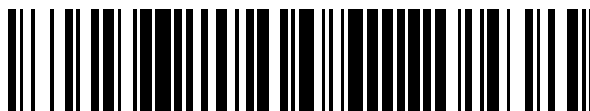


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 659 877**

51 Int. Cl.:

C07K 14/47 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **05.06.2001 E 10185365 (3)**

97 Fecha y número de publicación de la concesión europea: **15.11.2017 EP 2298799**

54 Título: **Un gen que codifica un homólogo de glicoproteína p humana multifármaco-resistente en el cromosoma 7p15-21 y usos del mismo**

30 Prioridad:

05.06.2000 US 208913 P

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

19.03.2018

73 Titular/es:

**THE BRIGHAM AND WOMEN'S HOSPITAL, INC.
(100.0%)
75 Francis Street
Boston, MA 02115, US**

72 Inventor/es:

**SAYEGH, MOHAMED H y
FRANK, MARKUS H**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

Observaciones:

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

ES 2 659 877 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

Un gen que codifica un homólogo de glicoproteína p humana multifármaco-resistente en el cromosoma 7p15-21 y usos del mismo

Campo de la Invención

La invención se refiere a secuencias genéticas que codifican proteínas que exhiben rasgos estructurales y funcionales característicos de los miembros de la familia de la glicoproteína P asociados con la multifármaco-resistencia del cáncer, funciones reguladoras inmunológicas, y funciones singulares en células madre pluripotentes humanas y otras células tisulares progenitoras. La invención abarca proteínas sustancialmente puras, tratamientos terapéuticos y usos diagnósticos relacionados con estas proteínas.

Antecedentes de la Invención

La glicoproteína P, una bomba de flujo de fármacos dependiente de adenosina-trifosfato (ATP), está sobreexpresada en células tumorales multifármaco-resistentes (MDR). La misma reduce la concentración intracelular de xenobióticos citotóxicos, reduciendo con ello la eficacia de muchos regímenes quimioterapéuticos del cáncer. La glicoproteína P pertenece a la superfamilia de transportadores activos ABC (casete de fijación de ATP), y está codificada por una familia de multigenes en los eucariotas superiores. Los miembros de la familia de glicoproteínas P de mamíferos pueden dividirse en tres clases. Las glicoproteínas P de clase I y clase II confieren multifármaco-resistencia, mientras que las proteínas de clase III no lo hacen.

En los humanos, la glicoproteína P está codificada por dos genes enlazados ("MDR1" y "MDR3") en el cromosoma 7q21.1. MDR3 funciona como una translocasa lipídica y las mutaciones en este gen están asociadas con colestasis intrahepática familiar. MDR1 confiere resistencia a fármacos en las células de ciertos cánceres. Además de estar sobreexpresada en las células del cáncer, la glicoproteína P MDR1 se expresa ampliamente en los tejidos humanos normales, predominantemente secretorios y absorbentes, donde la misma funciona en diversos procesos fisiológicos que incluyen diferenciación celular, proliferación celular y supervivencia celular. En estos tipos de células normales, la glicoproteína P funciona en la liberación o captura transmembranal o de xenobióticos y ciertos fármacos terapéuticos, pequeñas moléculas peptídicas, ciertos compuestos esteroidales, y fosfolípidos.

La glicoproteína P es expresada también por poblaciones de células linfoides de la médula ósea humana y la sangre periférica. Específicamente, se ha demostrado que la glicoproteína P se expresa en la membrana de células madre pluripotentes, monocitos, células dendríticas, linfocitos T CD4+ y CD8+, células agresoras naturales, y linfocitos B. En las células inmunológicas, la glicoproteína P funciona en el transporte de citoquinas y otras moléculas pequeñas, que son críticas para que se produzcan las respuestas inmunológicas fisiológicas. El bloqueo específico de la glicoproteína P puede suprimir la respuesta inmunológica a aloantígenos y antígenos nominales. Sin embargo, existe cierto grado de redundancia para la función de la glicoproteína P en estos tipos de células, que apunta a la existencia de moléculas afines adicionales no identificadas hasta ahora.

Las células madre pluripotentes y otras células progenitoras tisulares poseen también una actividad singular semejante a la glicoproteína P, caracterizada por acumulación intracelular reducida de colorantes fluorescentes, que permite el aislamiento específico de estos tipos de células para usos terapéuticos. No obstante, se cree que esta función no está mediada por la glicoproteína P MDR1, sino más bien por un miembro de la familia de glicoproteínas P afín, no identificado todavía.

A pesar del papel irrefutable de la glicoproteína P MDR1 en la multifármaco-resistencia del cáncer, los intentos de mejorar la quimioterapia por inhibición de esta proteína han alcanzado sólo un éxito limitado. Así, puede inferirse que existen proteínas homólogas que, como MDR1, son capaces de hacer las células resistentes a los agentes terapéuticos. Adicionalmente, puede inferirse que las proteínas MDR1 homólogas cumplen funciones semejantes a la glicoproteína P en los tejidos humanos fisiológicos, en particular en células del sistema inmunitario, células madre pluripotentes y células progenitoras tisulares, donde existe cualquier redundancia para la función de la glicoproteína P MDR1, o donde se sabe que la glicoproteína P MDR1 no promueve la actividad observada asociada a la glicoproteína P.

Sumario de la Invención

La invención está dirigida a un nuevo miembro de la familia de genes de glicoproteínas P humanas localizado en el cromosoma 7p15-2, que codifican proteínas que confieren el fenotipo multifármaco-resistente a las células tumorales y/o atienden a funciones fisiológicas críticas en los tejidos humanos normales.

Un examen de la estructura del nuevo gen indica que el mismo codifica dos mitades homólogas semiautónomas, cada una de ellas con sus propios dominios transmembranales y de fijación de ATP. Por remodelación alternativa y expresión diferencial de genes y/o modificaciones posteriores a la transcripción y posteriores a la traducción, el nuevo gen de glicoproteína P puede codificar varias glicoproteínas P distintas:

La proteína de SEQ ID NO: 1 (aminoácidos 1-659) está codificada por 14 exones (SEQ ID NO: 9) del DNA genómico humano del clon AC005060 en el cromosoma 7p15-21 y está constituida por 5 dominios transmembranales y un

dominio de fijación de ATP.

La proteína de SEQ ID NO: 2 (aminoácidos 1-812) está codificada por 19 exones (SEQ ID NO: 10) del DNA genómico humano de los clones contiguos AC002486 y AC005060 (AC002486 es el clon secuenciado a la izquierda del clon AC005060) en el cromosoma 7p15-21 y está constituida por 5 dominios transmembranales y dos dominios de fijación de ATP, el primero de los cuales está localizado en el lado N-terminal del dominio transmembranal #1, y el segundo en el lado C-terminal del dominio transmembranal #5 de la proteína, en el lado opuesto de la membrana plasmática. La proteína de SEQ ID NO: 2 puede expresarse también como resultado de la trans-remodelación del mRNA (SEQ ID NO: 9) que codifica la proteína de SEQ ID NO: 1 y el mRNA (SEQ ID NO: 11) que codifica la proteína de SEQ ID NO: 3 descrita más adelante. Adicionalmente, la proteína de SEQ ID NO: 2 puede expresarse como resultado del procesamiento posterior a la traducción de las proteínas de SEQ ID NO: 1 y NO: 3.

La proteína de SEQ ID NO: 3 (aminoácidos 1-131) está codificada por 6 exones (SEQ ID NO: 11) del DNA genómico humano del clon AC002486 en el cromosoma 7p15-21 y está constituida por un solo dominio de fijación de ATP, careciendo de dominios transmembranales.

La proteína de SEQ ID NO: 4 (aminoácidos 1-1058) está codificada por 20 exones (SEQ ID NO: 12) del DNA genómico humano de los clones contiguos AC002486 y AC005060 en el cromosoma 7p15-21 y está constituida por 8 dominios transmembranales y dos dominios de fijación de ATP, el primero de los cuales está localizado entre los dominios transmembranales #3 y #4, y el segundo en el lado C-terminal de los dominios transmembranales #8, en el lado opuesto de la membrana plasmática.

La proteína de SEQ ID NO: 5 (aminoácidos 1-1222) está codificada por 23 exones (SEQ ID NO: 13) de DNA genómico humano de los clones contiguos AC002486 y AC005060 en el cromosoma 7p15-21 y está constituida por 12 dominios transmembranales y dos dominios de fijación de ATP, el primero de los cuales está localizado entre los dominios transmembranales #7 y #8, y el segundo en el lado C-terminal del dominio transmembranal #12, en el lado opuesto de la membrana plasmática.

La proteína de SEQ ID NO: 6 (aminoácidos 1-1195) está codificada por 24 exones (SEQ ID NO: 14) de DNA genómico humano de los clones contiguos AC002486 y AC005060 en el cromosoma 7p15-21 y está constituida por 11 dominios transmembranales y dos dominios de fijación de ATP, el primero de los cuales está localizado entre los dominios transmembranales #6 y #7, y el segundo en el lado C-terminal del dominio transmembranal #11, en el lado opuesto de la membrana plasmática.

La proteína de SEQ ID NO: 7 (aminoácidos 1-541) está codificada por 10 exones (SEQ ID NO: 15) de DNA genómico humano del clon AC002486 en el cromosoma 7p15-21 y está constituida por 7 dominios transmembranales y un solo dominio de fijación de ATP, en el lado C-terminal del dominio transmembranal #7.

La proteína de SEQ ID NO: 8 (aminoácidos 1-514) está codificada por 11 exones (SEQ ID NO: 16) de DNA genómico humano del clon AC002486 en el cromosoma 7p15-21 y está constituida por 6 dominios transmembranales y un solo dominio de fijación de ATP, en el lado C-terminal del dominio transmembranal #6.

La multifármaco-resistencia del cáncer puede resultar de la expresión de cualquiera de las proteínas de SEQ ID NO: 1, NO: 2, NO: 3, NO: 4, NO: 5, NO: 6, NO: 7 y NO: 8. Las proteínas codificadas por el gen 7p15-21 de glicoproteína P de la presente invención pueden utilizarse como marcadores para identificación de células que exhiben probablemente multifármaco-resistencia y pueden servir como dianas en el diseño de nuevas terapias para los pacientes de cáncer. Debe entenderse que, a menos que se indique lo contrario, la referencia a la glicoproteína P de la presente invención también incluye también cualquiera de las proteínas de SEQ ID NO:1, NO:2, NO:3, NO:4, NO:5, NO:6, NO:7 y NO:8.

La glicoproteína P 7p15-21 confiere quimiorresistencia a múltiples agentes quimioterapéuticos, con inclusión del cisplatino, por mediación del eflujo celular de fármaco. Por tanto, el bloqueo específico de esta función de eflujo, por ejemplo mediante inhibición de anticuerpos monoclonales específicos, puede aumentar la acumulación intracelular de fármaco y, como resultado, la toxicidad de los fármacos y la destrucción de las células tumorales. Adicionalmente, dado que la glicoproteína P 7p15-21 es funcional en la proliferación de las células tumorales, el crecimiento de los tumores puede inhibirse terapéuticamente por administración de anticuerpos monoclonales bloqueantes específicos, incluso en ausencia de agentes quimioterapéuticos concurrentes. Entre las proteínas codificadas por el gen de la glicoproteína P 7p15-21, las proteínas de SEQ ID NO: 1, NO: 2, NO: 3, NO: 4, NO: 5 y NO: 6 son distintas de las proteínas de SEQ ID NO: 7 y NO: 8 en el sentido de que aquéllas se expresan selectivamente en ciertas células de cáncer pero no en tejidos normales no cancerosos. Adicionalmente, las proteínas de SEQ ID NO: 1, NO: 2, NO: 3, NO: 4, NO: 5 y NO: 6 se expresan preferentemente en aquellos cánceres que exhiben los grados más altos de quimiorresistencia a los fármacos quimioterapéuticos, tales como por ejemplo el melanoma humano maligno. Debido a su expresión selectiva en ciertos cánceres pero no en los tejidos normales, las proteínas de SEQ ID NO: 1, NO: 2, NO: 3, NO: 4, NO: 5 y NO: 6 pueden direccionarse terapéuticamente no sólo por inhibición del eflujo de fármacos citotóxicos o la inhibición de la proliferación tumoral por anticuerpos monoclonales específicos, sino también por medios adicionales, que incluyen la destrucción de las células específicas del tumor mediada por anticuerpos monoclonales específicos conjugados a toxinas celulares, o por administración terapéutica de preparaciones de vacuna específicas del antígeno a los pacientes afligidos.

Las proteínas de SEQ ID NO: 7 y NO: 8 codificadas por el gen 7p15-21 pueden expresarse también en ciertos tejidos humanos normales no cancerosos. La divulgación proporciona por tanto usos adicionales relacionados con la función de estas proteínas seleccionadas en tejidos fisiológicos. Entre dichos tejidos normales, las proteínas de SEQ ID NO: 7 y SEQ ID NO: 8 se expresan preferentemente a niveles altos en células madre pluripotentes y otras células tisulares progenitoras, donde aquéllas funcionan en el transporte transmembranal de xenobióticos y otras moléculas pequeñas. La divulgación proporciona así medios para detectar y enriquecer específicamente estas células madre y células progenitoras a partir de mezclas de células y preparaciones en las cuales están contenidas las mismas, por detección de las células con anticuerpos monoclonales específicos marcados.

Las proteínas de SEQ ID NO: 7 y NO: 8 se expresan también en cierto grado en la mayoría de otros tejidos humanos normales, con inclusión de células del sistema inmunitario tales como las células T, monocitos y células presentadoras de antígeno diferenciadas, donde aquéllas funcionan en el flujo de citoquinas y la captura de moléculas pequeñas con inclusión de péptidos y antígenos, cumpliendo así una función crítica para la integridad de las respuestas inmunitarias normales. Cuando se inhiben estas funciones, por ejemplo por bloqueo de anticuerpos monoclonales específicos, la respuesta inmunitaria normal puede modularse, lo cual se puede utilizar en la prevención y/o la terapia del rechazo de aloinjertos en el trasplante clínico de órganos, y también en diversas enfermedades autoinmunes tales como la artritis reumatoide y la esclerosis múltiple. Adicionalmente, cuando se expresa en células inmunitarias humanas y otros tejidos humanos tales como el endotelio de la barrera hematoencefálica y los epitelios del tracto gastrointestinal y el riñón, el bloqueo de la proteína puede emplearse además terapéuticamente para alterar de modo selectivo la captura y secreción, y por tanto la distribución farmacológica, la farmacocinética y la eficacia terapéutica de aquellos fármacos terapéuticos administrados por vía exógena que son sustratos de dichas proteínas.

En un primer aspecto, la invención está dirigida a proteínas sustancialmente puras que comprende la secuencia de aminoácidos de SEQ ID NO: 2. En otra realización, la proteína está constituida por SEQ ID NO: 2. El término "constituida esencialmente por" debe entenderse que abarca proteínas que tienen exactamente las mismas secuencias de aminoácidos, así como proteínas con secuencias insustancialmente diferentes, como se evidencia por el hecho de que poseen las mismas propiedades funcionales básicas. Una isoforma "sustancialmente purificada" es una que se ha separado de otros componentes biológicos acompañantes y comprenderá típicamente al menos 85% de una muestra, prefiriéndose porcentajes mayores. Están disponibles muchos medios para evaluar la pureza de una proteína en una muestra, con inclusión de análisis por electroforesis en gel de poliacrilamida, cromatografía y centrifugación analítica. Un método preferido para evaluar la pureza es mediante transferencia Western utilizando un anticuerpo dirigido contra epítopes de la glicoproteína P 7p15-21 de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 o SEQ ID NO: 8. La divulgación abarca también "péptidos MDR" que se definen en esta memoria como constituidos por un elemento de secuencia de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 o SEQ ID NO: 8, de al menos 10 y preferiblemente al menos 15 ó 20 residuos. Éstos pueden utilizarse en la generación de anticuerpos. Se estipula que un péptido MDR no puede tener una secuencia que sea la misma que cualquier serie de 10 a 15 residuos contiguos en la secuencia LSGGQKQRIAIARAL (SEQ ID NO: 17). Estas proteínas y péptidos MDR pueden administrarse también terapéuticamente a pacientes de cáncer afligidos con tumores que expresan glicoproteína P 7p15-21, como una vacuna tumoral para provocar una respuesta inmunitaria endógena dirigida contra estos tumores, a fin de dar como resultado la destrucción de las células específicas del tumor.

En otra realización, la invención está dirigida a un anticuerpo monoclonal producido por un proceso que comprende el paso de administrar a un animal hospedador una proteína codificada por SEQ ID NO: 2. La divulgación también proporciona anticuerpos obtenidos por la administración de una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 2. La proteína o el péptido deben administrarse al animal en una dosis suficiente para inducir la formación de anticuerpos. Los anticuerpos de la invención son monoclonales. Los anticuerpos se producen preferiblemente por inyección de una preparación farmacéuticamente aceptable en un ratón, seguido por fusión de células del bazo del ratón con células de mieloma utilizando métodos conocidos en la técnica. Los anticuerpos obtenidos deberían fijarse selectivamente a la proteína de SEQ ID NO: 2. En este contexto, fijación selectiva significa que un anticuerpo tiene al menos una afinidad 100 veces mayor para una o más de esas proteínas que para cualquier otra proteína encontrada normalmente en las células humanas.

