

(11) Número de Publicação: PT 959132 E

(51) Classificação Internacional:

C12N 15/57 (2006.01) C12N 9/48 (2006.01)

C12Q 1/68 (2006.01) G01N 33/574 (2006.01)

(12) FASCÍCULO DE PATENTE DE INVENÇÃO

(22) Data de pedido: 1997.10.13	(73) Titular(es): TAKARA BIO INC. 4-1, SETA 3-CHOME OTSU-SHI, SHIGA JP
(30) Prioridade(s): 1996.10.18 JP 2973299	
(43) Data de publicação do pedido: 1999.11.24	(72) Inventor(es): SHOHEI YOKOTA JP
(45) Data e BPI da concessão: 2006.08.24 012/2006	(74) Mandatário: JOSÉ EDUARDO LOPES VIEIRA DE SAMPAIO R DO SALITRE 195 RC DTO 1250-199 LISBOA PT

(54) Epígrafe: ÁCIDO NUCLEICO QUE CODIFICA UM RECEPTOR DO TIPO DE PROTEÍNA CINASE

(57) Resumo:

RESUMO

ÁCIDO NUCLEICO QUE CODIFICA UM RECEPTOR DO TIPO DE PROTEÍNA CINASE

A presente invenção tem por objecto providenciar um ácido nucleico que codifica um receptor de proteína cinase, em que o ácido nucleico tem uma duplicação em tandem numa sequência de nucleótidos de uma justamembrana e é útil para o diagnóstico da leucemia; um polipéptido codificado pelo ácido nucleico; um anticorpo capaz de se ligar especificamente a uma região codificada pelo ácido nucleico com uma duplicação em tandem que ocorre numa sequência de nucleótidos de uma justamembrana; um ácido nucleico capaz de se ligar especificamente ao ácido nucleico que tem uma duplicação em tandem ocorrendo numa sequência de nucleótidos de uma justamembrana; um processo para a detecção do ácido nucleico que codifica um receptor de proteína cinase; e um kit para o efeito. A presente invenção tem por objecto um ácido nucleico que codifica um receptor de proteína cinase, em que o ácido nucleico tem uma duplicação em tandem numa sequência de nucleótidos de uma justamembrana; um polipéptido codificado pelo ácido nucleico; um anticorpo capaz de se ligar especificamente à porção do polipéptido; um ácido nucleico capaz de se ligar especificamente ao ácido nucleico; um processo para a detecção do ácido nucleico; e um kit para a detecção.

DESCRIÇÃO

ÁCIDO NUCLEICO QUE CODIFICA UM RECEPTOR DO TIPO DE PROTEÍNA CINASE

ÂMBITO TÉCNICO

A presente invenção tem por objecto um ácido nucleico que codifica um receptor de proteína cinase, que tem uma duplicação em tandem numa sequência de nucleótidos de uma justamembrana, um polipéptido, um processo para a detecção do ácido nucleico anterior e um kit para a detecção.

TÉCNICA ANTERIOR

A proliferação e a diferenciação de células e as respostas das células a vários estímulos são estritamente reguladas por vários factores de crescimento. Estes factores de crescimento são conhecidos por actuarem por via dos receptores que são específicos dos factores de crescimento anteriores (Nicola, N. A., *Annu. Rev. Biochem.* **58**, 45, 1989; Lowenberg, B., *Blood* **81**, 281, 1991). Desses receptores, os receptores que contêm um domínio de tirosina cinase são classificados como receptores de tirosina cinases (RTCs).

Os RTCs compreendem uma região extracelular, uma região de transmembrana, assim como uma região intracelular contendo um domínio de tirosina cinase e uma justamembrana entre a região da transmembrana e o domínio de tirosina cinase e ainda são classificadas, de forma grosseira, em quatro tipos, de acordo com as características estruturais e a homologia das sequências de aminoácidos.

Os receptores do tipo I têm uma estrutura monomérica, com duas sequências repetidas ricas em cisteína na sua região extracelular e são exemplificados pelo receptor de EGF, HER2/neu e similares.

Os receptores do tipo II têm uma estrutura que compreende duas sub-unidades sendo uma α e a outra β , que se ligam, por via da ligação S-S, em que a cadeia α é uma região extracelular que contém uma sequência repetida rica em cisteína e em que a cadeia β tem uma região de transmembrana, uma justamembrana e um domínio de tirosina cinase. Exemplos são um receptor de insulina e um receptor de IGF-1.

Os receptores do tipo III têm uma estrutura monomérica contendo cinco sequências repetidas ricas em cisteína na sua região extracelular e dois domínios de tirosina interrompidos por uma inserção de cinase na sua região intracelular. Exemplos são o receptor de PDGF, *fms* (receptor de CSF-1), *kit* (receptor de SLF) e similares.

Os receptores do tipo IV assemelham-se aos receptores do tipo III, mas têm três sequências repetidas semelhantes a imunoglobulina e exemplificam-se por meio do receptor de FGF.

A tirosina cinase 3, semelhante a *fms* (daqui para a frente abreviada como FLT3; Matthews, W., *Cell* 65, 1143, 1991; Rosnet, O., *Genomics* 9, 380, 1991), que é expressa nas células leucémicas, etc, também referidas como cinase 2 do fígado fetal (CFF2, FLK2 na terminologia inglesa) ou STK1, é conhecida como receptor do tipo III (Small, B., *Proc. Natl. Acad. Sci. EUA* 91, 459, 1993; Lyman, S. D.,

Oncogene **8**, 815, 1993; Rosnet, O., *Blood* **82**, 1110, 1993; Agnes, F., *Gene* **145**, 283, 1994).

Nestas tirosina cinases receptoras, a agregação, tal como a dimerização, tem lugar após a ligação de um ligando, tal como um factor de crescimento, à região de extrapolação, resultando assim na activação da cinase. Embora nestas tirosina cinases, os seus ligandos tenham sido encontrados primeiro e só depois, na maioria dos casos, os seus receptores, há receptores dos quais permanecem desconhecidos os seus ligandos.

No que respeita a FLT3, que se fez notar no mecanismo de Proliferação das células hematopoiéticas do estroma, depois de se encontrar FLT3, encontrou-se o ligando de FLT3 (Lyman, S.D., *Cell* **75**, 1157, 1993; Hannum, C., *Nature* **368**, 643, 1994). Dado que o ligando de FLT3 é expresso em quase todas as células leucémicas, assume-se que as células Proliferam por meio de um mecanismo de estimulação de autocrina na leucemia (Meierhoff, G., *Leukemia* **9**, 1368, 1995). Também se referiu o ARNm de FLT3 como sendo expresso nas células leucémicas linfáticas e nas células leucémicas mielocíticas (Birg, F., *Blood* **80**, 2584, 1994; Da Silva, N., *Leukemia* **8**, 885, 1994; Brasel, K., *Leukemia* **9**, 1212, 1995; Drexler, H. G., *Leukemia* **10**, 588, 1996). Contudo, permanece desconhecida a forma como a expressão de ARNm de FLT3 se associa com a patologia da leucemia linfática e da leucemia mielocítica.

Clonou-se um ADNc de FLT3 humano e determinou-se uma sequência de nucleótido de ADNc [O. Rosnet et al., *Blood*, **82** (4), 1110 - 1119 (1993)]. A situação actual. Contudo, a situação actual é que ainda não se analisou bem a estrutura

e a função de FLT durante a diferenciação e as alterações malignas de células originais.

DESCRIÇÃO DA INVENÇÃO

De acordo com isto, um primeiro objecto da presente invenção consiste em providenciar um ácido nucleico que codifica uma proteína cinase receptora, em que o ácido nucleico tem uma duplicação em tandem numa sequência de nucleótido de uma justamembrana e é útil para o diagnóstico de leucemia e para providenciar um ácido nucleico que codifica a justamembrana anterior. Um segundo objecto da presente invenção consiste em providenciar um polipeptido que é codificado pelo ácido nucleico anterior. Um terceiro objecto da presente invenção consiste em Providenciar um anticorpo capaz de se ligar especificamente a uma porção codificada por um ácido nucleico que tem uma duplicação em tandem que ocorre numa sequência de nucleótido de uma justamembrana. Um quarto objecto da presente invenção consiste em providenciar um ácido nucleico capaz de se ligar especificamente a um ácido nucleico que tem uma duplicação em tandem que ocorre na sequência de nucleótido de uma justamembrana. Um quinto objecto da presente invenção consiste em Providenciar um Processo para a detecção do ácido nucleico que codifica uma Proteína cinase receptora e um respectivo kit.

Convencionalmente, tal como com FLT3, a mesma proteína cinase receptora é expressa, independentemente das espécies de células e da sua diferenciação [O. Rosnet et al., *Blood*, 82 (4), 1110 - 1119 (1993)]. Como resultado da investigação detalhada e de estudos intensivos da expressão de FLT3 em células leucémicas, contudo, os requerentes verificaram, surpreendentemente, que um gene de proteína cinase

receptora tinha uma nova duplicação em tandem numa justamembrana e verificaram que a duplicação em tandem anterior é somática e que a expressão de FLT3 que tem a duplicação em tandem anterior está associada com a malignidade da leucemia e a má consequência de Prognóstico e por isso a presente invenção foi completada.

De acordo com isto, a essência da presente invenção é a seguinte:

[1] um ácido nucleico que codifica uma Proteína cinase receptora em que o ácido nucleico tem uma duplicação em tandem numa sequência de nucleótido de uma justamembrana;

[2] o ácido nucleico de acordo com o item [1] anterior, em que a proteína cinase receptora é uma tirosina cinase receptora;

[3] o ácido nucleico de acordo com o item [2] anterior, em que a tirosina cinase receptora é Tirosina cinase 3 (FLT3) semelhante a FMS;

[4] o ácido nucleico de acordo com um qualquer dos itens [1] a [3] anteriores, em que o ácido nucleico compreende uma sequência de nucleótido que codifica uma sequência de aminoácido tal como se mostra em uma qualquer das SEQ ID N° 1 a 5 na lista de sequências numa justamembrana;

[5] o ácido nucleico de acordo com um qualquer dos itens [1] a [3] anteriores, em que o ácido nucleico compreende uma sequência de nucleótido que codifica uma sequência de aminoácido tal como se mostra em uma qualquer das SEQ ID N° 6 a 15 na lista de sequências numa justamembrana;

[6] um ácido nucleico que codifica um mutante de duplicação em tandem de FLT3 como se mostra em uma qualquer das SEQ ID N° 16 a 20 na lista de sequências;

[7] um ácido nucleico que compreende uma sequência de nucleótido que codifica um mutante de duplicação em tandem tal como se mostra em uma qualquer das SEQ ID N° 21 a 23 na lista de sequências ou um ácido nucleico capaz de hibridar em condições rigorosas, em que o ácido nucleico tem uma duplicação em tandem numa sequência de nucleótido codificando uma justamembrana;

[8] um ácido nucleico que tem uma duplicação em tandem, em que o ácido nucleico codifica uma sequência de aminoácido tal como se mostra em uma qualquer das SEQ ID N° 1 a 5 na lista de sequências;

[9] um ácido nucleico tal como se mostra em uma qualquer das SEQ ID N° 6 a 15 na lista de sequências ou um ácido nucleico capaz de hibridar em condições rigorosas, em que o ácido nucleico tem uma duplicação em tandem;

[10] um polipéptido codificado pelo ácido nucleico de acordo com um qualquer dos itens [1] a [9] anteriores;

[11] um polipéptido compreendendo uma sequência de aminoácido tal como se mostra em uma qualquer das SEQ ID N° 1 a 5 e 16 a 20 na lista de sequências;

[12] um polipéptido codificado por um ácido nucleico que tem uma duplicação em tandem numa sequência de nucleótido de uma justamembrana.

[13] um anticorpo capaz de se ligar especificamente a uma região codificada por um ácido nucleico que tem uma duplicação em tandem ocorrendo numa sequência de nucleótido de uma justamembrana de uma Proteína cinase receptora;

[14] um ácido nucleico capaz de se ligar especificamente a um ácido nucleico que tem uma duplicação em tandem ocorrendo numa sequência de nucleótido de uma justamembrana de uma Proteína cinase receptora;

[15] um processo para a detecção de um ácido nucleico que codifica uma Proteína cinase receptora e que tem uma duplicação em tandem ocorrendo numa sequência de nucleótido de uma justamembrana, compreendendo:

etapa (a): a preparação de uma amostra de ácido nucleico humano;

etapa (b): submissão da amostra de ácido nucleico obtida na etapa (a) a uma reacção de amplificação de genes para Providenciar um fragmento de ácido nucleico obtido por amplificação de uma região tendo uma duplicação em tandem numa justamembrana que pode ser encontrada num ácido nucleico que codifica uma Proteína cinase receptora; e

etapa (c): exame da presença da duplicação em tandem para o fragmento de ácido nucleico da etapa (b);

[16] o processo para a detecção, de acordo com o item [15] anterior, caracterizado pelo facto de o Processo ser utilizado no diagnóstico de M2, M4, ou M5 com base na classificação FAB (Francesa-Americana-Britânica) da leucemia mieloide aguda;

[17] um kit para a detecção de um ácido nucleico que codifica uma Proteína cinase receptora e que tem uma duplicação em tandem ocorrendo numa sequência de nucleótido de uma justamembrana, caracterizado pelo facto de o kit compreender iniciadores para a amplificação de uma região que tem uma duplicação em tandem, em que a região pode ser encontrada no gene da Proteína cinase receptora;

[18] o kit de acordo com o item [17] anterior, caracterizado pelo facto de esse kit ser utilizado no diagnóstico de M2, M4, ou M5 com base na classificação FAB (Francesa-Americana-Britânica) da leucemia mieloide aguda; e

[19] a utilização do ácido nucleico de acordo com um qualquer dos itens [1] a [9] anteriores para a detecção de um ácido nucleico que codifica uma Proteína cinase receptora e que tem uma duplicação em tandem ocorrendo numa sequência de nucleótido de uma justamembrana.

BREVE DESCRIÇÃO DOS DESENHOS

A figura 1 é uma vista que mostra a electroforese em gel de agarose do caso em que se realiza a RCP-TP com o ARN obtido de células leucémicas derivadas de pacientes com LMA, que funcionam como matrizes. Na figura, as as colunas 1 a 5 mostram, respectivamente, os resultados para pacientes que pertencem a M1, M2, M3, M4 e M5 (M34 pacientes) na classificação FAB, e as colunas 6 a 9 mostram respectivamente os resultados para M1, M2 (M155 pacientes), M3 e M4 (M162 pacientes).

A figura 2 é uma vista esquemática que mostra a duplicação em tandem no exão 11 e no exão 12 para M34, M155, M162, M810 e M839.

MELHOR MODO DE REALIZAÇÃO DA PRESENTE INVENÇÃO

A presente invenção será melhor explicada a seguir.

O ácido nucleico que codifica uma proteína cinase receptora da presente invenção e tem uma duplicação em tandem numa região que codifica uma justamembrana. O ácido nucleico da presente invenção que codifica uma proteína cinase é um ácido nucleico que codifica uma tirosina cinase, em particular uma tirosina cinase 3 (FLT3) semelhante a FMS. Para o diagnóstico da leucemia, utilizam-se ácidos nucleicos que codificam uma tirosina cinase 3 (FLT3) semelhante a FMS.

Na presente invenção, a justamembrana está presente entre a região da transmembrana e o domínio da cinase da Proteína cinase receptora e a justamembrana constrói uma região da membrana intracelular em conjunto com o domínio da cinase [O. Rosnet, et al., *Blood*, **82** (4), 1110 - 1119 (1993)].

Na presente invenção, duplicação em tandem refere-se a uma sequência de nucleótido de uma FLT3 em que uma porção completa ou uma porção parcial de um ácido nucleico que codifica uma justamembrana está repetida uma ou mais vezes na mesma orientação. As sequências de nucleótidos repetidas anteriores podem ser postas em linha uma directamente atrás da outra ou elas podem conter sequências de nucleótidos opcionais entre cada uma das sequências de nucleótidos repetidas. Além disso, o número de base duplicado não está

particularmente limitado. Além disso, as mutações de eliminação, substituição ou adição de uma ou mais bases pode existir numa porção de uma sequência de nucleótido entre as duplicações em tandem correspondentes. Na duplicação em tandem da presente invenção, a duplicação em tandem pode ser detectada como uma mutação em comprimento. Por exemplo, a duplicação em tandem está contida no ADNc comportando sequências de nucleótidos das SEQ ID NOs: 6 a 10 e no ADN genómico comportando sequências de nucleótidos das SEQ ID NOs: 11 a 15 como ácidos nucleicos codificando uma justamembrana.

Os ácidos nucleicos (ADNc ou ADN genómico) com uma duplicação em tandem de FLT3 que se encontraram de novo numa sequência de nucleótidos da justamembrana de FLT3 são designados por M34 (SEQ ID NOs: 6 e 11), M155 (SEQ ID NOs: 7 e 12), M162 (SEQ ID NOs: 8 e 13), M810 (SEQ ID NOs: 9 e 14) e M839 (SEQ ID NOs: 10 e 15), respectivamente e, na figura 2, mostra-se uma sua vista esquemática. Incidentalmente, é desejável que a duplicação em tandem na presente invenção tenha lugar numa secção. As sequências de aminoácidos codificadas pelas SEQ ID NOs: 6 a 10 estão indicadas nas SEQ ID NOs: 1 a 5.

O ácido nucleico da presente invenção diz respeito a um ácido nucleico do mutante de tirosina cinase 3 (FLT3) semelhante a FMS, em que o ácido nucleico tem a duplicação em tandem numa sequência de nucleótido de uma justamembrana. As sequências de aminoácidos de uma justamembrana com a duplicação em tandem estão indicadas, por exemplo, nas SEQ ID NOs: 1 a 5, como se mencionou antes e os ácidos nucleicos da presente invenção são os que compreendem as sequências de nucleótidos de FLT3 que codificam estas sequências de aminoácidos. Concretamente,

por exemplo, os ácidos nucleicos presentes são os que compreendem sequências de nucleótidos mostradas pelas SEQ ID NOS: 6 a 15. Mais particularmente, um ácido nucleico de um mutante com uma duplicação em tandem de FLT3 compreendendo um ácido nucleico de uma justamembrana inclui, por exemplo, ácidos nucleicos que codificam mutantes de duplicação em tandem de FLT3 que se mostra nas SEQ ID NOS: 16 a 20, mais concretamente, ácidos nucleicos que compreendem sequências de nucleótidos que codificam mutantes de duplicação em tandem de FLT3 que se mostra nas SEQ ID NOS: 21 a 25. Além disso, podem ser um ácido nucleico capaz de hibridar com o ácido nucleico anterior em condições severas e comportando uma duplicação em tandem numa sequência de nucleótido que codifica uma justamembrana.

