

(12) FASCÍCULO DE PATENTE DE INVENÇÃO

(22) Data de pedido: 2001.06.05	(73) Titular(es): THE BRIGHAM & WOMEN'S HOSPITAL, INC. 75 FRANCIS STREET BOSTON, MA 02115 US
(30) Prioridade(s): 2000.06.05 US 208913 P	
(43) Data de publicação do pedido: 2010.05.05	(72) Inventor(es): MOHAMED H SAYEGH US MARKUS H FRANK US
(45) Data e BPI da concessão: 2015.03.25 139/2015	(74) Mandatário: NUNO MIGUEL OLIVEIRA LOURENÇO RUA CASTILHO, Nº 50 - 9º 1269-163 LISBOA PT

(54) Epígrafe: **UM GENE QUE CODIFICA UM HOMÓLOGO DE GLICOPROTEÍNA P HUMANA DE RESISTÊNCIA A MÚLTIPLOS FÁRMACOS NO CROMOSSOMA 7P15-21 E SUAS TILIZAÇÕES**

(57) Resumo:

UM ANTICORPO PARA UMA GLICOPROTEÍNA P DA FAMÍLIA MDR LOCALIZADA NO CROMOSSOMA HUMANO 7P15-21 É ÚTIL PARA ENRIQUECER E CARACTERIZAR CÉLULAS ESTAMINAIS.

RESUMO

**"UM GENE QUE CODIFICA UM HOMÓLOGO DE GLICOPROTEÍNA P HUMANA
DE RESISTÊNCIA A MÚLTIPLOS FÁRMACOS NO CROMOSSOMA 7P15-21 E
SUAS UTILIZAÇÕES"**

Um anticorpo para uma glicoproteína P da família MDR localizada no cromossoma humano 7p15-21 é útil para enriquecer e caracterizar células estaminais.

DESCRIÇÃO

"UM GENE QUE CODIFICA UM HOMÓLOGO DE GLICOPROTEÍNA P HUMANA DE RESISTÊNCIA A MÚLTIPLOS FÁRMACOS NO CROMOSSOMA 7P15-21 E SUAS UTILIZAÇÕES"

Campo da Invenção

A invenção refere-se a sequências genéticas que codificam proteínas que apresentam particularidades estruturais e funcionais características de membros da família da glicoproteína P associadas a resistência de cancros a múltiplos fármacos, funções imunorreguladoras e funções únicas em células estaminais pluripotentes humanas e outras células progenitoras de tecido. A invenção abrange proteínas consideravelmente puras, tratamentos terapêuticos e utilizações em diagnóstico relacionadas com estas proteínas.

Antecedentes da Invenção

A glicoproteína P, uma bomba de efluxo de fármacos dependente de adenosina trifosfato (ATP), está sobre-expressa em células tumorais resistentes a múltiplos fármacos (MDR). Ela reduz a concentração intracelular de xenobióticos citotóxicos, diminuindo deste modo a eficácia de muitos regímenes quimioterapêuticos contra o cancro. A glicoproteína P pertence à superfamília ABC (cassete de ligação a ATP) de transportadores ativos e é codificada por uma família de múltiplos genes em eucariotas superiores. Os membros da família da glicoproteína P de mamíferos podem ser divididos em três classes. As glicoproteínas P de classe I e classe II conferem resistência a múltiplos fármacos enquanto as proteínas de classe III não.

Nos humanos, a glicoproteína P é codificada por dois genes ligados ("MDR1" e "MDR3") no cromossoma 7q21.1. O MDR3 funciona como uma translocase de lípidos e mutações neste gene estão associadas a colestase intra-hepática familiar. O MDR1 confere resistência a fármaco em determinadas células cancerosas. Além de estar sobre-expressa em células cancerosas, a glicoproteína P de MDR1 é expressa de modo disseminado em tecidos humanos, predominantemente secretores e absorptivos, normais onde atua em vários processos fisiológicos incluindo diferenciação celular, proliferação celular e sobrevivência celular. Nestes tipos de células normais, a glicoproteína P atua na libertação ou captação transmembranar de xenobióticos e determinados fármacos terapêuticos, moléculas de péptido pequeno, determinados compostos esteroides e fosfolípidos.

A glicoproteína P é também expressa por populações de células linfoides da medula óssea humana e de sangue periférico. Especificamente, foi demonstrado que a glicoproteína P é expressa na membrana de células estaminais pluripotentes, monócitos, células dendríticas, linfócitos T CD4+ e CD8+, células assassinas naturais e linfócitos B. Nas células imunes, a glicoproteína P atua no transporte de citocinas e outras moléculas pequenas, as quais são críticas para que ocorram as respostas fisiológicas imunes. O bloqueio específico da glicoproteína P pode suprimir a resposta imunológica a aloantígenos e antígenos nominais. No entanto, existe um grau de redundância para a função da glicoproteína P nestes tipos de células, que aponta para a existência de outras moléculas relacionadas, até agora não identificadas.

As células estaminais pluripotentes e outras células progenitoras de tecido também possuem uma atividade tipo

glicoproteína P única, caracterizada pela menor acumulação intracelular de corantes fluorescentes, o que permite o isolamento específico deste tipo de células para utilizações terapêuticas. No entanto, julga-se que esta função não seja mediada pela glicoproteína P MDR1, mas antes por um membro relacionado com a família da glicoproteína P, até agora não identificado.

Apesar do papel irrefutável da glicoproteína P MDR1 na resistência de cancros a múltiplos fármacos, tentativas para melhorar a quimioterapia inibindo esta proteína tem sido conseguida apenas com um sucesso limitado. Assim, pode inferir-se que existem proteínas homólogas que, como a MDR1, são capazes de tornar as células resistentes aos agentes terapêuticos. Além disso, pode deduzir-se que as proteínas homólogas de MDR1 desempenham funções tipo glicoproteína P em tecidos fisiológicos humanos, em particular em células do sistema imunitário, células estaminais pluripotentes e células progenitoras de tecido, onde há redundância da função da glicoproteína P MDR1 ou onde se sabe que a glicoproteína P MDR1 não promove a atividade associada a glicoproteína P observada.

Sumário da Invenção

A invenção é dirigida a um novo membro da família de genes da glicoproteína P humana localizado no cromossoma 7p15-2, que codifica proteínas as quais conferem o fenótipo resistente a múltiplos fármacos em células tumorais e/ou servem funções fisiológicas críticas em tecidos humanos normais.

Um exame da estrutura do novo gene indica que ele codifica duas metades homólogas semi-autónomas, cada uma com os seus

próprios domínios transmembranares e de ligação a ATP. Através da expressão diferencial de genes e excisão-união alternativa e/ou modificações pós-transcrição e pós-tradução, o novo gene de glicoproteína P pode codificar várias glicoproteínas P distintas:

A proteína de SEQ ID NO:1 (aminoácidos 1-659) é codificada por 14 exões (SEQ ID NO:9) de ADN genómico humano do clone AC005060 no cromossoma 7p15-21 e é constituída por 5 domínios transmembranares e um domínio de ligação a ATP.

A proteína de SEQ ID NO:2 (aminoácidos 1-812) é codificada por 19 exões (SEQ ID NO:10) de ADN genómico humano dos clones contíguos AC002486 e AC005060 (AC002486 é o clone sequenciado à esquerda do clone AC005060) no cromossoma 7p15-21 e é constituída por 5 domínios transmembranares e dois domínios de ligação a ATP, dos quais o primeiro está localizado no lado N-terminal do domínio transmembranar #1, e o segundo no lado C-terminal do domínio transmembranar #5 da proteína, do lado oposto da membrana do plasma. A proteína de SEQ ID NO:2 também pode ser expressa como um resultado de trans-excisão-união do ARNm (SEQ ID NO: 9) que codifica a proteína de SEQ ID NO:1 e do ARNm (SEQ ID NO:11) que codifica a proteína de SEQ ID NO:3 descrita a seguir. Além disso, a proteína de SEQ ID NO:2 pode ser expressa como o resultado de processamento pós-tradução das proteínas de SEQ ID NO:1 e NO: 3.

A proteína de SEQ ID NO:3 (aminoácidos 1-131) é codificada por 6 exões (SEQ ID NO:11) de ADN genómico humano do clone AC002486 no cromossoma 7p15-21 e é constituída por um domínio de ligação a ATP e sem domínios transmembranares.

A proteína de SEQ ID NO:4 (aminoácidos 1-1058) é codificada por 20 exões (SEQ ID NO:12) de ADN genómico humano dos

clones contíguos AC002486 e AC005060 no cromossoma 7p15-21 e é constituída por 8 domínios transmembranares e dois domínios de ligação a ATP, dos quais o primeiro está localizado entre os domínios transmembranares #3 e #4, e o segundo no lado C-terminal dos domínios transmembranares #8, do lado oposto da membrana do plasma.

A proteína de SEQ ID NO:5 (aminoácidos 1-1222) é codificada por 23 exões (SEQ ID NO:13) de ADN genómico humano dos clones contíguos AC002486 e AC005060 no cromossoma 7p15-21 e é constituída por 12 domínios transmembranares e dois domínios de ligação a ATP, dos quais o primeiro está localizado entre os domínios transmembranares #7 e #8, e o segundo no lado C-terminal do domínio transmembranar # 12, do lado oposto da membrana do plasma.

A proteína de SEQ ID NO:6 (aminoácidos 1-1195) é codificada por 24 exões (SEQ ID NO:14) de ADN genómico humano dos clones contíguos AC002486 e AC005060 no cromossoma 7p15-21 e é constituída por 11 domínios transmembranares e dois domínios de ligação a ATP, dos quais o primeiro está localizado entre os domínios transmembranares #6 e #7, e o segundo no lado C-terminal do domínio transmembranar #11, do lado oposto da membrana do plasma.

A proteína de SEQ ID NO:7 (aminoácidos 1-541) é codificada por 10 exões (SEQ ID NO:15) de ADN genómico humano do clone AC002486 no cromossoma 7p15-21 e é constituída por 7 domínios transmembranares e um domínio de ligação a ATP no lado C-terminal do domínio transmembranar #7.

A proteína de SEQ ID NO:8 (aminoácidos 1-514) é codificada por 11 exões (SEQ ID NO:16) de ADN genómico humano do clone AC002486 no cromossoma 7p15-21 e é constituída por 6

domínios transmembranares e um domínio de ligação a ATP no lado C-terminal do domínio transmembranar #6.

A resistência de cancros a múltiplos fármacos pode resultar da expressão de qualquer uma das proteínas de SEQ ID NO:1, NO:2, NO:3, NO:4, NO:5, NO:6 NO:7 e NO:8. As proteínas codificadas pelo gene 7p15-21 da glicoproteína P da presente invenção podem ser utilizadas como marcadores para identificar células que é provável que apresentem resistência a múltiplos fármacos e podem servir como alvos na conceção de novas terapias para doentes com cancro. Entender-se-á que, exceto quando indicado em contrário, a referência à glicoproteína P da presente invenção inclui também qualquer uma das proteínas de SEQ ID NO:1, NO:2, NO:3, NO:4, NO:5, NO:6 NO:7 e também NO:8.

