - K D G R Q Y E   | AnMNN4A |
| 261  |   | AnMNN4B |
| 976  | - N S P S G G S D N G F M E I E E D L D F A F                 | CaMNN4  |
| 296  | - - S N K G P E V S   | PaMNN4  |
| 860  |   | PpMNN4A |
| 652  |   | PpMNN4B |
| 763  |   | PpMNN4C |
| 777  |   | PpPNO1  |
| 1125 | N E D E E K K K N E E E E K K K Q E E K N K K N E D E E K K   | ScMNN4  |
| 346  |   | NcMNN4A |
| 217  |   | NcMNN4B |
| 311  |   | AnMNN4A |
| 261  |   | AnMNN4B |
| 997  |   | CaMNN4  |
| 303  |   | PaMNN4  |



**Figura 9 (hoja 13)**

---

|      |   |         |
|------|---|---------|
| 860  |   | PpMNN4A |
| 652  |   | PpMNN4B |
| 763  |   | PpMNN4C |
| 777  |   | PpPNO1  |
| 1155 | K Q E E E E K K K N E E E E K K K Q E E G H S N | ScMNN4  |
| 346  |   | NcMNN4A |
| 217  |   | NcMNN4B |
| 311  |   | AnMNN4A |
| 261  |   | AnMNN4B |
| 997  |   | CaMNN4  |
| 303  |   | PaMNN4  |