Além disso, o ácido nucleico da presente invenção diz respeito a um ácido nucleico de FLT3 que codifica uma justamembrana e que comporta uma duplicação em tandem, incluindo o ácido nucleico, por exemplo, um ácido nucleico com uma duplicação em tandem, em que o ácido nucleico codifica sequências de aminoácidos mostradas nas SEQ ID NOS: 1 a 5, concretamente, os ácidos nucleicos mostrados nas SEQ ID NOS: 6 a 15, ou um ácido nucleico que tem uma duplicação em tandem capaz de hibridar com esses ácidos nucleicos em condições rigorosas.

Aqui, hibridação em condições rigorosas refere-se à hibridação com os ácidos nucleicos, em que a hibridação compreende, por exemplo, a incubação do ácido nucleico - membrana imobilizada, com uma sonda a 50 °C, durante 12 a 20 horas em 6 X SSC, em que 1 X SSC indica NaCl 0,15 M, albumina de soro bovino (ASB) a 0,1 %, polivinil-

pirrolidona a 0,1 %, Ficoll 400 a 0,1 % e ADN de espermatozoides de salmão desnaturado, mas não se limita a estas condições.

Pode-se obter o ácido nucleico da presente invenção, por exemplo, pelo Processo que se segue.

Em primeiro lugar, detectam-se as células em que tem lugar a mutação do comprimento por meio da síntese do ADNc por meio de uma transcriptase reversa com ARN como matriz, sendo o ARN purificado de várias células patológicas, particularmente de células leucémicas, comportando por isso uma reacção de amplificação do ADN utilizando iniciadores que são dirigidos a uma região que codifica a justamembrana de uma Proteína cinase receptora desejada e comparando o comprimento dos fragmentos de ADN amplificados por meio de um Processo de electroforese. Depois, é possível identificar se uma mutação é ou não uma duplicação em tandem por meio da determinação de uma sequência de nucleótido do fragmento de ADN amplificado obtido.

Em seguida, pode-se obter o ADNc que codifica uma Proteína cinase receptora, tendo o ADNc a nova duplicação em tandem da presente invenção por síntese do ADNc por uma transcriptase reversa com o ARN obtido das células em que a duplicação em tandem tem lugar, comportando por isso a reacção de amplificação do ADN utilizando iniciadores que podem amplificar especificamente o ADNc de uma desejada Proteína cinase receptora.

O ácido nucleico da presente invenção pode também ser obtido utilizando o ADN genómico como matriz, o ADN genómico purificado a partir de células patológicas.

Na presente invenção, seleccionam-se as células de leucemia como as células patológicas e FLT3 é preferencialmente atingido como a Proteína cinase receptora.

Um gene de FLT3 compreende 21 exões e, alternativamente, a justamembrana é codificada nos 18 pb no lado de 3' do exão 10 e 117 pb no lado de 5' do exão 11 [O. Rosnet, et al., Oncogene, 6, 1641 - 1650 (1991)]. Os iniciadores que cobrem a região do exão e do exão 12 podem seleccionam-se como iniciadores utilizados na reacção de amplificação de ADN. Exemplos das sequência de nucleótidos estão indicados nas SEQ ID NOs: 26 e 27. Incidentalmente, o exão 12 e os 16 pb no lado 3' do exão 11 codificam uma porção parcial do domínio de tirosina cinase.

Os fragmentos de ADN amplificados, tal como se mostra nas SEQ ID NOs: 6 a 10 obtêm-se quando o ARN é utilizado como uma matriz para a reacção de amplificação do ADN e os fragmentos de ADN amplificados, tal como se mostra nas SEQ ID NOs: 11 a 15 obtêm-se quando se utiliza o ADN genómico como a matriz. Como resultado, confirma-se que estes fragmentos resultantes têm uma duplicação em tandem numa secção dentro do exão 11 ou nos exões 11 a 12.

Alternativamente, as sequências de nucleótidos do ADNc que codificam todo o comprimento de FLT3 e que têm a anterior duplicação em tandem numa secção estão indicadas nas SEQ ID NOs: 21 a 25.

O polipéptido da presente invenção é um polipéptido codificado pelos ácidos nucleicos anteriores. Concretamente, pode-se exemplificar com um polipéptido que compreende sequências de aminoácidos das SEQ ID NOs: 1 a 5

e os mutantes da duplicação em tandem de FLT3 estão indicados nas SEQ ID NOs: 16 a 20.

O polipéptido da presente invenção podem obter-se por purificação das células que expressam o polipéptido e podem ser também obtidos pela utilização de Processos de engenharia genética convencionais. Um mutante de duplicação em tandem de FLT3, por exemplo, pode obter-se por inserção dos ácidos nucleicos anteriores num vector de expressão apropriado e depois o Produto é expresso num hospedeiro apropriado. Além disso, um polipéptido apenas com uma justamembrana de uma Proteína cinase receptora da presente invenção pode obter-se por inserção de um fragmento de ADN que codifica apenas uma membrana próxima no vector de expressão anterior.

Além disso, o polipéptido da presente invenção pode ser expresso como uma Proteína de fusão. Por exemplo, para aumentar as quantidades de expressão de uma Proteína desejada, adiciona-se uma cadeia de péptido do terminal N derivada de outra Proteína ao terminal N da Proteína desejada ou adiciona-se uma cadeia de péptido apropriada ao terminal N ou ao terminal C da Proteína desejada para expressar o polipéptido resultante, de modo a que a purificação da Proteína desejada, utilizando um veículo de resina que tem afinidade com a cadeia de péptido, pode ser facilitada.

O presente polipéptido também engloba um polipeptido codificado por um ácido nucleico com uma duplicação em tandem numa sequência de nucleótido de uma membrana próxima, em que o polipéptido resulta de pelo menos uma eliminação, substituição ou adição de um ou mais resíduos de aminoácidos nas sequências de aminoácidos da presente

invenção, por exemplo, as SEQ ID NOs: 1 a 5, 16 a 20 na listagem de sequências. Por outras palavras, pode ser o caso quando não ocorrem mutações nas sequências de aminoácidos na região codificada pelos ácidos nucleicos duplicados em tandem e ocorrem eliminações, substituições ou adições de resíduos de aminoácidos noutras porções das sequências de aminoácidos; ou o caso em que a eliminação, substituição ou adição de resíduos de aminoácidos tem lugar nas sequências de aminoácidos da região codificada pelos ácidos nucleicos duplicados em tandem. A introdução de eliminações, substituições ou inserções dos resíduos de aminoácidos pode ser facilmente realizada pela introdução da mutação na sequência de ácidos nucleicos desejada, por um Processo que utiliza endonucleases de restrição, nucleases e similares ou um Processo que realiza a mutagêneses dirigida ao sítio [W. Ito, et al., Gene, 102, 67 - 70 (1991)] etc., incorporando assim a sequência de ácidos nucleicos num vector de expressão para expressar o Produto numa célula hospedeira apropriada.

Na presente invenção, o anticorpo refere-se a um anticorpo capaz de se ligar especificamente a uma região codificada por um ácido nucleico com uma duplicação em tandem a ocorrer numa sequência de nucleótidos de uma membrana próxima da Proteína cinase receptora. Para se obter o anticorpo, por exemplo, pode-se obter com um antissoro por imunização de animais com um péptido comportando as sequências de aminoácidos das SEQ ID NOs: 1 a 5 em conjunto com adjuvante, por um Processo convencional. Além disso, pode obter-se o anticorpo como um anticorpo monoclonal por um Processo descrito em G. Galfare, et al., Nature, 266, 550 - 552 (1997).

Na presente invenção, o ácido nucleico capaz de se ligar especificamente aos ácidos nucleicos com uma duplicação em tandem, que ocorre numa sequência de nucleótidos de uma membrana próxima da proteína cinase receptora, não está particularmente limitado e é exemplificado pelo ADN de estrutura helicoidal dupla, anti-senso, com uma duplicação em tandem ou o ARN correspondente ao ADN anti-senso.

O processo para a detecção de um ácido nucleico da presente invenção compreende as etapas seguintes:

etapa (a): obtenção de uma amostra de ácido nucleico humano;

etapa (b): submissão da amostra de ácido nucleico obtida na etapa (a) anterior a uma reacção de amplificação do gene para se obter um fragmento de ácido nucleico obtido por amplificação de uma região com uma duplicação em tandem na membrana próxima, em que a região pode ser encontrada num ácido nucleico que codifica uma Proteína cinase receptora; e

etapa (c): exame da presença da duplicação em tandem para o fragmento de ácido nucleico da etapa (b) anterior.

Primeiro, descreve-se a etapa (a). A amostra de ácido nucleico humano utilizável na presente invenção não está particularmente limitada, dado que se trata de um ácido nucleico que codifica uma proteína cinase receptora, o ácido nucleico que comporta uma duplicação em tandem numa sequência de nucleótido de uma membrana próxima, tal como ADN genómico, ADNc ou ARNm. A amostra de ácido nucleico humano pode ser preparada por processos de realização

convencionais conhecidos, incluindo, por exemplo, um processo descrito em *Molecular Cloning : A LABORATORY MANUAL*, 2ª eda. (T. Maniatis et al., Cold Spring Harbor Laboratory Press, publicado em 1989).

Em seguida descreve-se a etapa (b). A amostra de ácido nucleico e os iniciadores apropriados são utilizados para amplificar um ácido nucleico que codifica uma região contendo um sítio de mutação que pode ser encontrado numa justamembrana da proteína cinase receptora de interesse para se obter um fragmento de ácido nucleico desejado. Um Processo para realizar a reacção de amplificação de ADN utilizável nesta etapa não está particularmente limitado, na medida em que é um processo capaz de amplificar a região anterior e assim podem ser utilizados processos de amplificação de ácidos nucleicos, tais como o Processo de amplificação de ácidos nucleicos utilizando a RCP-TR, o processo de RCP ou as polimerases de ARN (patente de invenção japonesa depositada com os n.ºs Hei 2-5864 e Hei 7-203999) ou o processo de amplificação de substituição da estrutura helicoidal (publicação da patente de invenção japonesa examinada com o No. Hei 7-114718 e a patente de invenção japonesa depositada com o n.º Hei 7-88242). Entre eles, o processo de RCP-TR ou o processo da RCP é utilizado preferencialmente.

A região a ser amplificada, comportando uma duplicação em tandem numa justamembrana inclui, por exemplo, no caso de FLTR3, uma região contendo uma porção total ou parcial da região de 18 pb no lado 3' do exão 10 a 117 pb no lado de 5' do exão 11, sem estar particularmente limitada desde que a região contenha um sítio do exão 11.

Os iniciadores utilizados no processo de RCP-TR ou no processo de RCP não estão particularmente limitados desde que eles sejam iniciadores capazes de amplificar um fragmento de ADN que contém o sítio de mutação anterior. Concretamente, pode-se exemplificar um par de iniciadores com as sequências de nucleótidos tal como se mostra nas SEQ ID NOs: 26 e 27 da listagem de sequências. Além disso, as condições da RCP não são particularmente limitadas e podem utilizar-se condições de realização convencionais conhecidas para a reacção RCP.

Em terceiro lugar vai descrever-se a etapa (c). Nesta etapa, examina-se a presença da duplicação em tandem para o fragmento de ácido nucleico obtido na etapa (b). O processo para a detecção da presença da duplicação em tandem não é particularmente limitado é preferível utilizar um Processo de comparação dos comprimentos dos fragmentos de ADN amplificado por meio do processo de electroforese em gel de agarose.

Além disso, pode-se utilizar um processo de exame do polimorfismo da conformação da estrutura helicoidal simples (PCHS) para a detecção da mutação que é utilizável nesta etapa. O processo compreende o exame das diferenças de uma estrutura de ordem superior como as diferenças de mobilidade na electroforese, em que a estrutura de ordem superior está dependente da sequência de nucleótido em que o ADN de estrutura helicoidal simples é formado pela interacção intramolecular (*Proc. Natl. Acad. Sci. EUA*, 86: 2766 - 2770, 1989). A presença ou a ausência de mutação pode ser detectada submetendo o fragmento de ácido nucleico obtido na etapa (b) a electroforese nas condições descritas na publicação mencionada antes e comparando a sua

mobibilidade com a de um fragmento de ácido nucleico derivado de uma proteína cinase receptora normal.

Outros processos de detecção incluem um processo em que a etapa (c) anterior é alterada para outro Processo para a detecção da mutação. Para a detecção da mutação, pode-se utilizar um processo conhecido para a detecção da mutação, tal como um Processo de hibridação que utiliza um fragmento de ADN apropriado contendo um sítio de mutação como uma sonda, ou pelo Processo de DGGE [Val C. Sheffield et al., *Proc. Natl. Acad. Sci. EUA* 86, 232 - 236 (1989)]. Além disso, conhece-se um processo para a detecção da mutação utilizando uma proteína conhecida a MutS (patente de invenção japonesa depositada No. Hei 7-327698).

A mutação pode ser identificada pela sequenciação da sequência de nucleótidos para o fragmento de ADN em que uma mutação do comprimento é confirmada por meio do Processo mencionado antes. Para a sequenciação da sequência de nucleótidos, pode utilizar-se um processo convencionalmente usado, incluindo, por exemplo, um processo que compreende a clonagem de um fragmento de ADN modificado num vector apropriado e a determinação da sequência de nucleótidos utilizando um fragmento amplificado per se como uma matriz.

Tal como se descreveu antes, a presente invenção tem por objecto a utilização de um ácido nucleico da presente invenção descrito antes para a detecção de um ácido nucleico que codifica uma Proteína cinase receptora, em que o ácido nucleico tem uma duplicação em tandem numa sequência de nucleótidos de uma membrana próxima.

O processo para a detecção de um ácido nucleico da presente invenção pode ser utilizado no diagnóstico de M2,

M4, e M5 com base na classificação FAB (Francesa-Americana-Britânica), classificam-se os tipos patológicos de AML em seis classes como M1 (mieloblástica, sem potencial de maturação), M2 (mieloblástica, com potencial de maturação), M3 (promielocítica), M4 (mielomonocítica), M5 (monocítica) e M6 (eritroleucêmica) (Shin Rinsho Kensa Gishi Koza, 10, Ketsuekigaku, 75, Igaku-Shoin).

Entende-se que os pacientes que possuem um gene de FLT3 com a duplicação em tandem da presente invenção pertencem às classes M2, M4, e M5 anteriores e que os pacientes relapsos no sintoma até à morte mesmo com uma remissão transitente do sintoma, de modo que o seu prognóstico leva à consequência do mal. Por isso, de acordo com a presente invenção, providencia-se um processo para o exame útil para a análise patológica de AML.

Incidentalmente, dos pacientes anteriores, realiza-se a detecção da mutação utilizando o ADN genômico do mielócito de um paciente obtido no momento da remissão dos sintomas e, como resultado, não se encontram as duplicações em tandem numa membrana próxima.

Também o ácido nucleico da presente invenção que tem uma duplicação em tandem, serve como um marcador para o síndrome mielodisplástico (SMD), que desenvolve num pré-estádio da leucemia, AML com displasia e similares, assim como AML tal como se classifica com base na classificação FAB. O processo de detecção da presente invenção por isso é um processo para o exame que é útil para o julgamento patológico destas doenças.

Utilizando o processo de detecção anterior, pode-se providenciar um kit para a detecção de um ácido nucleico da

presente invenção. Concretamente, há um kit para a detecção de um ácido nucleico pelo processo de detecção descrito antes, codificando o ácido nucleico uma proteína cinase receptora comportando a duplicação em tandem na sequência de nucleótidos de uma membrana próxima, caracterizado pelo facto de o kit compreender iniciadores para a amplificação de uma região comportando uma duplicação em tandem, em que a região pode ser encontrada no gene da proteína cinase receptora.

O diagnóstico do AML anterior, etc. pode ser facilmente realizado utilizando esse kit.

Na presente invenção, um polipeptido codificado por um ácido nucleico comportando uma duplicação em tandem, tal como descrito antes pode ainda ser detectado pelas etapas que se mostram a seguir:

etapa 1: obtenção de uma amostra de Proteína humana; e

etapa 2: exame da presença da duplicação em tandem na sequência de nucleótidos de uma membrana próxima da amostra de Proteína na etapa 1 anterior.

Primeiro, descreve-se a etapa 1. Pode-se preparar a amostra de Proteína humana preparando uma Proteína de membrana a partir de uma célula que se assume que tenha o polipeptido da presente invenção aí expresso (por exemplo, célula leucémica, no caso de LFT3).

Em segundo lugar descreve-se a etapa 2. O processo para a detecção das mutações de duplicação em tandem não é particularmente limitado e pode ser realizado utilizando um anticorpo marcado capaz de se ligar especificamente à

membrana próxima codificada por um ácido nucleico com uma mutação de duplicação em tandem.

Esta etapa pode, por exemplo, ser realizada por um Processo que compreende a submissão da amostra de proteína obtida na etapa 1 a SDS-PAGE para separar as proteínas e, em seguida, a detecção da proteína desejada, pelo processo de imuno-mancha.

Num outro enquadramento da presente invenção providencia-se um processo para a regulação da proliferação, resposta imunitária e transmissão de informação do sinal das células leucémicas, células hematopoiéticas do estroma, etc. utilizando os ácidos nucleicos ou os polipéptidos anteriores ou ácidos nucleicos ou anticorpo capazes de se ligarem especificamente.

Entre eles, um enquadramento preferido inclui uma aplicação para imunoterapia de tumores. Convencionalmente, sabe-se que os péptidos específicos de tumores das Proteínas que expressam especificamente nas células de tumor servem como alvos das respostas imunitárias das células T às células de tumor. Num processo para a realização do presente pedido de patente de invenção, podem utilizar-se, por exemplo, as técnicas descritas nos relatos que se seguem. Concretamente, isolam-se as células CD4+T restritas a HLA-DR, as células reagem especificamente com o péptido resultando da substituição do 12º aminoácido, a glicina com outro aminoácido nas células T humanas (Jung, S., *J. Exp. Med.* 173, 273, 1991) e um LCT (linfócito citotóxico) reconhecendo um péptido que consiste em oito aminoácidos incluindo o sítio de mutação para a mutação do 61º aminoácido que pode ser derivado de um rato imunizado com um vírus de vacínia recombinante capaz de Produzir a

proteína ras, que tem mutação no 61º aminoácido (Skipper, J., *J. Exp. Med.* 177, 1493, 1993). Além disso, num rato imunizado com uma proteína ras mutante, solúvel, suprime-se a proliferação in vivo das células de tumores com a mesma mutação (Fenton, R.G., *J. Natl. Cancer Inst.* 85, 1294, 1993), e um LCT que exibe actividade citotóxica contra células de tumor que expressam a mesma ras mutante podem ser obtidas a partir de esplenócitos sensibilizados com o péptido mutante de ras (Peace, D.J., *J. Exp. Med.* 179, 473, 1994). Por outro lado, a proteína quimérica bcr-abl, que é muitas vezes detectada na leucemia mielocítica crónica, possui uma elevada actividade de tirosina cinase e desempenha um papel chave no início da leucemia e na Proliferação das células leucémicas. Imunizando com um péptido na vizinhança do sítio de fusão desta proteína de fusão, podem obter-se as células T reactivas a esta proteína de fusão (Chen, W., *Proc. Natl. Acad. Sci. EUA* 89, 1468, 1992). Além disso, o AND anti-senso ou o ARN correspondente ao gene de fusão é capaz de suprimir a proliferação de tumores que expressam este gene in vivo (Skorski, T., *Proc. Natl. Acad. Sci. EUA* 91, 4504, 1994).