A glicoproteína P do 7p15-21 confere quimiorresistência a múltiplos agentes quimioterapêuticos, incluindo cisplatina, através da mediação do efluxo celular do fármaco. Por isso, o bloqueio específico desta função de efluxo, por exemplo por meio da inibição específica de anticorpo monoclonal, pode aumentar a acumulação intracelular de fármaco e, em consequência, a toxicidade do fármaco e a morte de células tumorais. Além disso, uma vez que a glicoproteína P do 7p15-21 está funcional na proliferação celular de tumores, o crescimento tumoral pode ser inibido terapeuticamente pela administração de anticorpos monoclonais específicos de bloqueio, mesmo na ausência de agentes quimioterapêuticos simultâneos. De entre as proteínas codificadas pelo gene 7p15-21 da glicoproteína P, as proteínas de SEQ ID NO:1, NO:2, NO:3, NO:4, NO:5 e NO:6 são distintas das proteínas de SEQ ID NO:7 e NO:8 pelo facto de serem expressas seletivamente em determinadas células cancerosas mas não em tecidos não cancerosos, normais. Além disso, as proteínas de SEQ ID NO:1, NO:2, NO:3, NO:4, NO:5 e NO:6 são expressas

preferencialmente naqueles cancros que exibem os graus mais elevados de quimiorresistência a fármacos quimioterapêuticos, como por exemplo o melanoma maligno humano. Devido à sua expressão seletiva em determinados cancros mas não nos tecidos normais, as proteínas de SEQ ID NO:1, NO:2, NO:3, NO:4, NO:5 e NO:6 podem ser terapêuticamente visadas não só através da inibição do efluxo do fármaco citotóxico ou inibição da proliferação do tumor por anticorpos monoclonais específicos, mas também por meios adicionais, incluindo a morte celular específica de tumores mediada por anticorpos monoclonais específicos conjugados com toxinas celulares, ou pela administração terapêutica a doentes afetados de preparações de vacina específicas para antigénios de tumores.

As proteínas de SEQ ID NO:7 e NO:8 codificadas pelo gene 7p15-21 também podem ser expressas em determinados tecidos humanos normais, não cancerosos. Assim, a invenção proporciona utilizações adicionais no que se refere à função destas proteínas selecionadas em tecidos fisiológicos. Entre os tecidos normais, as proteínas de SEQ ID NO:7 e SEQ ID NO:8 são preferencialmente expressas em níveis elevados em células estaminais pluripotentes e outras células progenitoras de tecido, onde atuam no transporte transmembranar de xenobióticos e outras moléculas pequenas. Assim, a invenção proporciona meios para detetar especificamente e enriquecer estas células estaminais e células progenitoras de misturas e preparações de células nas quais elas estão contidas, por deteção das células com anticorpos monoclonais específicos marcados.

As proteínas de SEQ ID NO:7 e NO:8 também são expressas em certa medida na maioria dos outros tecidos humanos normais, incluindo nas células do sistema imunitário tais como células T, monócitos e células apresentadoras de antigénio

diferenciadas, onde elas funcionam no efluxo de citocinas e na captação de moléculas pequenas incluindo péptidos e antigénio, desempenhando desse modo um papel crítico para a integridade das respostas imunológicas normais. Quando estas funções são inibidas, por exemplo por bloqueio com anticorpos monoclonais específicos, a resposta imunológica normal pode ser modulada, o que pode ser utilizado na prevenção e/ou terapia da rejeição de aloenxerto no transplante clínico de órgãos, e também em várias doenças autoimunes como artrite reumatoide e esclerose múltipla. Além disso, quando expressa em células imunitárias humanas e noutros tecidos humanos como o endotélio da barreira hematoencefálica e os epitélios do aparelho gastrointestinal e do rim, o bloqueio da proteína pode ser ainda utilizado terapeuticamente para alterar seletivamente a captação e secreção, e desse modo a distribuição farmacológica, farmacocinética e eficácia terapêutica dos fármacos terapêuticos administrados exogenamente que são substratos das referidas proteínas.

Num primeiro aspeto, a invenção é dirigida a proteínas substancialmente puras consistindo essencialmente da sequência de aminoácidos 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 ou SEQ ID NO:8. O termo "consistindo essencialmente de" tem a intenção de incluir proteínas possuindo exatamente as mesmas sequências de aminoácido, assim como proteínas com sequências insubstancialmente diferentes, como evidenciado por possuírem as mesmas propriedades funcionais básicas. Uma isoforma "substancialmente purificada" é aquela que foi separada de outros componentes biológicos concomitantes e compreenderá tipicamente pelo menos 85% de uma amostra, sendo preferidas percentagens maiores. Estão disponíveis muitos meios para avaliar a pureza de uma proteína numa amostra, incluindo análise por eletroforese em gel de

poliacrilamida, cromatografia e centrifugação analítica. Um método preferido para avaliar a pureza é por transferência de Western utilizando um anticorpo dirigido contra epítomos da glicoproteína P do 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 ou SEQ ID NO:8. A invenção abrange também "péptidos MDR" os quais são aqui definidos como consistindo de um elemento de sequência 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 ou SEQ ID NO:8. de pelo menos 10 e preferencialmente pelo menos 15 ou 20 resíduos. Estes podem ser utilizados na produção de anticorpos. É estipulado que um péptido MDR não pode ter uma sequência que seja a mesma que qualquer conjunto de 10 a 15 resíduos contíguos na sequência LSGGQKQRIAIARAL (SEQ ID NO:17). Estas proteínas e péptidos MDR também podem ser administrados terapeuticamente a doentes com cancro afetados por tumores que expressam glicoproteína P do 7p15-21, como uma vacina para o tumor para desencadear uma resposta imunológica endógena dirigida contra estes tumores, para resultar em morte celular específica para tumores.

A invenção é dirigida a um anticorpo preparado por um processo compreendendo o passo de administração, a um hospedeiro animal, de uma proteína codificada pela 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 ou SEQ ID NO:8, ou um péptido MDR como descrito acima. A proteína ou péptido deveria ser administrado ao animal a uma dosagem suficiente para induzir a formação de anticorpo. Os anticorpos podem ser monoclonais ou policlonais. No último caso, os anticorpos são preferencialmente produzidos injetando uma preparação farmacêuticamente aceitável num ratinho, seguida da fusão de células de baço de ratinho com células de mieloma utilizando técnicas conhecidas na arte. Os anticorpos

obtidos deveriam ligar-se seletivamente às 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 ou SEQ ID NO:8. Neste contexto, ligação seletiva significa que um anticorpo tem pelo menos uma afinidade 100 vezes maior para uma ou mais destas proteínas do que para qualquer outra proteína normalmente encontrada em células humanas.

A invenção é também dirigida para um polinucleótido substancialmente puro consistindo essencialmente de uma sequência de nucleótidos que codifica as 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 ou SEQ ID NO:8, ou um péptido MDR. Preferencialmente, o polinucleótido consiste essencialmente da sequência 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 ou SEQ ID NO:16. A invenção inclui vetores de expressão compreendendo um elemento de codificação distinto que consiste destes polinucleótidos; e células hospedeiras transformadas com tais vetores. Um "elemento de codificação distinto" refere-se à porção de um vetor de expressão responsável por determinar a sequência de aminoácidos de uma proteína expressa. A invenção compreende todos esses elementos que produzem proteínas correspondentes às sequências de aminoácidos mostradas nas 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 ou SEQ ID NO:8, bem como outras proteínas possuindo essencialmente a mesma estrutura e função.

A invenção inclui uma proteína recombinante feita por células hospedeiras transformadas por um vetor de expressão como discutido acima. A proteína recombinante pode ser isolada utilizando técnicas correntes, incluindo cromatografia de afinidade com anticorpos contra epítopos

de glicoproteína P do 7p15-21. Preferencialmente, o polinucleótido utilizado nos vetores para expressão, como uma glicoproteína P recombinante, consiste essencialmente das sequências 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 ou SEQ ID NO:16. Podem ser utilizados como inibidores antimensageiros oligonucleótidos complementares 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 ou SEQ ID NO:16 e com pelo menos 15 nucleótidos de comprimento. Estes podem ser administrados a doentes submetidos a quimioterapia contra o cancro para aumentar a eficácia de fármacos citotóxicos. A transfeção *in vivo* de células é conhecida há muitos anos e pode ser efetuada utilizando vetores virais (ver *por exemplo*, U.S. 6,020,191); lipossomas (ver *por exemplo*, Nicolau, *Meth. Enzymol* 149:157-176 (1987)); ADN complexado com agentes que facilitam a captação celular (ver *por exemplo*, U.S. 5,264,618; WO 98/14431); ou ainda injetando simplesmente ADN nu (ver *por exemplo*, U.S. 5,693,622). Qualquer um destes procedimentos pode ser utilizado para administrar os oligonucleótidos antimensageiros da presente invenção.

A invenção é também dirigida a um método para determinar se uma célula cancerosa responderá a terapias que têm como objetivo a inversão da resistência a múltiplos fármacos medindo a expressão dos genes que codificam as 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 ou SEQ ID NO:8. Este método pode ser utilizado para detetar a existência do fenótipo resistente a múltiplos fármacos em células cancerosas ou para seguir o desenvolvimento de resistência a múltiplos fármacos ao longo do tempo, monitorizando alterações na expressão do gene em células cultivadas.

A invenção proporciona um método para determinar se um composto de ensaio inibe a resistência a múltiplos fármacos em células ocasionada por um gene que codifica 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 ou SEQ ID NO:8. Este método compreende a expressão de um gene que codifica um ou mais destes polipéptidos em células que, de outro modo, não são resistentes a múltiplos fármacos e a exposição destas células a um ou mais fármacos citotóxicos na presença de um composto de ensaio. A sobrevivência celular é medida após exposição e os resultados obtidos são comparados com aqueles de incubações realizadas essencialmente do mesmo modo mas na ausência do composto de ensaio. Concluiu-se que o composto de ensaio inibe a resistência a múltiplos fármacos se a sobrevivência celular for diminuída numa extensão significativa em incubações realizadas na presença do composto de ensaio relativamente à observada na sua ausência.

Descrição Detalhada da Invenção

A invenção é dirigida a um novo membro da família da glicoproteína P de proteínas relacionadas com resistência a fármacos, a sequências genéticas que codificam esta proteína, a métodos para determinar se uma célula cancerosa responderá a terapias que têm como objetivo a inversão de resistência a fármacos mediada pela glicoproteína P e a um método de pré-seleção de compostos de ensaios quanto à sua capacidade para inibir a resistência a múltiplos fármacos. O novo gene de glicoproteína P pode codificar as 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 ou SEQ ID NO:8.

Entender-se-á que a invenção inclui não só sequências idênticas às mostradas mas também sequências que são essencialmente iguais como evidenciado pela conservação das mesmas características estruturais e funcionais básicas. Por exemplo, podem ser utilizadas técnicas como a mutagênese de um locus específico para introduzir variações na estrutura de uma proteína. São divulgadas variações na glicoproteína P introduzidas por este ou por outros métodos semelhantes desde que a proteína resultante conserve as suas propriedades biológicas básicas, particularmente em relação à indução de resistência a múltiplos fármacos em células de mamíferos.

As sequências de ADN que codificam as proteínas da invenção podem ser obtidas de qualquer tecido ou fonte celular no qual são expressas. Por exemplo, linhas de células cultivadas podem ser manipuladas para expressar o gene de glicoproteína P utilizando técnicas recombinantes ou por exposição contínua a agentes quimioterapêuticos. Alternativamente, as sequências podem ser isoladas de células primárias obtidas de tumores.

Estão disponíveis muitos métodos para isolar sequências de ADN e podem ser adaptados para o isolamento do gene da glicoproteína P do cromossoma 7p15-21 (doravante "cromossoma 7p") (ver, por exemplo, Sambrook et al., Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Press (1989)). Por exemplo, um método consiste em rastrear uma biblioteca de ADNc que foi preparada por transcrição inversa de ARNm isolado de tecidos ou células que expressam o gene. A biblioteca pode ser preparada, por exemplo, de melanócito ou de tecido de testículo humano e as sondas para rastreio podem ser sintetizadas com base nas sequências mostradas na Listagem de Sequências. As sondas têm preferencialmente pelo menos

14 nucleótidos de comprimento e são selecionadas, de modo ótimo, de uma região que se julgue que seja única ao gene de glicoproteína P do cromossoma 7 p.

Como uma alternativa, a amplificação de uma sequência desejada pode ser conseguida pela reação em cadeia da polimerase ("PCR") de ARN transcrito inversamente. Os iniciadores para PCR podem ser construídos utilizando as sequências mostradas na 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 ou SEQ ID NO:16, e a confirmação da presença de ADNc da glicoproteína P do cromossoma 7p pode ser conseguida pela sequenciação de produtos de amplificação.

