



DESCRIÇÃO
DA
PATENTE DE INVENÇÃO

N.º 100 511

REQUERENTE: REGENERON PHARMACEUTICALS, INC., norte-americana (Estado de New York), com sede em 777 Old Saw Mill River Road, Tarrytown, NY 10591, Estados Unidos da América, FINN HALLBOOK, sueco, residente em Geijersgatan 48A, S-75231 Uppsala, Suécia, CARLOS FERNANDO IBANEZ MOLINER, espanhol, residente em Tangvagen 29, S-12638 Hagersten, Suécia e HAKAN BENGT PERSSON, sueco, residente em Vreta Gard, S-14743 Tumba, Suécia

EPÍGRAFE:

"Uso de neurotrofina-4, processo de diagnóstico baseado na expressão da mesma e produção de neurotrofina-4, e conjunto de diagnóstico"

INVENTORES: Nancy Ip, George D. Yancopoulos, Finn Hallbook, Carlos Fernando Ibanez Moliner e Hakan Bengt Persson

Reivindicação do direito de prioridade ao abrigo do artigo 4º da Convenção de Paris de 20 de Março de 1883.

Estados Unidos da América em 21 de Maio de 1991 sob o nº 07/703 450, em 12 de Julho de 1991 sob o nº 07/729 253, em 23 de Julho de 1991 sob o nº 07/734 422, em 28 de Agosto de 1991 sob o nº 07/751 356, em 20 de Setembro de 1991 sob o nº 07/762 674 e em 14 de Novembro de 1991 sob o nº 07/791 924

PATENTE N° 100 511

"Uso de neurotrofina-4, processo de diagnóstico baseado na expressão da mesma e produção de neurotrofina-4, e conjunto de diagnóstico"

RESUMO

O presente invento refere-se ao uso de neurotrofina-4 (NT-4) no fabrico de medicamentos para o tratamento de situações clínicas em que se verifica expressão de NT-4, a processos de diagnóstico baseados na expressão de NT-4, a conjuntos de diagnóstico para detectar o transporte retrógrado num neurónio motor compreendendo uma proteína NT-4, derivado ou fragmento peptídico marcados detectavelmente e a um processo de produção de NT-4 que compreende o crescimento de uma célula recombinante contendo uma molécula de ácido nucleico do invento, sob condições tais que a proteína prepro quimérica seja expressa e processada pela célula de modo a produzir a forma madura de NT-4.

MEMÓRIA DESCRITIVA

PROCESSOS TERAPÊUTICOS E DE DIAGNÓSTICO
BASEADOS NA EXPRESSÃO DE NEUROTROFINA-4

1. INTRODUÇÃO

O presente invento refere-se à neurotrofina-4 (NT-4), um novo membro caracterizado, da família de genes BDNF/NGF/NT-3 e a processos de diagnóstico e terapêuticos que utilizam a neurotrofina-4 no tratamento de desordens neurológicas.

2. ANTECEDENTES DO INVENTO

A família do factor de crescimento do nervo inclui o factor de crescimento do nervo β (NGF), o factor neurotrófico derivado de cérebro (BDNF), e a neurotrofina-3 (NT-3), também conhecido como factor neurotrófico derivado do hipocampo (HDNF). Esta família de proteínas desempenha um papel importante no sistema nervoso de vertebrados adultos ou em desenvolvimento, apoiando a sobrevivência neuronal.

Com base na sequência de aminoácidos da proteína NGF de ratinho (Angeletti, et al., 1973 Biochemistry, 12: 100-115) foram isoladas as sequências de ADN que codificam o NGF humano e de ratinho (Scott et al., 1983, Nature 302: 538-540; Ullrich et al., 1983 Nature 303: 821-825). A comparação entre NGF humano e de ratinho mostrou que a proteína é conservada entre os mamíferos e apoiando isto, isolaram-se actividades semelhantes à do NGF a partir de várias espécies (Harper e Thoenen, 1991 Ann. Rev. Pharmacol. Toxicol. 21: 205-229). Subsequentemente, sequências de ADN de NGF de boi (Meier et al., 1986, EMBO J. 5: 1489-1493); de pinto (Meier et al., EMBO J. 5: 1489-1493; Ebendal et al., 1986 EMBO J. 5: 1483-1487; Wion et al., 1986 FEBS Letters 203: 82-86); cobra (Selby et al., 1987, J. Neurosci. Res. 18: 293-298); rato (Whittemore et al., 1988, J. Neurosci. Res. 20: 403-410); e porquinho-da-índia (Schwarz et al., 1989, Neurochem. 52: 1203-1209) foram também determinadas. O factor neurotrófico derivado de cérebro (BDNF) foi inicialmente isolado de cérebro de porco (Barde et al., 1982,

EMBO J. 1: 549-553) e subsequentemente clonado a partir deste tecido como um ADNc (Leibrock et al., 1989, Nature 341: 149-152). O gene de NT-3 foi isolado de ratinho (Hohn et al., 1990, Nature 344: 339-341), de rato (Maisonpierre et al., 1990, Science 247: 1446-1451; Ernfors et al., 1990, Proc. Natl. Acad. Sci. USA 87: 5454-5458), e humano (Rosenthal et al., 1990, Neuron 4: 767-773), utilizando oligonucleótidos degenerados baseados na similaridade de sequência entre os outros dois factores. Os três factores apresentam aproximadamente 55% de similaridade de aminoácidos um com o outro, e a maior parte das diferenças de sequência estão presentes em cinco regiões que contêm motivos de aminoácidos característicos de cada proteína. Foi recentemente mostrado que a actividade neurotrófica in vitro de duas destas proteínas podia ser adquirida por combinações específicas destas regiões variáveis.

O NGF apoia o desenvolvimento e manutenção do simpático periférico e de neurónios sensoriais derivados da crista neural (revisto em Thoenen e Barde, 1980, Physiol. Rev., 60: 1284-1325; Levi-Montalcini, 1987, Science 237: 1154-1162). Não foi encontrada qualquer actividade de BDNF em neurónios simpáticos periféricos, mas este factor apoia in vivo a sobrevivência tanto de neurónios sensoriais derivados do placódio como da crista neural (Hofer e Barde, 1988, Nature, 331: 261-262). Os neurónios sensíveis a NT-3 in vivo permanecem sem ter sido identificados. No entanto, em gânglios explantados de pinto ou culturas neuronais dissociadas, in vitro, os três factores apoiam ambos os grupos de populações neuronais sobrepostas ou únicas, sugerindo que o NT-3 exerce também in vivo ambas as actividades neurotróficas específica ou sobreposta (Hohn et al., 1990, Nature, 344: 339-341; Maisonpierre et al., 1990, Science, 247: 1446-1451; Ernfors et al., 1990, Proc. Natl. Acad. Sci. USA, 87: 5454-5458; Rosenthal et al., 1990, Neuron 4: 767-773). Todos os três factores são expressos em grupos específicos de neurónios no cérebro, com os níveis de ARNm mais elevados para os três factores no hipocampo (Ayer-LeLievre et al., 1988, Science, 240: 1339-1341; Ernfors et al., 1990, Proc. Natl. Acad. Sci. USA, 87: 5454-5458; Ernfors et al., 1990, Neuron 5:

511-526; Watmore et al., 1991, Neurol. 109: 141-152; Hofer et al., 1990, EMBO J., 9: 2459-2464; Phillips et al., 1990, Science, 250: 290-294). No cérebro, o NGF revelou apoiar neurónios colinérgicos do prosencéfalo basal (revisão em Whitemore e Seiger, 1987, Brain Res., 434: 439-464; Thoenen et al., 1987, Rev. Physiol. Biochem. Pharmacol., 105: 145-178; Ebendal, 1989, Prog. Growth Factor Res. 1: 143-159) e BDNF revelou estimular a sobrevivência desses neurónios in vitro (Alderson et al., 1990, Neuron 5: 297-306).

Os efeitos das três proteínas são mediados pela sua interacção com receptores específicos presentes em células sensitivas. Foram isolados clones moleculares para os receptores de NGF (NGF-R) de rato, humano e de galinhas, e a análise da sequência de nucleótidos destes clones revelou que o NGF-R contém um domínio ligado à membrana plasmática, uma região citoplasmática e um domínio extracelular de extremidade amino, rico em cisteína (Johnson et al., 1986, Cell, 47: 545-554; Radeke et al., 1987, Nature, 325: 593-597; Large et al., 1989, Neuron 2: 1123-1134). O NGF-R apresenta uma similaridade de sequência baixa mas significativa em relação ao receptor para um factor da necrose tumoral (Schall et al., 1990, Cell, 61: 361-370) bem como em relação aos antigénios CD40 de superfície de linfócitos (Stamenkovic et al., 1989, EMBO J., 8: 1403-1410) e OX40 (Mallett et al., 1990, EMBO J., 9: 1063-1068). O NGF-R pode ocorrer em dois estados aparentes, conhecidos como estados de baixa e elevada afinidade (Sutter, et al., 1979, J. Biol. Chem., 254: 5972-5982; Landreth e Shooter, 1980, Proc. Natl. Acad. Sci. USA, 77: 4751-4755; Schechter e Bothwell, 1991, Cell 24: 867-874). O gene para o NGF-R parece codificar uma proteína que forma parte de ambos os estados de afinidade do receptor, baixa e elevada (Hampstead et al., 1989, Science, 243: 373-375), embora apenas o receptor de elevada afinidade tenha sido proposto para mediar a actividade biológica de NGF. Ambos, BDNF (Rodriguez-Tebar et al., 1990, Neuron, 4: 487-492) e NT-3 (Ernfors et al., 1990, Proc. Natl. Acad. Sci. USA, 87: 5454-5458) podem interagir com o NGF-R de baixa afinidade, sugerindo que o NGF-R de baixa afinidade possa estar, de um modo

ainda desconhecido, envolvido na mediação dos efeitos biológicos de todos os três factores.

No sistema nervoso em desenvolvimento, o NGF e o seu receptor revelaram ser sintetizados na área alvo e nos neurónios de resposta, respectivamente, no momento em que o axónio em crescimento atinge o seu alvo (Davies et al., 1987, Nature, 326: 353-358). Em concordância com isto, o nível de ARNm de NGF no embrião de pinto em desenvolvimento atinge um máximo no dia embrionário 8 (E8) (Ebendal e Persson, 1988, Development, 102: 101-106), que coincide com o tempo de inervação sensorial. Contudo, no pinto, o ARNm de NGF-R é expresso no seu máximo em etapas embrionárias precoces, anteriores à inervação neuronal (Ernfors et al., 1988, Neuron, 1, 983-996), e no embrião de pinto E8 foram detectados elevados níveis de ARNm de NGF-R no mesênquima, nos somitos e em células do tubo neural (Hallbook et al., 1990, Development, 108: 693-704; Heuer et al., 1990, Dev. Biol., 137: 287-304; Heuer 1990, Neuron, 5: 283-296). Esta observação, juntamente com o facto de que ARNm de NGF é expresso no embrião de pinto E3 em níveis relativamente elevados (Ebendal e Persson, 1988, Development, 102: 101-106), indica que o NGF pode desempenhar um papel no desenvolvimento precoce que é distinto da sua função como factor neurotrófico. Em concordância com esta possibilidade, o NGF revelou recentemente que controla a proliferação e diferenciação de células precursoras de estriamento de embrião de rato E14, em cultura (Cattaneo e McKay, 1990, Nature, 347: 762-765). No embrião de pinto o ARNm de BDNF e de NT-3 são expressos no seu máximo em E4,5, e o BDNF revelou controlar a diferenciação de células da crista neural de aves in vitro (Kalcheim e Gendreau, 1988, Dev. Brain Res., 41: 79-86).

Além disso, foi também apresentada evidência para uma função não neuronal do NGF. Os ainda não explicados níveis elevados de NGF encontrados na glândula submandibular de ratinho macho pode indicar outras funções para o NGF (Levi-Montalcini, 1987, Science, 237: 1154-1162). No rato adulto o NGF revelou induzir a síntese de ADN e estimular a secreção de IgM em

células B (Otten *et al.*, 1989, Proc. Natl. Acad. Sci. USA, **86**: 10059-10063). Adicionalmente, o NGF está presente em quantidade suficiente na próstata de porquinho-da-índia, de tal modo que Rubin e Bradshaw (1981, J. Neur. Res. **6**: 451-464) foram bem sucedidos no isolamento e na caracterização de NGF substancialmente puro, a partir deste tecido exócrino. O elevado nível de NGF na próstata de porco apoia a hipótese de que este factor neurotrófico funciona numa capacidade não neuronal ainda não compreendida (Bradshaw, 1978, Ann. Rev. Biochem. **47**: 191-216; Harper, *et al.*, 1979, Nature **279**: 160-162; Harper e Thoenen, 1980, J. Neurochem. **34**: 893-903).

Para além disso, o ARNm de NGF é expresso em espermatócitos e espermátides precoces em testículos de rato adulto (Ayer-LeLievre *et al.*, 1988, Proc. Natl. Acad. Sci. USA, **85**: 2628-2632), e a proteína NGF está presente em células germinais de todas as etapas desde espermatócitos até espermatozóides (Olson *et al.*, 1987, Cell Tissue Res., **248**: 275-286; Ayer-LeLievre *et al.*, 1988a, Proc. Natl. Acad. Sci. USA **85**: 2628-2632). ARNm de NGF-R foi também detectado em testículos de rato adulto, onde é expresso nas células de Sertoli sob controlo negativo da testosterona, e tem sido sugerido que nos testículos o NGF controla a meiose e a espermição (Persson *et al.*, 1990, Science, **247**: 704-707).

3. SUMÁRIO DO INVENTO

O presente invento refere-se a neurotrofina-4 (NT-4), um novo membro recentemente caracterizado da família de genes BDNF/NGF/NT-3.

O presente invento proporciona moléculas de ácido nucleico que codificam NT-4. Estas moléculas podem compreender uma sequência substancialmente como as apresentadas para NT-4 na Figura 1 [SEQ ID NO: 1 (NT-4, víbora), SEQ ID NO: 2 (NT-4, Xenopus)], Figura 4 (SEQ ID NO: 43), Figura 8 (SEQ ID NO: 49), Figura 14 (SEQ ID NO: 61), Figura 15 (SEQ ID NO: 63), Figura 17 (SEQ ID NO: 69), Figura 18 (SEQ ID NO: 75), Figura 20 (SEQ ID NO: 93) e Figura 21 (SEQ ID NO: 116) ou podem compreender uma

sequência que seja pelo menos cerca de setenta por cento homóloga com essa sequência.

O presente invento também proporciona uma proteína ou moléculas peptídicas que compreendem uma sequência substancialmente como as apresentadas para NT-4 na Figura 2 [SEQ ID NO: 22 (NT-4, víbora), SEQ ID NO: 23 (NT-4, Xenopus)], Figura 4 (SEQ ID NO: 44), Figura 8 (SEQ ID NO: 50), Figura 14 (SEQ ID NO: 62), Figura 15 (SEQ ID NO: 64), Figura 17 (SEQ ID NO: 70), Figura 18 (SEQ ID NO: 76), Figura 20 (SEQ ID NO: 94), ou Figura 21 (SEQ ID NO: 117) ou podem compreender uma sequência que seja pelo menos setenta por cento homóloga com essa sequência.

O presente invento proporciona, para além disso, a expressão de moléculas de NT-4 biologicamente activas em sistemas de procariotas e de eucariotas.

O presente invento proporciona, para além disso, a produção de NT-4 em quantidades suficientes para aplicações terapêuticas e de diagnóstico. Do mesmo modo, anticorpos anti-NT-4 podem ser utilizados em aplicações terapêuticas e de diagnóstico. Para a maior parte dos fins é preferível usar genes de NT-4 ou produtos de genes das mesmas espécies para fins terapêuticos e de diagnóstico, embora a utilidade de espécies cruzadas de NT-4 possa ser útil em concretizações específicas deste invento.

O presente invento proporciona, para além disso, aplicações terapêuticas e de diagnóstico baseadas na expressão de NT-4 pela revelação de níveis detectáveis de expressão de NT-4 em músculo esquelético humano, próstata, timo e testículos.

4. DESCRIÇÃO DAS FIGURAS

FIGURA 1. Alinhamentos de sequências de ADN dos fragmentos isolados que codificam NGF, BDNF, NT-3 e o novo factor neurotrófico NT-4 de espécies diferentes.

(A) Representação esquemática da molécula de preproNGF de ratinho. A caixa com tracejado indica a sequência de sinal

(SS), barras pretas indicam sítios de clivagem proteolítica e a caixa sombreada representa o NGF maduro. Regiões usadas para os iniciadores degenerados são indicadas por setas. O iniciador a montante da região de codificação, foi da lisina 50 até à treonina 56 e o iniciador a jusante inclui do triptofano 99 até ao ácido aspártico 105. A região amplificada compreende sequências de ADN do par de bases (bp) 168 até ao 294 nas moléculas de NGF maduro e em todos os membros da família do NGF até agora descritos, esta região está localizada num exão.