É possível, por isso, obter células T reactivas com uma proteína cinase receptora que compreende o péptido da presente invenção, em que o péptido é codificado por um ácido nucleico com uma duplicação em tandem que ocorre na sequência de nucleótidos de uma justamembrana e regula a Proliferação das células que expressam a cinase anterior por meio da imunização com o péptido anterior.

Também quando a presença da duplicação em tandem da presente invenção está envolvida na regulação da proliferação das células, é possível regular o mecanismo de

sinalização com ADN ou ARN anti-sentido para o gene anterior para regular a Proliferação das células.

Quando se liga um ligando a uma região extracelular, a Proteína cinase receptora sofre uma alteração conformacional para formar um dímero, resultando num aumento da actividade do domínio da cinase na região intracelular, pela qual uma auto-fosforilação ou fosforilação de um substrato da cinase anterior tem lugar. Nestas etapas, estão envolvidas várias moléculas de sinalização e a informação transmitida para dentro das células causa vários fenómenos biológicos, tal como a alteração morfológica das células, o movimento das células, a morfogénese, a proliferação das células, a alteração maligna, a diferenciação e a apoptose. As células da leucemia mielocítica aguda de elevada malignidade têm sido relatadas como possuindo uma forte afinidade com o ligando de FLT3 e promovem a proliferação das células (Piacibello, W., *Blood* **85**, 4105, 1995; Lisovsky, M., *Blood* **86**, 22a, 1995; McKenna, H., *J. Exp. Hematol.* **24**, 378, 1996; Debnel, U., *Leukemia* **10**, 261, 1996). Nas células que expressam o mutante duplicação em tandem de FLT3 da presente invenção, é expectável que o sistema para a sinalização do ligando de FLT3 está altamente activada. As células hematopoiéticas indiferenciadas que expressam o mutante são por isso Providenciadas como uma fonte de células hematopoiéticas indiferenciadas que possuem um forte potencial de proliferação. Comparando as células hematopoiéticas indiferenciadas com as células que expressam FLT3 normal, pode-se Providenciar os materiais e os processos apropriados para avaliar os vários fármacos.

Tal como se descreveu antes, utilizando um processo da presente invenção, pode-se aplicar ao exame e ao tratamento

de doenças das células sanguíneas, doenças das células hematopoiéticas indiferenciadas e outras doenças.

A presente invenção será descrita com mais detalhe a seguir por meio de exemplos, mas a presente invenção não fica limitada por esses exemplos.

Exemplo 1

1) Análise do modelo de expressão do gene de FLT3

Em 90 casos de pacientes com leucemia aguda (50 casos de crianças ALL, 30 casos de adultos AML), realizou-se a análise da expressão do gene de FLT3 por meio do processo de RCP-TR. Os iniciadores utilizados foram produzidos de modo a ter sequências de nucleótidos tal como se mostra nas SEQ ID NOs: 26 e 27 na Listagem de Sequências e para cobrir completamente e amplificar a região da transmembrana através da justamembrana. Utilizando o par de iniciadores anterior, o Produto de ADN amplificado resultante tem 366 pb de comprimento quando o FLT3 normal foi transcrito.

Extraíu-se o ARN total a partir do sangue periférico ou de mielócitos derivados do paciente anterior com um reagente de Trizol (fabricado pela LifeTech), seguida reacção de amplificação do ADN utilizando um kit de RCP-TR (fabricado pela Takara Shuzo Co., Ltd.) e um Thermal Cycler (fabricado pela Takara Shuzo Co., Ltd.) nas condições que se seguem. Sintetizou-se o ADNc a partir do ARN total utilizando uma transcriptase reversa. Numa mistura reacção de 50 µL contendo 1 µL do ADNc (equivalente a 40 ng de um ARN total), uma mistura de dNTP 200 µM, 1 x tampão de RCP, 2 U de Taq polimerase de ADN e 20 pmole de cada um dos iniciadores descritos antes, aqueceu-se a mistura

reaccional anterior e depois repetiu-se 35 vezes um ciclo térmico que consiste em 64 °C durante 30 segundos, 72 °C durante 45 segundos e 94 °C durante 30 segundos e depois tratou-se finalmente a 72 °C durante 5 minutos. Para verificar a qualidade do ARN, realizou-se a RCP-TR da mesma maneira, excepto no facto de se ter utilizado um par dos iniciadores mostrados nas SEQ ID NOs: 28 e 29 na listagem de sequências tendo a actina β como alvo. Os produtos de ADN amplificados assim obtidos foram submetidos a electroforese em gel de agarose a 2 a 3 % (fabricado pela FMC) contendo brometo de etídio e detectou-se por meio de irradiação UV. Um exemplo do modelo de electroforese está indicado na figura 1 e os resultados estão indicados no quadro 1.

Quadro 1.

Sub-tipo de FAB	Número de casos examinados	Número de expressão de ARNm positiva de FLT3 (%)	Comprimento da mutação (%)
AML total	30	22 (73 %)	5 (17 %)
M1	3	2 (67 %)	0
M2	9	7 (78 %)	1 (14 %)
M3	6	5 (83 %)	0
M4	5	4 (80 %)	2 (40 %)
M5	4	3 (75 %)	2 (50 %)
M6	1	1 (100 %)	0
ALL total	50	39 (78 %)	0
cALL	27	24 (89 %)	0
pre-B ALL	13	11 (85 %)	0
B-ALL	1	0	0
T-ALL	9	4 (44 %)	0

Verificou-se que, a partir do quadro 1 que ao Produto da transcrição do gene de FLT3 foram encontrados em 39 casos (78 %) dos 50 casos ALL e 22 casos (73 %) dos 30 casos AML. Entre eles, o Produto de ADN amplificado mais comprido do que os 366 pb esperados foi detectado em 5 casos (23 %) dos 22 casos AML positivos para FLT3, de tal modo que se observou uma mutação do comprimento do gene de FLT3. Incidentalmente, em quatro casos (M34, M155, M810, e M839), detectou-se a banda esperada de 366pb e uma banda mais comprida do que o esperado. Num dos casos (M162), a banda de 366 pb não foi detectada e só se detectou uma banda mais longa.

2) Análise da sequência de nucleótidos do Produto da mutação do comprimento do gene de FLT3.

Para examinar com mais detalhe as mutações do comprimento no gene nos 5 casos anteriores, o Produto de ADN amplificado foi purificado a partir de gel de agarose e as determinaram-se as sequências de nucleótidos das regiões do exão 11 e do exão 12. Como resultado, confirmou-se que estas mutações do comprimento resultaram das duplicações em tandem nas sequências de nucleótidos da justamembrana respectiva. Concretamente, encontrou-se uma duplicação em tandem de 39 pb ou de 60 pb dentro do exão 11, nos casos M34, M162, e M839; e encontrou-se uma duplicação em tandem de 26 pb incluindo uma inserção de 4 pb (GGCA) no caso M810. Além disso, verificou-se que o caso M155 tinha uma duplicação em tandem de 63 pb compreendendo os primeiros 16 pb do exão 12, imediatamente depois do exão 11, uma inserção do resíduo de citosina e os últimos 46 pb do exão 11. As sequências de nucleótidos obtidas estão indicadas nas SEQ ID NOs: 6 a 10 na lista de sequências e na Figura 2

apresenta-se uma vista esquemática das duplicações em tandem.

Caracteristicamente, estas duplicações em tandem ocorrem em secções e estas mutações reflectem-se nos polipéptidos actualmente expressos. As sequências de aminoácidos codificadas por estas sequências de nucleótidos estão indicadas nas nas SEQ ID NOs: 1 a 5.

3) Análise da sequência de nucleótido do ADN genómico

Analísaram-se os produtos de ADN amplificados obtidos por RCP com ADN genómico de FLT3 derivados de mielócitos dos 5 casos de pacientes anteriores como matrizes. Realizou-se a reacção de RCP sob condições de amplificação tal como o facto de se adicionar 2 U de Tag polimerase de ADN (Takara Shuzo Co., Ltd.) a um tampão de RCP contendo 50 ng de ADN genómico, 200 µM de uma mistura de dNTP e 20 pmole de cada um dos iniciadores, até preencher um volume total de 50 µL. Para os exões 10 ao exão 19, cada exão foi submetido individualmente à reacção de amplificação do ADN. A mutação do comprimento foi observada da mesma maneira que quando se analisou o ARNm no caso em que os exões 11 e 12 foram amplificados com iniciadores das SEQ ID NOs: 30 e 31, e os iniciadores das SEQ ID NOs: 32 e 33 na listagem de sequências como pares. Quando os produtos de RCP foram purificados utilizando QIAEXII (QIAGEN), clonaram-se no vector pCRTII (Invitrogen) e submeteram-se a uma análise da sequência de nucleótidos, obtendo-se resultados semelhantes aos da sequência de nucleótidos do ADNc (SEQ ID NOs: 21 a 25). Os resultados mostram-se, colectivamente na figura 2.

Exemplo 2

Análise de mutações na justamembrana de proteína cinase receptora e das suas relações patológicas

Para analisar as mutações numa justamembrana da Proteína cinase receptora e as suas relações patológicas, mostra-se, no quadro 1, a relação entre a classificação patológica dos sintomas e o gene de FLT3. Cinco casos mostraram a duplicação em tandem na sequência de nucleótidos de uma justamembrana pertencendo a M2 (mieloblástico, com potencial de maturação), M4 (mielomonocítico) ou M5 (monocítico), com base na classificação FAB e todos eles foram casos em que os pacientes recaíram em relação aos sintomas até à morte mesmo com remissão transiente dos sintomas.

As sequências de nucleótidos das regiões do exão que codificam o domínio de tirosina cinase foram também analisadas e não se encontraram mutações nestas regiões.

Por isso, foi sugerido que as duplicações em tandem são secções na região dos genes que codificam uma justamembrana de FLT3 estavam associadas com AML com o crescimento de monócitos que foi sugerido.

Também, dado que essas mutações do comprimento não foram detectadas nas amostras de ADN retiradas dos mielócitos recolhidos nestes três casos (M34, M162, e M810); no momento da remissão completa, verificou-se que a duplicação em tandem da presente invenção era uma mutação somática.

APLICABILIDADE INDUSTRIAL

De acordo com a presente invenção, pode providenciar-se uma nova proteína cinese receptora com uma mutação de duplicação em tandem na sequência de nucleótidos de uma justamembrana e a sua sequência de nucleótidos e informação sobre a sequência de aminoácidos. Além disso, podem-se providenciar diagnósticos patológicos, um processo para o exame de leucemia, etc. utilizando a presente invenção, um kit e um reagente para os exames relacionados. Além disso, pode providenciar-se um processo para a regulação e as condições de análise da proliferação e da diferenciação, alteração maligna, resposta imunitária e sinalização para células representadas por células hematopoiéticas indiferenciadas e células leucêmicas utilizando a presente invenção e um kit e um reagente relacionado com ele.

LISTAGEM DE SEQUÊNCIAS

(1) INFORMAÇÃO GERAL:

(i) REQUERENTE:

- (A) NOME: Takara Shuzo Co., Ltd.
- (B) RUA: 609, Takenaka-cho, Fushimi-ku., Kyoto-shi
- (C) CIDADE: Kyoto
- (E) PAÍS: Japão
- (F) CÓDIGO POSTAL (ZIP): 612

(ii) TÍTULO DA INVENÇÃO: Ácido nucleico codificando um receptor do tipo de Proteína cinase

(iii) NÚMERO DE SEQUÊNCIAS: 33

(iv) FORMA LISÍVEL EM COMPUTADOR:

- (A) TIPO DE MEIO: Floppy disk
- (B) COMPUTADOR: compatível com IBM PC
- (C) SISTEMA OPERATIVO: Dos/MS-DOS para PC
- (D) SOFTWARE: PatentIn Release #1.0, versão #1.25
- (EPO)

(v) DADOS SOBRE O PRESENTE PEDIDO DE PATENTE DE INVENÇÃO:

NÚMERO DO PRESENTE PEDIDO DE PATENTE DE INVENÇÃO: EP 97943188.9

(2) INFORMAÇÃO PARA A SEQ ID NO: 1:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 97 aminoácidos
- (B) TIPO: aminoácido

(C) ESTRUTURA HELICOIDAL: simples
(D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: SIM

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 1:

Gln	Phe	Arg	Tyr	Glu	Ser	Gln	Leu	Gln	Met	Val	Gln	Val	Tre	Gli	Ser
1				5					10					15	
Ser	Asp	Asn	Glu	Tyr	Phe	Tyr	Val	Glu	Ser	Gln	Leu	Gln	Met	Val	Gln
			20					25					30		
Val	Tre	Gli	Ser	Ser	Asp	Ser	Glu	Tyr	Phe	Tyr	Val	Asp	Phe	Arg	Glu
		35					40					45			
Tyr	Glu	Tyr	Asp	Leu	Lis	Trp	Glu	Phe	Pro	Arg	Glu	Asn	Leu	Gln	Phe
		50				55					60				
Gli	Lis	Val	Leu	Gli	Ser	Gli	Ala	Phe	Gli	Lis	Val	Met	Asn	Ala	Tre
65				70					75					80	
Ala	Leu	Glu	Leu	Ala	Lis	Gln	Glu	Ser	Gln	Ser	Arg	Leu	Pro	Ser	Lis
			85						90					95	

Cis

(2) INFORMAÇÃO PARA A SEQ ID NO: 2:

(1) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 100 aminoácidos
(B) TIPO: aminoácido
(C) ESTRUTURA HELICOIDAL: simples
(D) TOPOLOGIA: linear

Gln	Phe	Arg	Tir	Glu	Ser	Gln	Leu	Gln	Met	Val	Gln	Val	Tre	Gli	Ser
1				5				10					15		
Ser	Asp	Asn	Glu	Tir	Phe	Tir	Val	Asp	Phe	Arg	Glu	Tir	Glu	Tir	Asp
			20				25					30			
Leu	Lis	Trp	Glu	Phe	Asp	Phe	Arg	Glu	Tir	Glu	Tir	Asp	Leu	Lis	Trp
		35					40					45			
Glu	Phe	Pro	Arg	Glu	Asn	Leu	Glu	Phe	Gli	Lis	Val	Leu	Gli	Ser	Gli
	50					55					60				
Ala	Phe	Gli	Lis	Val	Met	Asn	Ala	Tre	Ala	Tir	Gli	Ile	Ser	Lis	Tre
65					70					75				80	
Gli	Val	Ser	Ile	Gln	Val	Ala	Val	Lis	Met	Leu	Lis				
				85					90						

(2) INFORMAÇÃO PARA A SEQ ID NO: 4:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 89 aminoácidos
- (B) TIPO: aminoácido
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: SIM

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 4:

Gln	Phe	Arg	Tir	Glu	Ser	Gln	Leu	Gln	Met	Val	Gln	Val	Tre	Gli	Ser
1				5				10					15		
Ser	Asp	Asn	Glu	Tir	Phe	Tir	Val	Asp	Phe	Arg	Glu	Tir	Glu	Tir	Asp
			20				25					30			
Leu	Lis	Trp	Glu	Phe	Pro	Arg	Glu	Asn	Trp	His	Lis	Trp	Glu	Phe	Pro
		35					40					45			
Arg	Glu	Asn	Leu	Glu	Phe	Gli	Lis	Val	Leu	Gli	Ser	Gli	Ala	Phe	Gli
	50					55					60				

(A) COMPRIMENTO: 296 pares de bases

(1) CARACTERÍSTICAS DA SEQUÊNCIA:

((2) INFORMAÇÃO PARA A SEQ ID NO: 6:

85	90	
65	70	75
50	55	60
35	40	45
20	25	30
5	10	15

(*) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 5:

(111) HIPOTÉTICA: SIM

(11) TIPO DE MOLÉCULA: péptido

(D) TOPOLOGIA: linear

(C) ESTRUTURA HELICOIDAL: simples

(B) TIPO: aminoácido

(A) COMPRIMENTO: 92 aminoácidos

(1) CARACTERÍSTICAS DA SEQUÊNCIA:

((2) INFORMAÇÃO PARA A SEQ ID NO: 5:

85	
65	70
50	75

- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 6:

```
CAATTAGGT ATGAAAGCCA GCTACAGATG GTACAGGTGA CCGGCTCCTC AGAATATGAG 60
TACTTCTACG TTGAAAGCCA GCTACAGATG GTACAGGTGA CCGGCTCCTC AGATATGAG 120
TACTTCTACG TTGATTTGAG AGAATATGAA TATGATCTCA AATGGGAGTT TCCAAGAGAA 180
AATTTAGAGT TTGGGAAGGT ACTAGGATCA GGTCCTTTTG GAAAAGTCAT GAACGCPACA 240
GCTTTGGAAT TAGCAAAACA GCAGTCTCA TCCAGGTTSC CGTCAAAATG CTGAAA 296
```

(2) INFORMAÇÃO PARA A SEQ ID NO: 7:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 300 pares de bases
- (B) TIPO: ácido NUCLEICO
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 7:

```
CAATTAGGT ATGAAAGCCA GCTACAGATG GTACAGGTGA CCGGCTCCTC AGATATGAG 60
```

TACTTCTACG TTGATTTTCAG AGRATATGAA TATGATCTCA AATGGGAGTT TCCAAGAGAA 120
 AATTTAGAGT TTGGCAAGGT ACTAGGATCC GAATATGATC TCAAAATGGCA GTTTCACAGA 180
 GAAATTTTAG ACATTGGGAA CGTACTTGGCA TCAGGTGCTT TTGGAAAAGT GATGAACGCA 240
 ACAGCTTATC GAATTACCA PACAGAGTC TCAATCCAGG TTGCGCTCAA AATGCTGAAA 300

(2) INFORMAÇÃO PARA A SEQ ID NO: 8:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 267 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 8:

CAGTTTAGGT ATGAAACCCA GCTACAGATG CTACAGGTGA CCGCTCTCTC ACATAATGAG 60
 TACTTCTACG TTGATTTTCAG AGRATATGAA TATGATCTCA AATGGGAGTT TGATTTCAGA 120
 GAATATGAAAT ATGATCTCAA ATGGGAGTTT CCAAGAGAAA ATTACAGIT TGGGAGGTA 180
 CTAGGATCAG GTGCTTTTGG AAGACTGATG AACGCAACAG CTTATGGAT TAGCAAAACA 240
 GGAGTCTCAA TCCAGTTTGC CGTCAAAATG CTGAAA 276

(2) INFORMAÇÃO PARA A SEQ ID NO: 9:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 267 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: NÃO

(x1) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 9:

```
CAATTTCAGT ATGAAAGCCA GCTACAGATG GTACAGGTGA CCGGCTCCTC AGATAATGAG 60
TACTTCTACG TTGATTTCAG AGAATATGAA TATGAICTCA AATGGGAGTT TCCAAGAGAA 120
AATTGCCACA AATGGGAGTT TCCAAGAGAA AATTAGACT TTGGGAGGT ACTAGCATCA 180
GTTGCTTTTG GAAAACIGAT GAACGCAACA GCTTATGGAA TTAGCAAAAC AGGAGTCTCA 240
ATCCAGGTTG CCGTCABAAT GGTGAAA 267
```

(2) INFORMAÇÃO PARA A SEQ ID NO: 10:

(1) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 276 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: NÃO

(x1) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 10:

```
CAATTTCAGT ATGAAAGCCA GCTACAGATG GTACAGGTGA CCGGCTCCTC AGATAATGAG 60
TACTTCTACG TTGATTTCAG AGGCTCTCTA GATAATGAGT ACTTCTACGT TGATTTCAGA 120
GAATAAGAAAT ATGATCTCTA ATGGGAGTTT CCAAGAGAAAT AATTAGACTT TGGGAGGTGA 180
CTAGCATCAG GTGCTTTTGG AAAAGTCATG AACGCAACAG GTTATGCAAT TAGCAFAACA 240
GGAGTCTCAA TCCAGGTTGC CGTCAAAATG CTGAAA 276
```

(2) INFORMAÇÃO PARA A SEQ ID NO: 11:

(1) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 386 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genômico)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 11:

```
CAATTGAGT ATGAAAGCCA GCTACAGATG GTACAGGTGA CCGGCTCCTC AGATAATGAG 60
TACTTCTACG TTGAAAGCCA GCTACAGATG GTACAGGTGA CCGGCTCCTC AGATAGTGAG 120
TACTTCTACG TTGATTTCAG AGAATATGAA TATCATCTCA AATGGGAGTT TCCACAGAGG 180
AATTTAGAGT TTGGTACGAA TGGAAATGTG CAAATGTTTC TCGAGCATTT CTTTCCATT 240
GCAAAATCTT TBAATGCAC GTACTCACCA TTTTCTTTG CAGGGAAGGT ACTAGGATCA 300
GCTGCTTTTG GAAAAGTGAAT GAACCCACCA GCTTTGGAAT TAGCABAACA GGAGTCTCAA 360
TCCAGGTTCG CGTCAAATG CTGAAA 386
```

(2) INFORMAÇÃO PARA A SEQ ID NO: 12:

(1) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 480 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genômico)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 12:

CAATTAGGT ATGAAAGCCA GCTACAGATG GTACAGGTGA CCGGCTCCTC AGATAATGAG 60
TACTTCTACG TTGATTTGAG AGAATATGAA TATGATCTCA AATGGGAGTY TCCAAGAGAA 120
AAATTAGAGT TTGGTAAGAA TGGAAATGTC CRAATGTTTC TCCAGCATTY CTTTCCCAAT 180
GGAAAATCTT TAAATGACAC GTACTACCA TTTTCTTTTC CAGCGAGGCT ACTAGGATCC 240
GAATATGATC TCAAATGGGA GTTTCBAGA GAAATTTAG AGTTTGCTCA GATGGGAATG 300
TGCCAAATGT TTCTGCACCA TTTCTTTTCC ATGGGAAAAT CTTTAAATG CAGTACTCA 360
CCATTTCTCT TTGCAGGGA GGTACIAGGA TCAGGTGCTT TTGAAAAGT GATGAACGCA 420
ACAGCTTATG GAATTAGCAA AACAGGAGTC TCARTOCAGG TTGCTCTCAA AATGCTGAAA 480

(2) INFORMAÇÃO PARA A SEQ ID NO: 13:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 366 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genômico)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 13:

CAATTAGGT ATGAAAGCCA GCTACAGATG GTACAGGTGA CCGGCTCCTC AGATAATGAG 60
TACTTCTACG TTGATTTGAG AGAATATGAA TATGATCTCA AATGGGAGTY TCCAAGAGAA 120
GAATATGAGT TTGGTAAGAA TGGAAATGTC CRAATGTTTC TCCAGCATTY CTTTCCCAAT 180
GGAAATGTCG AATGCTTTCT GCAGCATTTC TTTTCTATTG GAAATCTTTT AAAATCCACG 240
TACTCACCAT TTGCTTTTGC AAGGAAGGTA CTAGGATCAG GTGCTTTTGG AAAACTGATG 300
AAGCCAACAG CTTATGGGAT TAGCBAACA GGAGTCTCAA TCCAGGTTGC CGTCAAAATG 360
CTGAAA 366

(2) INFORMAÇÃO PARA A SEQ ID NO: 14:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 357 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genômico)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 14:

```
CAATTTAGGT ATGAAAGCCA GGTACAGATG GTACAGGTGA CCGGCTCCTC AGATAATGAG 60
TTCCTCTACG TTGATTTCAG AGAATATGAA TATGATCTCA AATCGGAGTT TCCAAGAGAA 120
AATTGGCACA AATCGGAGTT TCCAAGAGAA AATTIAGAGT TTGGTAAGAA TGGAAATGTGC 180
CAAACGTTTC TGCAGCATTT CTTTTCATT GGAAATCTT TAAATGCAC GTACTCACCA 240
TTTGTCTTTC CAGCGAAGGT ACTAGCATCA CCGCTTTTTC GAAAAGTCTT GAACGCCACA 300
CCTTATGGAA TTACCAAAAC AGCACTCTCA ATCAGGTTG CCGTCAAAAT GCTCBA 357
```

(2) INFORMAÇÃO PARA A SEQ ID NO: 15:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 366 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADN (genômico)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 15:

```
CAATTTAGGT ATGAAAGCCA GGTACAGATG GTACAGGTGA CCGGCTCCTC AGATAATGAG 60
TACTTCTACG TTGATTTCAG AGGCTCCTCA GATAATGAGT ACTTCTACGT TGATTTCAGA 120
```

GAATATGAAT ATGATCTCAA ATGGGAGTTT CCAAGAGAAA ATTACAGTT TGGTAAQAAAT 180
 GGAATSTGCC AATGTCTCT GCAGGCTTTC TTTCCTATC GAAATCTTT AATATGCACG 240
 TACTCACCAT TTGTCTTTC AGCGAAGCTA CTAGGATCAG GTCCTTTTGG AAAATGCAATG 300
 AATGCAACAG CTATCGAAT TAGCATAACA GGAGTCTCAA TCCAGGTTGC CTCGAAATG 360
 CTGAAA 366

(2) INFORMAÇÃO PARA A SEQ ID NO: 15:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 665 aminoácidos
- (B) TIPO: aminoácido
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 16:

Met	Pro	Ala	Leu	Ala	Arg	Asp	Ala	Gli	Tre	Val	Pro	Leu	Leu	Val	Val
1			5					10					15		
Phe	Ser	Ala	Met	Ile	Phe	Gli	Tre	Ile	Tre	Asn	Gln	Asp	Leu	Pro	Val
			20					25					30		
Ile	Lis	Cis	Val	Leu	Ile	Asn	His	Lis	Asn	Asn	Asp	Ser	Ser	Val	Gli
			35					40					45		
Lis	Ser	Ser	Ser	Tir	Pro	Met	Val	Ser	Glu	Ser	Pro	Glu	Asp	Leu	Gli
			50				55					60			
Cis	Ala	Leu	Arg	Pro	Gln	Ser	Ser	Gli	Tre	Val	Tir	Glu	Ala	Ala	Ala
65				70						75				80	
Val	Glu	Val	Asp	Val	Ser	Ala	Ser	Ile	Tre	Leu	Gln	Val	Leu	Val	Asp
			85							90				95	
Ala	Pro	Gli	Asn	Ile	Ser	Cis	Leu	Trp	Val	Phe	Lis	His	Ser	Ser	Leu
			100					105						110	

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

Tir	Ser	Ile	Ser	Lis	Fen	Cis	Asn	His	Lis	His	Gln	Pro	Gli	Glu	Tir	405	410	415
Ile	Fen	His	Ala	Glu	Asn	Asp	Asp	Ala	Gln	Fen	Tre	Lis	Met	Fen	Tre	420	425	430
Leu	Asn	Ile	Arg	Arg	Lis	Pro	Gln	Val	Leu	Ala	Glu	Ala	Ser	Ala	Ser	435	440	445
Gln	Ala	Ser	Cis	Fen	Ser	Asp	Gli	Tir	Pro	Leu	Pro	Ser	Trp	Tre	Trp	450	455	460
Lis	Lis	Cis	Ser	Asp	Lis	Ser	Pro	Asn	Cis	Tre	Glu	Glu	Ile	Tre	Glu	465	470	475
Gli	Val	Trp	Asn	Arg	Lis	Ala	Asn	Arg	Lis	Val	Fen	Gli	Gln	Trp	Val	485	490	495
Ser	Ser	Ser	Tre	Leu	Asn	Met	Ser	Glu	Ala	Ile	Lis	Gli	Fen	Leu	Val	500	505	510
Lis	Cis	Cis	Ala	Tir	Asn	Ser	Leu	Gli	Tre	Ser	Cis	Glu	Tre	Ile	Leu	515	520	525
Leu	Asn	Ser	Pro	Gli	Pro	Fen	Pro	Fen	Ile	Gln	Asp	Asn	Ile	Ser	Fen	530	535	540
Tir	Ala	Tre	Ile	Gli	Val	Cis	Leu	Leu	Fen	Ile	Val	Val	Leu	Tre	Leu	545	550	555
Leu	Ile	Cis	His	Lis	Tir	Lis	Lis	Gln	Fen	Arg	Tir	Glu	Ser	Gln	Leu	565	570	575
Gln	Met	Val	Gln	Val	Tre	Gli	Ser	Ser	Asp	Asn	Glu	Tir	Fen	Tir	Val	580	585	590
Glu	Ser	Gln	Leu	Gln	Met	Val	Gln	Val	Tre	Gli	Ser	Ser	Asp	Ser	Glu	595	600	605
Tir	Fen	Tir	Val	Asp	Fen	Arg	Glu	Tir	Glu	Tir	Asp	Leu	Lis	Trp	Glu	610	615	620
Fen	Pro	Arg	Glu	Asn	Leu	Glu	Fen	Gli	Lis	Val	Leu	Gli	Ser	Gli	Ala	625	630	635
Fen	Gli	Lis	Val	Met	Asn	Ala	Tre	Ala	Leu	Glu	Leu	Ala	Lis	Gln	Glu	645	650	655
Ser	Gln	Ser	Arg	Leu	Pro	Ser	Lis	Cis								660	665	

(2) INFORMAÇÃO PARA A SEQ ID NO: 17:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 994 aminoácidos
- (B) TIPO: aminoácido
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: SIM

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 17:

Met	Pro	Ala	Leu	Ala	Arg	Asp	Ala	Gli	Tre	Val	Pro	Leu	Leu	Val	Val
1			5					10					15		
Pen	Ser	Ala	Met	Ile	Pen	Gli	Tre	Ile	Tre	Asn	Gln	Asp	Leu	Pro	Val
			20				25					30			
Ile	Lis	Cis	Val	Leu	Ile	Asn	His	Lis	Asn	Asn	Asp	Ser	Ser	Val	Gli
		35					40				45				
Lis	Ser	Ser	Ser	Tir	Pro	Met	Val	Ser	Glu	Ser	Pro	Glu	Asp	Leu	Gli
	50					55					60				
Cis	Ala	Leu	Arg	Pro	Gln	Ser	Ser	Gli	Tre	Val	Tir	Glu	Ala	Ala	Ala
65					70					75				80	
Val	Glu	Val	Asp	Val	Ser	Ala	Ser	Ile	Tre	Leu	Gln	Val	Leu	Val	Asp
			85						90				95		
Ala	Pro	Gli	Asn	Ile	Ser	Cis	Leu	Trp	Val	Pen	Lis	His	Ser	Ser	Leu
			100					105					110		
Asn	Cis	Gln	Pro	His	Pen	Asp	Leu	Gln	Asn	Arg	Gli	Val	Val	Ser	Met
		115					120				125				
Val	Ile	Leu	Lis	Met	Tre	Glu	Tre	Gln	Ala	Gli	Glu	Tir	Leu	Leu	Pen
	130					135					140				
Ile	Gln	Ser	Glu	Ala	Tre	Asn	Tir	Tre	Ile	Leu	Pen	Tre	Val	Ser	Ile
145				150					155					160	
Arg	Asn	Tre	Leu	Leu	Tir	Tre	Leu	Arg	Arg	Pro	Tir	Pen	Arg	Lis	Met
			165					170					175		
Glu	Asn	Gln	Asp	Ala	Leu	Val	Cis	Ile	Ser	Glu	Ser	Val	Pro	Glu	Pro
			180					185					190		

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

Gli	Val	Trp	Asn	Arg	Lis	Ala	Asn	Arg	Lis	Val	Pen	Gli	Gln	Trp	Val
				485					490					495	
Ser	Ser	Ser	Tre	Leu	Asn	Met	Ser	Glu	Ala	Ile	Lis	Gli	Pen	Leu	Val
				500				505					510		
Lis	Cis	Cis	Ala	Tir	Asn	Ser	Leu	Gli	Tre	Ser	Cis	Glu	Tre	Ile	Leu
				515			520					525			
Leu	Asn	Ser	Pro	Gli	Pro	Pen	Pro	Pen	Ile	Gln	Asp	Asn	Ile	Ser	Pen
				530			535				540				
Tir	Ala	Tre	Ile	Gli	Val	Cis	Leu	Leu	Pen	Ile	Val	Val	Leu	Tre	Leu
545					550					555					560
Leu	Ile	Cis	His	Lis	Tir	Lis	Lis	Gln	Pen	Arg	Tir	Glu	Ser	Gln	Leu
				565					570					575	
Gln	Met	Val	Gln	Val	Tre	Gli	Ser	Ser	Asp	Asn	Glu	Tir	Pen	Tir	Val
				580					585				590		
Asp	Pen	Arg	Glu	Tir	Glu	Tir	Asp	Leu	Lis	Trp	Glu	Pen	Pro	Arg	Glu
				595			600					605			
Asn	Leu	Glu	Pen	Gli	Lis	Val	Leu	Gli	Ser	Glu	Tir	Asp	Leu	Lis	Trp
				610			615				620				
Glu	Pen	Pro	Arg	Glu	Asn	Leu	Glu	Pen	Gli	Lis	Val	Leu	Gli	Ser	Gli
625					630					635					640
Ala	Pen	Gli	Lis	Val	Met	Asn	Ala	Tre	Ala	Tir	Gli	Ile	Ser	Lis	Tre
				645					650					655	
Gli	Val	Ser	Ile	Gln	Val	Ala	Val	Lis	Met	Leu	Lis	Glu	Lis	Ala	Asp
				660					665					670	
Ser	Ser	Glu	Arg	Glu	Ala	Leu	Met	Ser	Glu	Leu	Lis	Met	Met	Tre	Gln
				675			680					685			
Leu	Gli	Ser	His	Glu	Asn	Ile	Val	Asn	Leu	Leu	Gli	Ala	Cis	Tre	Leu
				690			695				700				
Ser	Gli	Pro	Ile	Tir	Leu	Ile	Pen	Glu	Tir	Cis	Cis	Tir	Gli	Asp	Leu
705					710					715					720
Leu	Asn	Tir	Leu	Arg	Ser	Lis	Arg	Glu	Lis	Pen	His	Arg	Tre	Trp	Tre
				725					730					735	
Glu	Ile	Pen	Lis	Glu	His	Asn	Pen	Ser	Pen	Tir	Pro	Tre	Pen	Gln	Ser
				740					745					750	
His	Pro	Asn	Ser	Ser	Met	Pro	Gli	Ser	Arg	Glu	Val	Gln	Ile	His	Pro
				755			760							765	

Asp Ser Asp cin iie Ser gii Ien His gii Asn Ser Ien His Ser cin	770	775	780
Asp cin iie cin iie cin Asn cin Iis Arg Ien cin cin cin Asp	785	790	795
Ien Asn Val Ien Iie Ien cin Asp Ien cin Ien cin Iie cin Val	805	810	815
Aia Iis gii Met cin Ien cin Ien His Ser cin Val His Arg Asp	820	825	830
Ien Aia Aia Arg Asn Val Ien Val Iie His gii Iis Val Iis Iie	835	840	845
Cis Asp Ien gii Ien Aia Arg Asp Iie Met Ser Asp Ser Asn Iie Val	850	855	860
Val Arg gii Asn Aia Arg Ien Pro Val Iis Iie Met Aia Pro cin Ser	865	870	875
Ien Ien cin gii Iie Iie Iie Iie Ser Asp Val Iie Ser Iie gii	885	890	895
Iie Ien Ien Iie cin Iie Ien Ser Ien gii Val Asn Pro Iie Pro gii	900	905	910
Iie Pro Val Asp Aia Asn Ien Iie Iis Ien cin Asp gii Ien Iis	915	920	925
Met Asp cin Pro Ien Iie Aia Iie cin Iie Iie Iie Iie Met cin	930	935	940
Ser cis Iie Aia Ien Asp Ser Arg Iis Arg Pro Ser Ien Pro Asn Ien	945	950	955
Iie Ser Ien Ien gii cis Ien Ien Aia Asp Aia cin Aia Met Iie	965	970	975
Cin Asn Val Asp gii Arg Val Ser cin cis Pro His Iie Iie cin Asn	980	985	990

(2) INFORMAÇÃO PARA A SEQ ID NO: 18:

- (1) CARACTERÍSTICAS DA SEQUÊNCIA:
- (A) COMPRIMENTO: 986 aminoácidos
 - (B) TIPO: aminoácido
 - (C) ESTRUTURA SECUNDÁRIA: simples
 - (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: SIM