A expressão de proteína recombinante pode ser induzida numa célula hospedeira transformando-a com um vetor de expressão apropriado. O vetor deveria conter sinais de transcrição e tradução reconhecíveis pelo hospedeiro em conjunto com a sequência estrutural desejada, preferencialmente na forma de cadeia dupla, numa ligação operável. Por exemplo, a sequência de ADN da glicoproteína P deveria ser posicionada de modo a que as sequências reguladoras presentes no vetor controlem a síntese de ARNm e produzam a proteína possuindo a sequência desejada.

Preferencialmente, o ácido nucleico que codifica a glicoproteína P da invenção é expresso em células eucariotas, especialmente células de mamíferos. Tais células são capazes de promover as modificações pós-tradução necessárias para garantir que a proteína recombinante é estrutural e funcionalmente semelhante à proteína isolada, por exemplo, de células tumorais resistentes a múltiplos fármacos. Exemplos de células de mamíferos que se sabe que proporcionam modificação pós-tradução apropriada de proteínas clonadas incluem, *inter*

alia, células NIH-3T3, células CHO, células HeLA, células LM(tk-) e semelhantes. São preferencialmente utilizados promotores eucariotas que se sabe que controlam a expressão de genes recombinantes para conduzir a transcrição de ADN da glicoproteína P do cromossoma 7p e podem incluir aqueles do gene de metalotioneína I de ratinho, o promotor TK do vírus de Herpes, o promotor precoce de CMV e o promotor precoce de SV40. A transcrição também pode ser orientada por promotores procariotas, como aqueles capazes de reconhecer a polimerase T4, os promotores P_R e P_L de bacteriófago lambda, e os promotores *trp*, *recA*, choque térmico e *lacZ* de *E. coli*.

Os vetores de expressão podem ser introduzidos nas células hospedeiras por métodos como a precipitação com fosfato de cálcio, microinjeção, electroporação ou transferência viral e as células que expressam a sequência da proteína recombinante podem ser selecionadas por técnicas conhecidas na arte. A confirmação da expressão pode ser obtida através de amplificação por PCR das sequências de glicoproteína P utilizando iniciadores selecionados das sequências mostradas na 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 ou SEQ ID NO:16.

A proteína recombinante pode ser purificada utilizando técnicas correntes bem conhecidas na arte. Tais técnicas podem incluir filtração, precipitação, cromatografia e métodos eletroforéticos. A pureza pode ser avaliada realizando eletroforese em gel de poliacrilamida e visualizando as proteínas utilizando metodologias de revelação corrente. Também se pode efetuar a transferência de Western utilizando um anticorpo contra a glicoproteína P do cromossoma 7p.

A invenção é também dirigida a anticorpos gerados contra a glicoproteína P do cromossoma 7p. O processo de produção de tais anticorpos pode envolver a injeção da própria glicoproteína P do 7p num animal apropriado ou a injeção de péptidos antigénicos curtos feitos para corresponder a diferentes regiões da proteína. Estes péptidos deveriam ter pelo menos 5 aminoácidos de comprimento e deveriam ser, preferencialmente, selecionados de regiões que se julga que sejam únicas à glicoproteína P do 7p. Os métodos para gerar e detetar anticorpos são bem conhecidos na arte, e são ensinados por referências 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); e Campbell, "Monoclonal Antibody Technology", in Laboratory Techniques in Biochemistry and Molecular Biology, (1984).

O termo "anticorpo", como aqui utilizado, tem a intenção de incluir moléculas intactas, assim como fragmentos que conservem a sua capacidade para ligar o antigénio, como os fragmentos Fab e F(ab')₂. O termo "anticorpo" também é aqui definido como se referindo a anticorpos monoclonais e anticorpos policlonais. Os anticorpos policlonais são derivados dos soros de animais imunizados com um antigénio da glicoproteína P do cromossoma 7p. Os anticorpos monoclonais contra a proteína podem ser preparados utilizando tecnologia de hibridoma, como ensinado por referências como: Kohler, *et al.*, Nature 256:495 (1975); e Hammerling, *et al.*, in: Monoclonal Antibodies and T-Cell Hybridomas, Elsevier, N.Y., pp. 563-681 (1981). De uma maneira geral, esta tecnologia envolve a imunização de um animal imunocompetente, tipicamente um ratinho, com

glicoproteína P do cromossoma 7p intacta ou com um fragmento derivado daquele. Os esplénócitos são então extraídos do animal imunizado e são fundidos com células de mieloma adequadas, como células SP₂O. Depois disso, as células de hibridoma resultantes são seletivamente mantidas em meio HAT e depois clonadas por diluição restringida (Wands, *et al.*, *Gastroenterology* 80:225-232 (1981)). As células obtidas através desta seleção são depois avaliadas para identificar clones que segregam anticorpos capazes de ligar a glicoproteína P do cromossoma 7p.

Os anticorpos ou fragmentos de anticorpos da invenção podem ser utilizados para detetar a presença de glicoproteína P do cromossoma 7p em qualquer um de uma variedade de imunoensaios. Por exemplo, os anticorpos podem ser utilizados em radioimunoensaios ou em ensaios imunométricos, também conhecidos como ensaios "dois sítios" ou "sanduíche" (ver Chard, "An Introduction to Radioimmune Assay and related Techniques," in: Laboratory Techniques in Biochemistry and Molecular Biology, North Holland Publishing Co., NY (1978)). Num ensaio imunométrico típico, uma quantidade de anticorpo não marcado é ligada a um suporte sólido que é insolúvel no fluido a ser testado, como sangue, linfa, extratos celulares e semelhantes. Após a ligação inicial de antígeno ao anticorpo imobilizado, é adicionada uma quantidade do segundo anticorpo marcado de modo detetável (o qual pode ser o mesmo ou não ser o mesmo que o primeiro) para permitir a deteção e/ou quantificação de antígeno ligado (ver, por exemplo Radioimmune Assay Method, Kirkham, *et al.*, Ed. pp. 199-206, E&S Livingstone, Edinburgh (1970)). Na técnica são conhecidas muitas variantes destes tipos de ensaios e podem ser utilizadas para a deteção de glicoproteína P de 7p.

Os anticorpos contra a glicoproteína P do cromossoma 7p também podem ser utilizados em procedimentos de purificação (ver, em termos gerais, Dean et al., Affinity Chromatography, A Practical Approach, IRL Press (1986)). Tipicamente, o anticorpo é imobilizado numa matriz cromatográfica tal como Sepharose, 4B. A matriz é depois empacotada numa coluna e a preparação contendo a glicoproteína P do cromossoma 7p é feita passar através daquela em condições que promovem a ligação, *por exemplo*, condições de baixo teor de sal. A coluna é então lavada, a proteína é eluída utilizando tampão que promove a dissociação do anticorpo, *por exemplo*, tampão possuindo um pH ou concentração de sal alterado. A proteína eluída pode ser transferida para um tampão, *por exemplo* via diálise, e subsequentemente conservada ou utilizada diretamente. Os anticorpos também podem ser utilizados na transferência de Western para a deteção de glicoproteína P do cromossoma 7p numa amostra. Para estes tipos de ensaio, pode utilizar-se anticorpo que foi especificamente desenvolvido para reagir com a glicoproteína P do cromossoma 7p ou que reage com um epítopo da proteína.

A deteção da glicoproteína P do cromossoma 7p pode ser utilizada para determinar se as células tumorais são resistentes a múltiplos fármacos. Analogamente, a deteção de alterações na expressão de glicoproteína P pode ser útil para prever o desenvolvimento de resistência a múltiplos fármacos em células. O ADNc desta glicoproteína P pode ser útil na conceção de iniciadores para PCR de diagnóstico, conceção de sondas para transferência de Northern de diagnóstico, ensaios de proteção de RNase e para a conceção de oligonucleótidos antimensageiros complementares ao ADNc previsto para utilização em estratégias de abordagem seletiva de genes para a inversão da resistência a múltiplos fármacos. São idealizadas utilizações de

diagnóstico e terapêuticas *in vitro* e *in vivo* para sequências de nucleótidos antimensageiros relativamente à glicoproteína P do cromossoma 7p.

A sequência primária de aminoácidos e a estrutura proteica da glicoproteína P do cromossoma 7p podem ser utilizadas na produção de anticorpos monoclonais (mAbs) que podem ser utilizados no diagnóstico e terapia de cancro resistente a múltiplos fármacos. Por exemplo, péptidos sintéticos semelhantes às sequências de aminoácidos nativas de domínios extracelulares particulares como determinado por previsão de tipologia de membrana podem ser úteis para desenvolver mAbs inibidores orientados contra epítomos extracelulares da glicoproteína P do cromossoma 7p. Além disso, sequências sintéticas peptídicas de 10-20meros derivadas da sequência de aminoácidos primária não incluídas nas sequências de ansa extracelulares supramencionadas podem ser úteis no desenvolvimento de anticorpos monoclonais de diagnóstico específicos. Podem ser utilizados mAbs específicos em análise de diagnóstico por FACS, transferência de Western e imuno-histoquímica. Tais mAbs também podem ser usados em utilizações de diagnóstico *in vivo*, onde podem ser empregues mAbs conjugados com etiquetas para avaliar a carga tumoral, localização do tumor ou massa tumoral residual após quimioterapia ou terapia cirúrgica de tumores que expressam glicoproteína P do 7p15-21.

Também podem ser utilizados mAbs específicos para fins terapêuticos em doentes com cancro. Em particular, eles podem ser administrados para inverter a resistência de cancros a múltiplos fármacos em doentes administrados com agentes quimioterapêuticos que são substratos para efluxo pela glicoproteína P do 7p, *por exemplo*, cisplatina. Além disso, mAbs específicos podem ser terapeuticamente

utilizados em doentes com cancro para provocar a morte celular específica de tumores, quer administrados numa forma não conjugada, resultando em morte tumoral mediada imunologicamente, quer numa forma conjugada com uma toxina celular (por exemplo conjugado com iodo radioativo ou com toxinas químicas), resultando na morte celular direta e específica de tumores.

Também podem ser utilizados mAbs específicos para fins terapêuticos diferentes da resistência de cancros a múltiplos fármacos. Com base na função imunorreguladora prevista da glicoproteína P do 7p, estes mAbs podem ser administrados a doentes para prevenir e/ou tratar a rejeição do transplante de órgãos e também várias doenças autoimunes como artrite reumatoide e esclerose múltipla. Além disso, uma vez que as glicoproteínas P funcionam na captação, excreção e distribuição específica em relação ao tecido de uma diversidade de compostos farmacológicos e químicos, e têm sido implicadas em mecanismos de biodisponibilidade oral, função da barreira hematoencefálica e em mecanismos de excreção renal, hepática e biliar de vários fármacos, os mAbs específicos podem ser administrados terapêuticamente para alterar a farmacocinética e disponibilidade daqueles fármacos terapêuticos os quais são substratos da função de transporte mediada pela glicoproteína P do 7p.

As composições e métodos da presente invenção podem ter um número de utilizações além das descritas acima. Por exemplo, sabe-se que as células estaminais pluripotentes e células progenitoras de tecido como as células estaminais hematopoiéticas, células neuroprogenitoras e células progenitoras de músculo possuem atividades de efluxo tipo glicoproteína P para moléculas pequenas e corantes fluorescentes. A glicoproteína P do cromossoma 7p pode

desempenhar um papel no transporte de tais substratos, e podem assim servir como um marcador para o isolamento de tais células estaminais e células progenitoras via, por exemplo, análise de FACS. Além disso, uma vez que a glicoproteína P MDR1 parece estar envolvida na diferenciação celular, proliferação celular, sobrevivência celular e determinadas respostas imunológicas, é esperado que a glicoproteína P do cromossoma 7p, devido à sua homologia com a glicoproteína P MDR 1, também desempenhe um papel em tais funções fisiológicas. Assim, o gene da glicoproteína P do cromossoma 7p e as sequências proteicas podem ser úteis na modulação das ruturas fisiopatológicas destas funções relacionadas com o MDR.