(B) Alinhamento de sequências de nucleótidos para NGF, BDNF, NT-3 e NT-4 isolados a partir de espécies diferentes. Os fragmentos correspondem aos aminoácidos 57 a 98 no NGF maduro de ratinho. Bases idênticas estão marcadas por pontos. A numeração refere-se a nucleótidos nas sequências de NGF maduro de ratinho (Scott *et al.*, 1983, *Nature* 300: 538-540). SEQ ID NO: 1 (NT-4, víbora), SEQ ID NO: 2 (NT-4, Xenopus), SEQ ID NO: 3 (NGF, humano), SEQ ID NO: 4 (NGF, rato), SEQ ID NO: 5 (NGF, galinha), SEQ ID NO: 6 (NGF, víbora), SEQ ID NO: 7 (NGF, Xenopus), SEQ ID NO: 8 (NGF, salmão), SEQ ID NO: 9 (BDNF, humano), SEQ ID NO: 10 (BDNF, rato), SEQ ID NO: 11 (BDNF, galinha), SEQ ID NO: 12 (BDNF, víbora), SEQ ID NO: 13 (BDNF, Xenopus), SEQ ID NO: 14 (BDNF, salmão), SEQ ID NO: 15 (BDNF, raia), SEQ ID NO: 16 (NT-3, humano), SEQ ID NO: 17 (NT-3, rato), SEQ ID NO: 18 (NT-3, galinha), SEQ ID NO: 19 (NT-3, Xenopus), SEQ ID NO: 20 (NT-3, salmão), SEQ ID NO: 21 (NT-3, raia).

FIGURA 2. Alinhamento de sequências de aminoácidos deduzidas para NGF, BDNF, NT-3 e NT-4 de espécies diferentes. A numeração dos aminoácidos (código de letra única) é tirada a partir de NGF maduro de ratinho (Scott *et al.*, 1983, *Nature* 300: 538- 540). Os aminoácidos idênticos estão indicados com pontos. As posições que apresentam substituições de aminoácidos conservativos em todas as espécies variantes do mesmo factor estão sublinhadas. A linha a tracejado indica que a sequência correspondente não foi isolada. As barras apresentam regiões variáveis nas diferentes moléculas (R59 a S67 e D93 a A98). SEQ ID NO: 22 (NT-4, víbora), SEQ ID NO: 23 (NT-4, Xenopus), SEQ ID NO: 24 (NGF, humano), SEQ

ID NO: 25 (NGF, rato), SEQ ID NO: 26 (NGF, galinha), SEQ ID NO: 27 (NGF, víbora), SEQ ID NO: 28 (NGF, Xenopus), SEQ ID NO: 29 (NGF, salmão), SEQ ID NO: 30 (BDNF, humano), SEQ ID NO: 31 (BDNF, rato), SEQ ID NO: 32 (BDNF, galinha), SEQ ID NO: 33 (BDNF, víbora), SEQ ID NO: 34 (BDNF, Xenopus), SEQ ID NO: 35 (BDNF, salmão), SEQ ID NO: 36 (BDNF, raia), SEQ ID NO: 37 (NT-3, humano), SEQ ID NO: 38 (NT-3, rato), SEQ ID NO: 39 (NT-3, galinha), SEQ ID NO: 40 (NT-3, Xenopus), SEQ ID NO: 41 (NT-3, salmão), SEQ ID NO: 42 (NT-3, raia).

FIGURA 3. Filogenia deduzida de membros da família de NGF. Árvores filogenéticas apresentando evolução de NGF (A), BDNF (B), e NT-3 (C) foram construídas utilizando análises de sequências de nucleótidos. NT-3 de humano foi utilizado como ponto de referência em (A) e (B), NGF humano e BDNF humano foram utilizados em (C). A barra da escala em (A) representa um comprimento de ramo correspondente a uma marca de diferença relativa de 20. A mesma escala foi utilizada em (B) e em (C). (D) apresenta um filograma da relação evolutiva entre os diferentes membros da família de NGF. Os dados foram compilados a partir de sequências de aminoácidos deduzidas. A barra da escala representa um comprimento de ramo de 20. Todas as árvores apresentadas são desprovidas de raiz, de modo a que os ramos sejam medidos em relação um ao outro sem referência exterior. Abreviações: chi, galinha; hum, humano; sal, salmão; vip, víbora; xen Xenopus.

FIGURA 4. Sequência de NT-4 de Xenopus e Comparação com NGF, BDNF e NT-3.

(A) Um local de início de tradução potencial foi colocado numa caixa. Um local de clivagem de sinal putativo está indicado por uma seta marcada com SC. Aminoácidos dentro da sequência de sinal que são idênticos entre NT-4 de Xenopus e BDNF de porco e de rato estão indicados com estrelas. Uma sequência de consenso para N-glicosilação está sublinhada e a seta indica o início presumível da proteína madura NT-4. (SEQ ID NO: 43 e SEQ ID NO: 44)

(B) Comparação de sequência de aminoácidos (código de letra única) de NT-4 de Xenopus (SEQ ID NO: 45) com NGF de ratinho (Scott et al., 1983, Nature 300: 538-540) (SEQ ID NO: 46), BDNF de ratinho (Hofer et al., 1990, EMBO J. 9: 2459-2464) (SEQ ID NO: 47) e NT-3 de ratinho (Hohn et al., 1990, Nature 344: 339-341) (SEQ ID NO: 48). Substituições de aminoácidos idênticos comparados com a sequência de aminoácidos de NT-4 são apresentadas por pontos. As sequências que diferem entre NGF, BDNF e NT-3 também diferem na sequência da proteína NT- 4.

FIGURA 5. Expressão transiente da proteína NT-4 de Xenopus em células COS e a sua interação com NGF-R em células PC12.

(A) SDS-PAGE de meios condicionados de culturas de células COS marcadas in vivo (3×10^4 cpm aplicados em cada faixa) transfectadas com o gene de NGF de rato, um plasmídeo de controlo sem inserção, ou o gene de NT-4 de Xenopus. É apresentada uma auto-radiografia do gel seco após uma exposição durante a noite a película de raios X.

(B) Diluições seriadas de meio de células COS transfectadas, contendo quantidades iguais de proteínas NT-4 (círculos abertos), ou NGF (círculos fechados) foram testadas quanto à sua capacidade para substituir ^{125}I -NGF no seu receptor em células PC12. Foram realizados testes de ligação a 37°C utilizando ^{125}I -NGF $1.5 \times 10^9\text{M}$ e 1×10^4 células por ml. Meio de células simuladamente transfectadas não conseguiu substituir a ligação de ^{125}I -NGF das células PC12. Cada ponto representa a média $\pm$ SD de determinações em triplicado.

FIGURA 6. Estimulação de excrescência de neurites em gânglios embrionários de galinha.

(A, B, e C) Excrescência de neurites induzida em gânglios da raiz dorsal com proteína NT-4 recombinante (A), NGF recombinante (B) e proteína BDNF (C).

(D) A resposta de gânglios da raiz dorsal a meio condicionado de células simuladamente transfectadas.

(E e F) Estimulação de excrescência de neurites de gânglios

do simpático em resposta a NT-4 (E) ou NGF (F).

(G, H, e I) Gânglios nodosos estimulados com proteínas NT-4 recombinante (G), NT-3 (H), e BDNF (I). Todas as figuras são fotografias de microscópio de campo claro de gânglios após 1,5 dias em cultura.

FIGURA 7. Detecção de ARNm de NT-4 em diferentes tecidos de Xenopus.

(A) ARN Poli(A)⁺ (10 µg por poço) dos tecidos indicados de fêmea adulta de Xenopus, foi sujeito a electroforese num gel de agarose contendo formaldeído, transferido para um filtro de nitrocelulose e hibridado com um fragmento HindI de 500 bp do exão 3' do gene de NT-4 de Xenopus. Para comparação, o filtro foi também hibridado com um fragmento de PCR de 180 bp do gene de NGF de Xenopus (centro marcado da faixa, NGF). O filtro hibridado com a sonda de NT-4 foi exposto durante 2 dias; o filtro hibridado com a sonda de NGF foi exposto durante 2 semanas. Uma exposição prolongada de 2 semanas do filtro hibridado com a sonda de NT-4 não revelou ARNm de NT-4 em nenhum outro tecido senão o de ovário que inclui oócitos de diferentes etapas. A faixa marcada com CNS inclui cérebro e espinal medula.

(B) ARN Poli(A)⁺ (10 µg) de ovário de Xenopus foi analisado quanto à expressão dos quatro membros da família de NGF. Cada filtro foi hibridado com as sondas indicadas, obtidas por marcação de fragmentos de PCR dos seus respectivos genes de Xenopus. A localização dos fragmentos de PCR marcados no exão 3' dos seus genes é apresentada na Figura 1A. Os filtros foram lavados com elevada severidade e expostos a películas de raios X durante 5 dias.

FIGURA 8. Sequência de nucleótidos de NT-4 de Xenopus com locais de clivagem de endonucleases de restrição (SEQ ID NO: 49 e SEQ ID NO: 50).

FIGURA 9. Expressão de ARNm de NT-4 em ovário de Xenopus laevis.

Ovário de Xenopus laevis adulto foi seccionado num crióstato (secções de 14 μ m de espessura) e as secções foram depois hibridadas com os oligonucleótidos de mero 48 indicados, marcados com ^{35} S-dATP utilizando desoxinucleotidil-transferase terminal.

(A) Hibridação utilizando um oligonucleótido específico de ARNm de NT-4 de Xenopus com a sequência
5'CCCACAAGCTTGTGGCATCTATGGTCAGAGCCCTCACATAAGACTGTTTTGC3'.

(SEQ ID NO: 95)

(B) Hibridação utilizando um oligonucleótido de controlo de comprimento e conteúdo em G+C semelhantes. Após hibridação as secções foram lavadas em 1 x SSC a 55 °C seguido de exposição a películas de raios X durante 10 dias. São apresentadas na figura fotografias das películas de raios X reveladas. Note-se a marcação intensa sobre muitas células pequenas com a sonda de NT-4 e a ausência de marcação com a sonda de controlo. As setas apontam oócitos grandes (etapa VI) que não são marcados com nenhuma das duas sondas. Barra da escala, 2 mm.

FIGURA 10. Iluminação de campo claro de auto-radiografias de emulsão apresentando oócitos que expressam ARNm de NT-4 no ovário de Xenopus. Secções hibridadas com as sondas de ARNm de NT-4 de Xenopus, específicas (A,B) ou de controlo (C), como descrito na FIG. 9 foram cobertas com emulsão Kodak NTB2, expostas durante 5 semanas, reveladas e contrastadas à luz com violeta de cresilo. Esta figura apresenta micro-fotografias de campo claro das secções reveladas. Note-se no quadro A a marcação intensa de ARNm de NT-4 sobre oócitos de pequena dimensão (Etapas I e II) e a ausência de marcação sobre oócitos de grande dimensão (Etapas V e VI). O quadro B apresenta uma maior ampliação da área que se encontra dentro do rectângulo no quadro A. Note-se a intensa marcação do citoplasma dos oócitos de etapa II apresentados na fotografia.

(C) Nenhuma marcação pode ser observada usando a sonda de controlo. Abreviaturas: n, núcleo; fc, células do folículo;

pl, camada pigmentada. Barra da escala em A, 50 μm ; em B e em C, 15 μm .

FIGURA 11. Níveis de ARNm de NT-4 em oócitos em diferentes etapas da oógenese. Foram utilizadas autorradiografias de emulsão (apresentadas na figura 10) de secções hibridadas com a sonda específica de ARNm de NT-4 de Xenopus para contar o número de grãos numa unidade de área. A unidade de área escolhida foi cerca de um centésimo de um oócito de etapa I. Foram analisadas quinze unidades de área em 10 oócitos diferentes nas etapas indicadas. Barras de erro apresentam o S.D.

FIGURA 12. Análise de expressão de ARNm de NT-4 por transferência de Northern durante a oógenese em Xenopus laevis. Ovários de dois Xenopus laevis adultos foram dissecados e tratados com colagenase para remover células de folículo e libertar os oócitos. Os oócitos foram então agrupados nos grupos indicados seguindo as etapas descritas por Dumont, 1972, J. Morphol. 136: 153-180. O ovário total e as células de folículo libertadas foram também incluídos na análise. O ARN celular total foi então preparado e 40 μg /poço de ARN foi submetido a electroforese num gel de agarose a 1% contendo formaldeído. Este foi transferido para um filtro de nitrocelulose e hibridado com um fragmento HincII de 600 bp do exão 3' do gene de NT-4 de Xenopus. O filtro foi lavado com elevada severidade e exposto durante cinco dias a uma película de raios X. Note-se o decréscimo acentuado no nível de ARNm de NT-4 em oócitos nas etapas V e VI.

FIGURA 13. (A) A sequência de aminoácidos parcial de xNT-4 (SEQ ID NO: 51) indicando as posições onde os oligonucleótidos degenerados foram sintetizados e utilizados para iniciar a amplificação de ADN genómico humano e de rato através da reacção em cadeia com polimerase. As setas indicam oligonucleótidos representando oligonucleótidos degenerados com sentido ou anti-sentido. Um conjunto de oligonucleótidos degenerados para o iniciador 2Z representa os aminoácidos 184-189 de rBDNF (SEQ ID NO: 52). A sequência de aminoácidos parcial de NT-4 de Xenopus

representada é do aminoácido 167 - até ao aminoácido 223, como descrito na Figura 4, supra.

(B) Oligonucleótidos degenerados utilizados para clonagem de NT-4 humano e de rato. O oligonucleótido 3Z na Figura 13 é compreendido por uma mistura de 3Z e 3Z' por forma a permitir a degenerescência do codão da serina. 2Y (SEQ ID NO: 53), 2Z (SEQ ID NO: 54), 3Y (SEQ ID NO: 55), 3Z (SEQ ID NO: 56), 3'Z (SEQ ID NO: 57) e 4Z (SEQ ID NO: 58). (C) Caudas de clonagem para oligonucleótidos degenerados 3' (SEQ ID NO: 59) e 5' (SEQ ID NO: 60).

FIGURA 14. Sequência de ADN do fragmento isolado que codifica uma porção do NT-4 de rato (SEQ ID NO: 61). A estrutura de leitura aberta prevista para o péptido codificado pelo fragmento de ácido nucleico rNT-4 está representada pelo código de letra única (SEQ ID NO: 62). A sequência entre parênteses constitui parte do iniciador de PCR.

FIGURA 15. Sequência de ADN do fragmento isolado que codifica uma porção de NT-4 humano (SEQ ID NO: 63). A estrutura de leitura aberta prevista para o péptido codificado pelo fragmento de ácido nucleico hNT-4 está representada pelo código de letra única (SEQ ID NO: 64). A sequência entre parênteses constitui parte do iniciador de PCR.

FIGURA 16. Alinhamento de sequência de aminoácidos deduzidas a partir de neurotrofinas representativas. Os aminoácidos estão indicados usando o código de letra única. Aminoácidos idênticos estão indicados por pontos. As linhas a tracejado indicam uma inserção de 7 aminoácidos dentro da região conservada de ambos rNT-4 (SEQ ID NO: 62) e hNT-4 (SEQ ID NO: 64). xNT-4 (SEQ ID NO: 65), rNGF (SEQ ID NO: 66), rBDNF (SEQ ID NO: 67), rNT-3 (SEQ ID NO: 68). x=Xenopus, r=rato, h=humano. A sequência entre parênteses constitui parte do iniciador de PCR.

FIGURA 17. (A) Sequência de ADN de um fragmento isolado que codifica uma porção de NT-4 humano (SEQ ID NO: 69). O péptido

previsto codificado pelo fragmento de ácido nucleico hNT-4 de 192 bp é representado pelo código de letra única (SEQ ID NO: 70). A sequência entre parênteses constitui parte do iniciador de PCR.

(B) Sequência de oligonucleótidos do iniciador da extremidade 5', denominado hNT-4-5'' [contendo uma sequência (SEQ ID NO: 71) que codifica ETRCKA (SEQ ID NO: 72)], utilizada na amplificação primária de ADN genómico humano juntamente com o iniciador da extremidade 3', denominado 4Z (SEQ ID NO: 58) [contendo uma sequência de nucleótidos que codifica WIRIDT]. (C) Sequência de oligonucleótidos do iniciador da extremidade 5' utilizado para amplificar o produto de reacção do PCR primário. O iniciador, denominado hNT-4-5''' [contendo uma sequência (SEQ ID NO: 73) que codifica DNAEEG (SEQ ID NO: 74)] foi utilizado com o iniciador 3', 4Z (SEQ ID NO: 58), para obter um fragmento de 162 bp (mais os bp da cauda de clonagem). O fragmento de PCR de 162 bp foi depois utilizado numa reacção de separação de PCR utilizando o nosso fragmento de PCR anteriormente utilizado a montante (denominado 2YZ3Z) para gerar o fragmento único de 192 bp mais a cauda de clonagem apresentado em (A). Informação adicional da sequência de ácido nucleico prolongada em 3' foi obtida seguindo a subclonagem e a sequenciação deste fragmento.

FIGURA 18. Sequência de ADN da porção do clone fágico de genoma humano isolado 7-2 que codifica NT-4 humano (SEQ ID NO: 75). A proteína hNT-4 prevista codificada pelo clone genómico 7-2 é representada pelos símbolos de letra única para aminoácidos (SEQ ID NO: 76). A região em caixa representa o local de clivagem previsto da préproteína hNT-4. As setas indicam resíduos conservados na pré-sequência. A região sublinhada (N-R-S) representa uma sequência de consenso para n-glicosilação. A região em círculo representa a metionina de iniciação. O local de aceitação de união está localizado nos pares de bases 461-462 (AG) da SEQ ID NO: 75, representando a extremidade 3' do intrão.

FIGURA 19. Alinhamento de sequências de aminoácidos deduzidas a

partir de neurotrofinas representativas (SEQ ID NOS. 77-92). Os aminoácidos estão indicados utilizando o código de letra única. Aminoácidos idênticos aos codificados pelo clone fágico de genoma humano 7-2 (SEQ ID NO: 77) estão indicados com um asterisco. As linhas a tracejado representam interrupções em aminoácidos homólogos quando comparados com a proteína codificada pela SEQ ID NO: 77.