La invención también está dirigida a un polinucleótido sustancialmente puro constituido esencialmente por una secuencia de nucleótidos que codifica la proteína de SEQ ID NO: 2 o un péptido MDR de SEQ ID NO: 2. El polinucleótido puede estar constituido esencialmente por la secuencia de nucleótidos de SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15 o SEQ ID NO: 16. La invención incluye vectores de expresión que comprenden un elemento codificante distinto constituido por la secuencia polinucleotídica que codifica la proteína de SEQ ID NO: 2; y células hospedadoras transformadas con tales vectores. Un "elemento codificante distinto" hace referencia a la producción de un vector de expresión responsable de la determinación de la secuencia de aminoácidos de una proteína expresada. La invención comprende la totalidad de dichos elementos que producen proteínas correspondientes a la secuencia de aminoácidos que se muestra en SEQ ID NO: 2, así como otras proteínas que tienen sustancialmente la misma estructura y función.

La invención incluye una proteína recombinante producida por células hospedadoras transformadas por un vector de expresión como se ha expuesto anteriormente. La proteína recombinante puede aislarse utilizando técnicas estándar,

que incluyen cromatografía de afinidad con anticuerpos contra los epítopes de la glicoproteína P 7p15-21. El polinucleótido utilizado en los vectores para expresar una glicoproteína P recombinante de este tipo está constituido esencialmente por la secuencia de nucleótidos de SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15 o SEQ ID NO: 16. Oligonucleótidos complementarios a SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15 o SEQ ID NO: 16 y que tienen al menos 15 nucleótidos de longitud pueden utilizarse como inhibidores antisentido. Éstos pueden administrarse a pacientes sometidos a quimioterapia del cáncer a fin de aumentar la eficacia de los fármacos citotóxicos. La transfección *in vivo* de células se conoce desde hace muchos años y puede realizarse utilizando vectores virales (véase *v.g.*, U.S. 6.020.191); liposomas (véase *v.g.*, Nicolau, *Meth. Enzymol.* 149: 157-176 (1987)); DNA complejado con agentes que facilitan la captura celular (véase *v.g.*, U.S. 5.264.618; WO 98/14431); o incluso por simple inyección de DNA desnudo (véase *v.g.*, U.S. 5.693.622). Cualquiera de estos procedimientos puede utilizarse para suministrar los oligonucleótidos antisentido de la presente invención.

La divulgación también está dirigida a un método para determinar si una célula de cáncer responderá a las terapias orientadas a invertir la multifármaco-resistencia por medida de la expresión de los genes que codifican las proteínas de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 o SEQ ID NO: 8. Este método puede utilizarse para detectar la existencia del fenotipo multifármaco-resistente en células de cáncer o para rastrear el desarrollo de multifármaco-resistencia a lo largo del tiempo por monitorización de los cambios en la expresión génica en células cultivadas.

La divulgación también proporciona un método de determinación de si un compuesto de test inhibe la multifármaco-resistencia en células causada por un gen codificante de las proteínas de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 o SEQ ID NO: 8. Este método comprende expresar un gen que codifica uno o más de estos polipéptidos en células que por lo demás no son multifármaco-resistentes, y exponer estas células a uno o más fármacos citotóxicos en presencia de un compuesto de test. Se mide la supervivencia celular después de la exposición, y los resultados obtenidos se comparan con los de incubaciones realizadas esencialmente del mismo modo pero en ausencia del compuesto de test. Se llega a la conclusión de que el compuesto de test inhibe la multifármaco-resistencia si la supervivencia celular se reduce en un grado significativo en las incubaciones realizadas en presencia del compuesto de test con relación a la observada en su ausencia.

Descripción Detallada de la Invención

La invención está dirigida a un nuevo miembro de la familia de glicoproteínas P de una proteína relacionada con la resistencia a los fármacos y a secuencias genéticas que codifican esta proteína. La divulgación también se refiere a métodos de determinación de si una célula de cáncer responderá a las terapias dirigidas a invertir la resistencia a los fármacos mediada por la glicoproteína P, y a un método de cribado de compuestos de test respecto a su capacidad para inhibir multifármaco-resistencia. El nuevo gen de glicoproteína P codifica la proteína de SEQ ID NO: 2.

Debe entenderse que la invención abarca no sólo secuencias idénticas a las representadas, sino también secuencias que son esencialmente iguales como se demuestra por el hecho de que retienen las mismas características estructurales y funcionales básicas. Por ejemplo, pueden utilizarse técnicas tales como la mutagénesis orientada para introducir variaciones en la estructura de una proteína. La invención abarca variaciones en la glicoproteína P introducidas por este u otros métodos similares, con la condición de que la proteína resultante retiene sus propiedades biológicas básicas, en particularmente en lo que respecta a la inducción de multifármaco-resistencia en células de mamífero.

Las secuencias de DNA que codifican las proteínas de la invención pueden obtenerse a partir de cualquier fuente de tejido o célula en la cual se expresen las mismas. Por ejemplo, pueden manipularse genéticamente líneas de células cultivadas para expresar el gen de glicoproteína P utilizando técnicas recombinantes o por exposición continua a agentes quimioterapéuticos. Alternativamente, las secuencias pueden aislarse de las células primarias obtenidas de tumores.

Están disponibles muchos métodos para aislamiento de las secuencias de DNA, y pueden adaptarse para el aislamiento del gen de glicoproteína P del cromosoma 7p15-21 (en lo sucesivo "cromosoma 7p") (véase, *v.g.*, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press (1989)). Por ejemplo, un método consiste en cribar una biblioteca de cDNA que se ha preparado por transcripción inversa de mRNA aislado de tejidos o células que expresan el gen. La biblioteca puede prepararse a partir de, por ejemplo, tejido de melanocitos o de testículo humano, y pueden sintetizarse sondas para cribado basadas en las secuencias que se representan en el Listado de Secuencias. Las sondas tienen preferiblemente al menos 14 nucleótidos de longitud y se seleccionan óptimamente de una región que se cree es exclusiva para el gen de la glicoproteína P del cromosoma 7p.

Como alternativa, la amplificación de una secuencia deseada puede realizarse por la reacción en cadena de la polimerasa ("PCR") del RNA inverso transcrito. Iniciadores para la PCR pueden construirse utilizando las secuencias que se representan en SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15 o SEQ ID NO: 16, y la confirmación de la presencia del cDNA de la glicoproteína P del cromosoma 7p puede obtenerse por la secuenciación de los productos amplificados.

La expresión de proteína recombinante puede inducirse en una célula hospedadora por transformación de la misma con un vector de expresión apropiado. El vector debería contener señales de transcripción y traducción reconocibles por el hospedador junto con la secuencia estructural deseada, preferiblemente en forma bicatenaria, en un enlace operativo. Por ejemplo, la secuencia de DNA de la glicoproteína P debería estar posicionada de tal modo que las secuencias reguladoras presentes en el vector controlen la síntesis de mRNA y se produzca proteína que tenga la secuencia deseada.

El ácido nucleico que codifica la glicoproteína P de la invención se puede expresar en células eucariotas, especialmente en células de mamífero. Dichas células son capaces de promover las modificaciones posteriores a la traducción necesarias para asegurar que la proteína recombinante es estructural y funcionalmente la misma que la proteína aislada de, por ejemplo, células tumorales multifármaco-resistentes. Ejemplos de células de mamífero que se sabe proporcionan modificación apropiada posterior a la traducción de proteínas clonadas incluyen, *inter alia*, células N11i-3T3, células CHO, células HeLa, células LM(tk-), y análogas. Promotores eucariotas que se sabe controlan la expresión de genes recombinantes se utilizan preferiblemente para dirigir la transcripción del DNA de la glicoproteína P del cromosoma 7p, y pueden incluir el del gen de metalotioneína I del ratón, el promotor TK del Herpesvirus, el promotor temprano de CMV y el promotor temprano de SV40. La transcripción puede estar dirigida también por promotores procariotas, tales como los capaces de reconocer la polimerasa T4, los promotores P_R y P_L del bacteriófago lambda, y los promotores trp, recA, choque térmico y lacZ de *E. coli*.

Pueden introducirse vectores de expresión en las células hospedadoras por métodos tales como precipitación con fosfato de calcio, microinyección, electroporación o transferencia viral, y pueden seleccionarse células que expresan la secuencia de proteínas recombinante por métodos conocidos en la técnica. La confirmación de la expresión puede obtenerse por amplificación PCR de secuencias de glicoproteína P utilizando iniciadores seleccionados de las secuencias que se muestran en SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15 o SEQ ID NO: 16.

La proteína recombinante puede purificarse utilizando métodos estándar bien conocidos en la técnica. Dichos métodos pueden incluir filtración, precipitación, cromatografía y métodos electroforéticos. La pureza puede evaluarse por realización de la electroforesis en un gel de poliacrilamida y visualización de las proteínas utilizando metodología estándar de tinción. La transferencia Western puede realizarse también utilizando un anticuerpo para la glicoproteína P del cromosoma 7p.

La invención está dirigida también a anticuerpos generados contra la proteína P del cromosoma 7p de SEQ ID NO: 2. El proceso para producir tales anticuerpos puede implicar o bien inyectar la glicoproteína P 7p propiamente dicha en un animal apropiado o inyectar péptidos antigénicos cortos construidos para corresponder a diferentes regiones de la proteína. Estos péptidos deberían tener al menos 5 aminoácidos de longitud y preferiblemente, deberían seleccionarse de regiones que se consideran exclusivas de la glicoproteína P 7p. Métodos para generación y detección de anticuerpos son bien conocidos en la técnica, y se exponen en referencias tales como: Harlow, *et al.*, Antibodies, A Laboratory Manual, Cold Spring Harbor Laboratory, NY (1988); Klein, Immunology: The Science of Self-Nonself Discrimination, (1982); Kennett *et al.*, Monoclonal Antibodies and Hybridomas: A New Dimension in Biological Analyses, (1980); y Campbell, "Monoclonal Antibody Technology", en Laboratory Techniques in Biochemistry and Molecular Biology, (1984).

El término "anticuerpo", como se utiliza en esta memoria, debe entenderse que incluye moléculas intactas así como fragmentos que retienen su capacidad para fijar antígeno, tales como los fragmentos Fab y F(ab')₂. El término "anticuerpo" se define también en esta memoria como haciendo relación tanto a anticuerpos monoclonales como a anticuerpos policlonales. Los anticuerpos policlonales se derivan de los sueros de animales inmunizados con un antígeno de la glicoproteína P del cromosoma 7p. Los anticuerpos monoclonales para la proteína pueden prepararse utilizando la tecnología del hibridoma, como se expone en referencias tales como: Kohler, *et al.*, Nature **256**:495 (1975); y Hammerling, *et al.*, en: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981). En general, esta tecnología implica inmunizar un animal inmunocompetente, típicamente un ratón, sea con la glicoproteína P intacta del cromosoma 7p o con un fragmento derivado de la misma. Se extraen luego esplenocitos del animal inmunizado y se fusionan con células de mieloma adecuadas, tales como células SP₂O. Después de ello, las células de hibridoma resultantes se mantienen selectivamente en medio HAT y se clonan luego por dilución limitada (Wands, *et al.*, Gastroenterology, **80**: 225-232 (1981)). Las células obtenidas por dicha selección se ensayan luego para identificar clones que secreten anticuerpos capaces de fijar la glicoproteína P del cromosoma 7p.

Los anticuerpos o fragmentos de anticuerpos de la invención pueden utilizarse para detectar la presencia de la glicoproteína P del cromosoma 7p en cualquiera de una diversidad de inmunoensayos. Por ejemplo, pueden utilizarse anticuerpos en radioinmunoensayos o en ensayos inmunométricos, conocidos también como ensayos de "dos sitios" o "sándwich" (véase Chard, "An Introduction to Radioimmune Assay and related Techniques," en: Laboratory Techniques in Biochemistry and Molecular Biology, North Holland Publishing Co., NY (1978)). En un ensayo inmunométrico típico, una cantidad de anticuerpo sin marcar se fija un soporte sólido que es insoluble en el fluido sometido a test, tal como sangre, linfa, extractos celulares y análogos. Después de la fijación inicial del antígeno al anticuerpo inmovilizado, se añade una cantidad de un segundo anticuerpo marcado detectablemente (que puede ser o no el mismo que el primero) a fin de permitir la detección y/o cuantificación de antígeno fijado (véase, v.g. Radioimmune Assay Method, Kirkham, *et al.*, Ed. pp. 199-206, E&S y Livingstone, Edimburgo (1970)). Muchas variaciones de estos tipos de ensayos se conocen en la técnica y pueden emplearse para la detección de la glicoproteína P 7p.

Los anticuerpos para la glicoproteína P del cromosoma 7p pueden utilizarse también en procedimientos de purificación (véase en general, Dean *et al.*, *Affinity Chromatography, A Practical Approach*, IRL Press (1986)). Típicamente, el anticuerpo se inmoviliza en una matriz cromatográfica tal como Sepharosa, 4B. La matriz se introduce luego como relleno en una columna y la preparación que contiene la glicoproteína P del cromosoma 7p se hace pasar a su través en condiciones que promueven la fijación, v.g., condiciones bajas en sal. La columna se lava luego, y la proteína se eluye utilizando tampón que promueve la disociación del anticuerpo, v.g., un tampón que tiene un pH o concentración de sal alterados. La proteína eluida puede transferirse a un tampón, por ejemplo por diálisis, y almacenarse después de ello o utilizarse directamente. Pueden utilizarse también anticuerpos en transferencia Western para la detección de la glicoproteína P del cromosoma 7p en una muestra. Para estos tipos de ensayos, puede utilizarse un anticuerpo que o bien se ha desarrollado específicamente para reaccionar con la glicoproteína P del cromosoma 7p o que reacciona con un epítipo de la proteína.

La detección de la glicoproteína P del cromosoma 7p puede utilizarse para determinar si las células tumorales son multifármaco-resistentes. Análogamente, la detección de cambios en la expresión de la glicoproteína P puede ser útil en la predicción del desarrollo de multifármaco-resistencia en las células. El cDNA de esta glicoproteína P puede ser útil en el diseño de iniciadores para PCR de diagnóstico, diseño de sondas para transferencia Northern de diagnóstico, ensayos de protección de RNAsas, y para el diseño de los oligonucleótidos antisentido complementarios al cDNA predicho para uso en estrategias de direccionamiento de genes para la inversión de la multifármaco-resistencia. Se contemplan usos diagnósticos y terapéuticos tanto *in vitro* como *in vivo* para secuencias de nucleótidos antisentido para la glicoproteína P del cromosoma 7p.

La secuencia de aminoácidos primaria y la estructura proteínica de la glicoproteína P del cromosoma 7p pueden utilizarse en la producción de anticuerpos monoclonales (mAbs) que pueden emplearse en la diagnosis y terapia del cáncer multifármaco-resistente. Por ejemplo, péptidos sintéticos que se asemejan a secuencias nativas de aminoácidos de dominios extracelulares particulares como se determinan por predicción de topología de membrana pueden ser útiles para desarrollar mAbs inhibidores dirigidos contra epítopes extracelulares de la glicoproteína P del cromosoma 7p. Adicionalmente, secuencias peptídicas sintéticas de 10-20 meros derivadas de la secuencia primaria de aminoácidos no incluidas en las secuencias de bucle extracelular arriba mencionadas pueden ser útiles en el desarrollo de anticuerpos monoclonales específicos de diagnóstico. Pueden emplearse mAbs específicos en el análisis diagnóstico por FACS, transferencia Western, e inmunohistoquímica. Tales mAbs pueden emplearse también para usos diagnósticos *in vivo*, donde pueden utilizarse mAbs conjugados a un marcador para evaluar la carga tumoral, la localización del tumor o la masa residual tumoral después de quimioterapia o terapia quirúrgica de tumores que expresan la glicoproteína P 7p15-21.

Pueden utilizarse también mAbs específicos para propósitos terapéuticos en pacientes de cáncer. En particular, aquéllos pueden administrarse para invertir la resistencia multifármaco del cáncer en pacientes que reciben agentes quimioterapéuticos que son sustratos para el eflujo de glicoproteína P 7p, v.g., cisplatino. Adicionalmente, pueden utilizarse terapéuticamente mAbs específicos en pacientes de cáncer para destrucción de las células específicas del tumor, sea administrados en una forma no conjugada, que da como resultado la destrucción del tumor mediada por el sistema inmunitario, o en una forma conjugada con una toxina celular (por ejemplo conjugada a yodo reactivo o toxinas químicas), dando como resultado la destrucción directa de las células específicas del tumor.

Pueden utilizarse también mAbs específicos para propósitos terapéuticos distintos de la resistencia del cáncer a multifármacos. Basándose en la función inmunorreguladora predicha de la glicoproteína P 7p, estos mAbs pueden administrarse a pacientes a fin de prevenir y/o tratar el rechazo de trasplantes de órganos, así como diversas enfermedades autoinmunes tales como artritis reumatoide y esclerosis múltiple. Adicionalmente, dado que las glicoproteínas P funcionan en la captura, excreción y distribución específica de tejido de una diversidad de compuestos farmacológicos y químicos, y se han visto implicadas en mecanismos de biodisponibilidad oral, función de la barrera hematoencefálica y mecanismos de excreción renal, hepática y biliar de diversos fármacos, pueden administrarse mAbs específicos terapéuticamente para alterar la farmacocinética y la disponibilidad de dichos fármacos terapéuticos que son sustratos para la función de transporte mediada por la glicoproteína P 7p.