Exemplos

Uma vez que está presentemente a ser produzida informação nova de sequências genómicas a um ritmo elevado através do projeto do genoma humano, bases de dados contendo essa informação genómica contêm potencialmente sequências de membros da família de glicoproteína P não identificados até à data. Os membros da família da glicoproteína P mamífera partilham sequências de aminoácidos e epítomos de proteína característicos, e assumem conformações semelhantes. Assim, foi realizada uma pesquisa com base em homologia de proteína numa tentativa para identificar novos genes que codificam glicoproteína P. Foram utilizadas ferramentas bioinformáticas de análise de genes e análise de proteínas para caracterizar melhor a sequência de ácidos nucleicos e prever a estrutura proteica de genes candidatos identificados. Especificamente, foi utilizada a aplicação tblastn do National Center for Biotechnology Information (NCBI) para comparar as sequências de aminoácidos conservadas derivadas da estrutura conhecida da glicoproteína P MDR1 humana com a base de dados não

redundante de sequências de nucleótidos de *homo sapiens* da NCBI traduzida dinamicamente em todas as grelhas de leitura. A sequência assinatura comum aos membros da família de transportadores ABC, uma sequência de aminoácidos 15mero LSGGQKQRIAIARAL (SEQ ID NO:17), foi utilizada para identificar sequências de ADN genómico humano que codificam estruturas de proteínas homólogas. Também foram utilizadas sequências de aminoácidos hexaméricas conhecidas de três epítomos de ligação ao anticorpo monoclonal (mAb) específico contra a glicoproteína P.

Os clones de ADN genómico humano identificados do modo descrito acima foram examinados quanto à contaminação por vetores utilizando o programa VecScreen. Além disso, estes clones foram submetidos a mapeamento sistemático de homologia utilizando sequências de aminoácidos 20-meros contíguas sobrepostas derivadas da estrutura da proteína MDR1 humana e o programa de pesquisa tblastn. As sequências de ADN genómico candidatas que codificam sequências homólogas de aminoácidos foram comparadas em relação às sequências da grelha de leitura aberta (ORF) em cada clone de ADN utilizando o programa ORF Finder de NCBI (Altschul, *et al.*, *Nucleic Acids Res.* 25:3389-402 (1997)). As ORFs genómicas contendo sequências de ADN homólogas foram então analisadas utilizando o pacote de software NetGene2 para prever sítios intrão de excisão-união nos genes candidatos (Brunak *et al.*, *J. Mol. Biol.* 220:49-65 (1991)).

Uma sequência de ADNc foi gerada por transcrição linear conceptual de estruturas previstas de exão de ADN adjacentes. Utilizando esta abordagem foram identificados dois clones genómicos humanos adjacentes sobrepostos, CTA-367017 (AC002486, 79611 pares de bases de comprimento) e CTB-86D3 (AC005060, 120169 pares de bases de comprimento,

sequeenciados à direita) como fazendo parte de uma ilha desancorada de orientação desconhecida no cromossoma 7p15-21. Constatou-se que estes clones de sobreposição continham uma sequência genética que codificava um novo membro da família da glicoproteína P humana.

Para determinar se a estrutura prevista do gene era expressa em tecidos humanos, a sequência de ADNc gerada foi comparada com a base de dados não redundante de etiquetas de sequências humanas expressas (EST) dbest da NCBI como descrito por Altschul *et al*, e foram identificadas várias ESTs complementares a exões previstos do clone genómico AC002486. Foram então concebidos iniciadores da reação em cadeia da polimerase (PCR) com base na informação de sequência disponível na base de dados no National Center for Biotechnology Information (NCBI) e na análise bioinformática como descrito acima. Utilizando estes iniciadores oligonucleotídicos específicos para o gene e a técnica de PCR em ARN mensageiro total (ARNm) transcrito inversamente isolado de várias linhas de células humanas e tecidos humanos normais, incluindo a linha de células de melanoma humano G3361, a linha de células de cancro da mama MCF-7, a linha de células de carcinoma de células escamosas SCC25, a linha de células de leucemia U937 e células mononucleares de sangue periférico normal (PBMC), foram amplificadas as sequências de ADNc derivadas do novo gene 7p15-21 da glicoproteína P e os produtos de PCR foram subsequentemente sequeenciados utilizando o método dideoxilo de terminação de cadeia em ambas as cadeias.

A estrutura intrão-exão de vários produtos genéticos codificados pelo gene 7p15-21 da glicoproteína P foi determinada por comparação dos clones previstos e sequeenciados de ADNc com informação da sequência genómica do locus do gene 7p15-21 da glicoproteína P (clones

AC002486 e AC005060), como se mostra nas 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 e SEQ ID NO:16. As estruturas proteicas codificadas pelo novo gene 7p15-21 foram depois geradas por tradução conceptual de aminoácidos das sequências oligonucleotídicas previstas 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 e SEQ ID NO:16, como se mostra nas 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 e SEQ ID NO:8. Estas sequências de aminoácidos foram então comparadas com a sequência peptídica não redundante de NCBI em relação à homologia de sequência utilizando o programa blastp da NCBI. As sequências de aminoácidos previstas 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 e SEQ ID NO:8. também foram classificadas utilizando o Sistema Internacional de Classificação de Família de Proteínas PIR (Barker, et al., *Nucleic Acids Res.* 28:41-4 (2000); Huang et al., *Nucleic Acids Res.* 28:273-6 (2000)). As características funcionais potenciais das proteínas previstas foram determinadas por análise comparativa da composição primária de aminoácidos e utilizando o pacote de software TMHMM1.0 para prever a formação de hélices transmembranares em proteínas mamíferas (Sonnhammer et al., *Ismb* 6:175-82 (1998)).

O novo gene 7p15-21 da glicoproteína P pode codificar várias isoformas distintas da glicoproteína P as quais apresentam 68% de homologia de sequência com os MDR1 e MDR3 humanos. Um grau semelhante de homologia foi encontrado com as isoformas de ratinho e hamster destes genes humanos. A análise da sequência primária de aminoácidos sugere que a glicoproteína P do cromossoma 7p15-21 pode expressar o epítipo C32 e o epítipo de ligação ao mAB anti-glicoproteína P, mas não o epítipo C219 conservado em todas

as outras isoformas conhecidas de glicoproteína P (Georges, *et al.*, *Proc. Natl Acad Sci USA* 87:152-6 (1990)).

A previsão estrutural revelou que o gene 7p15-21 da glicoproteína P codifica isoformas de glicoproteína P que apresentam semelhanças estruturais mas também diferenças distintivas em comparação com membros conhecidos da família da glicoproteína P, como descrito por Georges *et al.* Por exemplo, a proteína de SEQ ID NO:2 contém dois domínios de ligação a ATP que estão localizados em lados opostos da membrana do plasma, proporcionando um domínio extracelular de ligação a ATP único que se prevê que ligue ATP extracelular. Com base nestas diferenças distintivas, prevê-se que a glicoproteína P do 7p15-21 esteja não só envolvida no efluxo de moléculas pequenas, mas que algumas das suas isoformas também são funcionais na captação dependente de energia de moléculas pequenas. O sistema de classificação PIR confirmou que a glicoproteína P do cromossoma 7p15-21 descoberta era um membro da família de proteínas de resistência a múltiplos fármacos e da família de superfamílias de homologia da cassete de ligação a ATP.

A análise por PCR utilizando iniciadores específicos para o gene demonstrou que o ADNc que codifica as proteínas de SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5 e SEQ ID NO:6, o qual em cada caso envolve exões codificados no clone genómico AC005060, era preferencialmente expresso em células de melanoma humano mas não na maioria dos outros cancros testados, ao contrário dos ADNcs que codificam as proteínas de SEQ ID NO:7 e SEQ ID NO:8, os quais se constatou que eram expressos na maioria dos cancros examinados e também em tecidos fisiológicos humanos. Isto realça que um subconjunto de produtos do gene da glicoproteína P do 7p15-21 pode ser seletivamente visado em

determinados cancros que apresentam graus particularmente elevados de quimiorresistência, como o melanoma humano.

Para avaliar a expressão e função da glicoproteína P do 7p15-21 e o efeito da modulação específica na função de transporte e quimiorresistência, gerou-se anticorpos policlonais contra os péptidos MDR CGTSLILNGEPGYTI (SEQ ID NO:18) e RFGAYLIQAGRMTPEGC (SEQ ID NO:19), correspondendo a epítomos distintos da ansa extracelular da glicoproteína P do 7p15-21, injetando ratinhos com estes péptidos antigénicos conjugados com a substância veículo KLH. Para avaliar a expressão superficial da glicoproteína P do 7p15-21 em células tumorais humanas foi realizada imunorrevelação indireta da superfície e citometria de fluxo monocromática de células colhidas de fresco. Para avaliar os efeitos da inibição da gp-P do 7p15-21 no efluxo de corante fluorescente mediado por gp-P, células tumorais foram incubadas com Ab policlonal anti-glicoproteína P do 7p15-21 seguido da adição de calceína-AM e medições subsequentes em série da fluorescência celular por citometria de fluxo.

Estes estudos demonstraram que a glicoproteína P é expressa em células tumorais, e que o epítomo RFGAYLIQAGRMTPEGC (SEQ ID NO:19) contido nas proteínas de SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5 e SEQ ID NO:6, é preferencialmente expresso em melanoma humano em níveis elevados, enquanto o epítomo CGTSLILNGEPGYTI (SEQ ID NO:18), também contido na SEQ ID NO:7 e SEQ ID NO:8, também é expresso noutros tipos de cancro e em células humanas normais. Os anticorpos contra o epítomo CGTSLILNGEPGYTI (SEQ ID NO:18) inibiram tanto a captação de corante como também o efluxo de corante em função do tipo de célula, indicando uma função dupla dos vários produtos do gene da glicoproteína P do 7p15-21 nestes processos distintos.

Estes anticorpos também aumentam a citotoxicidade celular da cisplatina em ensaios de morte celular específica em melanoma e também em cancro da mama entre outros, indicativo da sua utilidade terapêutica potencial no tratamento de doentes com cancro.

Sabe-se que determinados tipos de cancro exibem rearranjo cromossómico na região 7p15-21 e tais mutações podem estar associadas com o aparecimento do fenótipo MDR. Isto levanta a possibilidade de que o rearranjo de genes nestes cancros resulta potencialmente de formação de epissoma e cromossoma minúsculo duplo (DM) durante o processo de amplificação de gene da glicoproteína P do 7p15-21 sob stresses mutagénicos como a quimioterapia. Sabe-se que as células que expressam resistência a múltiplos fármacos mediada por MDR1 sofrem esses tipos de rearranjos cromossómicos e formação de cromossoma 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)). Assim, os produtos do gene da glicoproteína P do cromossoma 7p15-21 de SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:5 e SEQ ID NO:6 podem estar seletivamente sobre-expressos em determinadas células cancerosas, contribuindo desse modo para a resistência adquirida ao fármaco dessas células cancerosas enquanto permanecem silenciosos em células normais. Este padrão de expressão diferencial pode ser utilizado na deteção e inversão da resistência a múltiplos fármacos de células mamíferas tumorigénicas.

1. Uma proteína substancialmente pura que consiste essencialmente de uma sequência de aminoácidos selecionada do grupo que consiste em: 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; e SEQ ID NO:8.

2. Um péptido que consiste de um elemento de sequência derivado de qualquer um 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; ou SEQ ID NO:8; em que o referido péptido tem pelo menos 10 resíduos de comprimento, e com a condição de que a referida sequência não seja a mesma que quaisquer 10-15 aminoácidos contíguos da sequência LSGGQKQRIAIARAL.

3. Um anticorpo preparado por um processo que compreende o passo de administração da proteína ou péptido de qualquer uma das reivindicações 1-2 a um animal capaz de produzir o referido anticorpo, em que a referida proteína ou péptido é administrado a uma dosagem suficiente para induzir a formação do anticorpo no referido animal.