FIGURA 20. Sequência de ADN do fragmento isolado que codifica uma porção do clone fágico de genoma humano, 2-1 (SEQ ID NO: 93). A estrutura de leitura aberta prevista para o péptido codificado pelo fragmento de ácido nucleico isolado está representada pelo código de letra única (SEQ ID NO: 94).

FIGURA 21. Sequência de ADN do fragmento isolado que codifica uma porção do clone fágico de genoma humano, 4-2 (SEQ ID NO: 116). A estrutura de leitura aberta prevista para o péptido codificado pelo fragmento de ácido nucleico isolado está representada pelo código de letra única (SEQ ID NO: 117).

FIGURA 22. Análise da expressão de ARNm de NT-4 humano por transferência de Northern. ARNm específico de tecido de humano foi adquirido à Clontech. Foram fraccionados por electroforese ARN (10 μ g) num gel de 1% de agarose-formaldeído seguida de transferência por capilaridade para uma membrana de nylon (MagnaGraph, Micron Separations, Inc.) com 10X SSC (pH 7). Os ARN foram reticulados na membrana sob a acção de UV por exposição a luz ultravioleta (Stratalinker, Stratagene, Inc) e hibridados a 65 °C com uma sonda marcada radioactivamente (um fragmento XhoI-NotI de 680 bp contendo a região de codificação completa de HG7-2 NT-4 (ver Exemplo na Secção 9, infra) na presença de NaPO_4 0,5 M (pH 7), 1% de albumina de soro bovino (Fracção V, Sigma), 7% de SDS, EDTA 1 mM (Mahmoudi e Lin, 1989, Biotechniques 7: 331-333), e 100 μ g/ml de ADN de esperma de salmão desnaturado, tratado com ultra-sons (sonicated). O filtro foi lavado a 65 °C com 2X SSC, 0,1% de SDS e submetido a auto-radiografia durante a noite com um écran de intensificação (Cronex, DuPont) e película de raios X (XAR-5, Kodak) a -70 °C.

A coloração do gel com brometo de etídio demonstrou que tinham sido testados níveis equivalentes de ARN total para as diferentes amostras (como em Maisonpierre *et al.*, 1990, *Science* 247: 1446-1451).

Faixa 1: ARNm poli(A)⁺ de fígado fetal; Faixa 2: ARNm poli(A)⁺ de cérebro fetal; Faixa 3: ARNm poli(A)⁺ de próstata; Faixa 4: ARNm poli(A)⁺ de músculo; Faixa 5: ARNm poli(A)⁺ de intestino; Faixa 6: ARNm poli(A)⁺ de rim; Faixa 7: ARNm poli(A)⁺ de fígado; Faixa 8: ARNm poli(A)⁺ de baço; Faixa 9: ARNm poli(A)⁺ de timo; Faixa 10: ARNm poli(A)⁺ de ovário; Faixa 11: ARNm poli(A)⁺ de testículo; Faixa 12: ARNm poli(A)⁺ de placenta; Faixa 13: ARNm poli(A)⁺ de cérebro; Faixa 14: ARN total de cérebro.

FIGURA 23. Sobrenadantes de linhas de células COS transfectadas; Q1 (pCMX-HG7-2Q), N7 (pCMX-hNT-3/hNT-4) e X1 (pCMX-xNT4/hNT4) foram testados em volumes de 10 µl, 50 µl e 250 µl quanto a actividade promotora de neurites em explantes DRG. Foi utilizado como controlo um sobrenadante de uma linha de células COS simuladamente transfectada.

FIGURA 24. Os sobrenadantes de linhas de células COS, Q1 (pCMX-HG7-2Q) e M (pCMX-HG7-2M) foram testados quanto à sua actividade promotora de sobrevivência em células associadas a DRG. Os volumes testados variaram de 5 µl até 250 µl num volume total de 2 ml.

FIGURA 25. Culturas enriquecidas em neurónios motores isoladas a partir de embriões de rato E14 foram tratadas com duas diluições de sobrenadantes de células COS da linha celular M (pCMX-HG7-2M). A actividade biológica foi medida pela actividade da colina-acetiltransferase (CAT), como descrito em Fonnum, 1975, *J. Neurochem.* 24: 407-409. Foram apresentados como controlos tanto uma linha de células COS simuladamente transfectada (MOC COS), como um neurónio motor sem tratamento (C-NT).

5. DESCRIÇÃO DETALHADA DO INVENTO

O presente invento proporciona genes e proteínas de NT-4. Baseia-se, pelo menos em parte, na clonagem, caracterização e expressão do gene de NT-4.

Em particular, o presente invento proporciona moléculas de ácido nucleico recombinante que codificam NT-4. Estas moléculas compreendem uma sequência substancialmente como a apresentada na Figura 1 (SEQ ID NO: 1) para víbora, Figura 1 (SEQ ID NO: 2), Figura 4 (SEQ ID NO: 43) ou Figura 8 (SEQ ID NO: 49) para NT-4 de Xenopus, Figura 14 (SEQ ID NO: 61) para NT-4 de rato, Figura 15 (SEQ ID NO: 63), Figura 17 (SEQ ID NO: 69) ou Figura 18 (SEQ ID NO: 75) para NT-4 humano, Figura 20 (SEQ ID NO: 93) e Figura 21 (SEQ ID NO: 116) para uma sequência de tipo NT-4 humano, ou uma sequência que seja pelo menos cerca de setenta por cento homóloga com qualquer dessas sequências, na qual a homologia se refere a identidade de sequência (por exemplo, uma sequência que possua 70 por cento de homologia com uma segunda sequência partilha 70 por cento dos mesmos resíduos nucleotídicos com a segunda sequência).

Num aspecto particular, o presente invento, detalhado na Secção 8 do Exemplo e na Figura 15 (SEQ ID NO:63, SEQ ID NO:64) aqui apresentadas, é determinada a sequência de nucleótidos e de aminoácidos para uma porção de uma molécula de neurotrofina humana. Noutro aspecto do presente invento, detalhado na Secção 9 do Exemplo e na Figura 17 (SEQ ID NO:69, SEQ ID NO:70) e na Figura 18 (SEQ ID NO:75, SEQ ID NO:76) aqui apresentadas, é determinada a sequência de nucleótidos e de aminoácidos para a totalidade da molécula de neurotrofina humana. Em outro aspecto do presente invento detalhado na secção 9 do Exemplo e na Figura 20 (SEQ ID NO:93, SEQ ID NO:94) e na Figura 21 (SEQ ID NO:116, SEQ ID NO:117) aqui apresentadas, são detalhadas as sequências de nucleótidos e de aminoácidos para uma porção de clones dum fago genómico humano, 2-1 e 4-2, respectivamente, que são semelhantes mas não idênticos às sequências de nucleótidos e de aminoácidos descritas na Figura 18 (SEQ ID NO: 75). Embora esta molécula de neurotrofina humana seja aqui referida como

neurotrofina-4 humana, deve ser compreendido que esta molécula pode ser o homólogo humano da neurotrofina-4 de Xenopus aqui descrita ou, alternativamente, uma molécula de neurotrofina distinta porém homóloga. De modo semelhante, a molécula aqui referida como NT-4 de rato pode ser o homólogo de rato de NT-4 ou, alternativamente, uma molécula de neurotrofina distinta porém homóloga. Os processos e composições do presente invento não dependem de qualquer nomenclatura individual.

O presente invento também proporciona proteína NT-4 ou moléculas peptídicas substancialmente purificadas. Estas moléculas podem compreender uma sequência substancialmente como aqui apresentada na Figura 2, (SEQ ID NO: 1 e SEQ ID NO: 2), Figura 4 (SEQ ID NO: 44) Figura 8 (SEQ ID NO: 50), Figura 14 (SEQ ID NO: 62), Figura 15 (SEQ ID NO: 64) Figura 17 (SEQ ID NO: 70), Figura 18 (SEQ ID NO: 76), Figura 20 (SEQ ID NO: 94) e Figura 21 (SEQ ID NO: 117) para NT-4, ou uma sequência que seja pelo menos setenta por cento homóloga com qualquer destas sequências. Em concretizações adicionais específicas não limitantes do invento, uma proteína ou péptido substancialmente purificado compreende a sequência KCNPSGSTTR (SEQ ID NO: 96). Noutra concretização do invento, um péptido ou proteína substancialmente purificado compreende a sequência RGCRGVD (SEQ ID NO: 97). Ainda noutra concretização do invento, um péptido ou proteína substancialmente purificado compreende a sequência KQWIS (SEQ ID NO: 98). Ainda numa outra concretização do invento, um péptido ou proteína substancialmente purificado compreende a sequência KQSYVR (SEQ ID NO: 99). Ainda noutra concretização do invento, um péptido ou proteína substancialmente purificado compreende a sequência GPGXGGG (SEQ ID NO: 100), onde X representa um do grupo de 20 aminoácidos. Numa concretização afim do invento, um péptido ou proteína substancialmente purificado compreende a sequência GPGVGGG (SEQ ID NO: 101) ou GPGAGGG (SEQ ID NO: 102). Ainda noutra concretização do invento, um péptido ou proteína substancialmente purificado compreende a sequência ESAGE (SEQ ID NO: 103). Ainda noutra concretização do invento, um péptido ou proteína substancialmente purificado compreende a sequência

DNAEE (SEQ ID NO: 104).

As proteínas e péptidos do invento podem ser produzidas por síntese química utilizando técnicas padrão ou podem ser produzidas usando as moléculas de ácido nucleico do invento que codificam NT-4, utilizando sistemas de expressão eucarióticos ou procarióticos, conhecidos para um perito na arte, como os descritos no pedido de PCT, PCT/US90/04916, depositado em 29 de Agosto de 1990, publicado como WO 91/03569, que é aqui incorporado como referência na íntegra, ou como exemplificado infra (ver Secção 6.2.4., infra, e Figura 5) para expressão transiente em células COS.

O presente invento também proporciona o uso de NT-4 na promoção do crescimento e/ou da sobrevivência de células do sistema nervoso, em particular, mas não limitado a, células do gânglio da raiz dorsal ou de derivados de placódio neural (ver Secção 6.2.4., infra, e Figura 6, por exemplo).

O presente invento também proporciona porções de ácido nucleico ou sequência de aminoácidos de NT-4, substancialmente como se segue para NT-4 nas Figuras 1, 2, 4, 8, 14, 15, 17, 18, 20, ou 21 (SEQ ID NO listados supra) que não são idênticas a porções de BDNF, NGF, ou NT-3 substancialmente do mesmo tamanho.

O presente invento proporciona, para além disso, uma célula eucariótica ou procariótica que contém ácido nucleico recombinante que codifica NT-4 e que expressa proteína NT-4 recombinante. Numa concretização específica a célula é uma célula eucariótica, como uma célula COS. De acordo com isto, o presente invento também proporciona uma proteína ou péptido de NT-4 recombinante que é produzido por inserção na célula de ácido nucleico recombinante que codifica NT-4 (por ex., por transfecção, transdução, electroporação, micro-injecção, etc.) sob condições que permitam a expressão de NT-4 e posterior isolamento de NT-4 da célula.

Adicionalmente, o presente invento proporciona moléculas

produzidas por PCR utilizando, por exemplo, os seguintes oligonucleótidos como iniciadores:

5'CAGTATTTTACGAAACC (SEQ ID NO: 105) e

3'GTCTTGGTTGGCTTTACA (SEQ ID NO: 106) para NT-4 humano e

5'CAGTATTTTACGAGACG (SEQ ID NO: 107) e

3'CGATTGTTTGGCTTTACA (SEQ ID NO: 108) para NT-4 de rato, e

utilizando como molde qualquer ADN genómico ou ADNc adequado.

Numa concretização específica do invento, estes iniciadores

podem ser utilizados em conjugação com ADNc humano como molde

para produzir fragmentos do gene de NT-4 humano que sejam

adequados para clonagem.

A produção e uso de derivados, análogos, e péptidos afins com NT-4 são também considerados, e dentro do âmbito do presente invento. Tais derivados, análogos, ou péptidos que possuem a desejada actividade neurotrófica, imunogenicidade ou antigenicidade podem ser utilizados para, por exemplo, terapêutica, ou em imunoensaios, para imunização, etc. Derivados, análogos ou péptidos afins com NT-4 podem ser testados quanto à actividade desejada por processos conhecidos na arte.

Os derivados afins com NT-4, análogos e péptidos do invento podem ser produzidos por vários processos conhecidos na arte. As manipulações que resultam na sua produção podem ocorrer ao nível do gene ou da proteína. Por exemplo, o gene de NT-4 clonado pode ser modificado por qualquer uma das numerosas estratégias conhecidas na arte (Maniatis, T., 1982, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). A sequência de NT-4 pode ser clivada em locais apropriados com endonuclease(s) de restrição, seguida de posterior modificação enzimática, se desejada, isolada e ligada in vitro. Na produção do gene que codifica um derivado, análogo ou péptido afim com NT-4, deve ser tomado cuidado para assegurar que o gene modificado permanece na mesma estrutura de leitura aberta que a NT-4, não interrompido por sinais de paragem de tradução, na região do gene onde a actividade específica de NT-4 desejada está codificada.

Adicionalmente, o gene de NT-4 pode ser mutado in vitro ou in vivo, para criar e/ou destruir sequências de tradução, iniciação, e/ou terminação, ou para criar variações em regiões de codificação e/ou formar novos locais de endonuclease de restrição ou destruir outros pré-existentes, para facilitar a posterior modificação in vitro. Qualquer técnica para mutagênese conhecida na arte pode ser utilizada incluindo, mas não estando limitado a, mutagênese localmente dirigida in vitro (Hutchison, C., et al., 1978, J. Biol. Chem. 253: 6551), uso de ligantes TAB^R (Pharmacia), etc.

Como discutido infra, a prepro ou região madura de codificação de NT-4, pode ser utilizada para construir genes quiméricos baseados na neurotrofina. Por exemplo, genes de neurotrofina incluindo, mas não limitados a NGF, BDNF e NT-3 podem proporcionar a região prepro para a construção de genes quiméricos de neurotrofina e prepro/região de codificação madura de NT-4.

Manipulações da sequência NT-4 podem também ser realizadas ao nível da proteína. Qualquer uma das numerosas modificações químicas pode ser realizada por técnicas conhecidas incluindo, mas não estando limitadas a clivagem química específica por brometo de cianogénio, tripsina, quimiotripsina, papaína, protease V8, NaBH₄, acetilação, formilação, oxidação, redução, síntese metabólica na presença de tunicamicina, etc.

Adicionalmente, podem ser sintetizados quimicamente análogos e péptidos afins com NT-4. Por exemplo, um péptido correspondente a uma porção de NT-4 que medeia a actividade neurotrófica desejada pode ser sintetizado utilizando um sintetizador de péptidos.

O presente invento proporciona, para além disso, um processo de tratamento de desordens de fertilidade relacionadas com disfunção ovário/oócito. Como apresentado nos exemplos infra, em particular na Secção 7, a NT-4 está envolvida na maturação dos oócitos. A discussão da Secção 7.3 demonstra que a

NT-4 é produzida por oócitos, é concentrada preferencialmente nos oócitos imaturos em relação aos maduros, e parece desempenhar um papel na oógenese. A função putativa da proteína NT-4 no ovário parece estar acoplada a acontecimentos que ocorrem no oócito pré-vitelogénico e vitelogénico precoce/semi-precoce.

Tem sido estabelecido que outros membros da família de genes BDNF/NGF/NT-3, em particular o NGF, estão envolvidos na maturação meiótica (Nebrada et al., 1991, Science, 252: 558-563). Uma vez que a NT-4 tem exibido propriedades semelhantes a NGF (ver Secção 6, *infra*) pode ser usado como um factor envolvido na regulação do desenvolvimento do oócito. Estas propriedades do NT-4 podem ser exploradas para proporcionar um processo para tratamento de desordens de infertilidade e/ou outras disfunções dos ovários associadas com a oógenese.

Assim, de acordo com o invento, é proporcionado um processo de tratamento de desordens de infertilidade e/ou outras disfunções dos ovários compreendendo administrar uma quantidade terapêuticamente eficaz de NT-4 ou de um péptido afim com NT-4 num portador farmacêuticamente eficaz. Uma quantidade terapêuticamente eficaz é aquela que induz a maturação apropriada de um oócito e/ou ovulação. Por exemplo, uma dose terapêuticamente eficaz pode ser uma suficiente para manter níveis de NT-4 em circulação no soro a uma concentração de entre cerca de 1 a 100×10^{-10} M. O estabelecimento de doses eficazes adicionais está dentro da previsão de um perito na arte.

Doses eficazes de NT-4 ou de um péptido afim com NT-4 formuladas em portadores farmacológicos apropriados podem ser administradas por qualquer via apropriada incluindo, mas não limitada a, injeção (por ex., intravenosa, intraperitoneal, intramuscular, subcutânea, etc.), por absorção através de camadas epiteliais ou mucocutâneas (por ex., mucosa oral, mucosa rectal ou intestinal, etc.); etc.

Adicionalmente, a NT-4 ou um péptido de NT-4 pode ser utilizado em qualquer portador farmacológico adequado, ligado a um portador ou molécula alvo (por ex., anticorpo, hormona, factor de crescimento, etc.) e/ou incorporado em liposomas, microcápsulas, e preparação para libertação controlada anterior à administração in vivo.