Las composiciones y métodos de la presente invención pueden tener cierto número de usos en además de los arriba descritos. Por ejemplo, se sabe que las células madre pluripotentes y células progenitoras tisulares tales como células madre hematopoyéticas, células neuroprogenitoras y células progenitoras musculares poseen actividades de eflujo afines a la glicoproteína P para pequeñas moléculas y colorantes fluorescentes. La glicoproteína P del cromosoma 7p puede jugar un papel en el transporte de dichos sustratos, y puede servir por tanto como marcador para el aislamiento de tales células madre y células progenitoras mediante, por ejemplo, análisis FACS. Asimismo, dado que la glicoproteína P MDR1 parece estar implicada en la diferenciación celular, la proliferación celular, la supervivencia celular, y ciertas respuestas inmunitarias, la glicoproteína P del cromosoma 7p, debido a su homología con la glicoproteína P MDR1, se espera que juegue también un papel en dichas funciones fisiológicas. Así pues, las secuencias del gen de la glicoproteína P y de la proteína del cromosoma 7p pueden ser útiles en la modulación de las alteraciones patofisiológicas de estas funciones relacionadas con MDR.

Ejemplos

Dado que actualmente está produciéndose nueva información de secuencias genómicas a un ritmo rápido por el

proyecto del genoma humano, las bases de datos que contienen dicha información genómica contienen potencialmente secuencias de miembros no identificados hasta ahora de la familia de glicoproteínas P. Los miembros de la familia de glicoproteínas P de mamífero comparten secuencias de aminoácidos y epítopes de proteínas característicos, y asumen conformaciones similares. Por ello, se ha realizado una investigación basada en homología de proteínas en un intento de identificar nuevos genes codificantes de glicoproteínas P. Se han utilizado herramientas bioinformáticas de analítica de genes y analítica de proteínas para caracterizar ulteriormente la secuencia de ácido nucleico y la estructura predicha de las proteínas de los genes candidato identificados. Específicamente, se ha utilizado la aplicación tblastn del National Center for Biotechnology Information (NCBI) para comparar las secuencias de aminoácidos conservadas derivadas de la estructura conocida de la glicoproteína P MDR1 humana con la base de datos de la secuencia de nucleótidos de *homo sapiens* no redundante del NCBI traducida dinámicamente en todos los marcos de lectura. Se ha utilizado la secuencia de firma común a los miembros de la familia de transportadores ABC, una secuencia de aminoácidos 15-meros LSGGQKQRIAIARAL (SEQ ID NO: 17), para identificar secuencias de DNA genómico humanas codificantes de estructuras de proteínas homólogas. Se han empleado también secuencias conocidas de aminoácidos hexámeros de tres epítopes de fijación de anticuerpos monoclonales (mAb) específicos de glicoproteína P.

Los clones de DNA genómico humano identificados de la manera arriba descrita se cribaron respecto a contaminación con vectores utilizando el programa VecScreen. Adicionalmente, se sometieron estos clones a mapeado sistemático de homología utilizando secuencias de aminoácidos 20-meros solapantes contiguas derivadas de la estructura de la proteína humana MDR1 y el programa de búsqueda tblastn. Se compararon las secuencias de DNA genómico candidato que codificaban secuencias de aminoácidos homólogas con secuencias de marco de lectura abierto (ORF) predichas en cada clon de DNA utilizando el programa NCBI ORF Finder (Altschul, *et al.*, *Nucleic Acids Res.* 25: 3389-402 (1997)). Se analizaron luego las secuencias de DNA homólogas que contenían ORFs genómicos utilizando el paquete de software NetGene2 a fin de predecir sitios de remodelación de intrones en los genes candidato (Brunak *et al.*, *J. Biol.* 220: 49-65 (1991)).

Se generó una secuencia de DNA por transcripción conceptual lineal de estructuras predichas de exones de DNA adyacentes. Utilizando este enfoque, se identificaron dos clones genómicos humanos adyacentes solapantes, CTA-367017 (AC002486, de 79611 pares de bases de longitud) y CTB-86D3 (AC005060, de 120169 pares de bases de longitud, secuenciado a la derecha) como formando parte de una isla sin anclaje de orientación desconocida en el cromosoma 7p15-21. Se encontró que estos clones solapantes contienen una secuencia génica que codifica un nuevo miembro de la familia de glicoproteínas P humanas.

Con objeto de determinar si la estructura génica predicha se expresaba en tejidos humanos, la secuencia de cDNA generada se comparó con la base de datos humana NCBI dbest de marcadores de secuencia expresados (EST) no redundantes, como ha sido descrito por Altschul *et al.*, y se identificaron varios ESTs complementarios a exones predichos del clon genómico AC002486. Se diseñaron luego iniciadores de la reacción en cadena de polimerasa (PCR) basados en información de secuencias disponible en la base de datos del National Center for Biotechnology Information (NCBI) y el análisis bioinformático que se ha descrito arriba. Utilizando estos iniciadores oligonucleotídicos específicos de genes y la técnica PCR sobre RNA mensajero total (mRNA) transcrito inversamente aislado de varias líneas de células de cáncer humanas y tejidos humanos normales, con inclusión de la línea de células del melanoma humano G3661, la línea de células del carcinoma de mama MCF-7, la línea de células del carcinoma de células escamosas SCC25, la línea de células de leucemia U937, y células mononucleares de sangre periférica (PBMC) normales, se amplificaron secuencias de cDNA derivadas del nuevo gen de glicoproteína P 7p15-21 y los productos PCR se secuenciaron subsiguientemente utilizando el método de terminación de cadenas didesoxi en ambas cadenas.

La estructura intrón-exón de varios productos génicos codificados por el gen de glicoproteína P 7p15-21 se determinó por comparación de clones de cDNA predichos y secuenciados con información de secuencias genómica del locus del gen de glicoproteína P 7p15-21 (clones AC002486 y AC005060) como se muestra en SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15 y SEQ ID NO: 16. Se generaron luego estructuras de proteínas codificadas por el nuevo gen 7p15-21 por traducción conceptual de aminoácidos de las secuencias oligonucleotídicas predichas de SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15 y SEQ ID NO: 16, como se muestra en SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 y SEQ ID NO: 8. Estas secuencias de aminoácidos se compararon luego con la secuencia peptídica no redundante del NCBI respecto a homología de secuencia utilizando el programa blastp del NCBI. Las secuencias de aminoácidos predichas de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7 y SEQ ID NO: 8 se clasificaron también utilizando el Sistema de Clasificación de Familias de Proteínas PIR-International (Barker, *et al.*, *Nucleic Acids Res.* 28: 41-4 (2000); Huang *et al.*, *Nucleic Acids Res.* 28: 273-6 (2000)). Se determinaron las características funcionales potenciales de las proteínas predichas por análisis comparativo de la composición primaria de aminoácidos y por utilización del paquete de software TMHMM1.0 para predicción de la formación transmembranal de hélices en proteínas de mamífero (Sonnhammer *et al.*, *Ismb.* 6: 175-82 (1998)).

El nuevo gen 7p15-21 de glicoproteína P puede codificar varias isoformas de glicoproteína P distintas que exhiben 68% de homología de secuencia a la vez con MDR1 y MDR3 humanas. Se encontró un grado similar de homología con isoformas respectivas de ratón y hámster de estos genes humanos. El análisis de la secuencia primaria de aminoácidos sugiere que la glicoproteína P del cromosoma 7p15-21 puede expresar el epítipo de fijación de mAb C32 y anti-

glicoproteína P, pero no el epítipo C219 conservado en todas las restantes isoformas de glicoproteína P conocidas (Georges *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* **87**: 152-6 (1990)).

La predicción estructural reveló que el gen 7p15-21 de glicoproteína P codifica isoformas de glicoproteína P que exhiben similitudes estructurales pero también diferencias distintivas comparadas con miembros conocidos de la familia de glicoproteína P como se describe por Georges *et al.* Por ejemplo, la proteína de SEQ ID NO: 2 contiene dos dominios de fijación de ATP que están localizados en lados opuestos de la membrana plasmática, proporcionando un dominio de fijación de ATP extracelular único que se predice fijará ATP extracelular. Basándose en estas diferencias distintivas, se predice que la glicoproteína P 7p15-21 está implicada no sólo en eflujo de moléculas pequeñas, sino que algunas de sus isoformas son funcionales también en la captura de moléculas pequeñas dependiente de energía. El sistema de clasificación PIR confirmó que la glicoproteína P del cromosoma 7p15-21 descubierta es un miembro de la familia de proteínas de multifármaco-resistencia y la familia de superfamilias de homología de la casete de fijación de ATP.

El análisis PCR utilizando iniciadores específicos de genes demostró que el cDNA codificante de las proteínas de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 y SEQ ID NO: 6, que implica en cada caso exones codificados en el clon genómico AC005060, se expresaba preferentemente en células del melanoma humano pero no en la mayoría de los otros cánceres testados, al contrario que los cDNAs codificantes de las proteínas de SEQ ID NO: 7 y SEQ ID NO: 8, que se encontró se expresaban en la mayoría de los cánceres examinados y también en tejidos humanos fisiológicos. Esto pone de manifiesto que un subconjunto de productos del gen de glicoproteína P 7p15-21 pueden direccionarse selectivamente en ciertos cánceres que exhiben grados particularmente altos de quimiorresistencia, tales como el melanoma humano.

Para evaluar la expresión y función de la glicoproteína P 7p15-21 y el efecto de la modulación específica sobre la función de transporte y la quimiorresistencia, se generaron anticuerpos policlonales contra los péptidos MDR CGTSLILNGEPGYTI (SEQ ID NO: 18) y RFGAYLIQAGRMTPEGC (SEQ ID NO: 19), correspondientes a distintos epítopos de bucles extracelulares de la glicoproteína P 7p15-21, por inyección de ratones con estos péptidos antigénicos conjugados a la sustancia portadora KLH. A fin de evaluar la expresión superficial de la glicoproteína P 7p15-21 de las células tumorales humanas, se realizó una inmunotinción superficial indirecta y citometría de flujo monocolor de células recién cosechadas. Para valorar los efectos de la inhibición de P-gp 7p15-21 sobre el eflujo de tintes fluorescentes mediado por P-gp, se incubaron células tumorales con Ab policlona de anti-7p15-21 de glicoproteína P seguido por adición de calceína-AM y medidas seriadas subsiguientes de fluorescencia celular por citometría de flujo.

Estos estudios demostraron que la glicoproteína P se expresa en células tumorales, y que el epítipo RFGAYLIQAGRMTPEGC (SEQ ID NO: 19) contenido en las proteínas de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 y SEQ ID NO: 6, se expresa preferentemente en el melanoma humano a niveles altos, mientras que el epítipo CGTSLILNGEPGYTI (SEQ ID NO: 18) contenido también en SEQ ID NO: 7 y SEQ ID NO: 8, se expresa también en otros tipos de cánceres humanos y en células humanas normales. Los anticuerpos contra el epítipo CGTSLILNGEPGYTI (SEQ ID NO: 18) inhibían a la vez la captura de colorante y asimismo el eflujo de colorante dependiente del tipo de célula, indicando una función dual de los diversos productos génicos de la glicoproteína P 7p15-21 en estos procesos distintos. Estos anticuerpos aumentaban también la citotoxicidad celular del cisplatino en ensayos específicos de destrucción de células en el melanoma y asimismo en el cáncer de mama entre otros, indicativos de su utilidad terapéutica potencial en el tratamiento de pacientes de cáncer.

Se sabe que ciertos cánceres exhiben reordenación cromosómica en la región 7p15-21, y dichas mutaciones pueden asociarse con la aparición del fenotipo MDR. Esto aumenta la posibilidad de que la reordenación génica en estos cánceres sea potencialmente resultado de la formación de episomas y minicromosomas dobles (DM) durante el proceso de amplificación génica de la glicoproteína P 7p15-21 bajo tensiones mutagénicas tales como la quimioterapia. Se sabe que las células que expresan multifármaco-resistencia mediada por MDR1 sufren tales transposiciones cromosómicas y formación de cromosomas DM (Scehoenlein *et al.*, *Mol. Biol. Cell* **3**:507-20 (1992); Mickley *et al.*, *J. Clin. Invest.* **99**:1947-57 (1997); Knutsen *et al.*, *Genes Chromosomes Cancer* **23**:44-54 (1998)). Así pues, los productos génicos de la glicoproteína P del cromosoma 7p15-21 de SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 y SEQ ID NO: 6 pueden sobreexpresarse selectivamente en ciertas células de cáncer, contribuyendo con ello a la resistencia adquirida a los fármacos de dichas células de cáncer en tanto que se mantienen silenciosos en las células normales. Este patrón de expresión diferencial puede emplearse en la detección e inversión de la multifármaco-resistencia de las células tumorígenas de mamífero.

REALIZACIONES ESPECÍFICAS

Una proteína sustancialmente pura, constituida esencialmente por una secuencia de aminoácidos seleccionada del grupo constituido por SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7 y SEQ ID NO:8.

Un péptido constituido por un elemento de secuencia derivado de cualquiera de: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7 o SEQ ID NO:8; donde dicho péptido tiene una longitud de al menos 10 residuos y siempre que dicha secuencia no sea la misma que cualquiera de los 10-15 aminoácidos contiguos en la secuencia LSGGQKQRIAIARAL.

Un anticuerpo generado mediante un proceso que comprende el paso de administrar la proteína o el péptido de cualquiera de las reivindicaciones 1-2 a un animal capaz de producir dicho anticuerpo, donde dicha proteína o péptido se administra con una dosis suficiente para inducir la formación de anticuerpos en dicho animal.

Un anticuerpo que se une preferentemente a la proteína de la reivindicación 1.

Un polinucleótido sustancialmente puro que codifica la proteína o el péptido de la reivindicación 1 o de la reivindicación 2.

Un vector para expresar una glicoproteína P, que comprende un elemento codificante distinto constituido por el polinucleótido de la reivindicación 5.

Una célula hospedadora transformada con el vector de la reivindicación 6.

Una glicoproteína P recombinante producida por la célula hospedadora de la reivindicación 7.

El polinucleótido de la reivindicación 5, donde dicho polinucleótido tiene una secuencia de nucleótidos seleccionada del grupo constituido esencialmente por: SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15 y SEQ ID NO:16.

Un oligonucleótido que actúa como inhibidor antisentido de la expresión de la glicoproteína P, donde dicho oligonucleótido tiene una longitud de al menos 15 nucleótidos y está constituido por una secuencia complementaria a al menos 15 nucleótidos contiguos en cualquiera de: SEQ ID NO: 9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15 o SEQ ID NO:16.

Un vector para expresión de proteína, que comprende un elemento codificante distinto constituido por el polinucleótido de la reivindicación 9.

Una célula hospedadora transformada con el vector de la reivindicación 11.

Un método de determinación de si una célula de cáncer responderá a una terapia orientada a invertir la multifármaco-resistencia, que comprende el paso de medir la expresión de un gen que codifica una proteína seleccionada del grupo de: SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7 y SEQ ID NO:8.

El método de la reivindicación 13, donde dicha expresión se determina utilizando amplificación PCR de mRNA transcrito inversamente.

El método de la reivindicación 13, donde dicha expresión se determina utilizando el anticuerpo de la reivindicación 4.

Un método de determinación de si un compuesto de test inhibe la multifármaco-resistencia causada por un gen que codifica una proteína seleccionada del grupo constituido por SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7 y SEQ ID NO:8; donde dicho método comprende:

(a) expresar dicho gen en células que por lo demás no son multifármaco-resistentes;

(b) exponer dichas células a uno o más agentes citotóxicos en presencia de dicho compuesto de test;

(c) medir la supervivencia celular después de exposición de las células a dichos uno o más agentes citotóxicos y comparar los resultados obtenidos en el paso (b) con los de células incubadas esencialmente del mismo modo con dichos agentes citotóxicos pero en ausencia de dicho compuesto de test; y

(d) deducir que dicho compuesto de test inhibe la multifármaco-resistencia si la supervivencia celular se reduce en un grado significativo por incubación de células en presencia de dicho compuesto de test con relación a la supervivencia de las células en incubaciones realizadas en ausencia de dicho compuesto de test.

El método de la reivindicación 16, donde dicho gen está constituido esencialmente por una secuencia de nucleótidos seleccionada del grupo constituido por: SEQ ID NO: 9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13,

ES 2 659 877 T3

SEQ ID NO:14, SEQ ID NO:15 y SEQ ID NO:16.

LISTADO DE SECUENCIAS

<110> The Brigham and Women's Hospital, Inc.