4. Um anticorpo que se liga preferencialmente à proteína da reivindicação 1.

5. Um polinucleótido substancialmente puro que codifica a proteína ou péptido da reivindicação 1 ou reivindicação 2.

6. Um vetor para expressar uma glicoproteína P, compreendendo um elemento de codificação distinto que consiste no polinucleótido da reivindicação 5.

7. Uma célula hospedeira transformada com o vetor da reivindicação 6.

8. Uma glicoproteína P recombinante produzida pela célula hospedeira da reivindicação 7.

9. O polinucleótido da reivindicação 5, em que o referido polinucleótido tem uma sequência de nucleótidos selecionada do grupo que consiste essencialmente em: 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; e SEQ ID NO:16.

10. Um oligonucleótido que atua como um inibidor antimensageiro da expressão da glicoproteína P, em que o referido oligonucleótido tem pelo menos 15 nucleótidos de comprimento e consistem numa sequência complementar a pelo menos 15 nucleótidos contíguos em qualquer uma 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; ou SEQ ID NO:16.

11. Um vetor para expressar proteína, compreendendo um elementos de codificação distinto que consiste no polinucleótido da reivindicação 9.

12. Uma célula hospedeira transformada com o vetor da reivindicação 11.

13. Um método para determinar se uma célula cancerosa responderá a uma terapia com o objetivo de inverter a resistência a múltiplos fármacos, compreendendo o passo de medição da expressão de um gene que codifica uma proteína selecionada do grupo que consiste em: 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; e SEQ ID NO:8.

14. O método da reivindicação 13, em que a referida expressão é determinada utilizando amplificação por PCR de ARNm transcrito de forma inversa.

15. O método da reivindicação 13, em que a referida expressão é determinada utilizando o anticorpo da reivindicação 4.

16. Um método para determinar se um composto de ensaio inibe a resistência a múltiplos fármacos provocada por um gene que codifica uma proteína selecionada do grupo que consiste em: 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; e SEQ ID NO:8; compreendendo o referido método:

- (a) expressar o referido gene em células que, de outro modo, não são resistentes a múltiplos fármacos;
- (b) expor as referidas células a um ou mais agentes citotóxicos na presença do referido composto de ensaio;
- (c) medir a sobrevivência celular depois de expor as células aos referidos um ou mais agentes citotóxicos e comparar os resultados obtidos no passo (b) com aqueles de células incubadas essencialmente do mesmo modo com os referidos agentes citotóxicos mas na ausência do referido composto de ensaio; e
- (d) concluir que o referido composto de ensaio inibe a resistência a múltiplos fármacos se a sobrevivência celular diminuir numa medida significativa por incubação das células na presença do referido composto de ensaio relativamente à sobrevivência celular em incubações realizadas na ausência do referido composto de ensaio.

17. O método da reivindicação 16, em que o referido gene consiste essencialmente numa sequência de nucleótidos do grupo que consiste em: 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; ou SEQ ID NO:16.

LISTAGEM DE SEQUÊNCIAS

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

<120> Um Gene que Codifica um Homólogo de Glicoproteína P Humana de Resistência a Múltiplos Fármacos no Cromossoma 7p15-21 e Suas Utilizações

<130> 81994/268611

<160> 19

<170> PatentIn versão 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
 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

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
 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

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
 725 730 735

Ala Thr Ser Ala Leu Asp Asn Asp Ser Glu Lys Val Val Gln His Ala
740 745 750
Leu Asp Lys Ala Arg Thr Gly Arg Thr Cys Leu Val Val Thr His Arg
755 760 765
Leu Ser Ala Ile Gln Asn Ala Asp Leu Ile Val Val Leu His Asn Gly
770 775 780
Lys Ile Lys Glu Gln Gly Thr His Gln Glu Leu Leu Arg Asn Arg Asp
785 790 795 800
Ile Tyr Phe Lys Leu Val Asn Ala Gln Ser Val Gln
805 810

<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> Nota

<222> (66)..(66)

<223> Xaa na posição 66 representa qualquer L-aminoácido

<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
          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
 305 310 315 320
 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
 385 390 395 400
 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
 465 470 475 480
 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
 545 550 555 560
 Ala Gly Glu Ile Leu Thr Met Arg Leu Arg His Leu Ala Phe Lys Ala
 565 570 575
 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
 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> Nota

<222> (230).. (230)

<223> Xaa na posição 230 representa qualquer 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

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
 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
 690 695 700
 Val Ile Leu Gly Val Ile Cys Phe Val Ser Tyr Phe Met Gln Gly Leu
 705 710 715 720
 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
 785 790 795 800
 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
 865 870 875 880
 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
 945 950 955 960
 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
 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
 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 Tyr Thr
 195 200 205
 Gln Asn Leu Lys Asp Ala Lys Asp Phe Gly Ile Lys Arg Thr Ile Ala
 210 215 220
 Ser Lys Val Ser Leu Gly Ala Val Tyr Phe Phe Met Asn Gly Thr Tyr
 225 230 235 240
 Gly Leu Ala Phe Trp Tyr Gly Thr Ser Leu Ile Leu Asn Gly Glu Pro
 245 250 255
 Gly Tyr Thr Ile Gly Thr Val Leu Ala Val Phe Phe Ser Val Ile His
 260 265 270
 Ser Ser Tyr Cys Ile Gly Ala Ala Val Pro His Phe Glu Thr Phe Ala
 275 280 285
 Ile Ala Arg Gly Ala Ala Phe His Ile Phe Gln Val Ile Asp Lys Lys
 290 295 300
 Pro Ser Ile Asp Asn Phe Ser Thr Ala Gly Tyr Lys Pro Glu Ser Ile
 305 310 315 320
 Glu Gly Thr Val Glu Phe Lys Asn Val Ser Phe Asn Tyr Pro Ser Arg
 325 330 335
 Pro Ser Ile Lys Ile Leu Lys Gly Leu Asn Leu Arg Ile Lys Ser Gly
 340 345 350
 Glu Thr Val Ala Leu Val Gly Leu Asn Gly Ser Gly Lys Ser Thr Val
 355 360 365
 Val Gln Leu Leu Gln Arg Leu Tyr Asp Pro Asp Asp Gly Phe Ile Met
 370 375 380
 Val Asp Glu Asn Asp Ile Arg Ala Leu Asn Val Arg His Tyr Arg Asp
 385 390 395 400
 His Ile Gly Val Val Ser Gln Glu Pro Val Leu Phe Gly Thr Thr Ile
 405 410 415
 Ser Asn Asn Ile Lys Tyr Gly Arg Asp Asp Val Thr Asp Glu Glu Met
 420 425 430
 Glu Arg Ala Ala Arg Glu Ala Asn Ala Tyr Asp Phe Ile Met Glu Phe
 435 440 445

Pro Asn Lys Phe Asn Thr Leu Val Gly Glu Lys Gly Ala Gln Met Ser
 450 455 460
 Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu Val Arg Asn
 465 470 475 480
 Pro Lys Ile Leu Ile Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser Glu
 485 490 495
 Ser Lys Ser Ala Val Gln Ala Ala Leu Glu Lys Ala Ser Lys Gly Arg
 500 505 510
 Thr Thr Ile Val Val Ala His Arg Leu Ser Thr Ile Arg Ser Ala Asp
 515 520 525
 Leu Ile Val Thr Leu Lys Asp Gly Met Leu Ala Glu Lys Gly Ala His
 530 535 540
 Ala Glu Leu Met Ala Lys Arg Gly Leu Tyr Tyr Ser Leu Val Met Ser
 545 550 555 560
 Gln Asp Ile Lys Lys Ala Asp Glu Gln Met Glu Ser Met Thr Tyr Ser
 565 570 575
 Thr Glu Arg Lys Thr Asn Ser Leu Pro Leu His Ser Val Lys Ser Ile
 580 585 590
 Lys Ser Asp Phe Ile Asp Lys Ala Glu Glu Ser Thr Gln Ser Lys Glu
 595 600 605
 Ile Ser Leu Pro Glu Val Ser Leu Leu Lys Ile Leu Lys Leu Asn Lys
 610 615 620
 Pro Glu Trp Pro Phe Val Val Leu Gly Thr Leu Ala Ser Val Leu Asn
 625 630 635 640
 Gly Thr Val His Pro Val Phe Ser Ile Ile Phe Ala Lys Ile Ile Thr
 645 650 655
 Met Phe Gly Asn Asn Asp Lys Thr Thr Leu Lys His Asp Ala Glu Ile
 660 665 670
 Tyr Ser Met Ile Phe Val Ile Leu Gly Val Ile Cys Phe Val Ser Tyr
 675 680 685
 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
 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> Nota

<222> (230)..(230)

<223> Xaa na posição 230 representa qualquer L-aminoácido

<400> 7

```

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
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
 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
 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 Tyr Thr
 195 200 205
 Gln Asn Leu Lys Asp Ala Lys Asp Phe Gly Ile Lys Arg Thr Ile Ala
 210 215 220
 Ser Lys Val Ser Leu Gly Ala Val Tyr Phe Phe Met Asn Gly Thr Tyr
 225 230 235 240
 Gly Leu Ala Phe Trp Tyr Gly Thr Ser Leu Ile Leu Asn Gly Glu Pro
 245 250 255
 Gly Tyr Thr Ile Gly Thr Val Leu Ala Val Phe Phe Ser Val Ile His
 260 265 270
 Ser Ser Tyr Cys Ile Gly Ala Ala Val Pro His Phe Glu Thr Phe Ala
 275 280 285
 Ile Ala Arg Gly Ala Ala Phe His Ile Phe Gln Val Ile Asp Lys Lys
 290 295 300
 Pro Ser Ile Asp Asn Phe Ser Thr Ala Gly Tyr Lys Pro Glu Ser Ile
 305 310 315 320
 Glu Gly Thr Val Glu Phe Lys Asn Val Ser Phe Asn Tyr Pro Ser Arg
 325 330 335
 Pro Ser Ile Lys Ile Leu Lys Gly Leu Asn Leu Arg Ile Lys Ser Gly
 340 345 350
 Glu Thr Val Ala Leu Val Gly Leu Asn Gly Ser Gly Lys Ser Thr Val
 355 360 365
 Val Gln Leu Leu Gln Arg Leu Tyr Asp Pro Asp Asp Gly Phe Ile Met
 370 375 380
 Val Asp Glu Asn Asp Ile Arg Ala Leu Asn Val Arg His Tyr Arg Asp
 385 390 395 400
 His Ile Gly Val Val Ser Gln Glu Pro Val Leu Phe Gly Thr Thr Ile
 405 410 415
 Ser Asn Asn Ile Lys Tyr Gly Arg Asp Asp Val Thr Asp Glu Glu Met
 420 425 430
 Glu Arg Ala Ala Arg Glu Ala Asn Ala Tyr Asp Phe Ile Met Glu Phe
 435 440 445
 Pro Asn Lys Phe Asn Thr Leu Val Gly Glu Lys Gly Ala Gln Met Ser
 450 455 460
 Gly Gly Gln Lys Gln Arg Ile Ala Ile Ala Arg Ala Leu Val Arg Asn
 465 470 475 480
 Pro Lys Ile Leu Ile Leu Asp Glu Ala Thr Ser Ala Leu Asp Ser Glu
 485 490 495
 Ser Lys Ser Ala Val Gln Ala Ala Leu Glu Lys Asp Thr Pro Arg Tyr
 500 505 510
 Ser Phe

<211> 2066

<212> ADN

<213> Homo sapiens

<400> 9