Cada sequência de ADN de NT-4 de mamífero respectiva, pode ser utilizada como uma sonda marcada com ^{32}P para isolar um clone genómico e de ADNc respectivo através dos processos salientados na parte de Materiais e Métodos da Secção 8, infra. Os fragmentos do gene NT-4 humano e NT-4 de rato podem ser utilizados directamente (como sondas marcadas com ^{32}P) ou indirectamente (para deduzir uma estratégia de PCR como descrito infra) para isolar clones de ADNc ou genómicos de NT-4 de outros mamíferos, com base na natureza única da inserção do aminoácido 7 na região de codificação de rNT-4 e hNT-4, ou outros aspectos únicos da região de codificação de NT-4 humano ou de rato.

Qualquer gene de NT-4 de mamífero isolado através da informação revelada pela sequência de NT-4 humano e de rato pode ser utilizada nas várias manipulações discutidas para NT-4 de Xenopus, embora não seja limitada a estas. Por exemplo, as proteínas e péptidos de NT-4 de mamífero, subsequentes à caracterização do gene de comprimento total como discutido no Exemplo da Secção 9, podem ser produzidos utilizando as moléculas de NT-4 do respectivo mamífero num sistema de expressão procariótico ou eucariótico conhecido por um perito na arte, como os descritos no pedido de PCT, PCT/US90/04916, depositado em 29 de Agosto de 1990, publicado como WO91/03569, ou como exemplificado infra (ver Secção 6.2.4., supra, e Figura 5) para expressão transiente em células COS. Funções adicionais para NT-4 de mamífero, como descrito infra para NT-4 de Xenopus, incluem, mas não estão limitadas a: promoção de crescimento e/ou sobrevivência de células do sistema nervoso, em particular, mas não limitado a, células do gânglio da raiz dorsal ou derivados do placódio neural (ver Secção 6.2.4., e Figura 6, por exemplo), tratamento de desordens de fertilidade relacionadas com

disfunção ovário/oócito (ver Secção 7), tratamento de desordens de infertilidade e/ou outras disfunções do ovário associadas com a oogénese (ver Secção 6), tratamento de doenças do neurónio motor (ver Secção 10), tratamento de uma hiperplasia epitelial como a hipertrofia prostática benigna (ver Secção 10), tratamento da impotência enquanto relacionada com a função glandular da próstata (ver Secção 10) e, assim, as quantidades terapêuticamente eficazes de NT-4 de mamífero para o tratamento das ditas desordens como formulado em portadores farmacológicos adequados para proporcionar uma composição farmacêutica, podem ser administradas por qualquer via apropriada incluindo mas não limitada a injeção (por ex., intravenosa, intraperitoneal, intramuscular, subcutânea, etc.), por absorção através da camada epitelial ou mucocutânea (por ex., mucosa oral, mucosa rectal e intestinal, etc.); etc.

Adicionalmente, a NT-4 ou péptido de NT-4 de rato, humano ou de outro mamífero pode ser utilizado em qualquer portador farmacológico adequado, ligado a um portador ou molécula alvo (por ex., anticorpo, hormona, factor de crescimento, etc.) e/ou incorporado em liposomas, microcápsulas, e preparações de libertação controlada anterior à administração in vivo.

Adicionalmente, o presente invento, que se refere a ácidos nucleicos que codificam NT-4 e a proteínas, fragmentos peptídicos ou derivados a partir deles produzidos, bem como a anticorpos dirigidos contra proteína NT-4, péptidos ou derivados, pode ser utilizado para diagnosticar ou monitorizar a progressão de doenças e desordens do sistema nervoso que estão associadas com alterações no padrão de expressão de NT-4. Essas alterações podem ser um decréscimo ou aumento relativamente ao de pacientes normais, preferivelmente, ou de outras amostras retiradas do paciente, ou de amostras do mesmo paciente retiradas anteriormente.

Em várias concretizações do invento genes de NT-4 e sequências e subsequências de ácido nucleico afins, incluindo sequências complementares, podem ser utilizadas em testes de

hibridação para diagnóstico. As sequências de ácido nucleico de NT-4 ou suas subsequências, compreendendo cerca de 15 nucleótidos, podem ser usadas como sondas de hibridação. Testes de hibridação podem ser utilizados para detectar, prognosticar, diagnosticar ou monitorizar condições, desordens ou estados de doença associados com alterações nos níveis de NT-4. Por exemplo, os dados apresentados no Exemplo da Secção 10 revelam expressão de NT-4 humano específica no tecido em músculo esquelético bem como na glândula da próstata, timo e testículos. O nível de expressão de NT-4 humano no tecido muscular pode ser indicativo da presença ou ausência de degradação neuronal. Assim, ARNm poli(A)⁺ ou ARN total de uma amostra de tecido de um paciente pode ser testada quanto à presença de ARNm de NT-4 humano no tecido do músculo esquelético. Adicionalmente, os dados apresentados na Secção 10 do Exemplo revelam expressão específica de NT-4 de tecido na glândula da próstata humana. Sequências de ADN que codificam NT-4 ou uma porção deste, bem como proteína ou um péptido de NT-4, podem ser úteis como agente terapêutico para tratar doenças da próstata.

Num processo semelhante, os testes de diagnóstico podem ser imunoensaios. Assim, podem ser utilizados anticorpos em imunoensaios para quantificar o nível de NT-4 numa amostra de um paciente de modo a detectar, prognosticar, diagnosticar, ou monitorizar condições, desordens ou estados de doença associados com alterações nos níveis de NT-4.

Os imunoensaios que podem ser utilizados incluem, mas não estão limitados a sistemas de ensaio competitivo e não competitivo, utilizando técnicas tais como radioimunoensaio, ELISA (imunoensaio usando anticorpo ou antigénio ligado a enzimas), imunoensaios pela técnica de "sanduíche", reacções de precipitina, reacções de precipitina de difusão em gel, ensaios de imunodifusão, ensaios de aglutinação, ensaios de fixação do complemento, ensaios imuno-radiométricos, imunoensaios com marcadores fluorescentes, imunoensaios de proteína A, e ensaios de imuno-electroforese, para referir apenas alguns.

Fragmentos de anticorpo anti-NT-4 ou derivados contendo o domínio de ligação podem também ser utilizados nestes ensaios.

Fragmentos de anticorpo que contêm o idiótipo da molécula podem ser gerados por técnicas conhecidas. Por exemplo, tais fragmentos incluem mas não estão limitados a, fragmento $F(ab')_2$ que pode ser produzido por digestão com pepsina da molécula de anticorpo; fragmentos Fab' que podem ser obtidos por redução das pontes de dissulfureto deste e fragmentos Fab que podem ser obtidos pelo tratamento da molécula do anticorpo com papaína e um agente redutor.

São também proporcionados conjuntos de diagnóstico. Por exemplo, um tal conjunto pode incluir uma sonda específica para NT-4 em recipiente adequado. Numa concretização, a sonda é um anticorpo específico para NT-4. Noutra concretização, a sonda é um ácido nucleico (sonda molecular) capaz de hibridar com uma sequência de ácido nucleico de NT-4. A sonda pode ser detectavelmente marcada; alternativamente o conjunto pode ainda incluir um par de ligação específica marcado, para a sonda.

Os imunoensaios e ensaios de hibridação acima descritos podem também ser utilizados para quantificar níveis de NT-4 como indicadores de eficácia terapêutica, por comparação dos níveis em amostras de pacientes antes e depois do tratamento de uma desordem, particularmente, de uma doença de neurónios motores.

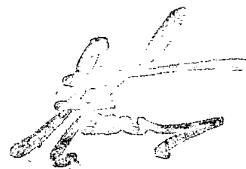
De uma maneira análoga, a expressão de ARNm de NT-4 humana em tecido muscular, conduz a processos de tratamento potenciais de desordens em neurónios motores que incluem a administração a um paciente com necessidade desse tratamento, de uma quantidade eficaz de um factor de NT-4 para apoiar a sobrevivência, crescimento e/ou diferenciação de neurónios motores. A expressão de ARNm de NT-4 em músculo humano sugere vias adicionais para o diagnóstico e tratamento de desordens neuronais.

O transporte axonal retrógrado pode ser característico de NT-4. O transporte retrógrado específico de NT-4 pode ser usado

para indicar se os neurónios estão a responder a NT-4 em estados normais ou de doença. Assim, o presente invento proporciona um processo de diagnóstico de desordens do sistema nervoso periférico relacionadas com NT-4 compreendendo a injeção de um péptido ou proteína NT-4 detectavelmente marcados num nervo periférico, e determinando se o péptido ou a proteína NT-4 marcado está ou não a ser transportado de um modo retrógrado, em que a falha de transporte retrógrado se correlaciona positivamente com a ausência de resposta a NT-4 e indica a presença de uma desordem do sistema nervoso periférico que está relacionada com NT-4. Processos análogos podem ser utilizados para diagnosticar uma desordem do sistema nervoso central. A avaliação do transporte retrógrado pode ser efectuada por qualquer processo conhecido na arte, incluindo mas não se limitando a MRI, CAT, ou varrimento de cintilação. Tais processos podem ser utilizados para identificar a localização de uma lesão do sistema nervoso, pois o transporte retrógrado deve diminuir substancialmente ao atingir a lesão.

O invento proporciona ainda conjuntos para avaliação de transporte retrógrado compreendendo uma proteína NT-4, um derivado ou fragmento detectavelmente marcados, num recipiente. Tal marcação pode ser com um isótopo radioactivo, ou outra marcação conhecida na arte.

O presente invento pode ser utilizado para tratar doenças e desordens do sistema nervoso que podem estar associadas com alterações no padrão da expressão de NT-4 ou podem beneficiar da exposição a NT-4 ou anticorpos anti-NT-4 (ou fragmentos deste contendo o domínio de ligação). Mostrámos que o NT-4 humano é expresso no músculo esquelético (Ver Exemplo da Secção 10, infra). Com base nesta verificação, o invento proporciona o tratamento de doenças em neurónios motores. Um vasto conjunto de desordens neurológicas pode afectar os neurónios motores. Os neurónios motores superiores, por exemplo, são predominantemente afectados por acidentes cerebrovasculares, neoplasmas, infecções e traumas. Os neurónios motores inferiores, ou células da córnea anterior, são secundariamente



afectadas por estes processos, mas adicionalmente são objecto de várias desordens em que a perda da célula da córnea anterior é o efeito primário, incluindo esclerose lateral amiotrópica, atrofia infantil e juvenil da espinal medula, poliomielite e síndrome pós-polio, neuropatias hereditárias motoras e sensoriais, e neuropatias motoras tóxicas (por exemplo, vincristina). As desordens dos neurónios motores que podem ser tratadas de acordo com o presente invento incluem mas não estão limitadas às precedentes. Processos de formulação e administração da proteína NT-4, derivados, fragmentos ou seus anticorpos que podem também ser utilizados, incluem mas não estão limitados aos descritos supra ou aos conhecidos na arte.

O invento pode também ser utilizado para tratar hipertrofia prostática benigna (BPH), uma condição comum mas ainda mal conhecida que ocorre principalmente em machos com mais de 50 anos de idade. A proliferação da próstata durante a BPH pode ser induzida por um factor de crescimento como a NT-4 através de um fenómeno de ansa autócrina. A síntese e excreção de NT-4 será seguida pelo transporte de NT-4 de volta à célula prostática por via de um receptor específico na membrana da célula prostática. Têm sido definidas ansas autócrinas para várias moléculas de factores de crescimento e linhas de células tumorais. Em alguns casos, estas ansas autócrinas foram definidas experimentalmente pelo uso de abordagens anti-sentido para a ruptura da ansa autócrina. Assim, uma aplicação terapêutica do presente invento inclui o uso de um ácido nucleico anti-sentido para NT-4 humana ou uma sua porção para inibir a tradução do ARNm de NT-4 na próstata, (para procedimentos que podem ser utilizados ver o pedido copendente da U.S. Nº. Série 07/728 784 depositado em 3 de Julho de 1991 e aqui incorporado na sua totalidade como referência). Por exemplo, um paciente que sofra de uma doença localizada na próstata, caracterizada pelo aumento da transcrição de um gene de NT-4 no tecido da próstata em relação aos níveis de transcrição do gene de NT-4 na próstata de pacientes normais, pode ser administrada uma quantidade eficaz de um oligonucleótido para tratar uma doença da próstata, preferencialmente a hipertrofia prostática benigna. O

oligonucleótido deve ter pelo menos 6 nucleótidos de comprimento e ser complementar a pelo menos uma porção do transcrito de ARN do gene de NT-4, e assim, ser capaz de hibridar com o transcrito de ARN. Adicionalmente, anticorpos anti-NT-4 podem ser utilizados para inibir a ligação de NT-4 ao seu receptor específico na membrana da célula da próstata. Uma quantidade terapeuticamente eficaz de um ácido nucleico anti-sentido ou de um anticorpo anti-NT-4 pode ser distribuída de qualquer forma supra descrita.

O invento pode também ser utilizado para tratar outras disfunções relacionadas com a próstata, especificamente a impotência. Tal doença pode ser o resultado directo ou indirecto de níveis inadequados de NT-4 na próstata. Assim, tanto a detecção da disfunção como o tratamento do paciente para a impotência por via da aplicação de uma quantidade terapeuticamente eficaz da proteína NT-4, ou um fragmento funcional ou derivado de NT-4 pode ser distribuída por qualquer um dos processos supra descritos.

O presente invento revela a detecção da expressão de NT-4 em tecido do timo humano. Assim, o invento pode também ser utilizado para tratar desordens imunológicas que afectam a transmissão neuromuscular, incluindo mas não estando limitado a miastenia grave, uma desordem auto-imunológica adquirida associada com o receptor da acetilcolina (AChR) nas pregas pós-sinápticas na junção neuromuscular. A doença manifesta-se como fraqueza e fadiga muscular devido ao bloqueio do AChR pós-sináptico ou membranas musculares por ligação de anticorpos específicos ao AChR. (Ver por exemplo, Drachman, 1983, Trends. Neurosci. 6:446-451). O tratamento de tais desordens neurológicas imunologicamente mediadas pode incluir aplicações terapêuticas da proteína NT-4 ou um fragmento funcional ou derivado de NT-4, distribuído por qualquer dos processos supra descritos.

O presente invento proporciona um processo de tratamento de desordens de neurónios motores compreendendo a administração,

a um paciente com necessidade de tal tratamento, de uma quantidade eficaz de uma proteína NT-4, derivado ou fragmento peptídico capaz de apoiar a sobrevivência, crescimento e/ou diferenciação de neurónios motores como demonstrado num sistema de cultura in vitro.

Em concretizações in vitro, quantidades eficazes de factor neurotrófico podem ser desejavelmente determinadas numa base de caso por caso, já que neurónios motores de diferentes fontes tissulares ou de espécies diferentes podem exhibir diferentes sensibilidades para o factor neurotrófico. Para qualquer cultura em particular, pode ser desejável a construção de uma curva de resposta a dose que correlaciona a concentração do factor neurotrófico e a resposta do neurónio motor. Para avaliar a sobrevivência, crescimento, e/ou diferenciação do neurónio motor, podem-se comparar neurónios motores expostos a uma proteína NT-4, derivado ou fragmento peptídico de NT-4 com neurónios motores não expostos a uma proteína NT-4, derivado ou fragmentos peptídicos de NT-4, utilizando por exemplo, corantes vitais para avaliar a sobrevivência, microscopia de contraste de fase e/ou coloração de neurofilamentos para medir a germinação de neurites, ou técnicas que medem a bioactividade de compostos associados a neurónios motores, tais como a colina actiltransferase (CAT), ou quaisquer outros processos conhecidos na arte. A actividade do CAT pode ser medida, por exemplo, por colheita e lise de neurónios motores tratados e não tratados numa solução Tris-HCl (pH 8,6) 20 mM contendo cerca de 1% de Triton X-100, removendo uma alíquota de alguns microlitros, e medindo a actividade de CAT utilizando, como substrato, 0,2 ml de [1- C] acetil-CoA, NaCl 300 mM, brometo de colina 8 mM, EDTA 20 mM, e neostigmina 0,1 mM em tampão NaH₂PO₄ (pH 7,4) 50 mM, utilizando o procedimento de micro-Fonnum como descrito em Fonnum, 1975, J. Neurochem. 24:407-409, aqui incorporado na sua totalidade como referência.