<120> Un Gen que Codifica un Homólogo de Glicoproteína P Humana Multifármaco-Resistente en el Cromosoma 7p15-21 y Usos del Mismo

<130> B0801.70347EP03

<160> 19

<170> PatentIn versión 3.0

<210> 1

<211> 659

<212> PRT

<213> Homo sapiens

<400> 1

Met Leu Ala Glu Lys Gly Ala His Ala Glu Leu Met Ala Lys Arg Gly
1 5 10 15
Leu Tyr Tyr Ser Leu Val Met Ser Gln Asp Ile Lys Lys Ala Asp Glu
20 25 30
Gln Met Glu Ser Met Thr Tyr Ser Thr Glu Arg Lys Thr Asn Ser Leu
35 40 45
Pro Leu His Ser Val Lys Ser Ile Lys Ser Asp Phe Ile Asp Lys Ala
50 55 60
Glu Glu Ser Thr Gln Ser Lys Glu Ile Ser Leu Pro Glu Val Ser Leu
65 70 75 80
Leu Lys Ile Leu Lys Leu Asn Lys Pro Glu Trp Pro Phe Val Val Leu
85 90 95
Gly Thr Leu Ala Ser Val Leu Asn Gly Thr Val His Pro Val Phe Ser
100 105 110
Ile Ile Phe Ala Lys Ile Ile Thr Met Phe Gly Asn Asn Asp Lys Thr
115 120 125
Thr Leu Lys His Asp Ala Glu Ile Tyr Ser Met Ile Phe Val Ile Leu
130 135 140
Gly Val Ile Cys Phe Val Ser Tyr Phe Met Gln Gly Leu Phe Tyr Gly
145 150 155 160

ES 2 659 877 T3

Arg Ala Gly Glu Ile Leu Thr Met Arg Leu Arg His Leu Ala Phe Lys
 165 170 175
 Ala Met Leu Tyr Gln Asp Ile Ala Trp Phe Asp Glu Lys Glu Asn Ser
 180 185 190
 Thr Gly Gly Leu Thr Thr Ile Leu Ala Ile Asp Ile Ala Gln Ile Gln
 195 200 205
 Gly Ala Thr Gly Ser Arg Ile Gly Val Leu Thr Gln Asn Ala Thr Asn
 210 215 220
 Met Gly Leu Ser Val Ile Ile Ser Phe Ile Tyr Gly Trp Glu Met Thr
 225 230 235 240
 Phe Leu Ile Leu Ser Ile Ala Pro Val Leu Ala Val Thr Gly Met Ile
 245 250 255
 Glu Thr Ala Ala Met Thr Gly Phe Ala Asn Lys Asp Lys Gln Glu Leu
 260 265 270
 Lys His Ala Gly Lys Ile Ala Thr Glu Ala Leu Glu Asn Ile Arg Thr
 275 280 285
 Ile Val Ser Leu Thr Arg Glu Lys Ala Phe Glu Gln Met Tyr Glu Glu
 290 295 300
 Met Leu Gln Thr Gln His Arg Asn Thr Ser Lys Lys Ala Gln Ile Ile
 305 310 315 320
 Gly Ser Cys Tyr Ala Phe Ser His Ala Phe Ile Tyr Phe Ala Tyr Ala
 325 330 335
 Ala Gly Phe Arg Phe Gly Ala Tyr Leu Ile Gln Ala Gly Arg Met Thr
 340 345 350
 Pro Glu Gly Met Phe Ile Val Phe Thr Ala Ile Ala Tyr Gly Ala Met
 355 360 365
 Ala Ile Gly Lys Thr Leu Val Leu Ala Pro Glu Tyr Ser Lys Ala Lys
 370 375 380
 Ser Gly Ala Ala His Leu Phe Ala Leu Leu Glu Lys Lys Pro Asn Ile
 385 390 395 400
 Asp Ser Arg Ser Gln Glu Gly Lys Lys Pro Asp Thr Cys Glu Gly Asn
 405 410 415
 Leu Glu Phe Arg Glu Val Ser Phe Phe Tyr Pro Cys Arg Pro Asp Val
 420 425 430
 Phe Ile Leu Arg Gly Leu Ser Leu Ser Ile Glu Arg Gly Lys Thr Val
 435 440 445
 Ala Phe Val Gly Ser Ser Gly Cys Gly Lys Ser Thr Ser Val Gln Leu
 450 455 460
 Leu Gln Arg Leu Tyr Asp Pro Val Gln Gly Gln Val Leu Phe Asp Gly
 465 470 475 480

ES 2 659 877 T3

Val Asp Ala Lys Glu Leu Asn Val Gln Trp Leu Arg Ser Gln Ile Ala
 485 490 495
 Ile Val Pro Gln Glu Pro Val Leu Phe Asn Cys Ser Ile Ala Glu Asn
 500 505 510
 Ile Ala Tyr Gly Asp Asn Ser Arg Val Val Pro Leu Asp Glu Ile Lys
 515 520 525
 Glu Ala Ala Asn Ala Ala Asn Ile His Ser Phe Ile Glu Gly Leu Pro
 530 535 540
 Glu Lys Tyr Asn Thr Gln Val Gly Leu Lys Gly Ala Gln Leu Ser Gly
 545 550 555 560
 Gly Gln Lys Gln Arg Leu Ala Ile Ala Arg Ala Leu Leu Gln Lys Pro
 565 570 575
 Lys Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Asn Asp Ser
 580 585 590
 Glu Lys Val Val Gln His Ala Leu Asp Lys Ala Arg Thr Gly Arg Thr
 595 600 605
 Cys Leu Val Val Thr His Arg Leu Ser Ala Ile Gln Asn Ala Asp Leu
 610 615 620
 Ile Val Val Leu His Asn Gly Lys Ile Lys Glu Gln Gly Thr His Gln
 625 630 635 640
 Glu Leu Leu Arg Asn Arg Asp Ile Tyr Phe Lys Leu Val Asn Ala Gln
 645 650 655
 Ser Val Gln

<210> 2
 < 211> 812
 < 212> PRT
 < 213> Homo sapiens

<400> 2
 Met Val Asp Glu Asn Asp Ile Arg Ala Leu Asn Val Arg His Tyr Arg
 1 5 10 15
 Asp His Ile Gly Val Val Ser Gln Glu Pro Val Leu Phe Gly Thr Thr
 20 25 30
 Ile Ser Asn Asn Ile Lys Tyr Gly Arg Asp Asp Val Thr Asp Glu Glu
 35 40 45
 Met Glu Arg Ala Ala Arg Glu Ala Asn Ala Tyr Asp Phe Ile Met Glu
 50 55 60
 Phe Pro Asn Lys Phe Asn Thr Leu Val Gly Glu Lys Gly Ala Gln Met
 65 70 75 80

ES 2 659 877 T3

Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu Val Arg
 85 90 95
 Asn Pro Lys Ile Leu Ile Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser
 100 105 110
 Glu Ser Lys Ser Ala Val Gln Ala Ala Leu Glu Lys Ala Ser Lys Gly
 115 120 125
 Arg Thr Thr Ile Val Val Ala His Arg Leu Ser Thr Ile Arg Ser Ala
 130 135 140
 Asp Leu Ile Val Thr Leu Lys Asp Gly Met Leu Ala Glu Lys Gly Ala
 145 150 155 160
 His Ala Glu Leu Met Ala Lys Arg Gly Leu Tyr Tyr Ser Leu Val Met
 165 170 175
 Ser Gln Asp Ile Lys Lys Ala Asp Glu Gln Met Glu Ser Met Thr Tyr
 180 185 190
 Ser Thr Glu Arg Lys Thr Asn Ser Leu Pro Leu His Ser Val Lys Ser
 195 200 205
 Ile Lys Ser Asp Phe Ile Asp Lys Ala Glu Glu Ser Thr Gln Ser Lys
 210 215 220
 Glu Ile Ser Leu Pro Glu Val Ser Leu Leu Lys Ile Leu Lys Leu Asn
 225 230 235 240
 Lys Pro Glu Trp Pro Phe Val Val Leu Gly Thr Leu Ala Ser Val Leu
 245 250 255
 Asn Gly Thr Val His Pro Val Phe Ser Ile Ile Phe Ala Lys Ile Ile
 260 265 270
 Thr Met Phe Gly Asn Asn Asp Lys Thr Thr Leu Lys His Asp Ala Glu
 275 280 285
 Ile Tyr Ser Met Ile Phe Val Ile Leu Gly Val Ile Cys Phe Val Ser
 290 295 300
 Tyr Phe Met Gln Gly Leu Phe Tyr Gly Arg Ala Gly Glu Ile Leu Thr
 305 310 315 320
 Met Arg Leu Arg His Leu Ala Phe Lys Ala Met Leu Tyr Gln Asp Ile
 325 330 335
 Ala Trp Phe Asp Glu Lys Glu Asn Ser Thr Gly Gly Leu Thr Thr Ile
 340 345 350
 Leu Ala Ile Asp Ile Ala Gln Ile Gln Gly Ala Thr Gly Ser Arg Ile
 355 360 365
 Gly Val Leu Thr Gln Asn Ala Thr Asn Met Gly Leu Ser Val Ile Ile
 370 375 380
 Ser Phe Ile Tyr Gly Trp Glu Met Thr Phe Leu Ile Leu Ser Ile Ala
 385 390 395 400

ES 2 659 877 T3

Pro Val Leu Ala Val Thr Gly Met Ile Glu Thr Ala Ala Met Thr Gly
 405 410 415
 Phe Ala Asn Lys Asp Lys Gln Glu Leu Lys His Ala Gly Lys Ile Ala
 420 425 430
 Thr Glu Ala Leu Glu Asn Ile Arg Thr Ile Val Ser Leu Thr Arg Glu
 435 440 445
 Lys Ala Phe Glu Gln Met Tyr Glu Glu Met Leu Gln Thr Gln His Arg
 450 455 460
 Asn Thr Ser Lys Lys Ala Gln Ile Ile Gly Ser Cys Tyr Ala Phe Ser
 465 470 475 480
 His Ala Phe Ile Tyr Phe Ala Tyr Ala Ala Gly Phe Arg Phe Gly Ala
 485 490 495
 Tyr Leu Ile Gln Ala Gly Arg Met Thr Pro Glu Gly Met Phe Ile Val
 500 505 510
 Phe Thr Ala Ile Ala Tyr Gly Ala Met Ala Ile Gly Lys Thr Leu Val
 515 520 525
 Leu Ala Pro Glu Tyr Ser Lys Ala Lys Ser Gly Ala Ala His Leu Phe
 530 535 540
 Ala Leu Leu Glu Lys Lys Pro Asn Ile Asp Ser Arg Ser Gln Glu Gly
 545 550 555 560
 Lys Lys Pro Asp Thr Cys Glu Gly Asn Leu Glu Phe Arg Glu Val Ser
 565 570 575
 Phe Phe Tyr Pro Cys Arg Pro Asp Val Phe Ile Leu Arg Gly Leu Ser
 580 585 590
 Leu Ser Ile Glu Arg Gly Lys Thr Val Ala Phe Val Gly Ser Ser Gly
 595 600 605
 Cys Gly Lys Ser Thr Ser Val Gln Leu Leu Gln Arg Leu Tyr Asp Pro
 610 615 620
 Val Gln Gly Gln Val Leu Phe Asp Gly Val Asp Ala Lys Glu Leu Asn
 625 630 635 640
 Val Gln Trp Leu Arg Ser Gln Ile Ala Ile Val Pro Gln Glu Pro Val
 645 650 655
 Leu Phe Asn Cys Ser Ile Ala Glu Asn Ile Ala Tyr Gly Asp Asn Ser
 660 665 670
 Arg Val Val Pro Leu Asp Glu Ile Lys Glu Ala Ala Asn Ala Ala Asn
 675 680 685
 Ile His Ser Phe Ile Glu Gly Leu Pro Glu Lys Tyr Asn Thr Gln Val
 690 695 700
 Gly Leu Lys Gly Ala Gln Leu Ser Gly Gly Gln Lys Gln Arg Leu Ala
 705 710 715 720
 Ile Ala Arg Ala Leu Leu Gln Lys Pro Lys Ile Leu Leu Leu Asp Glu

ES 2 659 877 T3

725	730	735
Ala Thr Ser 740	Ala Leu Asp Asn Asp Ser Glu Lys Val Val 745	Gln His Ala 750
Leu Asp Lys 755	Ala Arg Thr Gly Arg Thr Cys Leu Val Val 760	Thr His Arg 765
Leu Ser Ala 770	Ile Gln Asn Ala Asp Leu Ile Val Val 775	Leu His Asn Gly 780
Lys Ile Lys 785	Glu Gln Gly Thr His Gln Glu Leu Leu Arg Asn Arg Asp 790	800
Ile Tyr Phe 805	Lys Leu Val Asn Ala Gln Ser Val Gln 810	

<210> 3
 < 211> 131
 < 212> PRT
 < 213> Homo sapiens

<400> 3

Met	Val	Asp	Glu	Asn	Asp	Ile	Arg	Ala	Leu	Asn	Val	Arg	His	Tyr	Arg
1			5						10					15	
Asp	His	Ile	Gly	Val	Val	Ser	Gln	Glu	Pro	Val	Leu	Phe	Gly	Thr	Thr
			20					25					30		
Ile	Ser	Asn	Asn	Ile	Lys	Tyr	Gly	Arg	Asp	Asp	Val	Thr	Asp	Glu	Glu
		35					40					45			
Met	Glu	Arg	Ala	Ala	Arg	Glu	Ala	Asn	Ala	Tyr	Asp	Phe	Ile	Met	Glu
	50					55					60				
Phe	Pro	Asn	Lys	Phe	Asn	Thr	Leu	Val	Gly	Glu	Lys	Gly	Ala	Gln	Met
65					70					75					80
Ser	Gly	Gly	Gln	Lys	Gln	Arg	Ile	Ala	Ile	Ala	Arg	Ala	Leu	Val	Arg
			85						90					95	
Asn	Pro	Lys	Ile	Leu	Ile	Leu	Asp	Glu	Ala	Thr	Ser	Ala	Leu	Asp	Ser
			100					105						110	
Glu	Ser	Lys	Ser	Ala	Val	Gln	Ala	Ala	Leu	Glu	Lys	Asp	Thr	Pro	Arg
		115					120					125			
Tyr	Ser	Phe													
		130													

<210> 4
 < 211> 1058
 < 212> PRT
 < 213> Homo sapiens

<220>
 < 221> VARIANTE
 < 222> (66)..(66)
 < 223> /nota="Xaa representa cualquier L-aminoácido"

ES 2 659 877 T3

<400> 4

```

Met Val Ile Ser Leu Thr Ser Lys Glu Leu Ser Ala Tyr Ser Lys Ala
 1          5          10          15

Gly Ala Val Ala Glu Glu Val Leu Ser Ser Ile Arg Thr Val Ile Ala
 20          25          30

Phe Arg Ala Gln Glu Lys Glu Leu Gln Arg Ser Phe Leu Leu Asn Ile
 35          40          45

Thr Arg Tyr Ala Trp Phe Tyr Phe Pro Gln Trp Leu Leu Ser Cys Val
 50          55          60

Leu Xaa Phe Val Arg Tyr Thr Gln Asn Leu Lys Asp Ala Lys Asp Phe
 65          70          75          80

Gly Ile Lys Arg Thr Ile Ala Ser Lys Val Ser Leu Gly Ala Val Tyr
 85          90          95

Phe Phe Met Asn Gly Thr Tyr Gly Leu Ala Phe Trp Tyr Gly Thr Ser
 100         105         110

Leu Ile Leu Asn Gly Glu Pro Gly Tyr Thr Ile Gly Thr Val Leu Ala
 115         120         125

Val Phe Phe Ser Val Ile His Ser Ser Tyr Cys Ile Gly Ala Ala Val
 130         135         140

Pro His Phe Glu Thr Phe Ala Ile Ala Arg Gly Ala Ala Phe His Ile
 145         150         155         160

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

Gly Tyr Lys Pro Glu Ser Ile Glu Gly Thr Val Glu Phe Lys Asn Val
 180         185         190

Ser Phe Asn Tyr Pro Ser Arg Pro Ser Ile Lys Ile Leu Lys Gly Leu
 195         200         205

Asn Leu Arg Ile Lys Ser Gly Glu Thr Val Ala Leu Val Gly Leu Asn
 210         215         220

Gly Ser Gly Lys Ser Thr Val Val Gln Leu Leu Gln Arg Leu Tyr Asp
 225         230         235         240

Pro Asp Asp Gly Phe Ile Met Val Asp Glu Asn Asp Ile Arg Ala Leu