```

cgagcaaagg tcggactaca atcgtggttag cacaccgaet ttctactatt cgaagtgcag      60
atctgattgt gaccctaaag gatggaatgc tggcggagaa aggagcacat gctgaactaa      120
tggcaaaaacg aggtctatat tattcacttg tgatgtcaca ggatattaaa aaagctgatg      180
aacagatgga gtcaatgaca tattctactg aaagaaagac caactcactt cctctgcact      240
ctgtgaagag catcaagtca gacttcattg acaaggctga ggaatccacc caatctaaag      300
agataaagtct tctgaagtc tctctattaa aaattttaaa gttaaaacaag cctgaatggc      360
cttttgtggt tctggggaca ttggcttcctg ttctaaatgg aactgttcat ccagtatttt      420
ccatcatctt tgcaaaaatt ataacratgt ttggaaataa tgataaaaacc acattaaagc      480
atgatgcaga aatttatctc atgatattcg tcatcttggg tgttatttgc tttgtcagtt      540
atttcattga gggattattt tacggcagag caggggaaat ttaacgatg agattaagac      600
acttggcctt caaagccatg ttatatcagg atattgcctg gtttgatgaa aaggaaaaca      660
gcacaggagg ctigacaaca atattagcca tagatatagc acaaatccaa ggagcaacag      720
gttcagggat tggcgtctta acacaaaatg caactaacat gggactttca gttatcattt      780
cctttatata tggatgggag atgacattcc tgattctgag tattgctcca gtacttgccg      840
tgacagggaat gattgaaacc gcagcaatga ctggatttgc caacaagat aagcaagaac      900
ttaagcatgc tggaaagata gcaactgaag ctttggagaa tatacgtact atagtgtcat      960
taacaaggga aaaagccttc gagcaaatgt atgaagagat gcttcagact caacacagaa     1020
atacctogaa gaaagcacag attattggaa gctgttatgc attcagccat gcctttatat     1080
atcttgctta tgcagcaggg ttccgatttg gagcctatct aattcaagct ggacgaatga     1140
ccccagaggg catgttcata gtttttactg caattgcata tggagctatg gccatcggaa     1200
aaacgctcgt tttggctcct gaatatccca aagccaaatc gggggctgcg catctgtttg     1260
ccttggttga aaagaaacca aatatagaca gcgcagtcga agaagggaaa aagccagaca     1320
catgtgaagg gaatttagag ttccgagaag tctctttctt ctatccatgt ccccagatg     1380
tttcatcctt ccttggtcta tccctcagta ttgagcgagg aaagacagta gcatttgttg     1440
ggagcagcgg ctgtgggaaa agcactcttg tcaacttct gcagagactt tatgaccccg     1500
tgcaaggaca agtgcgtgtt gatggtgtgg atgcaaaaga attgaatgta cagtggctcc     1560
gttcccaaat agcaatcgtt cctcaagagc ctgtgctctt caactgcagc attgctgaga     1620

```

```

acatcgcta tggtgacaac agccgtgtgg tgccattaga tgagatcaaa gaagcogcaa 1680
atgcagcaaa tatccattct tttattgaag gtctccctga gaaatacaac acacaagttg 1740
gactgaaaag agcacagctt tctggcggee agaaacaaaag actagctatt gcaagggtct 1800
ttctccaaaa acccaaaatt ttattgttgg atgaggccac ttcagccctc gataatgaca 1860
gtgagaaggt gggtcagcat gcccttgata aagccaggac ggggaaggaca tgccctagtgg 1920
tcactcacag gctctctgca attcagaacg cagatttgat agtggttctg cacaatggaa 1980
agataaagga acaaggaact catcaagagc tcctgagaaa tcgagacata tattttaagt 2040
tagtgaatgc acagtcagtg cagtga 2066

```

<210> 10

<211> 2856

<212> ADN

<213> Homo sapiens

<400> 10

```

cctaattcct ctaatatctc tctgtgagcc taaccaata attatatatt acattctatt 60
gtctttctta tataactgca gaaagataaa taccactttg tttgttcttg taggttttct 120
ttagtgtaat ccatagcagt tattgcattg gacgagcagt ccttcattat tgataagaaa 180
ccagtatag ataacttttc cacagctgga tataaacctg aatccataga aggaactgtg 240
gaatttaaaa atgtttcttt caattatcca tcaagaccat ctatcaagat tctgaaaggt 300
ctgaatctca gaattaagtc tggagagaca gtgcctttg tgggtctcaa tggcagtgag 360
aagagtaagg tagtccagct tctgcagagg ttatatgac cggatgatgg ctttatcatg 420
gtggatgaga atgacatcag agctttaaat gtgcggcatt atcgagacca tattggagtg 480
gttagtcaag agcctgtttt gtgcgggacc accatcagta acaatatcaa gtatggacga 540
gatgatgtga ctgatgaaga gatggagaga gcagcaaggg aagcaaatgc gtatgatatt 600
atcatggagt ttctaataa atttaataca ttggtagggg aaaaaggagc tcaaatgagt 660
ggagggcaga aacagaggat cgcaattgct cgtgccttag ttcgaaacc caagattctg 720
attttagatg aggetacgtc tgccctggat tcagaaagca agtcagctgt tcaagctgca 780
ctggagaagg cgagcaaaag tcggactaca atcgtgtag cacaccgact ttctactatt 840
cgaagtgcag atttgattgt gaccctaaag gatggaatgc tggcggagaa aggagcacat 900
gctgaactaa tggcaaaacg aggtctatat tattcacttg tgatgtcaca ggatattaaa 960
aaagctgatg aacagatgga gtcaatgaca tattctactg aaagaaagac caactcactt 1020
cctctgcact ctgtgaagag catcaagtcx gacttcattg acaaggctga ggaatccacc 1080

```

```

caatctaaag agataagtc tctgaagtc tctctattaa aaatttttaa gttaaacaag 1140
cctgaatggc cttttgtggt tctggggaca ttggcttctg ttctaatgg aactgttcat 1200
ccagtatttt ccateatctt tgcaaaaatt ataaccatgt ttggaaataa tgataaaacc 1260
acattaaagc atgatgcaga aatttattcc atgatattcg tcattttggg tgttatttgc 1320
tttgtcagtt atttcatgca gggattatct tacggcagag caggggaaat tttaacgatg 1380
agattaagac acttggcctt caaagccatg ttatcagag atattgcctg gtttgatgaa 1440
aaggaaaaca gcacaggagg cttgacaaca atattagcca tagatatagc acaaattcaa 1500
ggagcaacag gttccaggat tggcgtctta acacaaaatg caactaacat gggactttca 1560
gttatcattt cctttatata tggatgggag atgacattcc tgattctgag tattgtctca 1620
gtacttgccg tgacaggaat gattgaaacc gcagcaatga ctggatttgc caacaaagat 1680
aagcaagaac ttaagcatgc tggaaagata gcaactgaag ctttgagaa tatacgtact 1740
atagtgtcat taacaaggga aaaagccttc gagcaaatgt atgaagagat gcttcagact 1800
caacacagaa atacctgaa gaaagcacag attattggaa gctgttatgc attcagccat 1860
gcctttatat attttgcta tgcagcaggg ttctgatttg gagcctatct aattcaagct 1920
ggacgaatga ccccagaggg catgttcata gtttttactg caattgcata tggagctatg 1980
gccatcggaa aaacgctcgt tttggctcct gaatattcca aagccaaatc gggggctgcg 2040
catctgtttg ccttgttggg aaagaaacca aatatagaca gccgcagtca agaagggaaa 2100
aagccagaca catgtgaagg gaatttagag ttctgagaag tctctttctt ctatccatgt 2160
cgcccagatg ttttcatcct ccgtggctta tccctcagta ttgagcgagg aaagacagta 2220
gcatttgtgg ggagcagcgg ctgtgggaaa agcacttctg ttcaacttct gcagagactt 2280
tatgaccccg tgcaaggaca agtgctgttt gatgggtgtg atgcaaaaga attgaatgta 2340
cagtggctcc gttcccaaatt agcaatcgtt cctcaagagc ctgtgctctt caactgcagc 2400
attgctgaga acatcgcta tggtgacaac agccgtgtgg tgccattaga tgagatcaaa 2460
gaagccgcaa atgcagcaaa tatccattct ttattgaag gtctccctga gaaatacaac 2520
acacaagttg gactgaaagg agcacagctt tctggcggcc agaaacaag actagctatt 2580
gcaagggctc ttctccaaaa acccaaaatt ttattgttgg atgaggccac ttcagccctc 2640
gataatgaca gtgagaaggg ggttcagcat gcccttgata aagccaggac ggggaaggaca 2700
tgcttagtgg tcaactcacag gctctctgca attcagaacg cagatttgat agtgggtctg 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 tatcactttg tttgttcctg taggttttct      120
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 tggcagtggt      360
aagagtacgg tagtccagct tctgcagagg ttatatgacg cggatgatgg ctttatcatg      420
gtggatgaga atgacatcag agctttaaat gtgcggcatt atcgagacca tattggagtg      480
gttagtcaag agcctgtttt gttcgggacc accatcagta acaatatcaa gtatggacga      540
gatgatgtga ctgatgaaga gatggagaga gcagcaaggg aagcaaatgc gtatgatttt      600
atcatggagt ttcttaataa atttaataca ttggtagggg aaaaaggagc tcaaatgagt      660
ggagggcaga aacagaggat cgcaattgct cgtgccttag ttcgaaacc caagattctg      720
attttagatg aggctacgtc tgccttggat tcagaaagca agtcagctgt tcaagctgca      780
ctggagaagg atacccccag gtattcattt tgacctaat tcacctcaag tggagaatcg      840
ctgacctga accagcgccc ttcgacagct ctggcccctc aaacctcacc ctgacctcct      900
gctgcctatg agctactgca catacctcaa ggccatatgc agttgtggcc ctgcacccaa      960
ttacactgaa tctaggaggg gagttggcag tggcgggtatg aaaaaccatt gaacagtttt     1020
ctcgatggcc tgactccctt ataaaccaga gccttcagac cccttacaag gcttaatggc     1080
acattttact ttgcatttgc ttggaagtga gttaagcggt tttttttctc taagaaaatc     1140
gcaggettct ttttttaaaa tgetgacttt atgga                                  1175

```

<210> 12

<211> 3177

<212> ADN

<213> Homo sapiens

<220>

<221> Nota

<222> (198)..(198)

<223> n na posição 198 representa qualquer nucleótido (A, T, C ou G)

<400> 12