(xii) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 18:

```
Met Pro Ala Leu Ala Arg Asp Ala Gln Tre Val Pro Leu Leu Val Val
1          5          10          15
Phe Ser Ala Met Ile Phe Gln Tre Ile Tre Asn Gln Asp Leu Pro Val
          20          25          30
Ile Lis Cys Val Leu Ile Asn His Lis Asn Asn Asp Ser Ser Val Gln
          35          40          45
Lis Ser Ser Ser Thr Pro Met Val Ser Glu Ser Pro Glu Asp Leu Gln
          50          55          60
Cys Ala Leu Arg Pro Gln Ser Ser Gln Tre Val Thr Glu Ala Ala Ala
65          70          75          80
Val Gln Val Asp Val Ser Ala Ser Ile Tre Leu Gln Val Leu Val Asp
          85          90          95
Ala Pro Gln Asn Ile Ser Cys Leu Trp Val Phe Lis His Ser Ser Leu
          100          105          110
Asn Cys Gln Pro His Phe Asp Leu Gln Asn Arg Gln Val Val Ser Met
          115          120          125
Val Ile Leu Lis Met Thr Glu Thr Gln Ala Gln Glu Thr Leu Leu Phe
          130          135          140
Ile Gln Ser Glu Ala Thr Asn Thr Thr Ile Leu Phe Thr Val Ser Ile
          145          150          155          160
Arg Asn Thr Leu Leu Thr Thr Leu Arg Arg Pro Thr Phe Arg Lis Met
          165          170          175
Glu Asn Gln Asp Ala Leu Val Cys Ile Ser Glu Ser Val Pro Glu Pro
          180          185          190
Ile Val Glu Trp Val Leu Cys Asp Ser Gln Gln Glu Ser Cys His Glu
          195          200          205
Glu Ser Pro Ala Val Val Lis Lis Glu Glu Lis Val Leu His Gln Leu
          210          215          220
Phe Gln Thr Asp Ile Arg Cys Cys Ala Arg Asn Glu Leu Gln Arg Glu
          225          230          235          240
```

Cis	Tre	Arg	Leu	Pen	Tre	Ile	Asp	Leu	Asn	Gln	Tre	Pro	Gln	Tre	Tre
				245					250					255	
Leu	Pro	Gln	Leu	Pen	Leu	Lis	Val	Gli	Glu	Pro	Leu	Trp	Ile	Arg	Cis
			260					265					270		
Lis	Ala	Val	His	Val	Asn	His	Gli	Pen	Gli	Leu	Tre	Trp	Glu	Leu	Glu
		275					280						285		
Asn	Lis	Ala	Leu	Glu	Glu	Gli	Asn	Tir	Pen	Glu	Met	Ser	Tre	Tir	Ser
	290					295					300				
Tre	Asn	Arg	Tre	Met	Ile	Arg	Ile	Leu	Pen	Ala	Pen	Val	Ser	Ser	Val
305				310						315					320
Ala	Arg	Asn	Asp	Tre	Gli	Tir	Tir	Tre	Cis	Ser	Ser	Ser	Lis	His	Pro
			325					330					335		
Ser	Gln	Ser	Ala	Leu	Val	Tre	Ile	Val	Gli	Lis	Gli	Pen	Ile	Asn	Ala
	340							345					350		
Tre	Asn	Ser	Ser	Glu	Asp	Tir	Glu	Ile	Asp	Gln	Tir	Glu	Glu	Pen	Cis
	355						360					365			
Pen	Ser	Val	Arg	Pen	Lis	Ala	Tir	Pro	Gln	Ile	Arg	Cis	Tre	Trp	Tre
	370					375					380				
Pen	Ser	Arg	Lis	Ser	Pen	Pro	Cis	Glu	Gln	Lis	Gli	Leu	Asp	Asn	Gli
385				390						395					400
Tir	Ser	Ile	Ser	Lis	Pen	Cis	Asn	His	Lis	His	Gln	Pro	Gli	Glu	Tir
			405						410				415		
Ile	Pen	His	Ala	Glu	Asn	Asp	Asp	Ala	Gln	Pen	Tre	Lis	Met	Pen	Tre
		420						425					430		
Leu	Asn	Ile	Arg	Arg	Lis	Pro	Gln	Val	Leu	Ala	Glu	Ala	Ser	Ala	Ser
	435						440					445			
Gln	Ala	Ser	Cis	Pen	Ser	Asp	Gli	Tir	Pro	Leu	Pro	Ser	Trp	Tre	Trp
	450					455						460			
Lis	Lis	Cis	Ser	Asp	Lis	Ser	Pro	Asn	Cis	Tre	Glu	Glu	Ile	Tre	Glu
465				470						475				480	
Gli	Val	Trp	Asn	Arg	Lis	Ala	Asn	Arg	Lis	Val	Pen	Gli	Gln	Trp	Val
			485					490					495		
Ser	Ser	Ser	Tre	Leu	Asn	Met	Ser	Glu	Ala	Ile	Lis	Gli	Pen	Leu	Val
		500						505					510		
Lis	Cis	Cis	Ala	Tir	Asn	Ser	Leu	Gli	Tre	Ser	Cis	Glu	Tre	Ile	Leu
	515						520						525		

Leu Asn Ser Pro Glt Pro Fen Pro Fen Ite Gln Asp Asn Ite Ser Fen	530	535	540
Tir Ala Tre Ite Glt Val Cts Leu Leu Ite Val Leu Tre Leu	545	550	555
Leu Ite Cts His Tir His Ite Gln Fen Arg Tir Gln Ser Gln Leu	565	570	575
Gln Met Val Gln Val Tre Glt Ser Ser Asp Asn Gln Tir Fen Tir Val	580	585	590
Asp Fen Arg Gln Tir Gln Tir Asp Leu His Trp Gln Fen Asp Fen Arg	595	600	605
Gln Tir Gln Tir Asp Leu His Trp Gln Fen Pro Arg Gln Asn Leu Gln	610	615	620
Fen Glt His Val Leu Glt Ser Glt Ala Fen Glt Tir Val Met Asn Ala	625	630	635
Tre Ala Tir Glt Ite Ser His Tre Glt Val Ser Ite Gln Val Ala Val	645	650	655
His Met Fen His Gln His Ala Asp Ser Ser Gln Arg Gln Ala Leu Met	660	665	670
Ser Gln Leu Tir Met Met Tre Gln Leu Glt Ser His Gln Asn Ite Val	675	680	685
Asn Leu Leu Glt Ala Cts Tre Leu Ser Glt Pro Tir Tre Ite Fen	690	695	700
Gln Tir Cts Cts Glt Asp Leu Leu Asn Tir Leu Arg Ser His Arg	705	710	715
Gln His Fen His Arg Tre Trp Tre Glt Ite Fen His Gln His Asn Fen	725	730	735
Ser Fen Tir Pro Tre Fen Gln Ser His Pro Asn Ser Ser Met Pro Glt	740	745	750
Ser Arg Gln Val Gln Ite His Pro Asp Ser Asp Gln Ite Ser Glt Leu	755	760	765
His Glt Asn Ser Fen His Ser Gln Asp Gln Ite Gln Tir Gln Asn Gln	770	775	780
Ite Arg Leu Gln Gln Gln Asp Leu Asn Val Leu Fen Gln Asp	785	790	795
Leu Leu Cts Ser Ala Tir Gln Val His Glt Met Gln Fen Leu Gln	805	810	815

Phe	Lis	Ser	Cis	Val	His	Arg	Asp	Leu	Ala	Ala	Arg	Asn	Val	Leu	Val
				820				825					830		
Tre	His	Gli	Lis	Val	Val	Lis	Ile	Cis	Asp	Phe	Gli	Leu	Ala	Arg	Asp
				835				840					845		
Ile	Met	Ser	Asp	Ser	Asn	Tir	Val	Val	Arg	Gli	Asn	Ala	Arg	Leu	Pro
				850				855					860		
Val	Lis	Trp	Met	Ala	Pro	Glu	Ser	Leu	Phe	Glu	Gli	Ile	Tir	Tre	Ile
				865				870					875		880
His	Ser	Asp	Val	Trp	Ser	Tir	Gli	Ile	Leu	Leu	Trp	Glu	Ile	Phe	Ser
				885					890					895	
Leu	Gli	Val	Asn	Pro	Tir	Pro	Gli	Ile	Pro	Val	Asp	Ala	Asn	Phe	Tir
				900				905					910		
Lis	Leu	Ile	Gln	Asn	Gli	Phe	Lis	Met	Asp	Gln	Pro	Phe	Tir	Ala	Tre
				915				920					925		
Gln	Glu	Ile	Tir	Ile	Ile	Met	Gln	Ser	Cis	Trp	Ala	Phe	Asp	Ser	Arg
				930				935					940		
Lis	Arg	Pro	Ser	Phe	Pro	Asn	Leu	Tre	Ser	Phe	Leu	Gli	Cis	Gln	Leu
				945				950					955		960
Ala	Asp	Ala	Glu	Glu	Ala	Met	Tir	Gln	Asn	Val	Asp	Gli	Arg	Val	Ser
				965					970					975	
Glu	Cis	Pro	His	Tre	Tir	Gln	Asn	Arg	Arg						
				980					985						

(2) INFORMAÇÃO PARA A SEQ ID NO: 19:

- (i) CARACTERÍSTICAS DA SEQUÊNCIA:
- (A) COMPRIMENTO: 983 aminoácidos
 - (B) TIPO: aminoácido
 - (C) ESTRUTURA HELICOIDAL: simples
 - (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: SIM

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 19:

Met Pro Ala Leu Ala Arg Asp Ala Gli Trp Val Pro Leu Leu Val Val
1 5 10 15
Phe Ser Ala Met Ile Phe Glu Trp Ile Trp Asn Gln Asp Leu Pro Val
20 25 30
Ile His Cys Val Leu Ile Asn His His Asn Asn Asp Ser Ser Val Gli
35 40 45
His Ser Ser Ser Thr Pro Met Val Ser Glu Ser Pro Glu Asp Leu Gli
50 55 60
Cys Ala Leu Arg Pro Gln Ser Ser Glu Trp Val Thr Glu Ala Ala Ala
65 70 75 80
Val Glu Val Asp Val Ser Ala Ser Ile Trp Leu Gln Val Leu Val Asp
85 90 95
Ala Pro Glu Asn Ile Ser Cys Leu Trp Val Phe His His Ser Ser Leu
100 105 110
Asn Cys Gln Pro His Phe Asp Leu Gln Asn Arg Glu Val Val Ser Met
115 120 125
Val Ile Leu His Met Thr Glu Thr Gln Ala Glu Glu Thr Leu Leu Phe
130 135 140
Ile Gln Ser Glu Ala Thr Asn Thr Thr Ile Leu Phe Thr Val Ser Ile
145 150 155 160
Arg Asn Thr Leu Leu Thr Thr Leu Arg Arg Pro Thr Phe Arg His Met
165 170 175
Glu Asn Gln Asp Ala Leu Val Cys Ile Ser Glu Ser Val Pro Glu Pro
180 185 190
Ile Val Glu Trp Val Leu Cys Asp Ser Gln Gly Glu Ser Cys His Glu
195 200 205
Glu Ser Pro Ala Val Val His His Glu Glu His Val Leu His Glu Leu
210 215 220
Phe Glu Thr Asp Ile Arg Cys Cys Ala Arg Asn Glu Leu Glu Arg Glu
225 230 235 240
Cys Thr Arg Leu Phe Thr Ile Asp Leu Asn Gln Thr Pro Gln Thr Thr
245 250 255
Leu Pro Gln Leu Phe Leu His Val Glu Glu Pro Leu Thr Ile Arg Cys
260 265 270
His Ala Val His Val Asn His Glu Phe Glu Leu Thr Thr Glu Leu Glu
275 280 285

Asn	Lis	Ala	Leu	Glu	Glu	Gli	Asn	Tir	Fen	Glu	Met	Ser	Tre	Tir	Ser
290							295					300			
Tre	Asn	Arg	Tre	Met	Ile	Arg	Ile	Leu	Fen	Ala	Fen	Val	Ser	Ser	Val
305					310					315					320
Ala	Arg	Asn	Asp	Tre	Gli	Tir	Tir	Tre	Cis	Ser	Ser	Ser	Lis	His	Pro
				325					330					335	
Ser	Gln	Ser	Ala	Leu	Val	Tre	Ile	Val	Gli	Lis	Gli	Fen	Ile	Asn	Ala
			340						345					350	
Tre	Asn	Ser	Ser	Glu	Asp	Tir	Glu	Ile	Asp	Gln	Tir	Glu	Glu	Fen	Cis
			355					360						365	
Fen	Ser	Val	Arg	Fen	Lis	Ala	Tir	Pro	Gln	Ile	Arg	Cis	Tre	Trp	Tre
			370					375						380	
Fen	Ser	Arg	Lis	Ser	Fen	Pro	Cis	Glu	Gln	Lis	Gli	Leu	Asp	Asn	Gli
385					390						395				400
Tir	Ser	Ile	Ser	Lis	Fen	Cis	Asn	Ris	Lis	His	Gln	Pro	Gli	Gln	Tir
					405						410				415
Ile	Fen	His	Ala	Glu	Asn	Asp	Asp	Ala	Gln	Fen	Tre	Lis	Met	Fen	Tre
				420					425					430	
Leu	Asn	Ile	Arg	Arg	Lis	Pro	Gln	Val	Leu	Ala	Glu	Ala	Ser	Ala	Ser
				435					440					445	
Gln	Ala	Ser	Cis	Fen	Ser	Asp	Gli	Tir	Pro	Leu	Pro	Ser	Trp	Tre	Trp
				450					455					460	
Lis	Lis	Cis	Ser	Asp	Lis	Ser	Pro	Asn	Cis	Tre	Glu	Gln	Ile	Tre	Glu
465							470							475	480
Gli	Val	Trp	Asn	Arg	Lis	Ala	Asn	Arg	Lis	Val	Fen	Gli	Gln	Trp	Val
					485						490				495
Ser	Ser	Ser	Tre	Leu	Asn	Met	Ser	Glu	Ala	Ile	Lis	Gli	Fen	Leu	Val
				500					505					510	
Lis	Cis	Cis	Ala	Tir	Asn	Ser	Leu	Gli	Tre	Ser	Cis	Glu	Tre	Ile	Leu
				515					520					525	
Leu	Asn	Ser	Pro	Gli	Pro	Fen	Pro	Fen	Ile	Gln	Asp	Asn	Ile	Ser	Fen
				530					535					540	
Tir	Ala	Tre	Ile	Gli	Val	Cis	Leu	Leu	Fen	Ile	Val	Val	Leu	Tre	Leu
545							550						555		560
Leu	Ile	Cis	His	Lis	Tir	Lis	Lis	Gln	Fen	Arg	Tir	Glu	Ser	Gln	Leu
					565						570				575

Gln Met Val Gln Val Tre Gli Ser Ser Asp Asn Glu Tir Fen Tir Val			
580	585	590	
Asp Fen Arg Glu Tir Glu Tir Asp Leu Lis Trp Glu Fen Pro Arg Glu			
595	600	605	
Asn Trp His Lis Trp Glu Fen Pro Arg Glu Asn Leu Glu Fen Gli Lis			
610	615	620	
Val Leu Gli Ser Gli Ala Fen Gli Lis Val Met Asn Ala Tre Ala Tir			
625	630	635	640
Gli Ile Ser Lis Tre Gli Val Ser Ile Gln Val Ala Val Lis Met Leu			
645	650	655	
Lis Glu Lis Ala Asp Ser Ser Glu Arg Glu Ala Leu Met Ser Glu Leu			
660	665	670	
Lis Met Met Tre Gln Leu Gli Ser His Glu Asn Ile Val Asn Leu Leu			
675	680	685	
Gli Ala Cis Tre Leu Ser Gli Pro Ile Tir Leu Ile Fen Glu Tir Cis			
690	695	700	
Cis Tir Gli Asp Leu Leu Asn Tir Leu Arg Ser Lis Arg Glu Lis Fen			
705	710	715	720
His Arg Tre Trp Tre Glu Ile Fen Lis Glu His Asn Fen Ser Fen Tir			
725	730	735	
Pro Tre Fen Gln Ser His Pro Asn Ser Ser Met Pro Gli Ser Arg Glu			
740	745	750	
Val Gln Ile His Pro Asp Ser Asp Gln Ile Ser Gli Leu His Gli Asn			
755	760	765	
Ser Fen His Ser Glu Asp Glu Ile Glu Tir Glu Asn Gln Lis Arg Leu			
770	775	780	
Glu Glu Glu Glu Asp Leu Asn Val Leu Tre Fen Glu Asp Leu Leu Cis			
785	790	795	800
Fen Ala Tir Gln Val Ala Lis Gli Met Glu Fen Leu Glu Fen Lis Ser			
805	810	815	
Cis Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val Tre His Gli			
820	825	830	
Lis Val Val Lis Ile Cis Asp Fen Gli Leu Ala Arg Asp Ile Met Ser			
835	840	845	
Asp Ser Asn Tir Val Val Arg Gli Asn Ala Arg Leu Pro Val Lis Trp			
850	855	860	

Met	Ala	Pro	Glu	Ser	Leu	Fen	Glu	Gli	Ile	Tir	Tre	Ile	Lis	Ser	Asp
865					870					875					880
Val	Trp	Ser	Tir	Gli	Ile	Leu	Leu	Trp	Glu	Ile	Fen	Ser	Leu	Gli	Val
				885					890					895	
Asn	Pro	Tir	Pro	Gli	Ile	Pro	Val	Asp	Ala	Asn	Fen	Tir	Lis	Leu	Ile
			900					905					910		
Gln	Asn	Gli	Fen	Lis	Met	Asp	Gln	Pro	Fen	Tir	Ala	Tre	Glu	Glu	Ile
	915						920					925			
Tir	Ile	Ile	Met	Gln	Ser	Cis	Trp	Ala	Fen	Asp	Ser	Arg	Lis	Arg	Pro
	930						935					940			
Ser	Fen	Pro	Asn	Leu	Tre	Ser	Fen	Leu	Gli	Cis	Gln	Leu	Ala	Asp	Ala
945				950					955					960	
Glu	Glu	Ala	Met	Tir	Gln	Asn	Val	Asp	Gli	Arg	Val	Ser	Glu	Cis	Pro
				965				970						975	
His	Tre	Tir	Gln	Asn	Arg	Arg									
							980								

(2) INFORMAÇÃO PARA A SEQ ID NO: 20:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 986 aminoácidos
- (B) TIPO: aminoácido
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: péptido

(iii) HIPOTÉTICA: SIM

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 20:

Met	Pro	Ala	Leu	Ala	Arg	Asp	Ala	Gli	Tre	Val	Pro	Leu	Leu	Val	Val
2				5					10					15	
Fen	Ser	Ala	Met	Ile	Fen	Gli	Tre	Ile	Tre	Asn	Gln	Asp	Leu	Pro	Val
				20				25					30		