```

ES 2 659 877 T3

	245	250	255
Asn Val Arg	His Tyr Arg Asp His Ile Gly Val Val Ser Gln Glu Pro 260	265	270
Val Leu Phe	Gly Thr Thr Ile Ser Asn Asn Ile Lys Tyr Gly Arg Asp 275	280	285
Asp Val Thr	Asp Glu Glu Met Glu Arg Ala Ala Arg Glu Ala Asn Ala 290	295	300
Tyr Asp Phe	Ile Met Glu Phe Pro Asn Lys Phe Asn Thr Leu Val Gly 305	310	315
Glu Lys Gly	Ala Gln Met Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile 325	330	335
Ala Arg Ala	Leu Val Arg Asn Pro Lys Ile Leu Ile Leu Asp Glu Ala 340	345	350
Thr Ser Ala	Leu Asp Ser Glu Ser Lys Ser Ala Val Gln Ala Ala Leu 355	360	365
Glu Lys Ala	Ser Lys Gly Arg Thr Thr Ile Val Val Ala His Arg Leu 370	375	380
Ser Thr Ile	Arg Ser Ala Asp Leu Ile Val Thr Leu Lys Asp Gly Met 385	390	395
Leu Ala Glu	Lys Gly Ala His Ala Glu Leu Met Ala Lys Arg Gly Leu 405	410	415
Tyr Tyr Ser	Leu Val Met Ser Gln Asp Ile Lys Lys Ala Asp Glu Gln 420	425	430
Met Glu Ser	Met Thr Tyr Ser Thr Glu Arg Lys Thr Asn Ser Leu Pro 435	440	445
Leu His Ser	Val Lys Ser Ile Lys Ser Asp Phe Ile Asp Lys Ala Glu 450	455	460
Glu Ser Thr	Gln Ser Lys Glu Ile Ser Leu Pro Glu Val Ser Leu Leu 465	470	475
Lys Ile Leu	Lys Leu Asn Lys Pro Glu Trp Pro Phe Val Val Leu Gly 485	490	495
Thr Leu Ala	Ser Val Leu Asn Gly Thr Val His Pro Val Phe Ser Ile 500	505	510
Ile Phe Ala	Lys Ile Ile Thr Met Phe Gly Asn Asn Asp Lys Thr Thr 515	520	525
Leu Lys His	Asp Ala Glu Ile Tyr Ser Met Ile Phe Val Ile Leu Gly 530	535	540
Val Ile Cys	Phe Val Ser Tyr Phe Met Gln Gly Leu Phe Tyr Gly Arg 545	550	555
Ala Gly Glu	Ile Leu Thr Met Arg Leu Arg His Leu Ala Phe Lys Ala 565	570	575

ES 2 659 877 T3

Met Leu Tyr Gln Asp Ile Ala Trp Phe Asp Glu Lys Glu Asn Ser Thr
 580 585 590

Gly Gly Leu Thr Thr Ile Leu Ala Ile Asp Ile Ala Gln Ile Gln Gly
 595 600 605

Ala Thr Gly Ser Arg Ile Gly Val Leu Thr Gln Asn Ala Thr Asn Met
 610 615 620

Gly Leu Ser Val Ile Ile Ser Phe Ile Tyr Gly Trp Glu Met Thr Phe
 625 630 635 640

Leu Ile Leu Ser Ile Ala Pro Val Leu Ala Val Thr Gly Met Ile Glu
 645 650 655

Thr Ala Ala Met Thr Gly Phe Ala Asn Lys Asp Lys Gln Glu Leu Lys
 660 665 670

His Ala Gly Lys Ile Ala Thr Glu Ala Leu Glu Asn Ile Arg Thr Ile
 675 680 685

Val Ser Leu Thr Arg Glu Lys Ala Phe Glu Gln Met Tyr Glu Glu Met
 690 695 700

Leu Gln Thr Gln His Arg Asn Thr Ser Lys Lys Ala Gln Ile Ile Gly
 705 710 715 720

Ser Cys Tyr Ala Phe Ser His Ala Phe Ile Tyr Phe Ala Tyr Ala Ala
 725 730 735

Gly Phe Arg Phe Gly Ala Tyr Leu Ile Gln Ala Gly Arg Met Thr Pro
 740 745 750

Glu Gly Met Phe Ile Val Phe Thr Ala Ile Ala Tyr Gly Ala Met Ala
 755 760 765

Ile Gly Lys Thr Leu Val Leu Ala Pro Glu Tyr Ser Lys Ala Lys Ser
 770 775 780

Gly Ala Ala His Leu Phe Ala Leu Leu Glu Lys Lys Pro Asn Ile Asp
 785 790 795 800

Ser Arg Ser Gln Glu Gly Lys Lys Pro Asp Thr Cys Glu Gly Asn Leu
 805 810 815

Glu Phe Arg Glu Val Ser Phe Phe Tyr Pro Cys Arg Pro Asp Val Phe
 820 825 830

Ile Leu Arg Gly Leu Ser Leu Ser Ile Glu Arg Gly Lys Thr Val Ala
 835 840 845

Phe Val Gly Ser Ser Gly Cys Gly Lys Ser Thr Ser Val Gln Leu Leu
 850 855 860

Gln Arg Leu Tyr Asp Pro Val Gln Gly Gln Val Leu Phe Asp Gly Val
 865 870 875 880

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

ES 2 659 877 T3

Val Pro Gln Glu Pro Val Leu Phe Asn Cys Ser Ile Ala Glu Asn Ile
 900 905 910
 Ala Tyr Gly Asp Asn Ser Arg Val Val Pro Leu Asp Glu Ile Lys Glu
 915 920 925
 Ala Ala Asn Ala Ala Asn Ile His Ser Phe Ile Glu Gly Leu Pro Glu
 930 935 940
 Lys Tyr Asn Thr Gln Val Gly Leu Lys Gly Ala Gln Leu Ser Gly Gly
 945 950 955 960
 Gln Lys Gln Arg Leu Ala Ile Ala Arg Ala Leu Leu Gln Lys Pro Lys
 965 970 975
 Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Asn Asp Ser Glu
 980 985 990
 Lys Val Val Gln His Ala Leu Asp Lys Ala Arg Thr Gly Arg Thr Cys
 995 1000 1005
 Leu Val Val Thr His Arg Leu Ser Ala Ile Gln Asn Ala Asp Leu
 1010 1015 1020
 Ile Val Val Leu His Asn Gly Lys Ile Lys Glu Gln Gly Thr His
 1025 1030 1035
 Gln Glu Leu Leu Arg Asn Arg Asp Ile Tyr Phe Lys Leu Val Asn
 1040 1045 1050
 Ala Gln Ser Val Gln
 1055

<210> 5

< 211> 1222

< 212> PRT

< 213> Homo sapiens

<220>

< 221> VARIANTE

< 222> (230)..(230)

< 223> /nota="Xaa representa cualquier L-aminoácido"

<400> 5

Met Ile Leu Gly Ile Leu Ala Ser Leu Val Asn Gly Ala Cys Leu Pro
 1 5 10 15
 Leu Met Pro Leu Val Leu Gly Glu Met Ser Asp Asn Leu Ile Ser Gly
 20 25 30
 Cys Leu Val Gln Thr Asn Thr Tyr Ser Phe Phe Arg Leu Thr Leu Tyr
 35 40 45

ES 2 659 877 T3

Tyr Val Gly Ile Gly Val Ala Ala Leu Ile Phe Gly Tyr Ile Gln Ile
 50 55 60

Ser Leu Trp Ile Ile Thr Ala Ala Arg Gln Thr Lys Arg Ile Arg Lys
 65 70 75 80

Gln Phe Phe His Ser Val Leu Ala Gln Asp Ile Gly Trp Phe Asp Ser
 85 90 95

Cys Asp Ile Gly Glu Leu Asn Thr Arg Met Thr Asp Ile Asp Lys Ile
 100 105 110

Ser Asp Gly Ile Gly Asp Lys Ile Ala Leu Leu Phe Gln Asn Met Ser
 115 120 125

Thr Phe Ser Ile Gly Leu Ala Val Gly Leu Val Lys Gly Trp Lys Leu
 130 135 140

Thr Leu Val Thr Leu Ser Thr Ser Pro Leu Ile Met Ala Ser Ala Ala
 145 150 155 160

Ala Cys Ser Arg Met Val Ile Ser Leu Thr Ser Lys Glu Leu Ser Ala
 165 170 175

Tyr Ser Lys Ala Gly Ala Val Ala Glu Glu Val Leu Ser Ser Ile Arg
 180 185 190

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

Leu Leu Asn Ile Thr Arg Tyr Ala Trp Phe Tyr Phe Pro Gln Trp Leu
 210 215 220

Leu Ser Cys Val Leu Xaa Phe Val Arg Tyr Thr Gln Asn Leu Lys Asp
 225 230 235 240

Ala Lys Asp Phe Gly Ile Lys Arg Thr Ile Ala Ser Lys Val Ser Leu
 245 250 255

Gly Ala Val Tyr Phe Phe Met Asn Gly Thr Tyr Gly Leu Ala Phe Trp
 260 265 270

Tyr Gly Thr Ser Leu Ile Leu Asn Gly Glu Pro Gly Tyr Thr Ile Gly
 275 280 285

Thr Val Leu Ala Val Phe Phe Ser Val Ile His Ser Ser Tyr Cys Ile
 290 295 300

Gly Ala Ala Val Pro His Phe Glu Thr Phe Ala Ile Ala Arg Gly Ala
 305 310 315 320

Ala Phe His Ile Phe Gln Val Ile Asp Lys Lys Pro Ser Ile Asp Asn
 325 330 335

Phe Ser Thr Ala Gly Tyr Lys Pro Glu Ser Ile Glu Gly Thr Val Glu
 340 345 350

Phe Lys Asn Val Ser Phe Asn Tyr Pro Ser Arg Pro Ser Ile Lys Ile
 355 360 365

ES 2 659 877 T3

Leu Lys Gly Leu Asn Leu Arg Ile Lys Ser Gly Glu Thr Val Ala Leu
 370 375 380
 Val Gly Leu Asn Gly Ser Gly Lys Ser Thr Val Val Gln Leu Leu Gln
 385 390 395 400
 Arg Leu Tyr Asp Pro Asp Asp Gly Phe Ile Met Val Asp Glu Asn Asp
 405 410 415
 Ile Arg Ala Leu Asn Val Arg His Tyr Arg Asp His Ile Gly Val Val
 420 425 430
 Ser Gln Glu Pro Val Leu Phe Gly Thr Thr Ile Ser Asn Asn Ile Lys
 435 440 445
 Tyr Gly Arg Asp Asp Val Thr Asp Glu Glu Met Glu Arg Ala Ala Arg
 450 455 460
 Glu Ala Asn Ala Tyr Asp Phe Ile Met Glu Phe Pro Asn Lys Phe Asn
 465 470 475 480
 Thr Leu Val Gly Glu Lys Gly Ala Gln Met Ser Gly Gly Gln Lys Gln
 485 490 495
 Arg Ile Ala Ile Ala Arg Ala Leu Val Arg Asn Pro Lys Ile Leu Ile
 500 505 510
 Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser Glu Ser Lys Ser Ala Val
 515 520 525
 Gln Ala Ala Leu Glu Lys Ala Ser Lys Gly Arg Thr Thr Ile Val Val
 530 535 540
 Ala His Arg Leu Ser Thr Ile Arg Ser Ala Asp Leu Ile Val Thr Leu
 545 550 555 560
 Lys Asp Gly Met Leu Ala Glu Lys Gly Ala His Ala Glu Leu Met Ala
 565 570 575
 Lys Arg Gly Leu Tyr Tyr Ser Leu Val Met Ser Gln Asp Ile Lys Lys
 580 585 590
 Ala Asp Glu Gln Met Glu Ser Met Thr Tyr Ser Thr Glu Arg Lys Thr
 595 600 605
 Asn Ser Leu Pro Leu His Ser Val Lys Ser Ile Lys Ser Asp Phe Ile
 610 615 620
 Asp Lys Ala Glu Glu Ser Thr Gln Ser Lys Glu Ile Ser Leu Pro Glu
 625 630 635 640
 Val Ser Leu Leu Lys Ile Leu Lys Leu Asn Lys Pro Glu Trp Pro Phe
 645 650 655
 Val Val Leu Gly Thr Leu Ala Ser Val Leu Asn Gly Thr Val His Pro
 660 665 670
 Val Phe Ser Ile Ile Phe Ala Lys Ile Ile Thr Met Phe Gly Asn Asn
 675 680 685
 Asp Lys Thr Thr Leu Lys His Asp Ala Glu Ile Tyr Ser Met Ile Phe

ES 2 659 877 T3

690	695	700
Val Ile Leu Gly Val Ile Cys Phe Val Ser Tyr Phe Met Gln Gly Leu 705	710	715
Phe Tyr Gly Arg Ala Gly Glu Ile Leu Thr Met Arg Leu Arg His Leu 725	730	735
Ala Phe Lys Ala Met Leu Tyr Gln Asp Ile Ala Trp Phe Asp Glu Lys 740	745	750
Glu Asn Ser Thr Gly Gly Leu Thr Thr Ile Leu Ala Ile Asp Ile Ala 755	760	765
Gln Ile Gln Gly Ala Thr Gly Ser Arg Ile Gly Val Leu Thr Gln Asn 770	775	780
Ala Thr Asn Met Gly Leu Ser Val Ile Ile Ser Phe Ile Tyr Gly Trp 785	790	795
Glu Met Thr Phe Leu Ile Leu Ser Ile Ala Pro Val Leu Ala Val Thr 805	810	815
Gly Met Ile Glu Thr Ala Ala Met Thr Gly Phe Ala Asn Lys Asp Lys 820	825	830
Gln Glu Leu Lys His Ala Gly Lys Ile Ala Thr Glu Ala Leu Glu Asn 835	840	845
Ile Arg Thr Ile Val Ser Leu Thr Arg Glu Lys Ala Phe Glu Gln Met 850	855	860
Tyr Glu Glu Met Leu Gln Thr Gln His Arg Asn Thr Ser Lys Lys Ala 865	870	875
Gln Ile Ile Gly Ser Cys Tyr Ala Phe Ser His Ala Phe Ile Tyr Phe 885	890	895
Ala Tyr Ala Ala Gly Phe Arg Phe Gly Ala Tyr Leu Ile Gln Ala Gly 900	905	910
Arg Met Thr Pro Glu Gly Met Phe Ile Val Phe Thr Ala Ile Ala Tyr 915	920	925
Gly Ala Met Ala Ile Gly Lys Thr Leu Val Leu Ala Pro Glu Tyr Ser 930	935	940
Lys Ala Lys Ser Gly Ala Ala His Leu Phe Ala Leu Leu Glu Lys Lys 945	950	955
Pro Asn Ile Asp Ser Arg Ser Gln Glu Gly Lys Lys Pro Asp Thr Cys 965	970	975
Glu Gly Asn Leu Glu Phe Arg Glu Val Ser Phe Phe Tyr Pro Cys Arg 980	985	990
Pro Asp Val Phe Ile Leu Arg Gly Leu Ser Leu Ser Ile Glu Arg Gly 995	1000	1005
Lys Thr Val Ala Phe Val Gly Ser Ser Gly Cys Gly Lys Ser Thr 1010	1015	1020

ES 2 659 877 T3

Ser Val Gln Leu Leu Gln Arg Leu Tyr Asp Pro Val Gln Gly Gln
 1025 1030 1035
 Val Leu Phe Asp Gly Val Asp Ala Lys Glu Leu Asn Val Gln Trp
 1040 1045 1050
 Leu Arg Ser Gln Ile Ala Ile Val Pro Gln Glu Pro Val Leu Phe
 1055 1060 1065
 Asn Cys Ser Ile Ala Glu Asn Ile Ala Tyr Gly Asp Asn Ser Arg
 1070 1075 1080
 Val Val Pro Leu Asp Glu Ile Lys Glu Ala Ala Asn Ala Ala Asn
 1085 1090 1095
 Ile His Ser Phe Ile Glu Gly Leu Pro Glu Lys Tyr Asn Thr Gln
 1100 1105 1110
 Val Gly Leu Lys Gly Ala Gln Leu Ser Gly Gly Gln Lys Gln Arg
 1115 1120 1125
 Leu Ala Ile Ala Arg Ala Leu Leu Gln Lys Pro Lys Ile Leu Leu
 1130 1135 1140
 Leu Asp Glu Ala Thr Ser Ala Leu Asp Asn Asp Ser Glu Lys Val
 1145 1150 1155
 Val Gln His Ala Leu Asp Lys Ala Arg Thr Gly Arg Thr Cys Leu
 1160 1165 1170
 Val Val Thr His Arg Leu Ser Ala Ile Gln Asn Ala Asp Leu Ile
 1175 1180 1185
 Val Val Leu His Asn Gly Lys Ile Lys Glu Gln Gly Thr His Gln
 1190 1195 1200
 Glu Leu Leu Arg Asn Arg Asp Ile Tyr Phe Lys Leu Val Asn Ala
 1205 1210 1215
 Gln Ser Val Gln
 1220

<210> 6

< 211> 1195

< 212> PRT

< 213> Homo sapiens

<400> 6

Met Ile Leu Gly Ile Leu Ala Ser Leu Val Asn Gly Ala Cys Leu Pro
 1 5 10 15
 Leu Met Pro Leu Val Leu Gly Glu Met Ser Asp Asn Leu Ile Ser Gly
 20 25 30