Numa concretização específica não limitante do invento, podem ser preparados neurónios motores, e cultivados in vitro, como se segue. Pelo menos uma porção da espinal medula, obtida

preferencialmente de um organismo embrionário como o rato, pode ser assepticamente obtida e separada do bolbo, gânglios sensoriais e meninges aderentes. Os segmentos ventrais da medula podem então ser isolados, uma vez que os neurónios motores estão localizados nas córneas ventrais (anteriores) da espinal medula. Os segmentos ventrais da medula podem ser dissecados em pedaços pequenos e incubados em cerca de 0,1% de tripsina e 0,01% de desoxiribonuclease tipo 1 em solução salina tamponada com fosfato (PBS) isento de cálcio e magnésio, a 37 °C durante 20 minutos. A solução de tripsina pode então ser removida, e as células podem ser lavadas e colocadas em meio fresco, tal como meio essencial mínimo de Eagle (MEM) a 45%, mistura nutriente F12 de Ham a 45%, soro fetal bovino inativado pelo calor a 5%, soro de cavalo inativado pelo calor a 5%, glutamina (2mM), penicilina G (0,5 U/ml), e estreptomicina (0,5 µg/ml). O tecido pode ser mecanicamente dissociado por trituração suave através de uma pipeta de Pasteur, e os sobrenadantes combinados e filtrados através de um filtro de nylon (por exemplo, Nitex, Tekto; 40 µm). A suspensão celular filtrada pode então ser fraccionada utilizando uma modificação do método apresentado por Schnaar e Shnaffner (1981, J. Neurosci. 1:204-217). Todos os passos são desejavelmente efectuados a 4 °C. Pode ser dissolvida metrizamida em meio F12:MEM (1:1) e pode ser estabelecido um gradiente descontínuo que consiste numa almofada de metrizamida a 18% (por exemplo, 0,5 ml), 3 ml de metrizamida a 17%, 3 ml de metrizamida a 12%, e 3 ml de metrizamida a 8%. A suspensão celular filtrada (por exemplo, 2,5 ml) pode ser disposta em camadas sobre o gradiente descontínuo e o tubo pode ser centrifugado a 2500 g durante 15 minutos utilizando um rotor oscilante (por exemplo, Sorvall HB4). A centrifugação pode resultar em três camadas de células: fracção I (na interface 0-8%), fracção II (na interface 8-12%) e fracção III (na interface 12-17%). A fracção I, enriquecida em neurónios motores, pode ser removida num volume pequeno (por exemplo, cerca de 1 ml) e lavado duas vezes com um meio definido isento de soro, tal como 50% de F12 e 50% de meio MEM suplementado com glutamina (2mM), insulina (5 µg/ml), transferrina (100 µg/ml), progesterona (20 nM), putrescina (100µM), e selenite de sódio

(30 nM, ver Bottenstein e Sato, 1979, Proc. Natl. Acad. Sci. USA 76: 514-517). A contagem das células viáveis pode então ser obtida por contagem em hemocitómetro na presença de azul de tripiano. A suspensão de células enriquecidas em neurónios motores pode então ser plaqueada a uma densidade de cerca de 100 000 células/cm² em poços de cultura de tecidos (preferencialmente de 6 mm) previamente revestidos com poli-L-ornitina (por exemplo, 10 µg/ml) e laminina (por exemplo, 10 µg/ml). Uma proteína NT-4, factor derivado ou peptídico, podem ser adicionados. Por exemplo, em concretizações específicas, a NT-4 pode ser adicionada para obter uma concentração final de cerca de 0,01 e 100 ng/ml, e preferencialmente cerca de 50 ng/ml. As culturas de neurónios motores podem ser mantidas em meio definido isento de soro a 37°C numa atmosfera de 95% de ar/ 5% de CO₂ e cerca de 100 % de humidade relativa.

Numa concretização adicional do invento, a sequência de ácido nucleico recombinante afim com NT-4, tal como a contida no bacteriófago HG7-2, HG4-2, e/ou HG2-1, pode ser utilizado para construir genes quiméricos de prepro/NT-4 maduro. Por exemplo, quando se deseja exprimir uma proteína NT-4 madura, um derivado ou fragmento peptídico in vivo ou in vitro, pode-se fundir a região prepro de um gene neurotrófico distinto com a região codificadora madura da sequência afim com NT-4. Os genes neurotróficos que podem proporcionar a região prepro incluem mas não estão limitados a NGF, BDNF, e NT-3. Tal construção quimérica pode promover o aumento da estabilidade do transcrito do ARNm quimérico em relação a um transcrito de ARNm de NT-4 do tipo selvagem, pode aumentar a eficiência da tradução ou pode produzir um molde mais adequado para o processamento proteolítico numa proteína ou fragmento peptídico de neurotrofina madura, biologicamente activa, aumentando assim a expressão. Qualquer perito na arte possui o conhecimento necessário para construir tais sequências quiméricas de ácido nucleico, dadas as sequências de ADN publicadas de outros genes de neurotrofina como NGF (Scott et al., 1983, Nature 302:538-540, Ullrich et al., 1983, Nature 302:821-825), BDNF

(Leibrock *et al.*, 1989, *Nature* 341:149-152) e NT-3 (Honh *et al.*, 1990, *Nature* 344:339-341; Maisonnier *et al.*, 1990, *Science* 247:1446-1451; Ernfors *et al.*, 1990, *Proc. Natl. Acad. Sci. USA* 87:5454-5458; Rosenthal *et al.*, 1990, *Neuron* 4: 767-773), bem como as indicações de estratégias para produzir uma junção de fusão (por exemplo, ver Darling *et al.*, 1983, *Cold Spring Harbor Symposium Quantitative Biology* 48:427-434; Edwards *et al.*, 1988, *J. Biol. Chem.* 263: 6810-6815; Suter *et al.*, 1991, *EMBO J.* 10:2395-2400). Numa outra concretização, construções quiméricas que fundem a região prepro de um ácido nucleico recombinante afim com NT-4, como as contidas no bacteriófago HG7-2, HG 4-2 e HG2-1, com as regiões maduras de outras neurotrofinas, podem também ser usadas para promover a expressão eficaz dessas outras neurotrofinas, como discutido supra.

O presente invento proporciona também processos de detecção ou medida da actividade de NT-4. Como descrito no Exemplo 12, verificámos que o trkB é um receptor funcional para NT-4. Com base nestas verificações, o invento proporciona processos para a detecção ou medida da actividade de NT-4 compreendendo a exposição de uma célula que expressa trkB a um agente de teste, e detectando ou medindo a ligação do agente de teste a trkB, em que a ligação específica a trkB se correlaciona positivamente com a actividade de NT-4 no agente teste. Numa concretização específica, a célula que expressa trkB é uma célula transfectada, tal como o fibroblasto 3T3, que expressa trkB recombinante, de tal forma que a sobrevivência da célula é dependente da exposição à neurotrofina-4 ou ao BDNF. Assim, a detecção da ligação do agente de teste pode ser efectuada por observação da sobrevivência destas células transfectadas.

6. EXEMPLO: ESTUDOS EVOLUTIVOS DA
FAMÍLIA DO FACTOR DE CRESCIMENTO DO
NERVO REVELAM UM NOVO MEMBRO ABUNDANTEMENTE
EXPRESSO NO OVÁRIO DE XENOPUS

6.1. MATERIAIS E MÉTODOS

6. 1. 1. PREPARAÇÃO DE ADN

O ADN genómico foi isolado por procedimentos padrão (Davis *et al.*, 1986, Basic Methods In Molecular Biology", Elsevier, New York) de leucócitos humanos e de fígado de rato Sprague-Dawley, de rã (Xenopus laevis) e de raia (Raja clavata). O ADN genómico foi também obtido de salmão (Salmão) e de cobra elefante (Vipera lebetina). O ADN foi precipitado com etanol, recolhido utilizando um gancho de vidro, lavado com 80% de etanol, seco e dissolvido em água até uma concentração final de 1 mg/ml. O ADN de salmão (Sigma, St. Louis, MO) foi dissolvido em água, extraído duas vezes com fenol e uma vez com clorofórmio, e precipitado com etanol.

6. 1. 2. REACÇÕES EM CADEIA COM POLIMERASE, CLONAGEM MOLECULAR E SEQUENCIACÃO DE ADN

Seis misturas separadas de oligonucleótidos mero 28 representando todos os códons possíveis correspondentes à sequência de aminoácidos KQYFYET (SEQ ID NO:110) (oligonucleótido-5') e WRFIRID (SEQ ID NO:111) (oligonucleótido-3') (Fig. 1A) foram sintetizadas num sintetizador de ADN da Applied Biosystems A381. O oligonucleótido-5' continha um local EcoRI sintético e o oligonucleótido-3' continha um local HindIII sintético (Knoth *et al.*, 1988, Nucl. Acids Res. 16:1093; Numberg *et al.*, 1989, J. Virology 63: 3240-3249). Cada mistura de oligonucleótidos foi então utilizada para iniciar a amplificação de 0,8 µg de ADN genómico utilizando a reacção em cadeia com polimerase (PCR) (Taq DNA polimerase, Promega) (Saiki *et al.*, 1985, Science 230:1350-1354). Os produtos de PCR foram digeridos com HindIII e EcoRI, analisados num gel de agarose a 2% e clonados no plasmídeo Bluescript KS⁺ (Stratagene, La Jolla, CA). O tamanho da região amplificada mais os iniciadores é de 179 pares de bases (bp) para o NGF e de 182 bp para o BDNF e NT-3. Como resultado de locais EcoRI internos em alguns casos, foram também isolados fragmentos mais pequenos de 144 bp e 95 bp. Os fragmentos de ADN clonados foram sequenciados utilizando o processo de terminação de cadeia por nucleótidos didesoxi

(Sanger *et al.*, 1977, Proc. Natl. Acad. Sci. USA 74:5463-5467) com a T7 DNA polimerase (Pharmacia, Uppsala). Foram sequenciados entre dois e vinte clones independentes para cada gene e espécie, e na totalidade foram sequenciados mais de 200 clones independentes.

Foram pesquisados aproximadamente 2 000 000 de clones provenientes de uma biblioteca genômica de Xenopus preparada por inserção de ADN genômico digerido com MboI no local BamHI do fago λ EMBL-3 utilizando procedimentos convencionais com um fragmento de PCR de 182 bp de NT-4 de Xenopus marcado com [α - ^{32} P]dCTP por tradução por corte para uma actividade específica de aproximadamente 5×10^8 cpm/ μ g. A hibridação foi efectuada em 4 x SSC (1 x SSC é NaCl 150 mM, citrato de sódio 15 mM (pH 7,0)), 40% de formamida, 1 x Solução de Denhardt, 10% de sulfato dextrano a 42 °C. Os filtros foram lavados a 55 °C em 0,1 x SSC, 0,1% SDS e expostos a filmes Kodak XAR-5 a -70 °C. Oito clones de fagos foram isolados, e um fragmento PstI de 1,5 kb de hibridação de um destes clones foi subclonado no plasmídeo pBS-KS (Stratagene). A sequência nucleotídica do fragmento subclonado foi determinada pelo processo de terminação de cadeia por nucleótidos didesoxi (Sanger *et al.*, 1977, Proc. Natl. Acad. Sci. USA 74: 5463-5467).

6. 1. 3. ANÁLISE COMPUTADORIZADA DOS DADOS DA SEQUÊNCIA

As comparações e alinhamentos das sequências de ADN e de aminoácidos apresentadas na Tabela I foram realizadas num computador VAX utilizando o software UWGCG (Devereux *et al.*, 1984, Nucl. Acid. Res. 72:387-395). Os resultados da comparação das sequências de aminoácidos utilizando os programas UWGCG estão apresentados como percentagem de semelhança de aminoácidos ou de identidade nucleotídica entre as sequências, tendo em conta as alterações de aminoácidos conservativos (Gribskov e Burgess, 1986, Nucl. Acid Res. 14:6745-6763; Schwartz e Dayhoff, 1979, "An Atlas of Protein Sequence and Structure", ed. Natl. Biomed. Res. Found., Washington D. C., pp. 353-358). Foi usada a "Phylogenetic Analysis Using Parsimony (versão 3.0f PAUP) para a construção dos filogramas (Felsenstein, 1988, Annu. Rev. Gene

22:521-555; Swofford e Olsen, 1990, em "Molecular Systematics", Hills and Moritz, Eds. Sunderland, M.A., Sinauer Assoc., Inc. pp 441-501). Investigações para as árvores mais prováveis foram realizadas utilizando tanto os algoritmos exaustivos como os heurísticos (permuta de ramos).

6. 1. 4. PRODUÇÃO DE PROTEÍNA RECOMBINANTE,
ENSAIO DE LIGAÇÃO A CÉLULAS PC12, E
ENSAIOS DE ACTIVIDADES NEUROTROFICAS

Para a expressão transiente de proteínas recombinantes em células COS, foram clonados fragmentos apropriados de ADN no vector pXM (Yang *et al.*, 1986, Cell 47:3-10). Para NT-4, o fragmento PstI de 1,5 kb sequenciado de Xenopus foi clonado em pXM, e para NGF foi usado um fragmento BstEII-PstI de 771 bp do exão 3' do gene de NGF de rato (Halbook *et al.*, 1988, Development 108:693-704). Para expressar a proteína BDNF, foi também sub-clonado em pXM um fragmento amplificado por PCR, contendo a sequência de codificação da prepro BDNF do gene de BDNF de ratinho (Hofer *et al.*, 1990. EMBO J. 9:2459-2464). Para NT-3, um clone de ADNC de rato de 1020 bp foi inserido em pXM (Ernfors *et al.*, 1990, Proc. Natl. Acad. Sci. USA 87:5454-5458).

Células COS (Gluzman, 1981, Cell 3:175-182) crescidas até cerca de 70% de confluência foram transfectadas com 25 µg de ADN plasmídico por placa de 100 mm utilizando o protocolo DEAE-dextrano-cloroquina (Luthman e Magnusson, 1983, Nucl. Acids Res. 17:1295-1305). As células transfectadas foram então crescidas em meio completo (DMEM com 10% de FCS), e meio condicionado foi recolhido 3 dias após a transfecção. Placas (35 mm) transfectadas em paralelo foram crescidas até à terceira noite após a transfecção na presença de 200 µCi/ml de [³⁵S]cisteína (Amersham, UK). Aliquotas (10-20 µl cada) de meios condicionados marcados in vivo foram analisadas por SDS-PAGE em geis de poliacrilamida a 13%. Os geis foram tratados com Enhance (New England Nuclear, Boston, MA), secos, e expostos a filmes Kodak XAR-5 com écrans de intensificação durante 24-48 horas a 80 °C. As auto-radiografias foram varridas num densitómetro Shimadzu, e as quantidades relativas das diferentes proteínas

recombinantes foram estimadas pelo cálculo da área correspondente a cada proteína em relação à área obtida com NGF de rato. A quantidade absoluta de proteína NGF de rato foi obtida por imunotransferência quantitativa dos meios condicionados utilizando padrões de NGF purificado de ratinho, e foi usada para determinar a concentração de proteína em amostras contendo outras proteínas recombinantes.

Para o teste de ligação das proteínas recombinantes às células PC12 (Greene e Tischler, 1976, *Proc. Natl. Acad. Sci. USA* 73:2424-2428), NGF de ratinho foi marcado com ^{125}I pelo processo da T-cloramina para uma actividade média de 7×10^7 cpm/ μg . A ligação em estado estacionário foi medida por ensaios de competição efectuados a 37°C ou a 0°C utilizando 1×10^4 células por ml, ^{125}I -NGF a $1,5 \times 10^{-9}$ M; e diluições seriadas de meios condicionados contendo quantidades equivalentes de NGF ou NT-4. Todos os componentes foram adicionados ao mesmo tempo, e as células foram recolhidas por centrifugação após se ter obtido o equilíbrio (1-3 horas de incubação). Experiências de controlo utilizando meio de células COS simuladamente transfectadas, mostraram que outras proteínas presentes no meio condicionado não tinham qualquer efeito sobre a ligação de ^{125}I -NGF às células PC12. A ligação não específica foi medida numa incubação paralela em que foi adicionado pelo menos um excesso de NGF não marcado de 1000 vezes. Todos os resultados foram corrigidos tendo em conta esta ligação não específica que foi sempre 10% inferior à ligação total.

As actividades biológicas das diferentes proteínas foram medidas pela capacidade dos meios condicionados das células COS transfectadas, contendo quantidades iguais de proteína recombinante, para estimular a excrescência de neurites a partir de explantes do simpático, gânglios da raiz dorsal e nodosos, de embriões de galinha E9 (Ebendal, 1984, "Organizing Principles of Neural Development, S. Sharms, ed., New York: Plenum Publishing Corp., pp. 93-107; Ebendal, 1989, "Use of Collagen Gels to Bioassay Nerve Growth Factor Activity In Nerve Growth Factors", R. A. Rush, ed. (Chichester: John Wiley & Sons, pp. 81-93).

Foram ensaiadas diluições seriadas de meio condicionado e foram contadas as excrescências fibrosas.

6. 1. 5. PREPARAÇÕES DE ARN E ANÁLISE POR TRANSFERÊNCIA

Os tecidos indicados de fêmea adulta de Xenopus foram dissecados e congelados em azoto líquido. O cérebro e a espinal medula foram combinados. Vários lóbulos do ovário foram dissecados, incluindo oócitos de diferentes etapas. As amostras de tecido congelado foram homogeneizadas em isotiocianato de guanidina 4 M, β -mercaptoetanol 0,1 M, citrato de sódio (pH 7,0) 0,025 M e homogeneizados três vezes durante 15 s com um Polytron. Cada homogeneizado foi aplicado em camada sobre uma almofada de 4 ml de CsCl em 0,025 M de citrato de sódio (pH 5,5) e centrifugados a 15 °C num rotor Beckman SW41 a 35 000 rpm durante 76 horas (Chirgwin et al., 1979, Biochemistry 78:5294-5299). O ARN poli(A)⁺ foi purificado por cromatografia de celulose-oligo(dT) (Aviv e Leder, 1972, Proc. Natl. Acad. Sci. USA 69:1408-1412) e o ARN recuperado foi quantificado espectrofotometricamente antes da utilização em análise de transferência ARN. O ARN poli(A)⁺ (10 μ g) de cada amostra foi submetido a electroforese num gel de agarose a 1% contendo 0,7% de formaldeído. Foi utilizada a transiluminação-UV do gel corado para confirmar que todas as amostras continham quantidades semelhantes de ARN intacto. O gel foi então transferido para um filtro de nitrocelulose. O filtro foi hibridado com as sondas de ADN indicadas. As sondas foram marcadas com [α -³²P]dCTP por tradução por corte com uma actividade específica de cerca de 5×10^5 cpm/ μ g, e a hibridação foi efectuada como descrito acima. Os filtros foram lavados em condições de elevada severidade (0,1 x SSC, 1% de SDS, 54 °C) e expostos a filmes Kodak XAR-5.