```

atggtcacatc cattgaccag taaggaatta agtgccatatt ccaaagctgg ggctgtggca      60
gaagaagtct tgtcatcaat ccgaacagtc atagcccttta gggcccagga gaaagaactt      120
caaaggtctt tcctttttaa tataacaaga tatgcttggt tttattttcc ccagtggcta      180
ctaagttgtg ttctgttntt tgtaaggtat acacagaatc tcaaagatgc aaaggatttt      240
ggcataaaaa ggactatagc ttcaaaagtg tctcttggtg ctgtgtactt ctttatgaat      300
ggaacctatg gaacttgcttt ttggtatgga acctccttga ttcttaatgg agaacctgga      360
tataccatcg ggactgttct tgctgttttc tttagtgtaa tccatagcag ttattgcatt      420
ggagcagcag tccctcactt tgaaaccttc gcaatagccc gaggagctgc ctttcatatt      480
ttccagggtta ttgataagaa acccagtata gataactttt ccacagctgg atataaacct      540
gaatccatag aaggaaactgt ggaattttaa aatgtttctt tcaattatcc atcaagacca      600
tctatcaaga ttctgaaagg totgaatctc agaattaagt ctggagagac agtcgccttg      660
gtcgggtctca atggcagtgga gaagagtacg gtagtcacgc ttctgcagag gttatatgat      720
ccggatgatg gctttatcat ggtggatgag aatgacatca gagcttttaa tgtgcggcat      780
tatcgagacc atattggagt ggttagtcaa gagcctgttt tgttcgggac caccatcagt      840
aacaatatca agtatggacg agatgatgtg actgatgaag agatggagag agcagcaagg      900
gaagcaaatg cgtatgattt tatcatggag ttctctaata aatttaatac attggtaggg      960
gaaaaaggag ctcaaatgag tggagggcag aaacagagga tcgcaattgc tegtgcctta     1020
gttcgaaaac ccaagattct gatttttagat gaggctacgt ctgccctgga ttcagaaagc     1080
aagtcagctg ttcaagctgc actggagaag gcgagcaaag gtcggactac aatcgtggta     1140
gcacaccgac tttctactat tgaagtga gatttgattg tgaccctaaa ggatggaatg     1200
ctggcggaga aaggagcaca tgctgaacta atggcaaaac gaggtctata ttattcactt     1260
gtgatgtcac aggatattaa aaaagctgat gaacagatgg agtcaatgac atattctact     1320
gaaagaaaga ccaactcact tcctctgcac tctgtgaaga gcatcaagtc agacttcatt     1380
gacaaggctg aggaatccac ccaatctaaa gagataagtc ttctgaagt ctctctatta     1440
aaaattttta agttaaacaa gcttgaatgg ccttttgtgg ttctggggac attggcttct     1500
gttctaaatg gaactgttca tccagtattt tccatcatct ttgcaaaaat tataaccatg     1560
tttggaaata atgataaaac cacattaaag catgatgcag aaatttatct catgatattc     1620
gtcatttttg gtgttatttg ctttgtcagt tatttcatgc agggattatt ttacggcaga     1680
gcaggggaaa ttttaacgat gagattaaga cacttggcct tcaaagccat gttatatcag     1740
gatattgctt ggtttgatga aaaggaaaaac 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 ctttaagcatg ctggaaagat agcaactgaa 2040
gctttggaga atatacgtac tatagtgtca ttaacaaggg aaaaagcctt cgagcaaattg 2100
tatgaagaga tgcttcagac tcaacacaga aatacctcga agaaagcaca gattattgga 2160
agctgttatg cattcagcca tgcctttata tattttgcct atgcagcagg gtttcgattt 2220
ggagcctatt taattcaagc tggacgaatg accccagagg gcatgttcat agtttttact 2280
gcaattgcat atggagctat ggccatcggg aaaacgctcg ttttggctcc tgaatattcc 2340
aaagccaaat cgggggctgc gcatctgttt gccttgttgg aaaagaaacc aaatatagac 2400
agccgcagtc aagaaggga aaagccagac acatgtgaag ggaatttaga gtttcgagaa 2460
gtctctttct tctatccatg tgcgccagat gttttcatcc tccgtggctt atccctcagt 2520
attgagcgag gaaagacagt agcatttgtg gggagcagcg gctgtgggaa aagcacttct 2580
gttcaacttc tgcagagact ttatgacccc gtgcaaggac aagtgtgtt tgatgggtgtg 2640
gatgcaaaaag aattgaatgt acagtggctc cgttcccaa tagcaatcgt tcctcaagag 2700
cctgtgtctt tcaactgcag cattgctgag aacatcgct atggtgacaa cagccgtgtg 2760
gtgccattag atgagatcaa agaagccgca aatgcagcaa atatccattc ttttattgaa 2820
ggctctccctg agaaatacaa cacacaagtt ggactgaaag gagcacagct ttctggcggc 2880
cagaaacaaa gactagctat tgcaagggtc cttctccaaa aacccaaaat tttattgttg 2940
gatgaggcca cttcagccct cgataatgac agtgagaagg tggttcagca tgcccttgat 3000
aaagccagga cgggaaggac atgcctagtg gtcactcaca 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> Nota

<222> (723)..(723)

<223> n na posição 723 representa qualquer nucleótido (A, T, C ou G)

<400> 13

ttccgctttg	ctgatggact	ggacatcaca	ctcatgatcc	tgggtatact	ggcatcactg	60
gtcaatggag	cctgccttcc	tttaatgcc	ctgggttttag	gagaaatgag	tgataacctt	120
attagtggat	gtctagtcca	aactaacaca	tactctttct	tcaggttgac	cctgtattat	180
gttggaatag	gtgttgctgc	cttgattttt	ggttacatac	agatttcctt	gtggattata	240
actgcagcac	gacagaccaa	gaggattcga	aaacagtttt	ttcattcagt	tttggcacag	300
gacatcggtt	ggtttgatag	ctgtgacatc	ggtgaaacta	acactcgcat	gacagacatt	360
gacaaaatca	gtgatgggat	tggagataag	attgctctgt	tgtttcaaaa	catgtctact	420
ttttcgattg	gcttggcagt	tgggttggtg	aagggctgga	aactcaccct	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	tttccccagt	ggctactaag	ttgtgttctg	720
ttntttgtaa	ggtatacaca	gaatctcaaa	gatgcaaagg	attttggcat	aaaaaggact	780
atagcttcaa	aagtgtctct	tgggtgctgt	tacttcttta	tgaatggaac	ctatggactt	840
gctttttggt	atggaacctc	cttgattctt	aatggagaac	ctggatatac	catcgggact	900
gttcttgctg	ttttctttag	tgtaatccat	agcagttatt	gcattggagc	agcagtcctt	960
cactttgaaa	ccttcgcaat	agcccgagga	gctgcctttc	atattttcca	ggttattgat	1020
aagaaaccca	gtatagataa	cttttcacac	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	ttaaatgtgc	ggcatfctcg	agaccatatt	1320
ggagtggtta	gtcaagagcc	tgttttggtc	gggaccacca	tcagtaacaa	tatcaagtat	1380
ggacgagatg	atgtgactga	tgaagagatg	gagagagcag	caaggggaagc	aatgcgtat	1440
gattttatca	tggagtttcc	taataaat	aatacattgg	taggggaaaa	aggagctcaa	1500
atgagtggag	ggcagaaaca	gaggatcgca	attgctcgtg	ccttagttcg	aaacccaag	1560
attctgattt	tagatgaggc	tacgtctgcc	ctggattcag	aaagcaagtc	agctgttcaa	1620
gctgcactgg	agaaggcgag	caaaggctcg	actacaatcg	tggtagcaca	ccgactttct	1680
actattcgaa	gtgcagattt	gattgtgacc	ctaaaggatg	gaatgctggc	ggagaaagga	1740
gcacatgctg	aactaatggc	aaaacgaggt	ctatattatt	cacttgatgat	gtcacaggat	1800
attaaaaaag	ctgatgaaca	gatggagtca	atgacatatt	ctactgaaag	aaagaccaac	1860

tcacttcctc tgcactctgt gaagagcatc aagtcagact tcattgacaa ggctgaggaa	1920
tccacccaat ctaaagagat aagtcttcct gaagtctctc tattaataat tttaaagtta	1980
aacaagcctg aatggccttt tgtggttctg gggacattgg cttctgttct aaatggaact	2040
gttcatccag tattttccat catctttgca aaaattataa ccatgtttgg aaataatgat	2100
aaaaccacat taaagcatga tgcagaaatt tattccatga tattegtcat tttgggtgtt	2160
atttgctttg tcagttatct catgcaggga ttattttaac gcagagcagg ggaaatttta	2220
acgatgagat taagacactt ggccttcaaa gccatgttat atcaggatat tgcctgggtt	2280
gatgaaaagg aaaacagcac aggaggcttg acaacaatat tagccataga tatagcacia	2340
attcaaggag caacagggtc caggattggc gtcttaacac aaaatgcaac taacatggga	2400
ctttcagtta tcatttcctt tatatatgga tgggagatga cattcctgat tctgagtatt	2460
gctccagtac ttgccgtgac aggaatgatt gaaaccgcag caatgactgg atttgccaac	2520
aaagataagc aagaacttaa gcatgctgga aagatagcaa ctgaagcttt ggagaatata	2580
cgtactatag tgtcattaac aagggaaaaa gccttcgagc aaatgtatga agagatgctt	2640
cagactcaac acagaaatac ctgaagaaa gcacagatta ttggaagctg ttatgcattc	2700
agccatgcct ttatatattt tgccatgca gcagggttc gatttgagc ctatttaatt	2760
caagctggac gaatgacccc agagggcatg ttcatagttt ttactgcaat tgcatatgga	2820
gctatggcca tgggaaaaac gctcgttttg gctcctgaat attccaaagc caaatcgggg	2880
gctgcgcac tgtttgcctt gttggaaaag aaaccaaata tagacagccg cagtcaagaa	2940
gggaaaaagc cagacacatg tgaagggaat tttagagttc gagaagtctc tttcttctat	3000
ccatgtcgcc cagatgtttt catcctccgt ggcttatccc tcagtattga gcgaggaaaag	3060
acagtagcat ttgtggggag cagcggctgt gggaaaagca cttctgttca acttctgcag	3120
agactttatg accccgtgca aggacaagtg ctgtttgatg gtgtggatgc aaaagaattg	3180
aatgtacagt ggctccgttc ccaaatagca atcgcttctc aagagcctgt gctcttcaac	3240
tgcagcattg ctgagaacat cgcctatggt gacaacagcc gtgtggtgcc attagatgag	3300
atcaaagaag cgcgaatgc agcaaatac cattctttta ttgaaggtct ccctgagaaa	3360
tacaacacac aagttggact gaaaggagca cagctttctg gcggccagaa acaaagacta	3420
gctattgcaa gggctcttct ccaaaaaccc aaaattttat tgttgatga ggccacttca	3480
gccctcgata atgacagtga gaaggtggtt cagcatgcc ttgataaagc caggacggga	3540
aggacatgcc tagtggtcac tcacaggctc tctgcaatc agaacgcaga tttgatagtg	3600
gttctgcaca atggaaagat aaaggaacaa ggaactcatc aagagctcct gagaaatcga	3660
gacatatatt ttaagttagt gaatgcacag tcagtgcagt ga	3702

<210> 14

<211> 3621

<212> ADN

<213> Homo sapiens

<400> 14

```

ttccgctttg ctgatggact ggacatcaca ctcatgatcc tgggtatact ggcacactg      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 ttggcacag      300
gacatcggct ggtttgatag ctgtgacatc ggtgaactta acactcgcac gacagacatt     360
gacaaaatca gtgatggat tggagataag attgctctgt tgtttcaaaa catgtctact     420
ttttcgattg gcctggcagt tggtttggtg aagggctgga aactcacctt agtgactcta     480
tccacgtctc ctcttataat ggcttcagcg gcagcatggt ctaggatggg catctcattg     540
accagtaagg aattaagtgc ctattccaaa gctggggctg tggcagaaga agtcttgctc      600
tcaatccgaa cagtcatagc ctttagggcc caggagaaag aacttcaaag gtatacacag     660
aatctcaaag atgcaaagga ttttggcata aaaaggacta tagcttcaaa agtgtctctt     720
ggtgctgtgt acttctttat gaatggaacc tatggacttg ctttttggtt tggaaacctc     780
ttgattctta atggagaacc tggatatacc atcgggactg ttcttgctgt tttcttttagt     840
gtaatccata gcagttattg cattggagca gcagtccttc actttgaaac cttcgcaata     900
gcccaggagg ctgcctttca tattttccag gttattgata agaaaccag tatagataac     960
ttttccacag ctggatataa acctgaatcc atagaaggaa ctgtggaatt taaaaatggt    1020
tctttcaatt atccatcaag accatctatc aagattctga aaggtctgaa totcagaatt    1080
aagtctggag agacagtcgc cttggtcggc ctcaatggca gtgggaagag tacggtagtc    1140
cagcttctgc agaggttata tgatccggat gatggcttta tcatggtgga tgagaatgac    1200
atcagagctt taaatgtgcg gcattatcga gaccatattg gagtggtag tcaagagcct    1260
gttttgttcg ggaccaccat cagtaacaat atcaagtatg gacgagatga tgtgactgat    1320
gaagagatgg agagagcagc aagggaaagc aatgcgtatg attttatcat ggagtttcct    1380
aataaattta atacattggg aggggaaaaa ggagctcaaa tgagtggagg gcagaaacag    1440
aggatcgcaa ttgctcgtgc cttagtccga aacccaaga ttctgatttt agatgaggct    1500
acgtctgccc tggattcaga aagcaagtca gctgttcaag ctgcactgga gaaggcgagc    1560