ES 2 659 877 T3

355	360	365
Val Gln Leu Leu Gln Arg 370	Leu Tyr Asp Pro Asp 375	Asp Gly Phe Ile Met 380
Val Asp Glu Asn Asp 385	Ile Arg Ala Leu Asn 390	Val Arg His Tyr Arg Asp 395 400
His Ile Gly Val Val Ser 405	Gln Glu Pro Val Leu Phe 410	Gly Thr Thr Ile 415
Ser Asn Asn Ile Lys Tyr 420	Gly Arg Asp Asp Val Thr 425	Asp Glu Glu Met 430
Glu Arg Ala Ala Arg Glu 435	Ala Asn Ala Tyr Asp 440	Phe Ile Met Glu Phe 445
Pro Asn Lys Phe Asn Thr 450	Leu Val Gly Glu Lys Gly 455	Ala Gln Met Ser 460
Gly Gly Gln Lys Gln Arg 465	Ile Ala Ile Ala Arg Ala 470	Leu Val Arg Asn 475 480
Pro Lys Ile Leu Ile Leu 485	Asp Glu Ala Thr Ser Ala 490	Leu Asp Ser Glu 495
Ser Lys Ser Ala Val Gln 500	Ala Ala Leu Glu Lys Ala 505	Ser Lys Gly Arg 510
Thr Thr Ile Val Val Ala 515	His Arg Leu Ser Thr Ile 520	Arg Ser Ala Asp 525
Leu Ile Val Thr Leu Lys 530	Asp Gly Met Leu Ala Glu 535	Lys Gly Ala His 540
Ala Glu Leu Met Ala Lys 545	Arg Gly Leu Tyr Tyr Ser 550	Leu Val Met Ser 555 560
Gln Asp Ile Lys Lys Ala 565	Asp Glu Gln Met Glu Ser 570	Met Thr Tyr Ser 575
Thr Glu Arg Lys Thr Asn 580	Ser Leu Pro Leu His Ser 585	Val Lys Ser Ile 590
Lys Ser Asp Phe Ile Asp 595	Lys Ala Glu Glu Ser Thr 600	Gln Ser Lys Glu 605
Ile Ser Leu Pro Glu Val 610	Ser Leu Leu Lys Ile Leu 615	Lys Leu Asn Lys 620
Pro Glu Trp Pro Phe Val 625	Val Leu Gly Thr Leu Ala 630	Ser Val Leu Asn 635 640
Gly Thr Val His Pro Val 645	Phe Ser Ile Ile Phe Ala 650	Lys Ile Ile Thr 655
Met Phe Gly Asn Asn Asp 660	Lys Thr Thr Leu Lys His 665	Asp Ala Glu Ile 670
Tyr Ser Met Ile Phe Val 675	Ile Leu Gly Val Ile Cys 680	Phe Val Ser Tyr 685

ES 2 659 877 T3

Phe Met Gln Gly Leu Phe Tyr Gly Arg Ala Gly Glu Ile Leu Thr Met
 690 695 700

Arg Leu Arg His Leu Ala Phe Lys Ala Met Leu Tyr Gln Asp Ile Ala
 705 710 715 720

Trp Phe Asp Glu Lys Glu Asn Ser Thr Gly Gly Leu Thr Thr Ile Leu
 725 730 735

Ala Ile Asp Ile Ala Gln Ile Gln Gly Ala Thr Gly Ser Arg Ile Gly
 740 745 750

Val Leu Thr Gln Asn Ala Thr Asn Met Gly Leu Ser Val Ile Ile Ser
 755 760 765

Phe Ile Tyr Gly Trp Glu Met Thr Phe Leu Ile Leu Ser Ile Ala Pro
 770 775 780

Val Leu Ala Val Thr Gly Met Ile Glu Thr Ala Ala Met Thr Gly Phe
 785 790 795 800

Ala Asn Lys Asp Lys Gln Glu Leu Lys His Ala Gly Lys Ile Ala Thr
 805 810 815

Glu Ala Leu Glu Asn Ile Arg Thr Ile Val Ser Leu Thr Arg Glu Lys
 820 825 830

Ala Phe Glu Gln Met Tyr Glu Glu Met Leu Gln Thr Gln His Arg Asn
 835 840 845

Thr Ser Lys Lys Ala Gln Ile Ile Gly Ser Cys Tyr Ala Phe Ser His
 850 855 860

Ala Phe Ile Tyr Phe Ala Tyr Ala Ala Gly Phe Arg Phe Gly Ala Tyr
 865 870 875 880

Leu Ile Gln Ala Gly Arg Met Thr Pro Glu Gly Met Phe Ile Val Phe
 885 890 895

Thr Ala Ile Ala Tyr Gly Ala Met Ala Ile Gly Lys Thr Leu Val Leu
 900 905 910

Ala Pro Glu Tyr Ser Lys Ala Lys Ser Gly Ala Ala His Leu Phe Ala
 915 920 925

Leu Leu Glu Lys Lys Pro Asn Ile Asp Ser Arg Ser Gln Glu Gly Lys
 930 935 940

Lys Pro Asp Thr Cys Glu Gly Asn Leu Glu Phe Arg Glu Val Ser Phe
 945 950 955 960

Phe Tyr Pro Cys Arg Pro Asp Val Phe Ile Leu Arg Gly Leu Ser Leu
 965 970 975

Ser Ile Glu Arg Gly Lys Thr Val Ala Phe Val Gly Ser Ser Gly Cys
 980 985 990

Gly Lys Ser Thr Ser Val Gln Leu Leu Gln Arg Leu Tyr Asp Pro Val
 995 1000 1005

ES 2 659 877 T3

Gln Gly Gln Val Leu Phe Asp Gly Val Asp Ala Lys Glu Leu Asn
 1010 1015 1020
 Val Gln Trp Leu Arg Ser Gln Ile Ala Ile Val Pro Gln Glu Pro
 1025 1030 1035
 Val Leu Phe Asn Cys Ser Ile Ala Glu Asn Ile Ala Tyr Gly Asp
 1040 1045 1050
 Asn Ser Arg Val Val Pro Leu Asp Glu Ile Lys Glu Ala Ala Asn
 1055 1060 1065
 Ala Ala Asn Ile His Ser Phe Ile Glu Gly Leu Pro Glu Lys Tyr
 1070 1075 1080
 Asn Thr Gln Val Gly Leu Lys Gly Ala Gln Leu Ser Gly Gly Gln
 1085 1090 1095
 Lys Gln Arg Leu Ala Ile Ala Arg Ala Leu Leu Gln Lys Pro Lys
 1100 1105 1110
 Ile Leu Leu Leu Asp Glu Ala Thr Ser Ala Leu Asp Asn Asp Ser
 1115 1120 1125
 Glu Lys Val Val Gln His Ala Leu Asp Lys Ala Arg Thr Gly Arg
 1130 1135 1140
 Thr Cys Leu Val Val Thr His Arg Leu Ser Ala Ile Gln Asn Ala
 1145 1150 1155
 Asp Leu Ile Val Val Leu His Asn Gly Lys Ile Lys Glu Gln Gly
 1160 1165 1170
 Thr His Gln Glu Leu Leu Arg Asn Arg Asp Ile Tyr Phe Lys Leu
 1175 1180 1185
 Val Asn Ala Gln Ser Val Gln
 1190 1195

<210> 7
 < 211> 541
 < 212> PRT
 < 213> Homo sapiens

<220>
 < 221> VARIANTE
 < 222> (230)..(230)
 < 223> /nota="Xaa representa cualquier L-aminoácido"

<400> 7
 Met Ile Leu Gly Ile Leu Ala Ser Leu Val Asn Gly Ala Cys Leu Pro
 1 5 10 15

ES 2 659 877 T3

Leu Met Pro Leu Val Leu Gly Glu Met Ser Asp Asn Leu Ile Ser Gly
 20 25 30
 Cys Leu Val Gln Thr Asn Thr Tyr Ser Phe Phe Arg Leu Thr Leu Tyr
 35 40 45
 Tyr Val Gly Ile Gly Val Ala Ala Leu Ile Phe Gly Tyr Ile Gln Ile
 50 55 60
 Ser Leu Trp Ile Ile Thr Ala Ala Arg Gln Thr Lys Arg Ile Arg Lys
 65 70 75 80
 Gln Phe Phe His Ser Val Leu Ala Gln Asp Ile Gly Trp Phe Asp Ser
 85 90 95
 Cys Asp Ile Gly Glu Leu Asn Thr Arg Met Thr Asp Ile Asp Lys Ile
 100 105 110
 Ser Asp Gly Ile Gly Asp Lys Ile Ala Leu Leu Phe Gln Asn Met Ser
 115 120 125
 Thr Phe Ser Ile Gly Leu Ala Val Gly Leu Val Lys Gly Trp Lys Leu
 130 135 140
 Thr Leu Val Thr Leu Ser Thr Ser Pro Leu Ile Met Ala Ser Ala Ala
 145 150 155 160
 Ala Cys Ser Arg Met Val Ile Ser Leu Thr Ser Lys Glu Leu Ser Ala
 165 170 175
 Tyr Ser Lys Ala Gly Ala Val Ala Glu Glu Val Leu Ser Ser Ile Arg
 180 185 190
 Thr Val Ile Ala Phe Arg Ala Gln Glu Lys Glu Leu Gln Arg Ser Phe
 195 200 205
 Leu Leu Asn Ile Thr Arg Tyr Ala Trp Phe Tyr Phe Pro Gln Trp Leu
 210 215 220
 Leu Ser Cys Val Leu Xaa Phe Val Arg Tyr Thr Gln Asn Leu Lys Asp
 225 230 235 240
 Ala Lys Asp Phe Gly Ile Lys Arg Thr Ile Ala Ser Lys Val Ser Leu
 245 250 255
 Gly Ala Val Tyr Phe Phe Met Asn Gly Thr Tyr Gly Leu Ala Phe Trp
 260 265 270
 Tyr Gly Thr Ser Leu Ile Leu Asn Gly Glu Pro Gly Tyr Thr Ile Gly
 275 280 285
 Thr Val Leu Ala Val Phe Phe Ser Val Ile His Ser Ser Tyr Cys Ile
 290 295 300
 Gly Ala Ala Val Pro His Phe Glu Thr Phe Ala Ile Ala Arg Gly Ala
 305 310 315 320
 Ala Phe His Ile Phe Gln Val Ile Asp Lys Lys Pro Ser Ile Asp Asn
 325 330 335

ES 2 659 877 T3

Phe Ser Thr Ala Gly Tyr Lys Pro Glu Ser Ile Glu Gly Thr Val Glu
 340 345 350
 Phe Lys Asn Val Ser Phe Asn Tyr Pro Ser Arg Pro Ser Ile Lys Ile
 355 360 365
 Leu Lys Gly Leu Asn Leu Arg Ile Lys Ser Gly Glu Thr Val Ala Leu
 370 375 380
 Val Gly Leu Asn Gly Ser Gly Lys Ser Thr Val Val Gln Leu Leu Gln
 385 390 395 400
 Arg Leu Tyr Asp Pro Asp Asp Gly Phe Ile Met Val Asp Glu Asn Asp
 405 410 415
 Ile Arg Ala Leu Asn Val Arg His Tyr Arg Asp His Ile Gly Val Val
 420 425 430
 Ser Gln Glu Pro Val Leu Phe Gly Thr Thr Ile Ser Asn Asn Ile Lys
 435 440 445
 Tyr Gly Arg Asp Asp Val Thr Asp Glu Glu Met Glu Arg Ala Ala Arg
 450 455 460
 Glu Ala Asn Ala Tyr Asp Phe Ile Met Glu Phe Pro Asn Lys Phe Asn
 465 470 475 480
 Thr Leu Val Gly Glu Lys Gly Ala Gln Met Ser Gly Gly Gln Lys Gln
 485 490 495
 Arg Ile Ala Ile Ala Arg Ala Leu Val Arg Asn Pro Lys Ile Leu Ile
 500 505 510
 Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser Glu Ser Lys Ser Ala Val
 515 520 525
 Gln Ala Ala Leu Glu Lys Asp Thr Pro Arg Tyr Ser Phe
 530 535 540

<210> 8

< 211> 514

< 212> PRT

< 213> Homo sapiens

<400> 8

Met Ile Leu Gly Ile Leu Ala Ser Leu Val Asn Gly Ala Cys Leu Pro
 1 5 10 15
 Leu Met Pro Leu Val Leu Gly Glu Met Ser Asp Asn Leu Ile Ser Gly
 20 25 30
 Cys Leu Val Gln Thr Asn Thr Tyr Ser Phe Phe Arg Leu Thr Leu Tyr
 35 40 45
 Tyr Val Gly Ile Gly Val Ala Ala Leu Ile Phe Gly Tyr Ile Gln Ile

ES 2 659 877 T3

```

gttccaggat tggcgtctta acacaaaatg caactaacat gggactttca gttatcattt 780
cctttatata tggatgggag atgacattcc tgattctgag tattgctcca gtacttgccg 840
tgacaggaat gattgaaacc gcagcaatga ctggatttgc caacaaagat aagcaagaac 900
ttaagcatgc tggaaagata gcaactgaag ctttggagaa tatacgtact atagtgtcat 960
taacaaggga aaaagccttc gagcaaatgt atgaagagat gcttcagact caacacagaa 1020
atacctcgaa gaaagcacag attattggaa gctgttatgc attcagccat gcctttatat 1080
attttgccta tgcagcaggg tttcgatttg gagcctatth aattcaagct ggacgaatga 1140
ccccagaggg catgttcata gtttttactg caattgcata tggagctatg gccatcggaa 1200
aaacgctcgt tttggctcct gaatattcca aagccaaatc gggggctgcg catctgtttg 1260
ccttgttgga aaagaaacca aatatagaca gccgcagtca agaagggaaa aagccagaca 1320
catgtgaagg gaatttagag tttcgagaag tctctttctt ctatccatgt cgcccagatg 1380
ttttcatcct ccgtggctta tccctcagta ttgagcgagg aaagacagta gcatttgtgg 1440
ggagcagcgg ctgtgggaaa agcacttctg ttcaacttct gcagagactt tatgaccccg 1500
tgcaaggaca agtgctgttt gatgggtgtg atgcaaaaga attgaatgta cagtggctcc 1560
gttcccaaat agcaatcgtt cctcaagagc ctgtgctctt caactgcagc attgctgaga 1620
acatcgccta tggtgacaac agccgtgtgg tgccattaga tgagatcaa gaagccgcaa 1680
atgcagcaaa tatccattct tttattgaag gtctccctga gaaatacaac acacaagttg 1740
gactgaaagg agcacagctt tctggcggcc agaaacaag actagctatt gcaagggctc 1800
ttctcaaaa acccaaaatt ttattgttgg atgaggccac ttcagccctc gataatgaca 1860
gtgagaaggt ggttcagcat gcccttgata aagccaggac ggaaggaca tgcctagtgg 1920
tactcacag gctctctgca attcagaacg cagatttgat agtggttctg cacaatggaa 1980
agataaagga acaaggaact catcaagagc tctgagaaa tcgagacata tattttaagt 2040
tagtgaatgc acagtcagtg cagtga 2066

```

<210> 10

< 211> 2856

< 212> ADN

< 213> Homo sapiens

<400> 10

```

cctaattcct ctaatatctc tctgtgagcc taaaccaata attatatatt acattctatt 60
gtctttctta tataactgca gaaagataaa taccactttg tttgttctg taggttttct 120

```

ES 2 659 877 T3

ttagtgtaat ccatagcagt tattgcattg gagcagcagt ccctcattat tgataagaaa	180
cccagtatag ataacttttc cacagctgga tataaacctg aatccataga aggaactgtg	240
gaatttaaaa atgtttcttt caattatcca tcaagaccat ctatcaagat tctgaaaggt	300
ctgaatctca gaattaagtc tggagagaca gtgcgcttgg tcggtctcaa tggcagtggg	360
aagagtacgg tagtccagct tctgcagagg ttatatgac cggatgatgg ctttatcatg	420
gtggatgaga atgacatcag agcttttaaat gtgcggcatt atcgagacca tattggagtg	480
gttagtcaag agcctgtttt gttcgggacc accatcagta acaatatcaa gtatggacga	540
gatgatgtga ctgatgaaga gatggagaga gcagcaaggg aagcaaatgc gtatgatttt	600
atcatggagt ttctaataa atttaataca ttggtagggg aaaagagagc tcaaatgagt	660
ggagggcaga aacagaggat cgcaattgct cgtgccttag ttcgaaacct caagattctg	720
attttagatg aggctacgtc tgccttgat tcagaaagca agtcagctgt tcaagctgca	780
ctggagaagg cgagcaagg tggactaca atcgtgtag cacaccgact ttctactatt	840
cgaagtgcag atttgattgt gaccctaag gatggaatgc tggcggagaa aggagcacat	900
gctgaactaa tggcaaacg aggtctatat tattcacttg tgatgtcaca ggatattaa	960
aaagctgatg aacagatgga gtcaatgaca tatttactg aaagaaagac caactcactt	1020
cctctgcact ctgtgaagag catcaagtc gacttcattg acaaggctga ggaatccacc	1080
caatctaag agataagtct tctgaagtc tctctattaa aaattttaaa gttaacaag	1140
cctgaatggc cttttgtggt tctggggaca ttggcttctg ttctaaatgg aactgttcat	1200
ccagtatfff ccatcatctt tgcaaaaatt ataacctgt ttggaaataa tgataaaacc	1260
acattaaagc atgatgcaga aatttattcc atgatattcg tcattttggg tgttatttgc	1320
tttgtcagtt atttcatgca gggattatft tacggcagag caggggaaat tttaacgatg	1380
agattaagac acttggcctt caaagccatg ttatatcagg atattgcctg gtttgatgaa	1440
aaggaaaaca gcacaggagg cttgacaaca atattagcca tagatatagc acaaattcaa	1500
ggagcaacag gttccaggat tggcgtctta acacaaaatg caactaacat gggactttca	1560
gttatcattt cctttatata tggatgggag atgacattcc tgattctgag tattgctcca	1620
gtacttgccg tgacaggaat gattgaaacc gcagcaatga ctggatttgc caacaaagat	1680
aagcaagaac ttaagcatgc tggaaagata gcaactgaag ctttggagaa tatacgtact	1740
atagtgtcat taacaaggga aaaagccttc gagcaaatgt atgaagagat gcttcagact	1800
caacacagaa atacctgaa gaaagcacag attattggaa gctgttatgc attcagccat	1860
gcctttatat attttgccta tgcagcaggg tttcgatttg ggcctatft aattcaagct	1920