Ile	Lis	Cis	Val	Leu	Ile	Asn	His	Lis	Asn	Asn	Asp	Ser	Ser	Val	Gli	35	40	45
Lis	Ser	Ser	Ser	Tir	Pro	Met	Val	Ser	Glu	Ser	Pro	Glu	Asp	Leu	Gli	50	55	60
Cis	Ala	Leu	Arg	Pro	Gln	Ser	Ser	Gli	Tre	Val	Tir	Glu	Ala	Ala	Ala	65	70	75
Val	Glu	Val	Asp	Val	Ser	Ala	Ser	Ile	Tre	Leu	Gln	Val	Leu	Val	Asp	85	90	95
Ala	Pro	Gli	Asn	Ile	Ser	Cis	Leu	Trp	Val	Fen	Lis	His	Ser	Ser	Leu	100	105	110
Asn	Cis	Gln	Pro	His	Fen	Asp	Leu	Gln	Asn	Arg	Gli	Val	Val	Ser	Met	115	120	125
Val	Ile	Leu	Lis	Met	Tre	Glu	Tre	Gln	Ala	Gli	Glu	Tir	Leu	Leu	Fen	130	135	140
Ile	Gln	Ser	Glu	Ala	Tre	Asn	Tir	Tre	Ile	Leu	Fen	Tre	Val	Ser	Ile	145	150	155
Arg	Asn	Tre	Leu	Leu	Tir	Tre	Leu	Arg	Arg	Pro	Tir	Fen	Arg	Lis	Met	165	170	175
Glu	Asn	Gln	Asp	Ala	Leu	Val	Cis	Ile	Ser	Glu	Ser	Val	Pro	Glu	Pro	180	185	190
Ile	Val	Glu	Trp	Val	Leu	Cis	Asp	Ser	Gln	Gli	Glu	Ser	Cis	Lis	Glu	195	200	205
Glu	Ser	Pro	Ala	Val	Val	Lis	Lis	Glu	Glu	Lis	Val	Leu	His	Glu	Leu	210	215	220
Fen	Gli	Tre	Asp	Ile	Arg	Cis	Cis	Ala	Arg	Asn	Glu	Leu	Gli	Arg	Glu	225	230	235
Cis	Tre	Arg	Leu	Fen	Tre	Ile	Asp	Leu	Asn	Gln	Tre	Pro	Gln	Tre	Tre	245	250	255
Leu	Pro	Gln	Leu	Fen	Leu	Lis	Val	Gli	Glu	Pro	Leu	Trp	Ile	Arg	Cis	260	265	270
Lis	Ala	Val	His	Val	Asn	His	Gli	Fen	Gli	Leu	Tre	Trp	Glu	Leu	Glu	275	280	285
Asn	Lis	Ala	Leu	Glu	Glu	Gli	Asn	Tir	Fen	Glu	Met	Ser	Tre	Tir	Ser	290	295	300
Tre	Asn	Arg	Tre	Met	Ile	Arg	Ile	Leu	Fen	Ala	Fen	Val	Ser	Ser	Val	305	310	315
																		320

Ala Arg Asn Asp Tre Gli Tir Tir Tre Cis Ser Ser Ser Lis His Pro			
	325	330	335
Ser Gln Ser Ala Leu Val Tre Ile Val Gli Lis Gli Fen Ile Asn Ala			
	340	345	350
Tre Asn Ser Ser Glu Asp Tir Glu Ile Asp Gln Tir Glu Glu Fen Cis			
	355	360	365
Fen Ser Val Arg Fen Lis Ala Tir Pro Gln Ile Arg Cis Tre Trp Tre			
	370	375	380
Fen Ser Arg Lis Ser Fen Pro Cis Glu Gln Lis Gli Leu Asp Asn Gli			
385	390	395	400
Tir Ser Ile Ser Lis Fen Cis Asn His Lis His Gln Pro Gli Glu Tir			
	405	410	415
Ile Fen His Ala Glu Asn Asp Asp Ala Gln Fen Tre Lis Met Fen Tre			
	420	425	430
Leu Asn Ile Arg Arg Lis Pro Gln Val Leu Ala Glu Ala Ser Ala Ser			
	435	440	445
Gln Ala Ser Cis Fen Ser Asp Gli Tir Pro Leu Pro Ser Trp Tre Trp			
	450	455	460
Lis Lis Cis Ser Asp Lis Ser Pro Asn Cis Tre Glu Glu Ile Tre Glu			
465	470	475	480
Gli Val Trp Asn Arg Lis Ala Asn Arg Lis Val Fen Gli Gln Trp Val			
	485	490	495
Ser Ser Ser Tre Leu Asn Met Ser Glu Ala Ile Lis Gli Fen Leu Val			
	500	505	510
Lis Cis Cis Ala Tir Asn Ser Leu Gli Tre Ser Cis Glu Tre Ile Leu			
	515	520	525
Leu Asn Ser Pro Gli Pro Fen Pro Fen Ile Gln Asp Asn Ile Ser Fen			
	530	535	540
Tir Ala Tre Ile Gli Val Cis Leu Leu Fen Ile Val Val Leu Tre Leu			
545	550	555	560
Leu Ile Cis His Lis Tir Lis Lis Gln Fen Arg Tir Glu Ser Gln Leu			
	565	570	575
Gln Met Val Gln Val Tre Gli Ser Ser Asp Asn Glu Tir Fen Tir Val			
	580	585	590
Asp Fen Arg Gli Ser Ser Asp Asn Glu Tir Fen Tir Val Asp Fen Arg			
	595	600	605

Glu Tir Glu Tir Asp Leu Lis Trp Glu Fen Pro Arg Glu Asn Leu Glu

610

615

620

Fen Gli Lis Val Leu Gli Ser Gli Ala Fen Gli Lis Val Met Asn Ala

625

630

635

Tre Ala Tir Gli Ile Ser Lis Tre Gli Val Ser Ile Gln Val Ala Val

645

650

655

Lis Met Leu Lis Glu Lis Ala Asp Ser Ser Glu Arg Glu Ala Leu Met

660

665

670

Ser Glu Leu Lis Met Met Tre Gln Leu Gli Ser His Glu Asn Ile Val

675

680

685

Asn Leu Leu Gli Ala Cis Tre Leu Ser Gli Pro Ile Tir Leu Ile Fen

690

695

700

Glu Tir Cis Cis Tir Gli Asp Leu Leu Asn Tir Leu Arg Ser Lis Arg

705

710

715

Glu Lis Fen His Arg Tre Trp Tre Glu Ile Fen Lis Glu His Asn Fen

725

730

735

Ser Fen Tir Pro Tre Fen Gln Ser His Pro Asn Ser Ser Met Pro Gli

740

745

750

Ser Arg Glu Val Gln Ile His Pro Asp Ser Asp Gln Ile Ser Gli Leu

755

760

765

His Gli Asn Ser Fen His Ser Gln Asp Gln Ile Gln Tyr Gln Asn Gln

770

775

780

Lis Arg Leu Glu Glu Glu Glu Asp Leu Asn Val Leu Tre Fen Glu Asp

785

790

795

Leu Leu Cis Fen Ala Tir Gln Val Ala Lis Gli Met Glu Fen Leu Glu

805

810

815

Fen Lis Ser Cis Val His Arg Asp Leu Ala Ala Arg Asn Val Leu Val

820

825

830

Tre His Gli Lis Val Val Lis Ile Cis Asp Fen Gli Leu Ala Arg Asp

835

840

845

Ile Met Ser Asp Ser Asn Tir Val Val Arg Gli Asn Ala Arg Leu Pro

850

855

860

Val Lis Trp Met Ala Pro Glu Ser Leu Fen Gln Gli Ile Tir Tre Ile

865

870

875

Lis Ser Asp Val Trp Ser Tir Gli Ile Leu Leu Trp Glu Ile Fen Ser

885

890

895

Leu Gli Val Asn Pro Tir Pro Gli Ile Pro Val Asp Ala Asn Fen Tir
 900 905 910
 Lis Leu Ile Gln Asn Gli Fen Lis Met Asp Gln Pro Fen Tir Ala Tre
 915 920 925
 Glu Glu Ile Tir Ile Ile Met Gln Ser Cys Trp Ala Fen Asp Ser Arg
 930 935 940
 Lis Arg Pro Ser Fen Pro Asn Leu Tre Ser Fen Leu Gli Cys Gln Leu
 945 950 955 960
 Ala Asp Ala Glu Glu Ala Met Tir Gln Asn Val Asp Gli Arg Val Ser
 965 970 975
 Glu Cys Pro His Tre Tir Gln Asn Arg Arg
 980 985

(2) INFORMAÇÃO PARA A SEQ ID NO: 21:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 2978 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 21:

ATGCCGGCGT TGGCGCGCGA CGCGGGCACC GTGCCGCTGC TCGTTGTTTT TTCTGCAATG 60
 ATATTTGGGA CTATTACAAA TCAAGATCTG CCTGTGATCA AGTGTGTTTT AATCAATCAT 120
 AAGAACAATG ATTCATCAGT GGGGAAGTCA TCATCATATC CCATGGTATC AGAATCCCCG 180
 GAAGACCTCG GGTGTGCGTT GAGACCCAG AGCTCAGGGA CAGTGTACGA AGCTGCCGCT 240
 GTGGAAGTGG ATGTATCTGC TTCCATCACA CTGCAAGTGC TGGTCGATGC CCCAGGGAAC 300
 ATTTCTGTGC TCTGGGTCTT TAAGCACAGC TCCCTGAATT GCCAGCCACA TTTTGATTTA 360
 CAAAACAGAG GAGTTGTTTC CATGGTCATT TTGAAAATGA CAGAAACCCA AGCTGGAGAA 420
 TACCTACTTT TTATTACAGAG TGAAGCTACC AATTACACAA TATTGTTTAC AGTGAGTATA 480
 AGAAATACCC TGCTTTACAC ATTAAGAAGA CCTTACTTTA GAAAAATGGA AAACCAGGAC 540

GCCCTGGTCT GCATATCTGA GAGCGTTCCA GAGCCGATCG TGGAAATGGCT GCTTTGCGAT 600
 TCACAGGGGG AAAGCTGTAA AGAAGAAAGT CCAGCTGTTG TTAAAAAGGA GGAAAAAGTG 660
 CTTTCATGAAT TATTTGGGAC GGACATAAGG TGCTGTGCCA GAAATGAACT GGGCAGGGAA 720
 TGCACCAGGC TGTTCACAAT AGATCTAAAT CAAACTCCTC AGACCACATT GCCACAATTA 780
 TTTCTTAAAG TAGGGGAACC CTTATGGATA AGGTGCAAAG CTGTTTCATGT GAACCATGGA 840
 TTCGGGCTCA CCTGGGAATT AGAAAAACAA GCACTCGAGG AGGGCAACTA CTTTGAGATG 900
 AGIACCTATT CAACAAACAG AACTATGATA CGGATTCTGT TTGCTTTTGT ATCATCAGTG 960
 GCAAGAAACG ACACCCGATA CTACACTTGT TCCTCTTCAA AGCATCCCAG TCAATCAGCT 1020
 TTSGTTACCA TCGTAGGAAA GGGATTTATA AATGCTACCA ATTCAAGTGA AGATTATGAA 1080
 ATTGACCAAT ATGAAGAGTT TTGTTTTTCT GTCAGGTTTA AAGCCTACCC ACAAATCAGA 1140
 TGTACGTGGA CCTTCTCTCG AAAATCATTI CCTTGTGAGC AAAAGGGTCT TGATAACGGA 1200
 TACAGCATAT CCAAGTTTTG CAATCATAAG CACCAGCCAG GAGAATATAT ATTCCATGCA 1260
 GAAAATGATG ATGCCCAATT TACCAAAATG TTCACGCTGA ATATAAGAAG GAAACCTCAA 1320
 GTGCTCGCAG AAGCATCGGC AAGTCAGGCG TCCTGTTTCT CGGATGGATA CCCATTACCA 1380
 TCTTGGA CCT GGAAGAAGTG TTCAGACAAG TCTCCCAACT GCACAGAAGA GATCACAGAA 1440
 GGAGTCTGGA ATAGAAAGGC TAACAGAAAA GTGTTTGGAC AGTGGGTGTC GAGCAGTACT 1500
 CTAAACATGA GTGAAGCCAT AAAAGGGTTC CTGGTCAAGT GCTGTGCATA CAATTCCTTT 1560
 GGCACATCTT GTGAGACGAT CCTTTTAAAC TCTCCAGGCC CCTTCCCTTT CATCCAAGAC 1620
 AACATCTCAT TCTATGCAAC AATTGGTGTT TGTCTCCTCT TCATTGTCTG TTTAACCTTG 1680
 CTAATTTGTC ACAAGTACAA AAAGCAATTT AGGTATGAAA GCCAGCTACA GATGGTACAG 1740
 GTGACCGGCT CCTCAGATAA TGAGTACTTC TACGTTGAAA GCCAGCTACA GATGGTACAG 1800
 GTGACCGGCT CCTCAGATAG TGAGTACTTC TACGTTGATT TCAGAGAATA TGAATATGAT 1860
 CTCAAATGGG AGTTTCCAAG AGAAAATTTA GAGTTTGGGA AGGTACTAGG ATCAGGTGCT 1920
 TTTGGAAAAG TGATGAACGC AACAGCTTTG GAATTAGCAA AACAGGAGTC TCAATCCAGG 1980
 TTGCCGTCAA AATGCTGAAA GAAAAAGCAG ACAGCTCTGA AAGAGAGGCA CTCATGTCAG 2040
 AACTCAAGAT GATGACCCAG CTGGGAAGCC ACGAGAATAT TGTGAACCTG CTGGGGGCGT 2100
 GCACACTGTC AGGACCAATT TACTTGATTT TTGAATACTG TTGCTATGGT GATCTTCTCA 2160
 ACTATCTAAG AAGTAAAGA GAAAAATTTT ACAGGACTTG GACAGAGATT TTCAAGGAAC 2220
 ACAATTTTCA TTTTACCCTT ACTTTCCAAT CACATCCAAA TTCCAGCATG CCTGGTTCAA 2280
 GAGAAGTTCA GATACACCCG GACTCGGATC AAATCTCAGG GCTTCATGGG AATTCATTTT 2340
 ACTCTGAAGA TGAAATTGAA TATGAAAACC AAAAAAGGCT GGAAGAAGAG GAGGACTTGA 2400
 ATGTGCTTAC ATTTGAAGAT CTTCTTTGCT TTGCATATCA AGTTGCCAAA GGAATGGAAT 2460
 TTCTGGAATT TAAGTCGTGT GTTCACAGAG ACCTGGCCGC CAGGAACGTG CTTGTCAACC 2520
 ACGGGAAAGT GGTGAAGATA TGTGACTTTG GATTGGCTCG AGATATCATG AGTGATTCCA 2580
 ACTATGTTGT CAGGGGCAAT GCCCGTCTGC CTGTAAAATG GATGGCCCCC GAAAGCCTGT 2640
 TTGAAGGCAT CTACACCATT AAGAGTGATG TCTGGTCATA TGGAAATATTA CTGTGGGAAA 2700
 TCTTCTCACT TGGTGTGAAT CCTTACCCTG GCATTCCGGT TGATGCTAAC TTCTACAAAC 2760
 TGATTCAAAA TGGATTTAAA ATGGATCAGC CATTTTATGC TACAGAAGAA ATATACATTA 2820
 TAATGCAATC CTGCTGGGCT TTTGACTCAA GGAAACGGCC ATCCTTCCCT AATTGACTT 2880
 CGTTTTTAGG ATGTCAGCTG GCAGATGCAG AAGAAGCGAT GTATCAGAAT GTGGATGGCC 2940

GTGTTTCGGA ATGTCCTCAC ACCTACC AAA ACAGCCGA

2978

(2) INFORMAÇÃO PARA A SEQ ID NO: 22

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 2982 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 22:

ATGCCGGCGT TGGCGCGCGA CGCGGGCACC GTGCCGCTGC TCGTTGTTTT TTCTGCAATG 60
ATATTTGGGA CTATTACAAA TCAAGATCTG CCTGTGATCA AGTGTGTTTT AATCAATCAT 120
AAGAACAATG ATTCATCAGT GGGGAAGTCA TCATCATATC CCATGGTATC AGAATCCCCG 180
GAAGACCTCG GGTGTGCGTT GAGACCCCGAG AGCTCAGGGA CAGTGTACGA AGCTGCCGCT 240
GTGGAAGTGG ATGTATCTGC TTCCATCACA CTGCAAGTGC TGGTCGATGC CCCAGGGAAC 300
ATTTCTGTGC TCTGGGTCTT TAAGCACAGC TCCCTGAATT GCCAGCCACA TTTTGATTTA 360
CAAAACAGAG GAGTTGTTTC CATGGTCATT TTGAAAATGA CAGAAACCCA AGCTGGAGAA 420
TACCTACTTT TTATTCAGAG TGAAGCTACC AATTACACAA TATTGTTTAC AGTGAGTATA 480
AGAAATACCC TGCTTTACAC ATTAAGAAGA CCTTACTTTA GAAAAATGGA AAACCAGGAC 540
GCCCTGGTCT GCATATCTGA GAGCGTTCCA GAGCCGATCG TGGAATGGGT GCTTTGCGAT 600
TCACAGGGGG AAAGCTGTAA AGAAGAAAGT CCAGCTGTTG TTAAAAAGGA GGAAAAAGTG 660