6. 2. RESULTADOS

Os fragmentos de ADN que codificam NGF, BDNF e NT-3 de humano, rato, cobra, rã, e peixe foram isolados utilizando a técnica de PCR com iniciadores degenerados de regiões conservadas nestas três proteínas, localizadas entre a lisina 50 e a treonina 56 para o iniciador a montante e entre o triptofano

99 até o ácido aspártico 105 para o iniciador a jusante (Fig. 1A). A região amplificada contém três dos seis resíduos de cisteínas e cobre aproximadamente um terço das moléculas maduras. Uma comparação da região amplificada nas moléculas já caracterizadas de NGF de diferentes espécies mostra que esta contém duas regiões variáveis, da arginina 59 à serina 67 e do ácido aspártico 93 à alanina 98. Uma região hidrofílica que se crê estar exposta na superfície da molécula (Bradshaw, 1978, Ann. Rev. Biochem. 47:191-216), bem como as regiões altamente conservadas da glicina 68 ao triptofano 76 e da treonina 85 à treonina 91 estão também incluídas na região amplificada. As moléculas de BDNF e NT-3 possuem um aminoácido extra entre as posições 94 e 95 da proteína NGF de ratinho, a qual está também incluída na região amplificada.

As sequências da molécula madura completa de proteína NGF, BDNF e NT-3 de ratinho, foram comparadas de modo a calcular qual a representatividade da região amplificada na molécula completa. As moléculas maduras completas mostram 65/57% de similaridade (similaridade de sequência de aminoácidos/identidade de sequência nucleotídica) entre NGF e BDNF, 70/61% de similaridade entre NGF e NT-3 e 68/58% de similaridade entre BDNF e NT-3. Quando se compara a região isolada neste estudo, a similaridade entre NGF e BDNF é de 62/53%, entre NGF e NT-3 é de 67/58% e entre BDNF e NT-3 é de 69/60%. Isto sugere fortemente que a região isolada neste estudo é representativa da molécula completa e que pode ser utilizada para monitorizar as relações evolutivas entre os diferentes factores. A comparação das sequências par a par foi efectuada (Tabela I), tendo em consideração as substituições de aminoácidos conservativos, utilizando a matriz de comparação de Schwartz e Dayhoff (1979, "In Atlas of Protein Sequence and Structure, M.O. Dayhoff, ed., Washington, D.C., Natl. Biomed. Res. Found., pp. 353-358). Além disso, a comparação das sequências de aminoácidos dadas abaixo e apresentadas na Tabela I indicam a percentagem de similaridade, não identidade. As árvores filogenéticas foram construídas utilizando análise parcimonica (Felsenstein, 1988, Ann. Rev. Genet. 22:521-565; Swofford e Olsen, 1990, "In Molecular

Systematics, D. M. Hills e C. Moritz, eds., Sunderland, MA:Sinauer Assoc., Inc., pp.411-501). Como abaixo apresentado, todos os fragmentos de ADN isolados com sequências de aminoácidos previstas relacionadas com as de NGF, BDNF, e NT-3 continham resíduos de cisteína conservados nas posições correctas. Isto foi usado como critério inicial para considerar uma sequência como membro da família de genes do factor de crescimento do nervo.

6. 2. 1. NGF, BDNF E HDNF/NT-3 SÃO ALTAMENTE
CONSERVADAS DURANTE A EVOLUÇÃO

6. 2. 1. 1. FACTOR DE CRESCIMENTO DO NERVO

A sequência nucleotídica (Fig. 1B [humano (SEQ ID NO:3), rato (SEQ ID NO:4), galinha (SEQ ID NO:5), víbora (SEQ ID NO:6), Xenopus (SEQ ID NO:7), salmão (SEQ ID NO:8)] e a sequência de aminoácidos prevista dos fragmentos isolados que codificam NGF estão altamente conservados do peixe ao humano (Fig. 2.[humano (SEQ ID NO:24), rato (SEQ ID NO:25), galinha (SEQ ID NO:26), víbora (SEQ ID NO:27), Xenopus (SEQ ID NO:28), salmão (SEQ ID NO:29)]. A maior parte das alterações de aminoácidos não conservativos foram encontradas nas regiões variáveis da arginina 59 à serina 67 e do ácido aspártico 93 à alanina 98 (Fig. 2). A similaridade entre as sequências de NGF de Xenopus e humano é 93/79% (Tabela I). NGF de Xenopus e de galinha são idênticas excepto para uma alteração conservativa da lisina 62 à arginina 62 (Fig. 2). As sequências de NGF de víbora e de salmão contêm, respectivamente, 11 e 19 aminoácidos diferentes (de 42), comparadas com NGF humano, enquanto todas as outras espécies apresentaram apenas quatro diferenças. Nenhuma das sequências de aminoácidos de NGF isoladas continha o resíduo aminoácido extra presente em BDNF e NT-3 entre o ácido glutâmico 94 e a lisina 95 das sequências de NGF humano.

As relações inter-espécies das diferentes sequências de NGF foram analisadas pela construção de uma árvore filogenética (Fig. 3A). A sequência de NGF de salmão parece ter divergido mais do que as sequências de NGF isoladas de outras espécies.

Nenhuma sequência de NGF pode ser isolada de raia utilizando a técnica de PCR descrita, sugerindo que as sequências de NGF de raia podem estar acima da tolerância de erros de emparelhamento dos iniciadores utilizados no nosso protocolo de PCR. Alternativamente, a ausência de NGF em peixes cartilaginosos pode implicar que o NGF aparecesse depois da separação do ramo levando à evolução de peixes ósseos (há cerca de 450 milhões de anos), mas antes dos anfíbios e vertebrados superiores terem evoluído a partir deste ramo (há cerca de 400 milhões de anos).

TABELA I

<u>NGF</u>	Hum	Rato	Gal	Víb	<u>Xen</u>	Sal
Humano	-	95	93	86	93	69
Rato	87	-	88	87	88	69
Galinha	80	73	-	90	100	74
Víbora	73	72	75	-	90	67
<u>Xenopus</u>	79	73	83	76	-	74
Salmão	60	61	60	51	53	-
Raia	X	X	X	X	X	X

NT-3

Hum	Rato	Gal	Xen	Sal	Raia
-	100	100	100	81	88
91	-	100	100	81	88
81	83	-	95	81	86
X	X	X	X	X	X
73	79	86	-	86	90
67	64	71	71	-	74
76	73	71	72	71	-

BDNF

Hum	Rato	Gal	Vib.	<u>Xen</u>	Sal	Raia
-	100	95	86	95	100	93
91	-	95	86	95	100	93
88	84	-	91	91	95	93
83	78	85	-	88	86	84
73	82	85	82	-	95	88
82	80	83	75	78	-	99
77	76	78	74	74	78	-

Identidades nucleotídicas e similaridades de aminoácidos foram calculadas com um computador VAX (pacote de software de UWGCG; Devereux *et al.*, 1984, Nucl. Acids Res. 12:389-395) de acordo com a matriz de comparação de Schwartz e Dayhoff (1979, Washington, D.C. Natl. Biomed. Res. Found. pp.353-358), tendo em consideração as alterações de aminoácidos conservativos. As figuras abaixo das diagonais apresentam a percentagem de identidade nucleotídica. As figuras acima das diagonais apresentam a percentagem de similaridade de aminoácidos. X indica que as sequências não foram isoladas daquelas espécies (NGF de raia e NT-3 de víbora). Hum, humano; Gal, galinha; Víb, víbora; Sal, salmão; Xen, Xenopus.

6. 2. 1. 2. FACTOR NEUROTROFICO DERIVADO DE CÉREBRO

As sequências de ADN similares à do BDNF humano foram encontradas em todas as espécies investigadas (Fig. 1B [SEQ ID NOS:1-21, supra listadas]). A similaridade nas sequências de aminoácidos e de nucleótidos, entre a raia, a espécie mais primitiva investigada, e o humano são de 93/77% (Tabela I). Apenas duas alterações não conservativas foram verificadas fora das regiões variáveis, enquanto dez alterações similares foram encontradas em duas regiões variáveis (Fig. 2). Em Xenopus, (SEQ ID NO:34) a leucina 90 é substituída por uma fenilalanina como resultado de uma mutação num único par de bases, C para T na primeira posição do codão, e em salmão (SEQ ID NO:35), o triptofano 77 é substituído por uma tirosina como resultado de uma mutação dupla, mudando o codão de TGG para TAT (Fig. 1B [SEQ ID NO:14]). Todas as sequências isoladas continham um resíduo de aminoácido extra na posição 96, comparadas com o NGF (Fig.2 [SEQ ID NO:24-29]). As sequências de BDNF de diferentes espécies apareceram como um grupo homogêneo de sequências quando analisadas pelo processo parcimonioso (Figura 3B).

6. 2. 1. 3. NEUROTOFINA-3

As sequências de nucleótidos e de aminoácidos previstas para NT-3 de humanos (SEQ ID NO:16 e 37), rato (SEQ ID NO:17 e 38), galinha (SEQ ID NO:18 e 39), Xenopus (SEQ ID NO:19 e 40), salmão (SEQ ID NO:20 e 41), e de raia são altamente similares

(Figs. 1B, Figura 2). A maior parte das alterações são mutações silenciosas que resultam de alterações na terceira posição dos códons, usualmente transições que preservam a estrutura da pirimidina ou purina do par de bases. Apenas alterações de aminoácidos não conservativos foram encontradas nas duas regiões variáveis e não se verificaram substituições de aminoácidos fora das duas regiões variáveis. A sequência de salmão não possui o Asp-94 que está presente em todas as outras moléculas de NT-3 (Fig. 2) e possui uma distância maior do ponto de ramificação na árvore filogenética do que as sequências de NT-3 das outras espécies (Figura 3C).

6. 2. 1. 4. UM NOVO MEMBRO DA FAMÍLIA DE GENES DO FACTOR DE CRESCIMENTO DO NERVO

Fragmentos adicionais de ADN foram isolados de víbora (SEQ ID NO:1) e Xenopus (SEQ ID NO:2), e as sequências previstas de aminoácidos (SEQ ID NO:22 e SEQ ID NO:23, respectivamente) revelaram que estes fragmentos continham todos os três resíduos de cisteína nas mesmas posições como no NGF, BDNF e NT-3 (Fig. 1B, Fig. 2). Uma comparação com as sequências de NGF, BDNF e NT-3 de Xenopus indicou que esta nova sequência está relacionada, mas não é idêntica, às sequências de outros membros da família do NGF. O gene que inclui esta sequência foi portanto denominado neurotrofina-4 (NT-4). A comparação das sequências de nucleótidos e de aminoácidos mostram que a NT-4 de Xenopus e de víbora têm uma similaridade de 91/73%. Esta similaridade está na mesma gama da encontrada entre NGF e BDNF de Xenopus e de víbora (Tabela I). Tal como para os outros membros da família do NGF, alterações de aminoácidos não conservativos foram apenas observadas nas duas regiões variáveis (Fig. 2).

6. 2. 1. 5. COMPARAÇÃO E FILOGENIA DOS MEMBROS DA FAMÍLIA DE GENES DO FACTOR DE CRESCIMENTO DO NERVO

Uma comparação das árvores filogenéticas para o NGF, BDNF e NT-3 mostraram ramos mais longos na árvore de NGF, indicando uma maior taxa de alterações evolutivas (Figuras 3A-3C). A relação de cada membro da família de NGF com os outros membros foi estudada pela construção de um filograma comparando as

sequências de aminoácidos deduzidas para os quatro membros da família. O filograma mostrou que o NGF está mais proximamente relacionado com NT-3 do que com BDNF e NT-4 (Figura 3D). NT-3 está tão relacionado com NGF como com BDNF. NT-4 está mais claramente relacionado com BDNF do que com os outros membros.

6. 2. 2. CARACTERÍSTICAS ESTRUTURAIS DA PROTEÍNA NT-4

De modo a permitir uma caracterização mais detalhada do gene de NT-4 e do seu produto genético, pesquisámos uma biblioteca genómica de Xenopus com o fragmento de PCR de NT-4 e isolámos um clone fágico contendo uma inserção de 16 kb. A partir desta inserção foi subclonado e sequenciado um fragmento PstI de 1,5 kb, Figura 4A (SEQ ID NO:43). A sequência nucleotídica continha uma estrutura de leitura aberta que codifica uma proteína de 236 aminoácidos (SEQ ID NO:44) que apresentou vários aspectos estruturais característicos dos outros membros da família do NGF. A extremidade amino da proteína NT-4 prevista contém uma sequência de sinal putativa de 18 aminoácidos em que uma região de 4 aminoácidos é idêntica às regiões correspondentes de BDNF no porco e no rato (Leibrock et al., 1989, Nature 341:149-152; Maisonpierre et al., 1990, Science 247:1446-1451). Segue-se um local de clivagem do sinal potencial, que é também idêntico ao que foi proposto para o BDNF (Figura 4A). Um local de clivagem potencial para uma proteína NT-4 madura de 123 aminoácidos é encontrado a seguir ao aminoácido 113 na proteína prepro-NT-4. Um único local de N-glicosilação previsto (Asn-Lys-Thr) está localizado 8 aminoácidos antes do local de clivagem putativo.

Uma comparação da proteína NT-4 madura com as proteínas BDNF, NT-3 e NGF maduras de ratinho revelaram 60%, 58% e 51% de identidade de aminoácidos, respectivamente. Todos os 6 resíduos de cisteína envolvidos na formação de pontes dissulfeto [Figura 4B (SEQ ID NO:45-48)] estão incluídos na proteína NT-4 madura. As regiões que são idênticas entre NGF, BDNF e NT-3 são também similares na proteína NT-4. A maioria das diferenças de sequências entre a proteína NT-4 e as outras três proteínas foram encontradas nas mesmas regiões variáveis previamente

identificadas nos outros membros da família.

6. 2. 3. LIGAÇÃO AO NGF-R E

ACTIVIDADE NEUROTROFICA DE NT-4

O fragmento PstI de 1,5 kb de Xenopus foi clonado no vector de expressão pXM (Yang et al., 1986, Cell 47:3-10) e transientemente expresso em células COS. SDS-PAGE de meios condicionados de células transfectadas marcada com [³⁵S]-cisteína apresentou uma proteína NT-4 com um Mr de 14K (Figura 5A). A proteína NGF produzida e marcada em placas paralelas migrou de alguma forma mais rapidamente do que a proteína NT-4. Esta diferença na mobilidade é devida provavelmente às variações na carga das duas proteínas. Diferenças de mobilidade semelhantes foram também observadas para proteínas NGF com tamanhos idênticos, de espécies diferentes.

Meios condicionados de células COS transfectadas contendo quantidades iguais de NGF de rato e proteína NT-4 de Xenopus foram testadas quanto à sua capacidade para competir para a ligação de NGF marcado com ¹²⁵I ao seu receptor nas células PC12. Os testes de ligação foram efectuados a 37 °C e sob condições em que 80% do ¹²⁵I-NGF associado às células é ligado ao NGF-R de baixa afinidade (Sutter et al., 1979, J. Biol. Chem. 254:3972). Foram necessárias concentrações semelhantes de NGF e NT-4 (6×10^{-10} M) para deslocar 50% do ¹²⁵I-NGF das células PC12, indicando que as duas proteínas se ligam ao NGF-R de baixa afinidade com uma afinidade semelhante (Figura 5B). A elevadas concentrações, a proteína NT-4 foi menos eficaz no deslocamento de ¹²⁵I-NGF, sugerindo que neste caso o ¹²⁵I-NGF remanescente associado às células foi ligado a receptores de elevada afinidade ou interiorizado. O facto de esta diferença poder não ser observada num ensaio paralelo efectuado a 0°C em que não ocorre mobilização ou interiorização da membrana sugere que a proteína NT-4 não é capaz de competir com NGF para a interiorização, um processo conhecido por ser exclusivamente mediado através dos receptores de elevada afinidade (Olender e Stach, 1980, J. Biol. Chem. 255:9338-9343; Bernd e Greene, 1984, J. Biol. Chem. 259:15509-15516; Hosang e Shooter, 1987, EMBO J.

6:1197-1202).

A proteína NT-4 transientemente expressa em células COS foi testada quanto à sua capacidade de promover o crescimento de neurites a partir de explantes de gânglios de embrião de pinto. Foi verificada uma clara estimulação da excrescência de neurites em explantes de gânglios da raiz dorsal de galinha (Figura 6A). A comparação das curvas de resposta a dose utilizando quantidades iguais de proteínas NT-4 e NGF revelaram que a actividade obtida com NT-4 era inferior àquela que se observava com NGF (Figuras 5A e 5B). As proteínas BDNF e NT-4 recombinantes estimulavam a excrescência de neurites em gânglios da raiz dorsal numa extensão semelhante (Figuras 6A e 6C). A proteína NT-4 provocou uma fraca, mas consistente, excrescência de neurites a partir de gânglios nodosos (Figura 6G), embora não se tivesse conseguido detectar em gânglios simpáticos (Figura 6E). Isto contrasta com o NGF, o qual estimula marcadamente a excrescência de neurites a partir de gânglios simpáticos. (Figura 6F), e com a NT-3, a qual apresentou uma clara actividade nos gânglios nodosos (Figura 6H). Tal como para NT-4, a actividade de promoção da excrescência de neurites do BDNF nos gânglios nodosos (Figura 6I) foi inferior à actividade observada com o NT-3.