ES 2 659 877 T3

```

ggacgaatga ccccagaggg catgttcata gtttttactg caattgcata tggagctatg 1980
gccatcggaa aaacgctcgt tttggctcct gaatattcca aagccaaatc gggggctgcg 2040
catctgtttg ccttgttgga aaagaaacca aatatagaca gccgcagtca agaagggaaa 2100
aagccagaca catgtgaagg gaatttagag ttcogagaag tctctttctt ctatccatgt 2160
cgcccagatg ttttcatcct ccgtggctta tccctcagta ttgagcgagg aaagacagta 2220
gcatttgtgg ggagcagcgg ctgtgggaaa agcacttctg ttcaacttct gcagagactt 2280
tatgaccccg tgcaaggaca agtgctgttt gatggtgtgg atgcaaaaaga attgaatgta 2340
cagtggctcc gttcccaaat agcaatcgtt cctcaagagc ctgtgctctt caactgcagc 2400
attgctgaga acatcgccta tggtgacaac agccgtgtgg tgccattaga tgagatcaaa 2460
gaagccgcaa atgcagaaa tatccattct tttattgaag gtctccctga gaaatacaac 2520
acacaagttg gactgaaaagg agcacagctt tctggcggcc agaaacaaag actagctatt 2580
gcaagggtc tcttccaaaa acccaaaatt ttattgttgg atgaggccac ttcagccctc 2640
gataatgaca gtgagaaggt ggttcagcat gcccttgata aagccaggac ggaaggaca 2700
tgccctagtgg tcaactcacag gctctctgca attcagaacg cagatttgat agtggttctg 2760
cacaatggaa agataaagga acaaggaact catcaagagc tcctgagaaa tcgagacata 2820
tattttaagt tagtgaatgc acagtcagtg cagtga 2856

```

<210> 11

< 211> 1175

< 212> ADN

< 213> Homo sapiens

<400> 11

```

cctaattcct ctaatatctc tctgtgagcc taaaccaata attatatatt acattctatt 60
gtctttctta tataactgca gaaagataaa taccactttg tttgttctg taggttttct 120
ttagtgaat ccatagcagt tattgcattg gagcagcagt ccctcattat tgataagaaa 180
cccagtatag ataacttttc cacagctgga tataaacctg aatccataga aggaactgtg 240
gaatttaaaa atgtttcttt caattatcca tcaagaccat ctatcaagat tctgaaaggt 300
ctgaatctca gaattaagtc tggagagaca gtcgccttgg tcggtctcaa tggcagtggg 360
aagagtacgg tagtccagct tctgcagagg ttatatgatc cggatgatgg ctttatcatg 420
gtggatgaga atgacatcag agctttaaat gtgcggcatt atcgagacca tattggagtg 480
gttagtcaag agcctgtttt gttcgggacc accatcagta acaatatcaa gtatggacga 540

```

ES 2 659 877 T3

```

gatgatgtga ctgatgaaga gatggagaga gcagcaaggg aagcaaatgc gtatgatttt 600
atcatggagt ttctaataa atttaataca ttggtagggg aaaaaggagc tcaaatgagt 660
ggagggcaga aacagaggat cgcaattgct cgtgccttag ttcgaaaccc caagattctg 720
atTTTTtagatg aggctacgct tgccctggat tcagaaagca agtcagctgt tcaagctgca 780
ctggagaagg atacccccag gtattcattt tgacctaat tcaacctcaag tggagaatcg 840
ctgaccttga acccagcggc ttcgacagct ctggcccctc aaacctcacc ctgacctcct 900
gctgcctatg agctactgca catacctcaa ggccatatgc agttgtggcc ctgcaccaa 960
ttacactgaa tctaggaggg gagttggcag tggcggtatg aaaaaccatt gaacagtttt 1020
ctcgatggcc tgactccctt ataaaccaga gccttcagac cccttacaag gcttaatggc 1080
acattttact ttgcatttgc ttggaagtga gttaagcgtt tttttttctc taagaaaatc 1140
gcaggcttct ttttttaaaa tgctgacttt atgga 1175

```

```

<210> 12
< 211> 3177
< 212> ADN
< 213> Homo sapiens

```

```

<220>
< 221> característica_misc
< 222> (198)..(198)
< 223> /nota="n representa cualquier nucleótido (A, T, C o G)"

```

```

<400> 12
atggtcacatc cattgaccag taaggaatta agtgcctatt ccaaagctgg ggctgtggca 60
gaagaagtct tgtcatcaat ccgaacagtc atagccttta gggcccagga gaaagaactt 120
caaaggtcct tccttttaaa tataacaaga tatgcttggc tttattttcc ccagtggcta 180
ctaagttgtg ttctgttntt tgtaaggtat acacagaatc tcaaagatgc aaaggatttt 240
ggcataaaaa ggactatagc ttcaaaagtg tctcttggcg ctgtgtactt ctttatgaat 300
ggaacctatg gacttgcttt ttggatgga acctccttga ttcttaatgg agaacctgga 360
tataccatcg ggactgttct tgctgttttc tttagtgtaa tccatagcag ttattgcatt 420
ggagcagcag tccctcaact tgaaaccttc gcaatagccc gaggagctgc ctttcatatt 480

```

ES 2 659 877 T3

ttccaggtta ttgataagaa acccagtata gataactttt ccacagctgg atataaacct	540
gaatccatag aaggaactgt ggaatttaaa aatgtttctt tcaattatcc atcaagacca	600
tctatcaaga ttctgaaagg tctgaatctc agaattaagt ctggagagac agtcgccttg	660
gtcggctca atggcagtgg gaagagtacg gtagtccagc ttctgcagag gttatatgat	720
ccggatgatg gctttatcat ggtggatgag aatgacatca gagctttaa tgtgcggcat	780
tatcgagacc atattggagt ggttagtcaa gagcctgttt tgttcgggac caccatcagt	840
aacaatatca agtatggacg agatgatgtg actgatgaag agatggagag agcagcaagg	900
gaagcaaatg cgtatgattt tatcatggag tttcctaata aatttaatac attggtaggg	960
gaaaaaggag ctcaaatgag tggagggcag aaacagagga tcgcaattgc tcgtgcctta	1020
gttcgaaacc ccaagattct gatttttagat gaggctacgt ctgccctgga ttcagaaagc	1080
aagtcagctg ttcaagctgc actggagaag gcgagcaaag gtcggactac aatcgtggta	1140
gcacaccgac tttctactat tcgaagtgca gatttgattg tgaccctaaa ggatggaatg	1200
ctggcggaga aaggagcaca tgctgaacta atggcaaac gaggtctata ttattcactt	1260
gtgatgtcac aggatattaa aaaagctgat gaacagatgg agtcaatgac atattctact	1320
gaaagaaaga ccaactcact tcctctgcac tctgtgaaga gcatcaagtc agacttcatt	1380
gacaaggctg aggaatccac ccaatctaaa gagataagtc ttctgaagt ctctctatta	1440
aaaattttaa agttaaaca gctgaatgg ccttttgtgg ttctggggac attggcttct	1500
gttctaaatg gaactgttca tccagtattt tccatcatct ttgcaaaaat tataaccatg	1560
tttgaaaata atgataaaac cacattaag catgatgcag aaatttattc catgatattc	1620
gtcattttgg gtgttatttg ctttgtcagt tatttcatgc agggattatt ttacggcaga	1680
gcaggggaaa ttttaacgat gagattaaga cacttggcct tcaaagccat gttatatcag	1740
gatattgcct ggtttgatga aaaggaaaac agcacaggag gcttgacaac aatattagcc	1800
atagatatag cacaaattca aggagcaaca ggttccagga ttggcgtctt aacacaaaat	1860
gcaactaaca tgggactttc agttatcatt tcctttatat atggatggga gatgacattc	1920
ctgattctga gtattgctcc agtacttgcc gtgacaggaa tgattgaaac cgcagcaatg	1980
actggatttg ccaacaaaga taagcaagaa cttaagcatg ctggaaagat agcaactgaa	2040
gctttggaga atatacgtac tatagtgca ttaacaaggg aaaagcctt cgagcaaatg	2100
tatgaagaga tgcttcagac tcaacacaga aatacctcga agaaagcaca gattattgga	2160
agctgttatg cattcagcca tgcccttata tattttgcct atgcagcagg gtttcgattt	2220
ggagcctatt taattcaagc tggacgaatg accccagagg gcatgttcat agtttttact	2280
gcaattgcat atggagctat ggccatcgga aaaacgctcg ttttggctcc tgaatattcc	2340

ES 2 659 877 T3

```

aaagccaaat cgggggctgc gcatctgttt gccttggtgg aaaagaaacc aaatatagac 2400
agccgcagtc aagaaggaa aaagccagac acatgtgaag ggaatttaga gtttcgagaa 2460
gtctctttct tctatccatg tcgcccagat gttttcatcc tccgtggctt atccctcagt 2520
attgagcgag gaaagacagt agcatttggtg gggagcagcg gctgtgggaa aagcacttct 2580
gttcaacttc tgcagagact ttatgacccc gtgcaaggac aagtgtgtgt tgatggtgtg 2640
gatgcaaaag aattgaatgt acagtggctc cgttcccaa tagcaatcgt tcctcaagag 2700
cctgtgtctt tcaactgcag cattgctgag aacatcgctt atggtgacaa cagccgtgtg 2760
gtgccattag atgagatcaa agaagccgca aatgcagcaa atatccattc ttttattgaa 2820
ggtctccctg agaaatacaa cacacaagtt ggactgaaag gagcacagct ttctggcggc 2880
cagaaacaaa gactagctat tgcaagggtc cttctccaaa aacccaaaat tttattgttg 2940
gatgaggcca cttcagccct cgataatgac agtgagaagg tggttcagca tgcccttgat 3000
aaagccagga cgggaaggac atgcctagtg gtcaactaca ggctctctgc aattcagaac 3060
gcagatttga tagtggttct gcacaatgga aagataaagg aacaaggaac tcatcaagag 3120
ctcctgagaa atcgagacat atattttaag ttagtgaatg cacagtcagt gcagtga 3177

```

<210> 13
 < 211> 3702
 < 212> ADN
 < 213> Homo sapiens

<220>
 < 221> característica_misc
 < 222> (723)..(723)
 < 223> /nota="n representa cualquier nucleótido (A, T, C o G)"

```

<400> 13
ttccgctttg ctgatggact ggacatcaca ctcatgatcc tgggtatact ggcatcactg 60
gtcaatggag cctgccttcc tttaatgcca ctggtttttag gagaaatgag tgataacctt 120
attagtgat gtctagtcca aactaacaca tactctttct tcaggttgac cctgtattat 180
gttggaatag gtgttgctgc cttgattttt ggttacatac agatttcctt gtggattata 240
actgcagcac gacagaccaa gaggattcga aacagtttt ttcattcagt tttggcacag 300
gacatcggct ggtttgatag ctgtgacatc ggtgaactta aactcgcac gacagacatt 360

```

ES 2 659 877 T3

gacaaaatca gtgatggtat tggagataag attgctctgt tgtttcaaaa catgtctact	420
ttttcgattg gcctggcagt tggtttggtg aagggtgga aactcacct agtgactcta	480
tccacgtctc ctcttataat ggcttcagcg gcagcatggt ctaggatggt catctcattg	540
accagtaagg aattaagtgc ctattccaaa gctggggctg tggcagaaga agtcttgca	600
tcaatccgaa cagtcatagc cttagggcc caggagaaag aacttcaaag gtctttcctt	660
ttaaatataa caagatatgc ttggttttat tttccccagt ggctactaag ttgtggtctg	720
ttntttgtaa ggtatacaca gaatctcaa gatgcaaagg attttgcat aaaaaggact	780
atagctcaa aagtgtctct tgggtgctg tacttcttta tgaatggaac ctatggactt	840
gctttttggt atggaacctc cttgattctt aatggagaac ctggatatac catcgggact	900
gttcttgctg ttttcttag tgtaatccat agcagttatt gcattggagc agcagtcctt	960
cactttgaaa ccttcgcaat agcccagga gctgcctttc atattttcca ggttattgat	1020
aagaaacca gtatagataa cttttccaca gctggatata aacctgaatc catagaagga	1080
actgtggaat ttaaaaatgt ttctttcaat tatccatcaa gaccatctat caagattctg	1140
aaaggtctga atctcagaat taagtctgga gagacagtcg ccttggtcgg tctcaatggc	1200
agtgggaaga gtacggtagt ccagcttctg cagaggttat atgatccgga tgatggcttt	1260
atcatggtgg atgagaatga catcagagct ttaaattgtc ggcattatcg agaccatatt	1320
ggagtggta gtcaagagcc tgtttgttc gggaccacca tcagtaacaa tatcaagtat	1380
ggacgagatg atgtgactga tgaagagatg gagagagcag caaggggaagc aatgcgtat	1440
gattttatca tggagtttc taataaattt aatacattgg taggggaaaa aggagctcaa	1500
atgagtggag ggcagaaaca gaggatcgca attgctcgtg ccttagttcg aaacccaag	1560
attctgattt tagatgagcc tacgtctgcc ctggattcag aaagcaagtc agctgttcaa	1620
gctgcactgg agaaggcgag caaaggctcg actacaatcg tggtagcaca ccgactttct	1680
actattcgaa gtgcagattt gattgtgacc ctaaaggatg gaatgctggc ggagaaagga	1740
gcacatgctg aactaatggc aaaacgaggt ctatattatt cacttgtgat gtcacaggat	1800
attaaaaaag ctgatgaaca gatggagtca atgacatatt ctactgaaag aaagaccaac	1860
tcacttcctc tgcactctgt gaagagcatc aagtcagact tcattgaca ggctgaggaa	1920
tccaccaat ctaaagagat aagtcttct gaagtctctc tattaataa tttaaagtta	1980
aacaagcctg aatggccttt tgtggttctg gggacattgg cttctgttct aaatggaact	2040
gttcatccag tattttccat catctttgca aaaattataa ccatgtttgg aaataatgat	2100
aaaaccacat taaagcatga tgcagaaatt tattccatga tattcgtcat tttgggtggt	2160
atgtgcttg tcagttattt catgcagga ttattttacg gcagagcagg ggaaatttta	2220

ES 2 659 877 T3

```

acgatgagat taagacactt ggccctcaaa gccatgttat atcaggatat tgccctggttt 2280
gatgaaaagg aaaacagcac aggaggcttg acaacaatat tagccataga tatagcacia 2340
attcaaggag caacaggttc caggattggc gtcttaacac aaaatgcaac taacatggga 2400
ctttcagtta tcatttcctt tatatatgga tgggagatga cattcctgat tctgagtatt 2460
gtcccagtac ttgccgtgac aggaatgatt gaaaccgcag caatgactgg atttgccaac 2520
aaagataagc aagaacttaa gcatgctgga aagatagcaa ctgaagcttt ggagaatata 2580
cgtactatag tgtcattaac aagggaaaaa gccttcgagc aaatgtatga agagatgctt 2640
cagactcaac acagaaatc ctcgaagaaa gcacagatta ttggaagctg ttatgcattc 2700
agccatgcct ttatatattt tgccatgca gcagggtttc gatttggagc ctatttaatt 2760
caagctggac gaatgacccc agagggcatg ttcatagttt ttactgcaat tgcatatgga 2820
gctatggcca tcggaaaaac gctcgttttg gctcctgaat attccaaagc caaatcgggg 2880
gctgcgcatc tgtttgccct gttggaaaag aaaccaaata tagacagccg cagtcaagaa 2940
gggaaaaagc cagacacatg tgaagggaaat ttagagtttc gagaagtctc tttcttetat 3000
ccatgtcgcc cagatgtttt catcctccgt ggcttatecc tcagtattga gcgaggaaag 3060
acagtagcat ttgtggggag cagcggctgt gggaaaagca cttctgttca acttctgcag 3120
agactttatg accccgtgca aggacaagtg ctgtttgatg gtgtggatgc aaaagaattg 3180
aatgtacagt ggctccgttc ccaaatagca atcgttctc aagagcctgt gctcttcaac 3240
tgcagcattg ctgagaacat cgcctatggt gacaacagcc gtgtggtgcc attagatgag 3300
atcaaagaag ccgcaaatgc agcaaatatc cattctttta ttgaaggtct cctgagaaa 3360
tacaacacac aagttggact gaaaggagca cagctttctg gcggccagaa acaaagacta 3420
gctattgcaa gggctcttct ccaaaaaacc aaaattttat tgttgatga ggccacttca 3480
gcctcgata atgacagtga gaagtggtt cagcatgcc ttgataaagc caggacggga 3540
aggacatgcc tagtggtcac tcacaggctc tctgcaattc agaacgcaga tttgatagtg 3600
gttctgcaca atggaaagat aaaggaaca ggaactcatc aagagctcct gagaaatcga 3660
gacatatatt ttaagttagt gaatgcacag tcagtgcagt ga 3702