CTTTCATGAAT TATTTGGGAC GSACATAAGG TGCTGTGCCA GAAATGAACT GGGCAGGGAA 720
 TGCACCAGGC TGTTCACAAT AGATCTAAAT CAAACTCCTC AGACCACATT GCCACAATTA 780
 TTTCTTAAAG TAGGGGAACC CTTATGGATA AGGTGCAAAG CTGTTTCATGT GAACCATGGA 840
 TTCGGGGCTCA CCTGGGAATT AGAAAACAAA GCACTCGAGG AGGGCAACTA CTTTGAGATG 900
 AGTACCTATT CAACAAACAG AACTATGATA CGGATTCTGT TTGCTTTTGT ATCATCACTG 960
 GCAAGAAACG ACACCGGATA CTACACTTGT TCCTCTTCAA AGCATCCCAG TCAATCAGCT 1020
 TTGGTTACCA TCGTAGGAAA GGGATTTATA AATGCTACCA ATTCAAGTGA AGATTATGAA 1080
 ATTGACCAAT ATGAAGAGTT TTGTTTTTCT GTCAGGTTTA AAGCCTACCC ACAAATCAGA 1140
 TGTACGTGGA CTTTCTCTCG AAAATCATTT CTTGTGAGC AAAAGGGTCT TCATAACGGA 1200
 TACAGCATAT CCAAGTTTTG CAATCATAAG CACCAGCCAG GAGAATATAT ATTCCATGCA 1260
 GAAAATGATG ATGCCCAATT TACCAAAATG TTCACGCTGA ATATAAGAAG GAAACCTCAA 1320
 GTGCTCGCAG AAGCATCGGC AAGTCAGGCG TCCTGTTTCT CGGATGGATA CCCATTACCA 1380
 TCTTGGACCT GGAAGAAGTG TTCAGACAAG TCTCCCAACT GCACAGAAGA GATCACAGAA 1440
 GGAGTCTGGA ATAGAAAGGC TAACAGAAAA GTGTTTGGAC AGTGGGTGTC GAGCAGTACT 1500
 CTAAACATGA GTGAAGCCAT AAAAGGGTTC CTGGTCAAGT GCTGTGCATA CAATTCCCTT 1560
 GGCACATCTT GTGAGACGAT CCTTTTAAAC TCTCCAGGCC CCTTCCCTTT CATCCAAGAC 1620
 AACATCTCAT TCTATGCAAC AATTGGTGTT TGTCTCCTCT TCATTGTCGT TTTAACCTTG 1680
 CTAATTTGTC ACAAGTACAA AAAGCAATTT AGGTATGAAA GCCAGCTACA GATGGTACAG 1740
 GTGACCGGCT CCTCAGATAA TGAGTACTTC TACGTTGATT TCAGAGAATA TGAATATGAT 1800
 CTCAAATGGG AGTTTCCAAG AGAAAATTTA GAGTTTGGGA AGGTACTAGG ATCCGAATAT 1860
 GATCTCAAAT GGGAGTTTCC AAGAGAAAAT TTAGAGTTTG GGAAGGTACT AGGATCAGGT 1920

GCTTTTGGAA AAGTGATGAA CGCAACAGCT TATGGAATTA GCAAAACAGG AGTCTCAATC 1980
CAGGTTGCCG TCAAAATGCT GAAAGAAAAA GCAGACAGCT CTGAAAGAGA GGCACTCATG 2040
TCAGAACTCA AGATGATGAC CCAGCTGGGA AGCCACGAGA ATATTGTGAA CCTGCTGGGG 2100
GCGTGACAC TGTCAGGACC AATTACTTG ATTTTGAAT ACTGTTGCTA TGGTGATCTT 2160
CTCAACTATC TAAGAAGTAA AAGAGAAAAA TTTCACAGGA CTTGGACAGA GATTTTCAAG 2220
GAACACAATT TCAGTTTTTA CCCCACTTTC CAATCACATC CAAATCCAG CATGCCTGGT 2280
TCAAGAGAAG TTCAGATACA CCCGGACTCG GATCAAATCT CAGGGCTTCA TGGGAATTCA 2340
TTTCACTCTG AAGATGAAAT TGAATATGAA AACCAAAAAA GGCTGGAAGA AGAGGAGGAC 2400
TTGAATGTGC TTACATTGTA AGATCTTCTT TGCTTTGCAT ATCAAGTTGC CAAAGGAATG 2460
GAATTTCTGG AATTTAAGTC GTGTGTTTAC AGAGACCTGG CCGCCAGGAA CGTGCTTGTC 2520
ACCCACGGGA AAGTGGTGAA GATATGTGAC TTTGGATTGG CTCGAGATAT CATGAGTGAT 2580
TCCAACATATG TTGTCAGGGG CAATGCCCGT CTGCCTGTAA AATGGATGGC CCCCAGAAAGC 2640
CTGTTTGAAG GCATCTACAC CATTAGAGT GATGTCTGGT CATATGGAAT ATTACTGTGG 2700
GAAATCTTCT CACTTGGTGT GAATCCTTAC CCTGGCATTG CGGTTGATGC TAACTTCTAC 2760
AAACTGATTC AAAATGGATT TAAATGGAT CAGCCATTTT ATGCTACAGA AGAAATATAC 2820
ATTATAATGC AATCCTGCTG GGCTTTTGAC TCAAGGAAAC GGCCATCCTT CCCTAATTTG 2880
ACTTCGTTTT TAGGATGTCA GCTGGCAGAT GCAGAAGAAG CGATGTATCA GAATGTGGAT 2940
GGCCGTGTTT CGGAATGTCC TCACACCTAC CAAAACAGGC GA 2982

(2) INFORMAÇÃO PARA A SEQ ID NO: 23:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

(A) COMPRIMENTO: 2958 pares de bases

- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: não

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 23:

```

ATGCCCCGGT TGGCGCGCGA CGCGGGCACC GTGCGGCTGC TCCTTGTTTT TTCTGCAATG   60
ATATTTGGGA CTATTACAAA TCAAGATCTG CCTGTGATCA AGTGTGTTTT AATCAATCAT  120
AAGAACAATG ATTCATCAGT GGGGAAGTCA TCATCATATC CCATGGTATC AGAATCCCCG  180
GAAGACCTCG GGTGTGCGTT GAGACCCAG AGCTCAGGGA CAGTGTACGA AGCTGCCGCT  240
GTGGAAGTGG ATGTATCTGC TTCCATCACA CTGCAAGTGC TGGTCGATGC CCCAGGGAAC  300
ATTTCTGTGC TCTGGGTCTT TAAGCACAGC TCCCTGAATT GCCAGCCACA TTTTGATTTA  360
CAAAACAGAG GAGTTGTTTC CATGGTCATT TTGAAAATGA CAGAAACCCA AGCTGGAGAA  420
TACCTACTTT TTATTCAGAG TGAAGCTACC AATTACACAA TATTGTTTAC AGTGAGTATA  480
AGAAATACCC TGCTTTACAC ATTAAGAAGA CCTTACTTTA GAAAAATGGA AAACCAGGAC  540
GCCCTGGTCT GCATATCTGA GAGCGTTCCA GAGCCGATCG TGGAATGGGT GCTTTGCGAT  600
TCACAGGGGG AAAGCTGTAA AGAAGAAAGT CCAGCTGTTG TTAAAAAGGA GGAAAAAGTG  660
CTTCATGAAT TATTTGGGAC GGACATAAGG TGCTGTGCCA GAAATGAACT GGGCAGGGAA  720
TGCACCAGGC TGTTACAAT AGATCTAAAT CAAACTCCTC AGACCACATT GCCACAATTA  780
TTTCTTAAAG TAGGGGAACC CTTATGGATA AGGTGCAAAG CTGTTTATGT GAACCATGGA  840
TTCGGGCTCA CCTGGGAATT AGAAAACAAA GCACTCGAGG AGGGCAACTA CTTTGAGATG  900

```

AGTACCTATT CAACAAACAG AACTATGATA CGGATTCTGT TTGCTTTTGT ATCATCAGTG 960
 GCAAGAAACG ACACCGGATA CTACACTTGT TCCTCTTCAA AGCATCCCAG TCAATCAGCT 1020
 TTGGTTACCA TCGTAGGAAA GGGATTTATA AATGCTACCA ATTCAAGTGA AGATTATGAA 1080
 ATTGACCAAT ATGAAGAGTT TPGTTTTTCT GTCAGGTTTA AAGCCTACCC ACAAATCAGA 1140
 TGTACGTGGA CCTTCTCTCG AAAATCATT TCTTGTGAGC AAAAGGGTCT TGATAACGGA 1200
 TACAGCATAT CCAAGTTTTG CAATCATAAG CACCAGCCAG GAGAATATAT ATTCCATGCA 1260
 GAAAATGATG ATGCCCAATT TACCAAATG TTCACGCTGA ATATAAGAAG GAAACCTCAA 1320
 GTGCTCGCAG AAGCATCGGC AAGTCAGGCG TCCTGTTTCT CGGATGGATA CCCATTACCA 1380
 TCTTGGACCT GGAAGAAGTG TTCAGACAAG TCTCCCAACT GCACAGAAGA GATCACAGAA 1440
 GGAGTCTGGA ATAGAAAGGC TAACAGAAAA GTGTTTGGAC AGTGGGTGTC GAGCAGTACT 1500
 CTAAACATGA GTGAAGCCAT AAAAGGGTTC CTGGTCAAGT GCTGTGCATA CAATTCCCTT 1560
 GGCACATCTT GTGAGACGAT CCTTTTAAAC TCTCCAGGCC CCTTCCCTTT CATCCAAGAC 1620
 AACATCTCAT TCTATGCAAC AATTGGTGTT TGTCTCCTCT TCATTGTCGT TTTAACCTTG 1680
 CTAATTTGTC ACAAGTACAA AAAGCAATTT AGGTATGAAA GCCAGCTACA GATGGTACAG 1740
 GTGACCGGCT CCTCAGATAA TGAGTACTTC TACGTTGATT TCAGAGAATA TGAATATGAT 1800
 CTCAAATGGG AGTTTGATTT CAGAGAATAT GAATATGATC TCAAATGGGA GTTTCCAAGA 1860
 GAAAATTTAG AGTTTGGGAA GGTACTAGGA TCAGGTGCTT TTGGAAAAGT GATGAACGCA 1920
 ACAGCTTATG GAATTAGCAA AACAGGAGTC TCAATCCAGG TTGCCGTCAA AATGCTGAAA 1980
 GAAAAAGCAG ACAGCTCTGA AAGAGAGGCA CTCATGTCAG AACTCAAGAT GATGACCCAG 2040
 CTGGGAAGCC ACGAGAATAT TGTGAACCTG CTGGGGGCGT GCACACTGTC AGGACCAATT 2100

TACTTGATTT TTGAATACTG TTGCTATGGT GATCTTCTCA ACTATCTAAG AAGTAAAAGA 2160
GAAAAATTTC ACAGGACTTG GACAGAGATT TTCAAGGAAC ACAATTTTCAG TTTTACCCC 2220
ACTTTCCAAT CACATCCAAA TTCCAGCATG CCTGGTTCAA GAGAAGTTCA GATACACCCG 2280
GACTCGGATC AAATCTCAGG GCTTCATGGG AATTCATTTC ACTCTGAAGA TGAAATTGAA 2340
TATGAAAACC AAAAAAGGCT GGAAGAAGAG GAGGACTTGA ATGTGCTTAC ATTTGAAGAT 2400
CTTCTTTGCT TTGCATATCA AGTTGCCAAA GGAATCGAAT TTCTGGAATT TAAGTCGTGT 2460
GTTACACAGAG ACCTGGCCCG CAGGAACGTG CTGTCAACC ACGGGAAAGT GGTGAAGATA 2520
TGTGACTTTG GATTGGCTCG AGATATCATG AGTGATTCCA ACTATGTTGT CAGGGGCAAT 2580
GCCCGTCTGC CTGTAAAATG GATGGCCCCC GAAAGCCTGT TTGAAGGCAT CTACACCATT 2640
AAGAGTGATG TCTGGTCATA TGGAATATTA CTGTGGGAAA TCTTCTCACT TGGTGTGAAT 2700
CCTTACCCTG GCATTCCGGT TGATGCTAAC TTCTACAAAC TGATTCAAAA TGGATTTAAA 2760
ATGGATCAGC CATTTTATGC TACAGAAGAA ATATACATTA TAATGCAATC CTGCTGGGCT 2820
TTTGA CTCAA GGAAACGGCC ATCCTTCCCT AATTTGACTT CGTTTTTAGG ATGTCAGCTG 2880
GCAGATGCAG AAGAAGCGAT GTATCAGAAT GTGGATGGCC GTGTTTCGGA ATGTCCTCAC 2940
ACCTACCAAA ACAGGCGA 2958

(2) INFORMAÇÃO PARA A SEQ ID NO: 24:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 2949 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 24:

ATGCCGGCGT TGGCGCGCGA CGCGGGCACC GTGCCGCTGC TCGTTGTTTT TTCTGCAATG 60
ATATTTGGGA CTATTACAAA TCAAGATCTG CCTGTGATCA AGTGTGTTTT AATCAATCAT 120
AAGAACAATG ATTCATCAGT GGGGAAGTCA TCATCATATC CCATGGTATC AGAATCCCCG 180
GAAGACCTCG GGTGTGCGTT GAGACCCAG AGCTCAGGGA CAGTGACGA AGCTGCCGCT 240
GTGGAAGTGG ATGTATCTGC TTCCATCACA CTGCAAGTGC TGGTCGATGC CCCAGGGAAC 300
ATTCCTGTGC TCTGGGTCTT TAAGCACAGC TCCCTGAATT GCCAGCCACA TTTTGATTTA 360
CAAAACAGAG GAGTTGTTTC CATGGTCATT TTGAAAATGA CAGAAACCCA AGCTGGAGAA 420
TACCTACTTT TTATTCAGAG TGAAGCTACC AATTACACAA TATTGTTTAC AGTGAGTATA 480
AGAAATACCC TGCTTTACAC ATTAAGAAGA CCTTACTTTA GAAAAATGGA AAACCAGGAC 540
GCCCTGGTCT GCATATCTGA GAGCGTTCCA GAGCCGATCG TGGAATGGGT GCTTTGCGAT 600
TCACAGGGGG AAAGCTGTAA AGAAGAAAGT CCAGCTGTTG TTAAAAAGGA GGAAAAAGTG 660
CTTCATGAAT TATTTGGGAC GGACATAAGG TGCTGTGCCA GAAATGAACT GGGCAGGGAA 720
TGCACCAGGC TGTTACAAT AGATCTAAAT CAAACTCCTC AGACCACATT GCCACAATTA 780
TTTCTTAAAG TAGGGGAACC CTTATGGATA AGGTGCAAAG CTGTTTCATGT GAACCATGGA 840
TTCGGGCTCA CCTGGGAATT AGAAAACAAA GCACTCGAGG AGGGCAACTA CTTTGAGATG 900
AGTACCTATT CAACAAACAG AACTATGATA CGGATTCTGT TTGCTTTTGT ATCATCAGTG 960
GCAAGAAACG ACACCGGATA CTACACTTGT TCCTCTTCAA AGCATCCCAG TCAATCAGCT 1020
TTGGTTACCA TCGTAGGAAA GGGATTTATA AATGCTACCA ATTCAAGTGA AGATTATGAA 1080

ATTGACCAAT ATGAAGAGTT TTGTTTTTCT GTCAGGTTTA AAGCCTACCC ACAAATCAGA 1140
 TGTACGTGGA CCTTCTCTCG AAAATCATTT CCTTGTGAGC AAAAGGGTCT TGATAACGGA 1200
 TACAGCATAT CCAAGTTTTG CAATCATAAG CACCAGCCAG GAGAATATAT ATTCCATGCA 1260
 GAAAATGATG ATGCCCAATT TACCAAAATG TTCACGCTGA ATATAAGAAG GAAACCTCAA 1320
 GTGCTCGCAG AAGCATCGGC AAGTCAGGCG TCCTGTTTCT CGGATGGATA CCCATTACCA 1380
 TCTTGGACCT GGAAGAAGTG TTCAGACAAG TCTCCCAACT GCACAGAAGA GATCACAGAA 1440
 GGAGTCTGGA ATAGAAAGGC TAACAGAAAA GTGTTTGGAC AGTGGGTGTC GAGCAGTACT 1500
 CTAAACATGA GTGAAGCCAT AAAAGGGTTC CTGGTCAAGT GCTGTGCATA CAATTCCCTT 1560
 GGCACATCTT GTGAGACGAT CCTTTTAAAC TCTCCAGGCC CCTTCCCTTT CATCCAAGAC 1620
 AACATCTCAT TCTATGCAAC AATTGGTGTT TGTCTCCTCT TCATTGTCGT TTTAACCTTG 1680
 CTAATTTGTC ACAAGTACAA AAAGCAATTT AGGTATGAAA GCCAGCTACA GATGGTACAG 1740
 GTGACCGGCT CCTCAGATAA TGAGTACTTC TACGTTGATT TCAGAGAATA TGAATATGAT 1800
 CTCAAATGGG AGTTTCCAAG AGAAAATTGG CACAAATGGG AGTTTCCAAG AGAAAATTTA 1860
 GAGTTTGGGA AGGTACTAGG ATCAGGTGCT TTTGGAAAAG TGATGAACGC AACAGCTTAT 1920
 GGAATTAGCA AAACAGGAGT CTCAATCCAG GTTGCCGTCA AAATGCTGAA AGAAAAAGCA 1980
 GACAGCTCTG AAAGAGAGGC ACTCATGTCA GAACTCAAGA TGATGACCCA GCTGGGAAGC 2040
 CACGAGAATA TTGTGAACCT GCTGGGGGCG TGCACACTGT CAGGACCAAT TTAATTGATT 2100
 TTTGAATACT GTTGCTATGG TGATCTTCTC AACTATCTAA GAAGTAAAAG AGAAAAATTT 2160
 CACAGGACTT GGACAGAGAT TTTCAAGGAA CACAATTTCA GTTTTTACCC CACTTTCCAA 2220
 TCACATCCAA ATTCCAGCAT GCCTGGTTCA AGAGAAGTTC AGATACACCC GGAICTGGAT 2280
 CAAATCTCAG GGCTTCATGG GAATTCATTT CACTCTGAAG ATGAAATTGA ATATGAAAAC 2340

CAAAAAAGGC TGGAAGAAGA GGAGGACTTG AATGTGCTTA CATTTGAAGA TCTTCTTTGC 2400
 TTTGCATATC AAGTTGCCAA AGGAATGGAA TTTCTGGAAT TTAAGTCGTG TGTTACAGA 2460
 GACCTGGCCG CCAGGAACGT GCTTGTACC CACGGGAAAG TGGTGAAGAT ATGTGACTTT 2520
 GGATTGGCTC GAGATATCAT GAGTGATTCC AACTATGTTG TCAGGGGCAA TGCCCGTCTG 2580
 CCTGTAAAAT GGATGGCCCC CGAAAGCCTG TTTGAAGGCA TCTACACCAT TAAGAGTGAT 2640
 GTCTGGTCAT ATGGAATATT ACTGTGGGAA ATCTTCTCAC TTGGTGTGAA TCCTTACCCT 2700
 GGCATTCCGG TTGATGCTAA CTTCTACAAA CTGATTCAA ATGGATTTAA AATGGATCAG 2760
 CCATTTTATG CTACAGAAGA AATATACATT ATAATGCAAT CCTGCTGGGC TTTTGAATCA 2820
 AGGAAACGGC CATCCTTCCC TAATTTGACT TCGTTTTTAG GATGTCAGCT GGCAGATGCA 2880
 GAAGAAGCGA TGTATCAGAA TGTGGATGGC CGTGTTCGG AATGTCCTCA CACCTACCAA 2940
 AACAGGCGA 2949

(2) INFORMAÇÃO PARA A SEQ ID NO: 25:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 2958 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: dupla
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: ADNc a ARNm

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 25:

ATGCCGGCGT TGGCGCGCGA CGCGGGCACC GTGCCGCTGC TCGTTGTTTT TTCTGCAATG 60

ATATTTGGGA CTATTACAAA TCAAGATCTG CCTGTGATCA AGTGTGTTTT AATCAATCAT 120
 AAGAACAATG ATTCATCAGT GGGGAAGTCA TCATCATATC CCATGGTATC AGAATCCCCG 180
 GAAGACCTCG GGTGTGCGTT GAGACCCAG AGCTCAGGGA CAGTGTACGA AGCTGCCGCT 240
 GTGGAAGTGG ATGTATCTGC TTCCATCACA CTGCAAGTGC TGGTCGATGC CCCAGGGAAC 300
 ATTTCTGTG TCTGGGTCTT TAAGCACAGC TCCCTGAATT GCCAGCCACA TTTTGATTTA 360
 CAAAACAGAG GAGTTGTTTC CATGGTCATT TTGAAAATGA CAGAAACCCA AGCTGGAGAA 420
 TACCTACTTT TTATTCAGAG TGAAGCTACC AATTACACAA TATTGTTTAC AGTGAGTATA 480
 AGAAATACCC TGCTTTACAC ATTAAGAAGA CTTACTTTA GAAAAATGGA AAACCAGGAC 540
 GCCCTGGTCT GCATATCTGA GAGCGTTCCA GAGCCGATCG TGAATGGGT GCTTTGCGAT 600
 TCACAGGGGG AAAGCTGTAA AGAAGAAAGT CCAGCTGTTG TTA AAAAGGA GGAAAAGTG 660
 CTTTATGAAT TATTTGGGAC GGACATAAGG TGCTGTGCCA GAAATGAACT GGGCAGGGAA 720
 TGCACCAGGC TGTTACAAT AGATCTAAAT CAACTCCTC AGACCACATT GCCACAATTA 780
 TTTCTTAAAG TAGGGGAACC CTTATGGATA AGGTGCAAAG CTGTTTATGT GAACCATGGA 840
 TTCGGGCTCA CCTGGGAATT AGAAAACAAA GCACTCGAGG AGGGCAACTA CTTTGAGATG 900
 AGTACCTATT CAACAAACAG AACTATGATA CGGATTCTGT TTGCTTTTGT ATCATCAGTG 960
 GCAAGAAACG ACACCGGATA CTACACTTGT TCCTCTTCAA AGCATCCCAG TCAATCAGCT 1020
 TTGGTTACCA TCGTAGGAAA GGGATTTATA AATGCTACCA ATTCAAGTGA AGATTATGAA 1080
 ATTGACCAAT ATGAAGAGTT TTGTTTTTCT GTCAGGTTTA AAGCCTACCC ACAAATCAGA 1140
 TGTACGTGGA CCTTCTCTCG AAAATCATTT CCTTGTGAGC AAAAGGGTCT TGATAACGGA 1200
 TACAGCATAT CCAAGTTTTG CAATCATAAG CACCAGCCAG GAGAATATAT ATTCCATGCA 1260
 GAAATGATG ATGCCCAATT TACCAAAATG TTCACGCTGA ATATAAGAAG GAAACCTCAA 1320

GTGCTCGCAG AAGCATCGGC AAGTCAGGCG TCCTGTTTCT CGGATGGATA CCCATTACCA 1380
 TCTTGGACCT GGAAGAAGTG TTCAGACAAG TCTCCCAACT GCACAGAAGA GATCACAGAA 1440
 GGAGTCTGGA ATAGAAAGGC TAACAGAAAA GTGTTTGGAC AGTGGGTGTC GAGCAGTACT 1500
 CTAAACATGA GTGAAGCCAT AAAAGGGTTC CTGGTCAAGT GCTGTGCATA CAATTCCCTT 1560
 GGCACATCTT GTGAGACGAT CCTTTTAAAC TCTCCAGGCC CCTTCCCTT CATCCAAGAC 1620
 AACATCTCAT TCTATGCAAC AATTGGTGTT TGTCTCCTCT TCATTGTCGT TTTAACCCCTG 1680
 CTAATTTGTC ACAAGTACAA AAAGCAATTT AGGTATGAAA GCCAGCTACA GATGGTACAG 1740
 GTGACCGGCT CCTCAGATAA TGAGTACTTC TACGTTGATT TCAGAGGCTC CTCAGATAAT 1800
 GAGTACTTCT ACGTTGATTT CAGAGAATAT GAATATGATC TCAAATGGGA GTTTCCAAGA 1860
 GAAAATTTAG AGTTTGGGAA GGTACTAGGA TCAGGTGCTT TTGGAAAAGT GATGAACGCA 1920
 ACAGCTTATG GAATTAGCAA AACAGGAGTC TCAATCCAGG TTGCCGTCAA AATGCTGAAA 1980
 GAAAAAGCAG ACAGCTCTGA AAGAGAGGCA CTCATGTCAG AACTCAAGAT GATGACCCAG 2040
 CTGGGAAGCC ACGAGAATAT TGTGAACCTG CTGGGGGCGT GCACACTGTC AGGACCAATT 2100
 TACTTGATTT TTGAATACTG TTGCTATGGT GATCTTCTCA ACTATCTAAG AAGTAAAAGA 2160
 GAAAATTTT ACAGGACTTG GACAGAGATT TTCAAGGAAC ACAATTTTCA TTTTACCCC 2220
 ACTTTCCAAT CACATCCAAA TTCCAGCATG CCTGGTTCAA GAGAAGTTCA GATACACCCG 2280
 GACTCGGATC AAATCTCAGG GCTTCATGGG AATTCATTT ACTCTGAAGA TGAAATTGAA 2340
 TATGAAAACC AAAAAAGGCT GGAAGAAGAG GAGGACTTGA ATGTGCTTAC ATTTGAAGAT 2400
 CTTCTTTGCT TTGCATATCA AGTTGCCAAA GGAATGGAAT TTCTGGAATT TAAGTCGTGT 2460
 GTTCACAGAG ACCTGGCCGC CAGGAACGTG CTTGTCACCC ACGGGAAAGT GGTGAAGATA 2520

TGTGACTTTG GATTGGCTCG AGATATCATG AGTGATTCCA ACTATGTTGT CAGGGGCAAT 2580

GCCCGTCTGC CTGTAAAATG GATGGCCCCC GAAAGCCTGT TTGAAGGCAT CTACACCATT 2640

AAGAGTGATG TCTGGTCATA TGGAATATTA CTGTGGGAAA TCTTCTCACT TGGTGTGAAT 2700

CCTTACCCTG GCATTCCGGT TGATGCTAAC TTCTACAAAC TGATTCAAAA TGGATTTAAA 2760

ATGGATCAGC CATTTTATGC TACAGAAGAA ATATACATTA TAATGCAATC CTGCTGGGCT 2820

TTTGACTCAA GGAAACGGCC ATCCTTCCCT AATTGACTT CGTTTTTAGG ATGTCAGCTG 2880

GCAGATGCAG AAGAAGCGAT GTATCAGAAT GTGGATGGCC GTGTTTCGGA ATGTCCTCAC 2940

ACCTACCAAA ACAGGCGA 2958

(2) INFORMAÇÃO PARA A SEQ ID NO: 26:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: outros ácidos nucleicos (ADN sintético)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 265:

TGTCGAGCAG TACTCTAAAC A 21

(2) INFORMAÇÃO PARA A SEQ ID NO: 27:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 22 pares de bases

- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: outros ácidos nucleicos (ADN sintético)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 27:

ATCCTAGTAC CTTCCCAAAC TC

22

((2) INFORMAÇÃO PARA A SEQ ID NO: 28:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 18 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: outros ácidos nucleicos (ADN sintético)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 28:

CTTCCTGGGC ATGGAGTC

18

(2) INFORMAÇÃO PARA A SEQ ID NO: 29:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases

- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: outros ácidos nucleicos (ADN sintético)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 29:

CGCTCAGGAG GAGCAATGAT

20

(2) INFORMAÇÃO PARA A SEQ ID NO: 30:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 19 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: outros ácidos nucleicos (ADN sintético)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 30:

CAATTAGGT ATGAAAGCC

19

(2) INFORMAÇÃO PARA A SEQ ID NO: 31:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 19 pares de bases

- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: outros ácidos nucleicos (ADN sintético)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 31:

CAAACCTCTAA ATTTTCTCT

19

(2) INFORMAÇÃO PARA A SEQ ID NO: 32:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 21 pares de bases
- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: outros ácidos nucleicos (ADN sintético)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 32:

TGTCTTTGCA GGGAAGGTTA C

21

(2) INFORMAÇÃO PARA A SEQ ID NO: 33:

(i) CARACTERÍSTICAS DA SEQUÊNCIA:

- (A) COMPRIMENTO: 20 pares de bases

- (B) TIPO: ácido nucleico
- (C) ESTRUTURA HELICOIDAL: simples
- (D) TOPOLOGIA: linear

(ii) TIPO DE MOLÉCULA: outros ácidos nucleicos (ADN sintético)

(iii) HIPOTÉTICA: NÃO

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 335:

GTACCTTTCA GCATTTTGAC

20

Lisboa, 27 de Outubro de 2006

REIVINDICAÇÕES

1. Molécula de ácido nucleico de um mutante com uma duplicação em tandem que codifica uma tirosina cinase 3 (FLT3) semelhante a FMS, caracterizada pelo facto de a referida molécula de ácido nucleico ter uma sequência de nucleótidos que corresponde a:

(a) uma mutação de uma duplicação em tandem na sequência de aminoácidos da justamembrana de FLT3 ou

(b) uma mutação de uma duplicação em tandem na sequência de nucleótidos de uma região que compreende o exão 11 ou os exões 11 a 12 de FLT3.

2. Molécula de ácido nucleico caracterizada pelo facto de comportar uma sequência de nucleótidos que corresponde a:

(a) uma mutação de uma duplicação em tandem na sequência de aminoácidos da justamembrana de FLT3 ou

(b) uma mutação de uma duplicação em tandem na sequência de nucleótidos uma região que compreende ao exão 11 ou aos exões 11 a 12 de FLT3.

3. Molécula de ácido nucleico capaz de hibridar especificamente, em condições severas, com uma molécula de ácido nucleico que corresponde a:

(a) uma mutação de uma duplicação em tandem na sequência de aminoácidos da justamembrana de FLT3 ou

(b) uma mutação de uma duplicação em tandem na sequência de nucleótidos de uma região que compreende o

exão 11 ou os exões 11 a 12 de FLT3 no ácido nucleico de acordo com a reivindicação 1.

4. Polipéptido caracterizado pelo facto de ser codificado pela molécula de ácido nucleico de acordo com a reivindicação 1 ou a molécula de ácido nucleico de acordo com a reivindicação 2.
5. Célula caracterizada pelo facto de expressar a molécula de ácido nucleico de acordo com a reivindicação 1 ou a molécula de ácido nucleico de acordo com a reivindicação 2 ou o polipéptido de acordo com a reivindicação 4.
6. Anticorpo caracterizado pelo facto de ser específico para o polipéptido de acordo com a reivindicação 4.
7. Processo para a detecção da molécula de ácido nucleico de acordo com a reivindicação 1 ou a molécula de ácido nucleico de acordo com a reivindicação 2, caracterizado pelo facto de compreender as etapas de:
 - (a) submissão de uma amostra de ácido nucleico de um ser humano a uma reacção de amplificação de genes, em que se amplifica um fragmento de ácido nucleico que compreende o exão 11 ou os exões 11 a 12 do gene de tirosina cinase 3 (FLT3) semelhante a FMS e tendo uma mutação de uma duplicação em tandem na justamembrana, que pode ser encontrada no gene de FLT3;
 - (b) detecção da presença da mutação de uma duplicação em tandem no fragmento de ácido nucleico da etapa (a).
8. Processo de acordo com a reivindicação 7, caracterizado pelo facto de a etapa (b) ser realizada por comparação do

fragmento de ácido nucleico amplificado obtido na etapa (a) com uma sequência derivada de uma FLT3 normal, detectando assim a presença da mutação de uma duplicação em tandem na justamembrana.

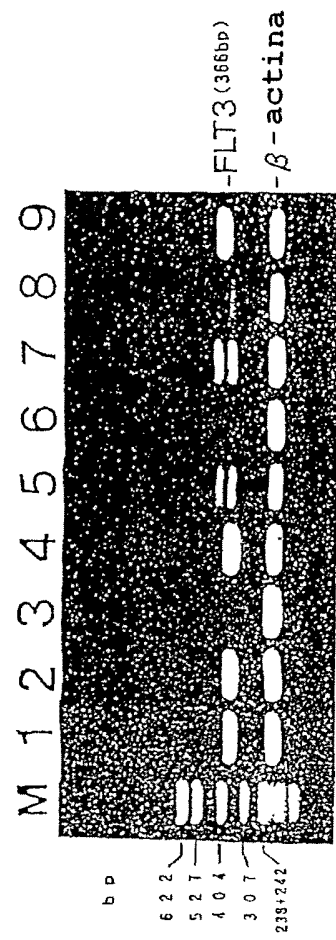
9. Processo de acordo com a reivindicação 7 ou 8, caracterizado pelo facto de, na etapa (b), se utilizar uma mutação do comprimento como um índice da mutação de uma duplicação em tandem.
10. Processo de acordo com uma qualquer das reivindicações 7 a 9, caracterizado pelo facto de a reacção de amplificação do gene na etapa (a) ser realizada com um par de iniciadores seleccionado no grupo que consiste nas: SEQ ID N°S 26 e 27, SEQ ID N°S 30 e 31 e SEQ ID N°S 32 e 33.
11. Processo de acordo com uma qualquer das reivindicações 7 a 10, caracterizado pelo facto de a mutação da duplicação em tandem não se encontrar num gene de um indivíduo normal.
12. Kit para o processo de acordo com uma qualquer das reivindicações 7 a 11, caracterizado pelo facto de o referido kit compreender iniciadores para a amplificação de uma região que compreende o exão 11 ou os exões 11 a 12 do gene de FLT3.
13. Kit de acordo com a reivindicação 12, caracterizado pelo facto de os referidos iniciadores se seleccionaram no grupo que consiste nas: SEQ ID N°S 26 e 27, SEQ ID N°S 30 e 31 e SEQ ID N°S 32 e 33.

14. Utilização de uma célula hemetopoiética indiferenciada que expressa uma molécula de ácido nucleico que codifica um polipéptido de mutante com uma duplicação em tandem e uma célula que expressa FLT3 normal, para avaliação de um fármaco para o exame e tratamento de uma doença das células sanguíneas ou uma doenças das célula hemetopoiéticas indiferenciadas.
15. Utilização, de acordo com a reivindicação 14, caracterizada pelo facto de a referida molécula de ácido nucleico ter uma mutação de duplicação em tandem numa sequência de nucleótidos de uma justamembrana.
16. Utilização de uma célula hemetopoiética indiferenciada que expressa a molécula de ácido nucleico, de acordo com a reivindicação 1 ou a molécula de ácido nucleico de acordo com a reivindicação 2 ou o polipéptido de acordo com a reivindicação 4 e uma célula que expressa FLT3 normal para avaliação de um fármaco para o exame e tratamento de uma doença das células sanguíneas ou uma doenças das célula hemetopoiéticas indiferenciadas.
17. Utilização de uma molécula de ácido nucleico, de acordo com a reivindicação 1 ou da molécula de ácido nucleico de acordo com a reivindicação 2, caracterizada pelo facto de se destinar à preparação de um fármaco para regular a proliferação, a resposta imunitária ou a transmissão da informação do sinal de uma célula sanguínea ou de uma célula hematopoiética indiferenciada.
18. Utilização da molécula de ácido nucleico, de acordo com a reivindicação 1 ou da molécula de ácido nucleico de acordo com a reivindicação 2 ou do polipéptido de acordo com a reivindicação 4, caracterizada pelo facto de se

destinar à preparação de material utilizado para a avaliação patológica do síndrome mielodiplásico (SMD) ou de leucemia.

19. Composição farmacêutica caracterizada pelo facto de compreender a molécula de ácido nucleico, de acordo com a reivindicação 1 e/ou a molécula de ácido nucleico de acordo com a reivindicação 2 e/ou a molécula de ácido nucleico de acordo com a reivindicação 3 e/ou o polipéptido de acordo com a reivindicação 4 e/ou o anticorpo de acordo com a reivindicação 12 e, eventualmente, um veículo e/ou um diluente aceitáveis sob o ponto de vista farmacêutico.
20. Kit para a detecção do polipéptido de acordo com a reivindicação 4, caracterizado pelo facto de o kit compreender o anticorpo de acordo com a reivindicação 6.

Lisboa, 27 de Outubro de 2006



F I G . 1

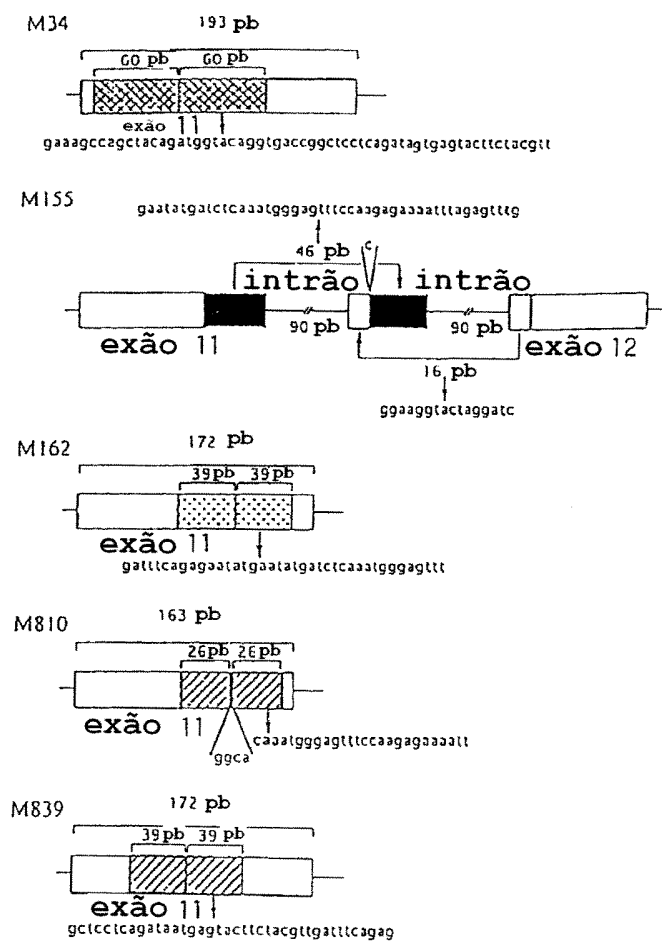


FIG. 2