6. 2. 4. EXPRESSÃO DO ARNm DE NT-4 EM DIFERENTES TECIDOS DE XENOPUS

Foi preparado ARN poliadenilado a partir de 11 tecidos diferentes de Xenopus e utilizado para análise por transferência de Northern. A hibridação com a sonda de NT-4 de Xenopus revelou elevados níveis de dois transcritos de NT-4 de 2,3 kb e 6,0 kb no ovário (Figura 7A). Em contraste, o nível de ARNm de NT-4 estava abaixo do limite de detecção em todos os outros tecidos analisados. A hibridação com a sonda de NGF de Xenopus mostrou um ARNm de NGF de 1,3 kb no coração (Figura 7A) e no cérebro. No entanto, a quantidade de ARNm de NGF nestes tecidos era na ordem de 100 vezes inferior ao nível de ARNm de NT-4 no ovário. Foi também detectado ARNm de NGF no ovário, embora a quantidade de ARNm de NGF fosse aproximadamente 100 vezes inferior ao nível de

ARNm de NT-4 neste tecido (Figura 7B). Os níveis de ARNm NT-3 e de BDNF no ovário estavam ambos abaixo do limite de detecção (Figura 7B).

6. 3. DISCUSSÃO

Usámos a reacção em cadeia com polimerase (PCR) em combinação com iniciadores de oligonucleótidos degenerados para isolar os genes para diferentes membros da família do NGF a partir de espécies diferentes. Uma comparação das sequências nucleotídicas e de aminoácidos das proteínas completas maduras NGF, BDNF e NT-3 revelou similaridades que são as mesmas que se obtiveram por comparação da região dos genes analisados neste estudo. Assim, esta região parece ser representativa do resto do gene e pode desta forma ser utilizada para estudar a conservação evolutiva da proteína completa madura.

Os genes de NGF, BDNF e NT-3 de diferentes espécies incluem regiões que mostram identidade completa entre peixes e mamíferos, bem como regiões de baixa similaridade. Uma comparação de sequências de NGF de diferentes espécies com as sequências correspondentes de BDNF ou NT-3 mostrou que o gene de NGF é menos conservado em vertebrados do que os de BDNF e BDNF/NT-3. Os dois últimos genes parecem ser igualmente conservados em todas as espécies estudadas, excepto em salmão, em que o NT-3 é menos conservado do que o BDNF. Neste contexto, é interessante especular acerca do facto de que o relógio molecular parece acelerar-se em alguns ramos, notavelmente o de NGF, e não em outros. Geralmente acredita-se que existe uma força selectiva que preserva a estrutura terciária correcta da proteína (Dickerson, 1971, J. Mol. Evol. 1:26-45; Kimura e Ohta, 1974, Proc. Natl. Acad. Sci. USA 71:2848-2852). A diferença na conservação evolutiva dos três factores sugere que tem havido uma maior pressão selectiva sobre os genes de BDNF e NT-3 do que sobre o gene do NGF. Foram propostas alterações ambientais para levar a alterações na pressão selectiva alterando o óptimo de desempenho de um produto genético específico (Kimura, 1983, em "Evolution of Genes and Proteins", pp.208-233). Neste contexto, é possível que as alterações evolutivas mais extensas observadas

no NGF quando comparado com BDNF e NT-3, reflectam o facto de que a função do NGF se tenha modificado mais durante a evolução. Estudos de estrutura-função do NGF mostraram que esta molécula pode tolerar alterações estruturais consideráveis sem perda ou modificação do seu perfil de actividade, sugerindo que o grau mais baixo da conservação evolutiva do NGF possa ser devida a uma estrutura mais estável desta proteína, que é assim menos facilmente perturbada por substituições. Outra explicação possível é a de que as regiões do genoma onde os genes para os diferentes factores estão localizados possuem diferentes taxas de mutação geral. Foram apresentadas taxas diferentes de mutação para regiões do genoma não codificadores (Wolfe *et al.*, 1989, *Nature* 337:283-285) mas isto é menos claro se puder levar a um número aumentado de alterações em regiões codificadoras.

O NGF e a NT-3 de salmão são notavelmente mais diferentes quando comparados com estas moléculas em outras espécies. Alguns aminoácidos, incluindo a treonina 82 e histidina-treonina-fenilalanina na posição 85 a 87 em NGF, bem como a ausência do aminoácido entre as posições 94 e 95 (comparado com as duas outras proteínas), são características consistentes da proteína NGF. O facto da sequência isolada de salmão conter todos estes motivos específicos de NGF argumenta que esta não é um membro adicional da família, mas que representa o NGF de salmão. Em contraste com todas as outras sequências de NT-3 estudadas, o NT-3 de salmão não possui o aminoácido na posição 95. Uma vez que o aminoácido extra está presente na NT-3 de raia, é provável que o ancestral comum da raia e salmão tivesse uma sequência de NT-3 ancestral que incluísse o aminoácido extra na posição 95. Desta forma, as alterações na molécula de NT-3 de salmão devem ter ocorrido depois desta separação do gene do ancestral comum. A maior parte das alterações nos aminoácidos da sequência de salmão estão nas mesmas regiões que variam, com um menor grau, também nas outras espécies, sugerindo fortemente que as sequências isoladas de NGF e NT-3 de salmão não são pseudogenes. A maior divergência de NGF e NT-3 de salmão, comparadas com outras espécies, reflecte provavelmente o elevado grau de expansão evolutiva dos peixes ósseos.

Os resultados neste estudo indicam que a família do NGF existia provavelmente há 500 milhões de anos nos peixes primitivos, que eram os ancestrais dos actuais vertebrados superiores. A família de genes pode ter sido formada por duplicação genética, que se crê ser o mecanismo mais comum através do qual novos genes evoluem (Li, W., 1983, em "Evolution of Genes and Proteins", pp.14-37). Duplicações de genes funcionais podem ter sido facilitadas, uma vez que toda a informação necessária para a síntese de uma proteína biologicamente activa está contida num exão 3' (Hallbook *et al.*, 1988, Mol. Cell. Biol. 8:452-456; Leibrock *et al.*, 1989, Nature 341:149-152; Hohn *et al.*, 1990, Nature 344:339-441). A formação da família envolveu várias duplicações de genes (Figura 3D).

Uma vez que a NT-4 está mais proximamente relacionada com BDNF do que com NT-3 ou NGF, parece que a NT-4 e o BDNF foram formados a partir de um gene ancestral comum. No entanto, uma vez que não foi distinguida a partir dos resultados apresentados nenhuma molécula do tipo progenitor para todos os quatro factores, a relação evolutiva do ancestral putativo BDNF/NT-4 com os ancestrais de NGF e NT-3 não pode ser definitivamente estabelecida. A topologia dos filogramas utilizando dados de espécies diferentes está geralmente de acordo com a relação evolutiva consensual entre espécies diferentes. No entanto, para ambos de NGF e BDNF, as sequências de galinha mostram um ramo mais precoce no filograma do que o esperado. A comparação de NT-4, NGF e BDNF de víbora e *Xenopus* revelou que as sequências de NT-4 nestas espécies têm 11 substituições de aminoácidos, comparados com 9 e 8 substituições em NGF e BDNF, respectivamente. Isto sugere que nestas espécies, a NT-4 divergiu com uma taxa que é comparável a, ou mais rápida que, a taxa de divergência de NGF ou BDNF.

As substituições de aminoácidos altamente conservados na molécula de NGF não eliminam a actividade biológica, mas em muitos casos afectam a quantidade de proteína produzida, indicando que existem outros constrangimentos para além da actividade biológica, tais como a estabilidade da proteína, que

pode ser importante para a conservação da proteína NGF. Adicionalmente, o facto de todos os membros da família do NGF poderem interagir com o NGF-R de baixa afinidade sugere que a conservação completa de certas regiões nestes factores pode ser devida a constrangimentos nestes genes para reter proteínas que possam interagir com o NGF-R. Os mecanismos básicos e estratégias para a ontogenia precoce do embrião são semelhantes em todos os vertebrados e envolvem presumivelmente genes que são conservados em todos os vertebrados. A conservação evolutiva dos factores neurotróficos é dessa forma consistente com a noção de que são importantes no desenvolvimento embrionário precoce em muitas espécies diferentes.

O hipocampo contém os níveis mais elevados de ARNm de NGF, BDNF, e NT-3 no cérebro (Ernfors et al., 1990, J. Dev. Neurosci. 2:57-66). É uma estrutura altamente especializada derivada do arquipálpio que inicialmente apareceu nos cérebros de anfíbios e de répteis. O hipocampo de mamíferos é importante para a memória, aprendizagem e funções cognitivas conhecidas por estarem associadas com elevada plasticidade neuronal (Crutcher e Collins, 1982, Science 227:67-68). Estas exigências podem ter produzido uma pressão selectiva durante a filogenia para mecanismos de promoção de plasticidade, possivelmente mediados por factores neurotróficos. No entanto, os resultados neste estudo mostram claramente que o evento da duplicação dos genes para os factores neurotróficos precedeu largamente a formação do hipocampo. Esta evidência indica que os factores neurotróficos não evoluíram como consequência da formação do hipocampo e apoia a noção de que a plasticidade neuronal nesta região do cérebro é, pelo menos em parte, devida a estas moléculas.

A organização do sistema nervoso dos vertebrados primitivos, isto é, peixes cartilaginosos, apresenta algumas semelhanças básicas com o sistema nervoso de vertebrados superiores. Os nervos cranianos e o sistema nervoso somático sensorial e autónomo em peixes cartilaginosos são em geral semelhantes aos dos vertebrados superiores (Young, J.Z., 1981, The Life of Vertebrates, New York, Oxford University Press). É

desta forma provável que os princípios das interações neurotróficas sejam os mesmos tanto nos vertebrados primitivos como nos superiores. A conservação evolutiva dos factores neurotróficos semelhantes a NGF também nos vertebrados primitivos sugere que estes factores evoluíram inicialmente em invertebrados e foram depois adaptados para funcionar no desenvolvimento do sistema nervoso dos vertebrados.

O nosso estudo da conservação evolutiva da família do NGF levou ao isolamento de um novo membro desta família, denominado neurotrofina-4 ou NT-4, fragmentos de PCR do gene de NT-4 foram isolados de Xenopus e víbora, e um clone genómico foi subsequentemente isolado a partir de Xenopus. A análise da sequência nucleotídica deste clone revelou uma estrutura de leitura aberta para uma proteína de 236 aminoácidos, a qual mostrou várias características estruturais que se assemelham às dos outros três membros da família do NGF. Isto inclui a presença de uma sequência de sinal putativa de extremidade amino e um potencial local de N-glicosilação próximo de um local de clivagem proteolítica que prevê uma proteína NT-4 madura de 123 aminoácidos. O tamanho da proteína NT-4 madura é 4 aminoácidos mais longa do que a de BDNF e NT-3, e 5 aminoácidos mais longa do que a proteína madura NGF. Na proteína NT-4 madura, todos os 6 resíduos de cisteína envolvidos na formação das pontes dissulfeto são conservados. A proteína NT-4 difere dos outros membros da família nas mesmas regiões que variam entre as sequências dos outros três membros da família. Como para NGF, BDNF, e NT-3, a proteína completa prepro-NT-4 é codificada num único exão. Assim, tanto a organização do gene como as características estruturais da proteína prevista indicam que o gene de NT-4 é um membro adicional da família do NGF. O facto do gene de NT-4 ter sido isolado tanto de répteis como de anfíbios sugere que este está presente em várias espécies diferentes.

Tanto BDNF como NT-3 apresentaram interacção com NGF-R de baixa afinidade (Rodriguez-Tebar et al., 1990, Neuron 4:487-492; Ernfors et al., 1990, Proc. Natl. Acad. Sci. USA 87:5454-5458).

A proteína NT-4 de Xenopus deslocou ^{125}I -NGF do seu receptor de baixa afinidade em células PC12, indicando que o quarto membro desta família pode também interagir com o NGF-R de baixa afinidade. A comparação das curvas de deslocamento obtidas a 37 °C e a 0 °C sugere que a proteína NT-4 não pode competir para a ligação ao NGF-R de elevada afinidade. A proteína codificada pelo gene do NGF-R de baixa afinidade parece fazer parte tanto dos receptores de baixa como de elevada afinidade (Hempstead et al., 1989, Science 243:373-375). O mecanismo pelo qual dois receptores cineticamente diferentes são formados a partir do mesmo gene de receptor não é conhecido, embora tenha sido proposto que os dois estados podem ser produzidos pela formação de um complexo entre o domínio citoplasmático do receptor e uma proteína intracelular (Radeke et al., 1987, Nature 325:593-597; Meakin e Shooter, 1991, Neuron 6:153-163). Alternativamente, uma cadeia de receptor de elevada afinidade pode ser codificada por um gene separado e, em semelhança ao receptor da interleuquina-2 (Hatakeyama et al., 1989, Science 244:551-556) e ao receptor do factor de crescimento derivado das plaquetas (Matsui et al., 1989, Science 243:800-804), as duas cadeias de receptor podem formar um dímero que constitui o receptor de elevada afinidade. O facto de todos os quatro membros da família do NGF poderem interagir com o NGF-R de baixa afinidade sugere que o estado de baixa afinidade do NGF-R pode estar, de uma forma ainda não conhecida, envolvida na mediação dos efeitos biológicos de todos estes factores. Neste contexto, é interessante notar que foi mostrado que o gene de NGF-R de baixa afinidade era expresso em muitos tecidos tanto de origem neuronal como não neuronal, não conhecidos por responderem a NGF. Isto inclui o mesênquima, células do tubo neural e somitos no embrião de pinto precoce (Halbrook et al., 1990, Development 108:693-704; Heuer et al., 1990a, Dev. Biol. 137:287-304; Heuer et al., 1990b, Neuron 5:283-296), assim como desenvolver e regenerar neurónios motores da espinal medula (Ernfors et al., 1989, Neuron 2:1605-1613; Ernfors et al., 1991, J. Dev. Neurosci. 9:57-66). Será portanto de interesse investigar se a proteína NT-4 é de importância funcional em qualquer desses tecidos ou populações neuronais.

A actividade neurotrófica da proteína NT-4 foi ensaiada em explantes de gânglios embrionários de pinto, e como para os outros três membros da família do NGF, a proteína NT-4 mostrou uma clara estimulação da excrescência de neurites a partir de gânglios da raiz dorsal. No entanto, quando comparado com NGF, a proteína NT-4 mostrou uma actividade inferior nos gânglios da raiz dorsal. Tanto BDNF como como NT-3 induzem prontamente a excrescência de neurites em explantes de gânglios nodosos, embora a resposta com NT-3 fosse consistentemente mais forte do que com BDNF. O NGF estimula fortemente a excrescência de neurites em gânglios simpáticos, e a NT-3 possui também actividade nestes gânglios, ainda que seja muito menor do que a de NGF (Maisonpierre et al., 1990, Science 247:1446-1451; Ernfors et al., 1990, Proc. Natl. Acad. Sci. USA 87:5454-5458). A NT-4 mostrou uma actividade mais fraca em gânglios nodosos comparada com a NT-3 e nenhuma actividade em gânglios simpáticos. O espectro da actividade biológica da NT-4 em explantes de gânglios periféricos assemelha-se à de BDNF, o que está de acordo com o facto de que a NT-4 é estruturalmente semelhante ao BDNF.

A análise por transferência de Northern de 11 tecidos diferentes de Xenopus apresentou níveis elevados de NT-4 no ovário, enquanto que o nível de ARNm de NT-4 estava abaixo do limite de detecção em todos os outros tecidos examinados. Dois ARNm de NT-4 de 2,3 kb e 6,0 kb foram encontrados nos oócitos. A presença de dois transcritos do mesmo gene foi previamente observada para o BDNF, em cujo caso dois ARNm de 1,4 kb e 4,0 kb estão presentes no cérebro de rato (Leibrock et al., 1989, Nature 341:149-152; Maisonpierre et al., 1990, Science 247:1446-1451; Ernfors et al., 1990a, Proc. Natl. Acad. Sci. USA 87:5454-5458). A hibridação com uma sonda de NGF de Xenopus revelou ARNm de NGF no coração de Xenopus, provavelmente como resultado da expressão de ARNm de NGF em tecidos alvo para inervação neuronal. O nível de ARNm de NGF no coração era, no entanto, mais do que 100 vezes inferior ao nível de ARNm de NT-4 no ovário. Uma vez que o elevado nível de ARNm de NT-4 no ovário não se correlaciona com a inervação neuronal, parece improvável

que a proteína NT-4 tenha apenas uma função neurotrófica neste caso. Em vez disso, a expressão abundante de ARNm de NT-4 no ovário de Xenopus implica uma função não neurotrófica adicional e importante para a proteína NT-4. Foi também detectado ARNm de NGF no ovário de Xenopus embora com níveis cerca de 100 vezes inferiores aos do ARNm de NT-4; Não foram detectados ARNm de BDNF e NT-3 neste tecido.

Foram descritos ARNm para dois factores de crescimento como ARNm maternos em oócitos de Xenopus. Um destes ARNm codifica uma proteína com forte similaridade com o factor de crescimento de fibroblastos básico (Kimelman e Kirschner, 1987, Cell 51:869-877); o outro ARNm codifica uma proteína homóloga do factor de crescimento transformante α (Weeks e Melton, 1987, Cell 51:861-867). Foi sugerido que estes factores funcionassem como morfogénios para a formação do mesoderma e subsequente indução deste tecido no tubo neural. No rato, estudos de hibridação in situ revelaram ARNm de NT-3 no epitélio de folículos secundários e terciários, e foi sugerido um papel para o NT-3 na oogénese (Ernfors et al., 1990, Neuron 5:511-526).