```

<210> 14

< 211> 3621

< 212> ADN

< 213> Homo sapiens

ES 2 659 877 T3

<400> 14

ttccgctttg ctgatggact ggacatcaca ctcatgatcc tgggtatact ggcactactg	60
gtcaatggag cctgccttcc tttaatgcc a ctggttttag gagaaatgag tgataacctt	120
attagtggat gtctagtcca aactaacaca tactctttct tcaggttgac cctgtattat	180
gttggaatag gtgttgctgc cttgatTTTT ggttacatac agatttcctt gtggattata	240
actgcagcac gacagaccaa gaggattcga aaacagtttt ttcattcagt tttggcacag	300
gacatcggct ggtttgatag ctgtgacatc ggtgaactta aactcgcac gacagacatt	360
gacaaaatca gtgatggat tggagataag attgctctgt tgtttcaaaa catgtctact	420
ttttcgattg gcctggcagt tggtttgggtg aagggtgga aactcacctt agtgactcta	480
tccacgtctc ctcttataat ggcttcagcg gcagcatggt ctaggatggt catctcattg	540
accagtaagg aattaagtgc ctattccaaa gctggggctg tggcagaaga agtcttgtca	600
tcaatccgaa cagtcatagc ctttagggcc caggagaaag aacttcaaag gtatacacag	660
aatctcaaag atgcaaagga ttttggcata aaaaggacta tagcttcaa agtgtctctt	720
ggtgctgtgt acttctttat gaatggaacc tatggacttg ctttttggta tggaacctcc	780
ttgattctta atggagaacc tggatatacc atcgggactg ttcttgctgt tttctttagt	840
gtaatccata gcagttattg cattggagca gcagtcctc actttgaaac cttcgcaata	900
gcccgaggag ctgcctttca tattttccag gttattgata agaaaccag tatagataac	960
ttttccacag ctggatataa acctgaatcc atagaaggaa ctgtggaatt taaaaatgtt	1020
tctttcaatt atccatcaag accatctatc aagattctga aaggctgaa tctcagaatt	1080
aagtctggag agacagtcgc cttggctcgt ctcaatggca gtgggaagag tacggtagtc	1140
cagcttctgc agaggttata tgatccggat gatggcttta tcatgggtgga tgagaatgac	1200
atcagagctt taaatgtgcg gcattatcga gaccatattg gagtggttag tcaagagcct	1260
gttttgttcg ggaccacat cagtaacaat atcaagtatg gacgagatga tgtgactgat	1320
gaagagatgg agagagcagc aagggaagca aatgcgtatg attttatcat ggagtttctt	1380
aataaattta atacattggt aggggaaaaa ggagctcaa tgagtggagg gcagaaacag	1440
aggatcgsaa ttgctcgtgc cttagtctga aacccaaga ttctgatttt agatgaggct	1500
acgtctgccc tggattcaga aagcaagtca gctgttcaag ctgcactgga gaaggcgagc	1560
aaaggtcgsa ctacaatcgt ggtagcacac cgactttcta ctattcgaag tgcagatttg	1620
attgtgacct taaaggatgg aatgctggcg gagaaaggag cacatgctga actaatggca	1680
aaacgaggtc tatattatc acttgtgatg tcacaggata ttaaaaaagc tgatgaacag	1740
atggagtcaa tgacatattc tactgaaaga aagaccaact cacttcctct gcaactctgtg	1800

ES 2 659 877 T3

```

aagagcatca agtcagactt cattgacaag gctgaggaat ccaccaatc taaagagata 1860
agtcttcctg aagtctctct attaaaaatt ttaaagttaa acaagcctga atggcctttt 1920
gtggttctgg ggacattggc ttctgttcta aatggaactg ttcattccagt attttccatc 1980
atctttgcaa aaattataac catgtttgga aataatgata aaaccacatt aaagcatgat 2040
gcagaaattht attccatgat attcgtcatt ttgggtgtta tttgctttgt cagttatttc 2100
atgcagggat tattttacgg cagagcaggg gaaatthttaa cgatgagatt aagacacttg 2160
gccttcaaag ccatgttata tcaggatatt gcctggtttg atgaaaagga aaacagcaca 2220
ggaggcttga caacaatatt agccatagat atagcacaaa ttcaaggagc aacaggttcc 2280
aggattggcg tcttaacaca aaatgcaact aacatgggac tttcagttat catttccttt 2340
atatatggat gggagatgac attcctgatt ctgagtattg ctccagtact tgccgtgaca 2400
ggaatgattg aaaccgcagc aatgactgga tttgcccaaca aagataagca agaacttaag 2460
catgctggaa agatagcaac tgaagctttg gagaatatac gtactatagt gtcattaaca 2520
agggaaaaag ccttcgagca aatgtatgaa gagatgcttc agactcaaca cagaaatacc 2580
tcgaagaaag cacagattat tggaagctgt tatgcattca gccatgcctt tatatatttt 2640
gcctatgcag cagggtttgc atttggagcc tatttaattc aagctggagc aatgacccca 2700
gagggcatgt tcatagtttt tactgcaatt gcatatggag ctatggccat cggaaaaacg 2760
ctcgttttgg ctctgaata ttccaaagcc aaatcggggg ctgcgcatct gtttgccttg 2820
ttggaaaaaga aaccaaatat agacagccgc agtcaagaag ggaaaaagcc agacacatgt 2880
gaaggggaatt tagagtttgc agaagtctct ttcttctatc catgtcgccc agatgttttc 2940
atcctccgtg gcttatccct cagtattgag cgaggaaaga cagtagcatt tgtggggagc 3000
agcggctgtg ggaaaagcac ttctgttcaa ctctgcaga gactttatga ccccgtgcaa 3060
ggacaagtgc tgtttgatgg tgtggatgca aaagaattga atgtacagtg gctccgttcc 3120
caaatagcaa tcgttctca agagcctgtg ctcttcaact gcagcattgc tgagaacatc 3180
gcctatggtg acaacagccg tgtggtgcca ttagatgaga tcaaagaagc cgcaaatgca 3240
gcaaatatcc attcttttat tgaaggtctc cctgagaaat acaacacaca agttggactg 3300
aaaggagcac agctttctgg cggccagaaa caaagactag ctattgcaag ggctcttctc 3360
caaaaaccca aaatthtatt gttggatgag gccacttcag ccctcgataa tgacagtgag 3420
aaggtggttc agcatgccct tgataaaagcc aggacgggaa ggacatgcct agtggtcact 3480
cacaggctct ctgcaattca gaacgcagat ttgatagtgg ttctgcacaa tggaaagata 3540
aaggaacaag gaactcatca agagctcctg agaaatcgag acatatattt taagttagtg 3600
aatgcacagt cagtgcagtg a 3621

```

<210> 15

< 211> 2021

< 212> ADN

< 213> Homo sapiens

<220>

< 221> característica_misc

< 222> (723)..(723)

< 223> /nota="n representa cualquier nucleótido (A, T, C o G)"

ES 2 659 877 T3

<400> 15

```

ttccgctttg ctgatggact ggacatcaca ctcatgatcc tgggtatact ggcatcactg      60
gtcaatggag cctgccttcc tttaatgcc a ctggttttag gagaaatgag tgataacctt     120
attagtggat gtctagtcca aactaacaca tactctttct tcaggttgac cctgtattat     180
gttggaatag gtgttgctgc cttgattttt ggttacatac agatttcctt gtggattata     240
actgcagcac gacagaccaa gaggattcga aaacagtttt ttcattcagt tttggcacag     300
gacatcggct ggtttgatag ctgtgacatc ggtgaaacta acactcgcac gacagacatt     360
gacaaaatca gtgatggat tggagataag attgctctgt tgtttcaaaa catgtctact     420
ttttcgattg gcctggcagt tggtttggtg aagggctgga aactcacctt agtgactcta     480
tccacgtctc ctcttataat ggcttcagcg gcagcatggt ctaggatggt catctcattg     540
accagtaagg aattaagtgc ctattccaaa gctggggctg tggcagaaga agtcttgtca     600
tcaatccgaa cagtcatagc ctttagggcc caggagaaag aacttcaaag gtctttcctt     660
ttaaatataa caagatatgc ttggttttat tttcccagc ggctactaag ttgtgttctg     720
ttntttgtaa ggtatacaca gaatctcaa gatgcaaagg attttggcat aaaaggact     780
atagcttcaa aagtgtctct tgggtgctgt tacttcttta tgaatggaac ctatggactt     840
gctttttggt atggaacctc cttgattctt aatggagaac ctggatatac catcgggact     900
gttcttgctg ttttctttag tgtaatccat agcagttatt gcattggagc agcagtcctt     960
cactttgaaa ccttcgcaat agcccagga gctgcctttc atattttcca ggttattgat    1020
aagaaaccca gtatagataa cttttccaca gctggatata aacctgaatc catagaagga    1080
actgtggaat ttaaaaatgt ttctttcaat tatccatcaa gaccatctat caagattctg    1140
aaaggtctga atctcagaat taagtctgga gagacagtcg ccttggtcgg tctcaatggc    1200
agtgggaaga gtacggtagt ccagcttctg cagaggttat atgatccgga tgatggcttt    1260
atcatggtgg atgagaatga catcagagct ttaaattgtc ggcattatcg agaccatatt    1320
ggagtgttta gtcaagagcc tgttttgttc gggaccacca tcagtaacaa tatcaagtat    1380
ggacgagatg atgtgactga tgaagagatg gagagagcag caagggaaagc aaatgcgtat    1440
gattttatca tggagtttcc taataaattt aatacatggg taggggaaaa aggagctcaa    1500
atgagtggag ggcagaaaca gaggatcgca attgctcgtg ccttagttcg aaacccaag    1560
attctgattt tagatgagcg tacgtctgcc ctggattcag aaagcaagtc agctgttcaa    1620
gctgcactgg agaaggatac ccccaggtat tcattttgac ctaatttcac ctcaagtgga    1680
gaatcgctga ccttgaacca gcgcccttgc acagctctgg ccctcaaac ctcaccctga    1740
cctcctgctg cctatgagct actgcacata cctcaagcc atatgcagtt gtggccctgc    1800
accaaattac actgaatcta gggggggagt tggcagtggc ggtatgaaaa accattgaac    1860
agttttctcg atggcctgac tcccttataa accagagcct tcagaccctt tacaaggctt    1920
aatggacat tttactttgc atttgcttgg aagtgagtta agcgtttttt tttctctaag    1980
aaaatcgcag gcttcttttt ttaaaatgct gactttatgg a                                2021

```

<210> 16

< 211> 1940

< 212> ADN

< 213> Homo sapiens

ES 2 659 877 T3

<400> 16

```

ttccgctttg ctgatggact ggacatcaca ctcatgatcc tgggtatact ggcactactg      60
gtcaatggag cctgccttcc tttaatgcc a ctggtttttag gagaaatgag tgataacctt     120
attagtggat gtctagtcca aactaacaca tactctttct tcaggttgac cctgtattat     180
gttggaatag gtgttgetgc cttgattttt ggttacatac agatttcctt gtggattata     240
actgcagcac gacagaccaa gaggattcga aaacagtttt ttcattcagt tttggcacag     300
gacatcggct ggtttgatag ctgtgacatc ggtgaaacta aactcgcac gacagacatt     360
gacaaaatca gtgatggat tggagataag attgctctgt tgtttcaaaa catgtctact     420
ttttcgattg gcctggcagt tggtttggtg aagggctgga aactcacctt agtgactcta     480
tccacgtctc ctcttataat ggcttcagcg gcagcatggt ctaggatggt catctcattg     540
accagtaagg aattaagtg c tattcctaaa gctggggctg tggcagaaga agtcttgtca     600
tcaatccgaa cagtcatagc ctttagggcc caggagaaag aacttcaaag gtatacacag     660
aatctcaaag atgcaaagga ttttggcata aaaaggacta tagcttcaa agtgtctctt     720
ggtgctgtgt acttctttat gaatggaacc tatggacttg ctttttggtg tggaacctcc     780
ttgattctta atggagaacc tggatatacc atcgggactg ttcttgctgt tttcttttagt     840
gtaatccata gcagttattg cattggagca gcagtcctc actttgaaac cttcgcaata     900
gcccgaggag ctgcctttca ttttttcag gttattgata agaaaccag tatagataac     960
ttttccacag ctggatataa acctgaatcc atagaaggaa ctgtggaatt taaaaatgtt    1020
tctttcaatt atccatcaag accatctatc aagattctga aaggtctgaa tctcagaatt    1080
aagtctggag agacagtcgc cttggctcgt ctcaatggca gtgggaagag tacggtagtc    1140
cagcttctgc agaggttata tgatccggat gatggcttta tcatggtgga tgagaatgac    1200
atcagagcct taaatgtgcg gcattatcga gaccatattg gagtggttag tcaagagcct    1260
gttttgttcg ggaccacat cagtaacaat atcaagtatg gacgagatga tgtgactgat    1320
gaagagatgg agagagcagc aagggaaagca aatgcgatg attttatcat ggagtttcct    1380
aataaattta atacattggt aggggaaaaa ggagctcaa tgagtggagg gcagaaacag    1440
aggatcgcaa ttgctcgtgc cttagttcga aacccaaga ttctgatttt agatgaggct    1500
acgtctgccc tggattcaga aagcaagtca gctgttcaag ctgcactgga gaaggatacc    1560
cccaggtatt ctttttgacc taatttcacc tcaagtggag aatcgctgac cttgaaccag    1620
cgcccttcga cagctctggc ccctcaaacc tcacctgac ctctgctgc ctatgagcta    1680
ctgcacatac ctcaaggcca tatgcagttg tggcctgca ccaaattaca ctgaatctag    1740
gaggggagtt ggcagtgccg gtatgaaaaa ccattgaaca gttttctcga tggcctgact    1800
cccttataaa ccagagcctt cagaccctt acaaggetta atggcacatt ttactttgca    1860
tttgcttggg agtgagttaa gcgttttttt ttctctaaga aaatcgagg cttctttttt    1920
taaaatgctg actttatgga                                     1940

```

<210> 17

<211> 15

<212> PRT

<213> Homo sapiens

ES 2 659 877 T3

<400> 17

Leu Ser Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu
1 5 10 15

<210> 18

< 211> 15

< 212> PRT

< 213> Secuencia artificial

<220>

< 221> fuente

< 222> (1)..(15)

< 223> /nota="péptido sintético"

<400> 18

Cys Gly Thr Ser Leu Ile Leu Asn Gly Glu Pro Gly Tyr Thr Ile
1 5 10 15

<210> 19

< 211> 17

< 212> PRT

< 213> Secuencia artificial

<220>

< 221> fuente

< 222> (1)..(17)

< 223> /nota="péptido sintético"

<400> 19

Arg Phe Gly Ala Tyr Leu Ile Gln Ala Gly Arg Met Thr Pro Glu Gly
1 5 10 15

Cys

REIVINDICACIONES

1. Una proteína sustancialmente pura que comprende la secuencia de aminoácidos de SEQ ID NO: 2.
2. La proteína de la reivindicación 1, constituida por SEQ ID NO: 2.
- 5 3. Un anticuerpo monoclonal que se puede obtener administrando la proteína de la reivindicación 1 o la reivindicación 2 y donde el anticuerpo monoclonal se une selectivamente a la proteína constituida por la secuencia de aminoácidos de SEQ ID NO: 2.
4. Un polinucleótido sustancialmente puro que codifica la proteína de la reivindicación 1 o la reivindicación 2.
5. Un vector para expresar una glicoproteína P, que comprende un elemento codificante distinto constituido por el polinucleótido de la reivindicación 4.
- 10 6. Una célula hospedadora transformada con el vector de la reivindicación 5.
7. Una glicoproteína P producida por la célula hospedadora de la reivindicación 6.
8. Un método para obtener células madre que son positivas para la glicoproteína P codificada por el aminoácido de SEQ ID NO: 2, que comprende aislar células madre de una muestra de células madre utilizando el anticuerpo de la reivindicación 3, donde dichas células madre no son células madre embrionarias humanas.
- 15 9. Una población aislada de células madre que son positivas para la glicoproteína P codificada por el aminoácido de SEQ ID NO: 2 y que se han aislado utilizando el anticuerpo de la reivindicación 3, donde dichas células madre no son células madre embrionarias humanas.
10. La célula madre de la reivindicación 9, para su uso en terapia.
- 20 11. Un método de enriquecimiento de células madre a partir de una mezcla de células, que comprende poner en contacto una mezcla de células con el anticuerpo de la reivindicación 3, y aislar las células madre unidas al anticuerpo a partir de la mezcla celular, donde dichas células madre no son células madre embrionarias humanas.
12. Un método de caracterización de una célula madre, que comprende poner en contacto una célula madre con el anticuerpo de la reivindicación 3 para determinar si la célula madre es positiva para glicoproteína P codificada por la secuencia de aminoácidos de SEQ ID NO: 2.