```

```

aaaggctcga ctacaatcgt ggtagcacac cgactttcta ctattcgaag tgcagatttg 1620
attgtgaccc taaaggatgg aatgctggcg gagaaaggag cacatgctga actaatggca 1680
aaacgaggtc tatattattc acttgtgatg tcacaggata ttaaaaaagc tgatgaacag 1740
atggagtcaa tgacatattc tactgaaaga aagaccaact cacttcctct gcactctgtg 1800
aagagcatca agtcagactt cattgacaag gctgaggaat ccaccaatc taaagagata 1860
agtcttcctg aagtctctct attaaaaatt ttaaagttaa acaagcctga atggcctttt 1920
gtggttcttg ggacattggc ttctgttcta aatggaactg ttcattccagt attttccatc 1980
atctttgcaa aaattataac catgtttgga aataatgata aaaccacatt aaagcatgat 2040
gcagaaatit attccatgat attcgtcatt ttgggtgtta tttgctttgt cagttatttc 2100
atgcagggat tattttacgg cagagcaggg gaaattttta 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 tttgccaaca aagataagca agaacttaag 2460
catgctggaa agatagcaac tgaagctttg gagaatatac gtactatagt gtcattaaca 2520
agggaaaaag ccttcgagca aatgtatgaa gagatgcttc agactcaaca cagaaatacc 2580
tcgaagaaaag cacagattat tggaagctgt tatgcattca gccatgcctt tatatatattt 2640
gcctatgcag cagggtttcg atttgagacc tatttaattc aagctggacg aatgacccca 2700
gagggcatgt tcatagtttt tactgcaatt gcatatggag ctatggccat cggaaaaacg 2760
ctcgttttgg ctctgaata ttccaaagcc aaatcggggg ctgcgcattc gtttgccctg 2820
ttggaaaaga aaccaaatat agacagccgc agtcaagaag ggaaaaagcc agacacatgt 2880
gaaggggaatt tagagtttcg agaagtctct ttcttctatc catgtcgccc agatgttttc 2940
atcctccgtg gcttatccct cagtattgag cgaggaaaga cagtagcatt tgtggggagc 3000
agcggctgtg ggaaaagcac ttctgttcaa ctctctgcaga gactttatga ccccggtcaa 3060
ggacaagtgc tgtttgatgg tgtggatgca aaagaattga atgtacagtg gctccgttcc 3120
caaatagcaa tcgttcctca agagcctgtg ctcttcaact gcagcattgc tgagaacatc 3180
gcctatggtg acaadagccg tgtggtgcc aatagatgaga tcaaagaagc cgcaaatgca 3240
gcaaatatcc attcttttat tgaaggctc cctgagaaat acaacacaca agttggactg 3300
aaaggagcac agctttcttg cggccagaaa caaagactag ctattgcaag ggctcttctc 3360
caaaaaccca aaattttatt gttggatgag gccacttcag ccctcgataa tgacagtgag 3420

```

```

aaggtaggttc agcatgccct tgataaagcc aggacgggaa ggacatgcct agtggtcact 3480
cacaggtctct ctgcaattca gaacgcagat ttgatagtgg ttctgcacaa tggaaagata 3540
aaggaacaag gaactcatca agagctcctg agaaatcgag acatatatatt taagttagt 3600
aatgcacagt cagtgcagt a 3621

```

<210> 15

<211> 2021

<212> ADN

<213> Homo sapiens

<220>

<221> Nota

<222> (723)..(723)

<223> n na posição 723 representa qualquer nucleótido (A, T, C ou G)

<400> 15

```

ttccgctttg ctgatggact ggacatcaca ctcatgatcc tgggtatact ggcatcactg 60
gtcaatggag cctgccttcc tttaatgcc a ctgggttttag 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 acactcgcat gacagacatt 360
gacaaaatca gtgatggat tggagataag attgctctgt tgtttcaaaa catgtctact 420
ttttcgattg gcttggcagt tggtttggg aagggctgga aactcacctt agtgactcta 480
tccacgtctc ctcttataat ggcttcagcg gcagcatgtt ctaggatggc catctcattg 540
accagtaagg aattaagtgc ctattccaaa gctggggctg tggcagaaga agtcttgtca 600
tcaatccgaa cagtcatagc ctttagggcc caggagaaag aacttcaaag gtctttcctt 660
ttaaataata caagatatgc ttggttttat tttccccagt ggctactaag ttgtgttctg 720
ttntttgtaa ggtatacaca gaatctcaa gatgcaaagg attttggcat aaaaaggact 780
atagcttcaa aagtgtctct tgggtctgtg tacttcttta tgaatggaac ctatggactt 840
gctttttggg atggaacctc cttgattctt aatggagaac ctggatatac catcgggact 900
gttcttgctg ttttcttttag tgtaatccat agcagttatt gcattggagc agcagtcctt 960

```



```

cactttgaaa ccttcgcaat agcccgagga gctgccttcc 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 ttaaatgtgc ggcattatcg agaccatatt 1320
ggagtgggta gtcaagagcc tgttttgttc gggaccacca tcagtaacaa tatcaagtat 1380
ggacgagatg atgtgactga tgaagagatg gagagagcag caaggaagc aaatgcgtat 1440
gattttatca tggagtttcc taataaattt aatacattgg taggggaaaa aggagctcaa 1500
atgagtggag ggcagaaaca gaggatcgca attgctcgtg ccttagttcg aaacccaag 1560
attctgattt tagatgaggc tacgtctgcc ctggattcag aaagcaagtc agctgttcaa 1620
gctgcactgg agaaggatac cccaggtat tcattttgac ctaatttcac ctcaagtgga 1680
gaatcgctga ccttgaacca gcgcccttcg acagctctgg cccctcaaac ctcacctga 1740
cctectgctg cctatgagct actgcacata cctcaaggcc atatgcagtt gtggccctgc 1800
accaaattac actgaatcta ggaggggagt tggcagtggc ggtatgaaaa accattgaac 1860
agttttctcg atggcctgac tcccttataa accagagcct tcagaccctt tacaaggctt 1920
aatggcacat tttactttgc atttgcttgg aagtgagtta agcgtttttt tttctctaag 1980
aaaatcgag gcttcttttt ttaaaatgct gactttatgg a 2021

```

<210> 16

<211> 1940

<212> ADN

<213> Homo sapiens

<400> 16

```

ttccgctttg ctgatggact ggacatcaca ctcatgatec tgggtatact ggcattcactg 60
gtcaatggag cctgccttcc tttaatgcc a ctgggttttag gagaatgag tgataacctt 120
attagtggat gtctagtcca aactaacaca tactctttct tcaggttgac cctgtattat 180
gttggaatag gtgttgctgc cttgattttt ggttacatac agatttcctt gtggattata 240
actgcagcac gacagaccaa gaggattcga aaacagtttt ttcatcagtt tttggcacag 300
gacatcggct ggtttgatag ctgtgacatc ggtgaactta acactcgcat gacagacatt 360
gacaaaatca gtgatggat tggagataag attgctctgt tgtttcaaaa catgtctact 420
ttttcgattg gcctggcagt tggtttgggt aagggctgga aactcacctt agtgacteta 480

```

```

tccaggtctc ctcttataat ggcttcagcg gcagcatggt ctaggatggt catctcattg 540
accagtaagg aattaagtgc ctattccaaa gctggggctg tggcagaaga agtcttgta 600
tcaatccgaa cagtcatagc ctttagggcc caggagaaag aacttcaaag gtatacacag 660
aatctcaaag atgcaaagga ttttggcata aaaaggacta tagcttcaaa agtgtctctt 720
ggtgctgtgt acttctttat gaatggaacc tatggacttg ctttttggtg tggaaacctcc 780
ttgattctta atggagaacc tggatatacc atcgggactg ttcttgctgt tttctttagt 840
gtaatccata gcagttattg cattggagca gcagtccctc actttgaaac cttcgcaata 900
gcccaggagag ctgcctttca tattttccag gttattgata agaaacctcag tatagataac 960
ttttccacag ctggatataa acctgaatcc atagaaggaa ctgtggaatt taaaaatggt 1020
tctttcaatt atccatcaag accatctatc aagattctga aaggtctgaa tctcagaatt 1080
aagtctggag agacagtcgc cttggtcggt ctcaatggca gtgggaagag tacggtagtc 1140
cagcttctgc agaggttata tgatccggat gatggcttta tcatggtgga tgagaatgac 1200
atcagagctt taaatgtgcg gcattatcga gaccatattg gagtggttag tcaagagcct 1260
gttttggtcg ggaccaccat cagtaacaat atcaagtatg gacgagatga tgtgactgat 1320
gaagagatgg agagagcagc aagggaagca aatgcgtatg attttatcat ggagtttctt 1380
aataaattta atacattggt aggggaaaaa ggagctcaaa tgagtggagg gcagaaacag 1440
aggatcgcaa ttgctcgtgc cttagtctga aaccocaaga ttctgatttt agatgaggct 1500
acgtctgccc tggattcaga aagcaagtca gctgttcaag ctgcaactgga gaaggatacc 1560
cccaggtatt catthtgacc taatttcacc tcaagtggag aatcgtgac cttgaaccag 1620
cgcccttoga cagctctggc cctcaaacc tcacctgac ctctgctgc ctatgagcta 1680
ctgcacatac ctcaaggcca tatgcagttg tggccctgca ccaaattaca ctgaatctag 1740
gaggggagtt ggcagtggcg gtatgaaaaa ccattgaaca gttttctcga tggcctgact 1800
cccttataaa ccagagcctt cagaccctt acaaggetta atggcacatt ttactttgca 1860
tttgcttga agtgagttaa gcgtttttt ttctctaaga aaatcgcagg cttctttttt 1920
taaaatgctg actttatgga 1940

```

<210> 17

<211> 15

<212> PRT

<213> Homo sapiens

<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> 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> 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

Lisboa, 12 de Junho de 2015

REIVINDICAÇÕES

1. Método de enriquecimento de células estaminais a partir de uma mistura de células, que compreende pôr uma mistura de células em contacto com um anticorpo monoclonal específico que reage especificamente com uma glicoproteína P de 7p que tem uma sequência de proteína selecionada do grupo que consiste em SEQ ID NO:1, NO:2, NO:3, NO:4, NO:5, NO:6, NO:7 e NO:8 e isolar as células estaminais ligadas ao anticorpo a partir da mistura de células em que as células estaminais não são células estaminais embrionárias humanas.
2. Método de obtenção de células estaminais positivas para a glicoproteína P de 7p que compreende:
isolar as células positivas para a glicoproteína P de 7p a partir de uma amostra de células estaminais, em que as células positivas para a glicoproteína P de 7p são isoladas utilizando um anticorpo monoclonal específico contra a glicoproteína P de 7p que tem uma sequência de proteína selecionada do grupo que consiste em SEQ ID NO:1, NO:2, NO:3, NO:4, NO:5, NO:6, NO:7 e NO:8
e em que as células estaminais não são células estaminais embrionárias humanas.
3. Método da reivindicação 2, em que as células positivas para a glicoproteína P de 7p são isoladas utilizando um anticorpo imobilizado contra a glicoproteína P de 7p.
4. Método da reivindicação 3, em que o anticorpo está imobilizado em esférulas.

5. Método da reivindicação 2, em que as células de glicoproteína P de 7p são isoladas utilizando análise por FACS.

6. Método de caracterização de uma célula estaminal, que compreende

pôr uma célula estaminal em contacto com um anticorpo monoclonal específico dirigido contra a glicoproteína P de 7p que tem uma sequência de proteína seleccionada do grupo que consiste em SEQ ID NO:1, NO:2, NO:3, NO:4, NO:5, NO:6, NO:7 e NO:8 para determinar se a célula estaminal é positiva para a glicoproteína P de 7p,

em que as células estaminais não são células estaminais embrionárias humanas.

Lisboa, 12 de Junho de 2015