7. EXEMPLO: IDENTIFICAÇÃO DE CÉLULAS QUE EXPRESSAM ARNm DE NT-4 NO OVÁRIO DE XENOPUS LAEVIS POR HIBRIDAÇÃO IN SITU

7. 1. MATERIAIS E MÉTODOS

7.1.1. ISOLAMENTO, MANIPULAÇÃO E CULTURA DE OÓCITOS, EMBRIÕES E CÉLULAS DE XENOPUS

Rãs machos e fêmeas X. laevis foram mantidos no laboratório a 19 °C. Depois de anestesiados por imersão dos animais em 0,25% de metanossulfonato de tricaina (Sandoz, Suíça), os lóbulos do ovário foram removidos cirurgicamente, lavados com Hepes salino de Barth modificado (MBSH) (Gurdon e Wickens, 1983, Methods Enzymol. 101:370-86) e dissociados por incubação durante a noite a 20 °C em MBSH isento de cálcio e

contendo 2 mg/ml de collagenase. A separação do extracto dos oócitos pré-vitelogénicos e vitelogénicos foi obtida por sedimentação diferencial, e os oócitos foram depois separados manualmente sob um microscópio de dissecação nas várias classes de desenvolvimento descritas por Dumont (1972, supra).

Embriões de clivagem sincronómica foram obtidos por fertilização in vitro como essencialmente descrito por Newport e Kirchner (1982).

Células de rim de Xenopus A₆ foram cultivadas em meio de Leibowitz L₁₅ diluído com água destilada 60:40 (v/v) e suplementado com Hepes pH 7,35 10 mM, hipoxantina 10 μ M (Sigma), glutamina 4 mM e 10% de soro fetal bovino (Gibco) a 20 °C. As culturas foram equilibradas com ar e mantidas no escuro.

7.1.2. HIBRIDAÇÃO IN SITU

Ovários frescos congelados de rãs Xenopus laevis adultas foram seccionados (14 μ) num crióstato (Leitz, Alemanha) e as secções foram descongeladas em lâminas pré-tratadas com poli-L-lisina (50 μ g/ml), sendo posteriormente fixadas em formalina a 10% durante 30 min. e lavadas duas vezes em PBS. A desidratação foi efectuada numa série graduada de etanol incluindo 5 min. de incubação em clorofórmio e posterior secagem das lâminas ao ar. Dois oligonucleótidos de mero 53, um específico para ARNm de NT-4 de Xenopus (5'CCCACAAGCTTGTGGCATCTATGGTCAGAGCCCTCACATAAGACTGTTTTC3' [SEQ ID NO:109]) e o outro, como um controlo, específico para ARNm de BDNF de galinha (correspondendo aos aminoácidos 61 a 77 da proteína BDNF madura de galinha (Hallbook et al., 1991 Neuron 6:845-58 [contido na SEQ ID NOS:11 e 32]), foram marcados na extremidade 3' com α^{35} S-dATP usando desoxiribonucleotidil transferase terminal (IBI, New Haven) com uma actividade específica de aproximadamente 1×10^9 cpm/ μ . A hibridação foi efectuada a 42 °C durante 16 horas em formamida a 50%, 4x SSC, 1x solução de Denhardt, 1% de sarcosil, NaPO₄ (pH 7,0) 0,02M, 10% de sulfato dextrano, 0,5 mg/ml de ARNt de levedura, DTT 0,06M 0,1 mg/ml de ADN de esperma de salmão fragmentado e 1×10^7

cpm/ml de sonda oligonucleotídica marcada com ^{35}S . As secções foram subsequentemente enxaguadas, lavadas 4 vezes (15 min. cada) a 55 °C em 1 x SSC, enxaguadas com água, desidratadas numa série graduada de etanol e secas ao ar. As secções foram expostas à película de raios-X, seguidas por revestimento com a emulsão fotográfica NTB-3 da Kodak (diluída 1:1 em água), expostas durante 5-6 semanas a -20 °C, e desenvolvidas, fixadas e contrastadas com violeta de cresilo.

7. 1. 3. ANÁLISE DAS TRANSFERÊNCIAS DE ARN

As amostras indicadas foram homogeneizadas em isotiocianato de guanidina 4M, β -mercaptoetanol 0,1 M, citrato de sódio 0,025 M (pH 7,0) e homogeneizadas 3 vezes durante 15 segundos com um Polytrone. Cada homogeneizado foi aplicado em camadas sobre uma almofada de 4 ml de CsCl 5,7 M em citrato de sódio 0,025 M (pH 5,5) e centrifugado a 15°C num rotor Beckman SW41 a 35 000 rpm durante 16 horas (Chirgwin *et al.*, 1979, Biochemistry 78:5294-5299). ARN poliadenilado (Poli(A)+ARN) foi purificado por cromatografia de oligo(dT)-celulose (Aviv e Leder, 1972, PNAS 69:1408-1412) e a recuperação de ARN (40 μg) foi quantificada espectrofotometricamente antes de usar na análise de transferência de ARN. O ARN celular total (40 μg) ou, quando indicado, o poli(A)+ARN (5 μg) de cada amostra foi submetido a electroforese num gel de agarose a 1% contendo 0,7% de formaldeído. A transiluminação-UV do gel corado foi utilizada para confirmar que todas as amostras continham quantidades semelhantes de ARN intacto. O gel foi então transferido para um filtro de nitrocelulose. O filtro foi então hibridado com um fragmento HincII de 350 bp do exão 3' do gene de NT-4 de Xenopus (Hallbook *et al.*, 1991, Neuron 6:845-858). O fragmento foi marcado com α - ^{32}P -dCTP por tradução por corte com uma actividade específica de cerca de 5×10^8 cpm/ μg e a hibridação efectuou-se como descrito (Ernfors *et al.*, 1988, Neuron 1:983-96). Os filtros foram lavados com elevada severidade (0,1xSSC, 0,1% de SDS, 54 °C) e expostos a películas AR-5 da Kodak a -70 °C.

7. 2. RESULTADOS

Secções de tecido oriundo de ovário de Xenopus laevis



adultas foram hibridadas com uma sonda oligonucleotídica marcada com ^{35}S -dATP específica para ARNm de NT-4 de Xenopus. Como controlo para a especificidade da hibridação, secções adjacentes foram hibridadas com uma sonda oligonucleotídica do mesmo comprimento e conteúdo em GC complementar ao ARNm do factor neurotrófico derivado de cérebro (BDNF) de galinha. A sonda específica para o ARNm de NT-4 revelou uma intensa marcação entre muitas células pesquisadas através do ovário com um tamanho (50-400 μm de diâmetro) correspondente aos oócitos em etapas precoces da oogénese (Fig. 9A). Nenhum ARNm de NT-4 pôde ser detectado entre oócitos maduros, oócitos na etapa VI pós-vitelogénica (setas na Fig. 9A). A sonda de controlo específica de ARNm de BDNF de galinha não marcou quaisquer células no ovário de Xenopus.

A análise de auto-radiografias de emulsão das secções hibridadas revelou uma marcação intensa entre o citoplasma de oócitos com um diâmetro de 50-200 μm (Fig. 10A e 10B) correspondente aos oócitos na etapa I de acordo com Dumont, 1972, supra. A sonda específica para o ARNm de NT-4 também marcou oócitos com diâmetro superior correspondentes às etapas II a IV, embora a intensidade da marcação entre estas células fosse inferior à verificada entre os oócitos na etapa I. De acordo com a análise de baixa amplificação de iluminações de campo escuro (Fig. 9), as auto-radiografias de emulsão não apresentaram qualquer marcação entre os oócitos maduros nas etapas V e VI. Nenhuma marcação foi verificada entre quaisquer células depois da hibridação com a sonda controlo para BDNF (Fig. 10C). Para permitir uma determinação mais detalhada do nível de ARNm de NT-4 durante a oogénese, foi contado o número de grãos por uma unidade de área arbitrariamente escolhida. A unidade de área escolhida correspondeu aproximadamente a um centésimo da área da secção transversal de um oócito na etapa I. O resultado desta análise mostrou que a intensidade da marcação entre oócitos na etapa I era de 1,7 e 4,3 vezes superior à dos oócitos nas etapas II/III e IV, respectivamente (Fig. 11). O número de grãos por unidade de área entre os oócitos nas etapas V e VI não foi significativamente superior ao nível do ruído de

fundo da marcação.

7. 2. 1 ANÁLISE POR TRANSFERÊNCIA DE NORTHERN DA EXPRESSÃO DE ARNm DE NT-4 DURANTE A OOGÊNESE E DESENVOLVIMENTO PRECOCE DE XENOPUS

Uma quantidade fixa de ARN celular total (40 μ g) preparado a partir de etapas diferentes de oócitos bem como de uma fracção enriquecida de células de folículo foi analisada por transferências de Northern utilizando uma sonda específica para NT-4 de Xenopus (Hallbook *et al.*, 1991, *supra*). De acordo com os resultados da hibridação *in situ*, os níveis mais elevados de transcritos de NT-4 com tamanhos de 2,3 kb e 6,0 kb estavam presentes nos oócitos mais pequenos (etapas I e II) (Fig. 12). O nível de ARNm de NT-4 declinou abruptamente nos oócitos das etapas mais maduras V e VI. Foi observado um fraco sinal de hibridação na preparação das células de folículo que foi provavelmente devido a contaminação com um pequeno número de oócitos nas etapas I e II. O mesmo resultado foi obtido quando uma quantidade fixa de ARN poliadenilado (5 μ g) foi analisado a partir das diferentes amostras apresentadas na Fig. 12.

Os resultados da análise da expressão de ARNm de NT-4 no ovário mostrou que o ARNm de NT-4 está restrito aos oócitos imaturos. Para testar a possibilidade da expressão de ARNm de NT-4 ser induzida após a fertilização, o nível de ARNm de NT-4 foi determinado em embriões de Xenopus em desenvolvimento por transferência de Northern de ARN poliadenilado. Foi encontrado um nível baixo de ARNm de NT-4 em células somáticas de rim A₆ de Xenopus em cultura que foram também incluídas na análise. No entanto, não pôde ser detectado ARNm de NT-4 em embriões precoces desde o início das divisões por clivagem até à etapa de nêurula.

7. 3. DISCUSSÃO

A expressão abundante de ARNm de NT-4 no ovário de Xenopus (Hallbook *et al.*, 1991 *supra*) indica que este membro da família de NGF desempenha um papel na oogénese e/ou na embriogénese precoce. A localização das células que expressam ARNm de NT-4 no

ovário proporciona discernir sobre a função putativa da proteína NT-4 no ovário. Em anfíbios, assim como em todos os outros vertebrados, a fertilização do ovo desencadeia um período de rápida clivagem celular. Este acontecimento é controlado por uma classe de ARNm maternos solúveis expressos durante a oogénese e armazenados no ovo infertilizado para subsequente desenvolvimento (Davidson, 1986, *Gene Activity in Early Development* (New York, Academic Press)). Esta classe de ARNm maternos inclui dois factores de crescimento, o factor de crescimento do fibroblasto básico (Kimelman e Kirchner, 1987, *Cell* 51:869-77) e o factor- β de crescimento transformante (Weeks e Melton, 1987, *Cell*, 51:861-67), assim como vários protooncogenes tal como o c-myc (Godeau *et al.*, 1986, *EMBO J.* 5:3571-77); (Vriz *et al.*, 1989, *EMBO J.* 8:4091-97), c-fos (Mohun *et al.*, 1989, *Development*, 107:835-46), ras (Andeol *et al.*, 1990, *Dev. Biol.* 139:24-34), ets-2 (Chen *et al.*, 1990, *Science*, 250:1416-1418) e c-mos (Sagata *et al.*, 1988, *Nature* 335:519-25). Oócitos de *Xenopus* na etapa VI madura são apanhados na profase da meiose I e tanto c-mos (Sagata *et al.*, 1988) como ets-2 (Chen *et al.*, 1990), mostraram funcionar durante a re-iniciação da divisão meiótica. A verificação de níveis elevados de ARNm de NT-4 em oócitos na etapa I e II mas de um nível diminuído abaixo do limite de detecção tanto das transferências de Northern como da hibridação *in situ* em oócitos nas etapas V e VI, sugere fortemente que o ARNm de NT-4 não pertence à classe dos ARNm maternos. Este resultado é também argumento contra o papel da proteína NT-4 na re-iniciação da divisão meiótica ou na embriogénese precoce. De acordo com isto, a adição de proteína NT-4 recombinante a oócitos imaturos na etapa VI não resultou na indução da clivagem das vesículas germinais *in vitro* e não foi detectado ARNm de NT-4 em embriões precoces de *Xenopus*. Em vez disso, a função putativa da proteína NT-4 no ovário parece estar acoplada a acontecimentos que ocorrem em oócitos pré-vitelogénicos e vitelogénicos semi-precoces. Tanto NGF (Ayer-LeLievre *et al.*, 1988 *PNAS* 85:2628-2632), como o receptor de NGF de baixa afinidade de 75 kD (Persson *et al.*, 1990, *Science*, 247:704-707) e o componente trkA de elevada afinidade do receptor de NGF (J.P. Merlo e H. Persson, não publicado) são

expressos nos testes em que o NGF mostrou recentemente estimular a síntese de ADN no início da meiose (Parvinen et al., 1991, enviado para publicação). Assim, parece que as neurotrofinas não funcionam apenas como factores neurotróficos mas desempenham também um papel importante nos tecidos reprodutivos.

8. EXEMPLO: ISOLAMENTO E CARACTERIZAÇÃO DE FRAGMENTOS DE ÁCIDO NUCLEICO QUE CODIFICAM NT-4 DE MAMÍFEROS

8.1. MATERIAIS E MÉTODOS

8.1.1. PREPARAÇÃO DE ADN

ADN genómico foi isolado como descrito em 6.1. 1, supra.

8.1.2. REACÇÕES EM CADEIA COM POLIMERASE, CLONAGEM MOLECULAR E SEQUENCIACÃO DE ADN

Misturas de oligonucleótidos de mero 34 (incluindo a cauda) representando todos os codões possíveis correspondentes às sequências de aminoácidos QYFFET (contida na SEQ ID NO:51) e QYFYET (SEQ ID NO:52) (oligonucleótido 5') e WISECK, CKAKQS E WIRIDT (cada uma contida na SEQ ID NO:51) (oligonucleótido 3') (Fig. 13) foram sintetizadas, com ligantes, como descrito em 6.1.2. supra. Em conjunto, 2Y (derivado de xNT-4 [SEQ ID NO:50]) e 2Z (derivado de BDNF/NT-3 [SEQ ID NO:51]) representam todas as sequências conhecidas para neurotrofinas de todas as espécies nesta região. Uma primeira amplificação tanto de ADN genómico de rato como de humano foi efectuada com a Taq polimerase (Cetus) com ciclos de 1 minuto a 95°C, 2 minutos a 43°C e 2 minutos a 72°C. Uma alíquota das reacções de PCR primárias foi então reamplificada utilizando os mesmos iniciadores da primeira amplificação, ou novos oligonucleótidos iniciadores degenerados e complementares que vão resultar num desvio esperado de tamanho. Os produtos de PCR do procedimento de reamplificação foram purificados como se segue: as bandas com o tamanho esperado foram purificadas em gel, reamplificadas, e purificadas em coluna utilizando as colunas "primerase" da Stratagene. Estas foram então digeridas completamente com EcoRI e SalI,

analisadas e repurificadas utilizando colunas Primerase (Stratagene) e ligadas em Bluescript KS(-) digerido com EcoRI-XhoI. Os transformantes foram seleccionados para um pBS-KS contendo uma inserção de tamanho aproximado ao previsto. Os fragmentos clonados foram sujeitos a uma análise da sequência de ADN como descrito em 6.1.2, supra.

8.1.3 ISOLAMENTO CLONES GENÓMICOS DE COMPRIMENTO TOTAL E DE CLONES DE ADNc QUE CODIFICAM rNT-4 E hNT-4

Uma biblioteca de ADNc de ovário humano em λ GT-10 foi obtida da Clontech. Uma biblioteca de ADNc de hipocampo humano em λ :ZAPII foi obtida da Stratagene. Uma biblioteca de ADN genómico humano em EMBL3/SP6/T7 foi obtida de Clontech. Uma biblioteca de ADNc de cérebro de rato em λ -ZAP foi obtida da Stratagene.

O isolamento de clones de NT-4 pode ser efectuado como se segue:

Uma inserção clonada codificando o fragmento rNT-4 (Fig. 14 [SEQ ID NO:61]) ou o fragmento hNT-4 (Fig.15 [SEQ ID NO:63]) são marcadas por PCR para uma actividade específica de aproximadamente 5×10^8 cpm/ng. A hibridação foi efectuada numa solução de hibridação consistindo em 0,5 mg/ml de ADN de esperma de salmão a 60 °C. Os filtros foram lavados a 60 °C em 2xSSC, 0,1% SDS e expostos a uma película XAR-5 da Kodak a -70 °C. Alternativamente, oligonucleótidos cuja sequência corresponde exactamente à da neurotrofina de mamífero desejada, podem ser utilizados para produzir sondas (por exemplo, marcação pela quinase) e podem ser utilizados para seleccionar as mesmas bibliotecas pelos processos convencionais. Os fagos positivos foram purificados em placa e infectaram a multiplicidade baixa, uma estirpe apropriada de E. coli em caldo líquido como descrito por Maniatis et al., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York). Os fagos GT-10 e EMBL3/SP6/T7 foram preparados como se segue: as culturas foram incubadas durante a noite a 37 °C com agitação constante. A suspensão nocturna é tornada 1 M em NaCl e 8% em

PEG, bem misturada e incubada durante a noite a 4 °C para precipitar o bacteriófago. O bacteriófago foi precipitado via centrifugação, e ressuscitado em tampão TM (Tris-HCl 10 mM, pH 7,5; MgCl₂ 10 mM), aplicado em camadas sobre um gradiente descontínuo de CsCl e centrifugado à velocidade, e durante o tempo apropriados para se ligar o bacteriófago. O bacteriófago é removido, transferido para um tubo Eppendorf fresco e lisado pela adição de 1 volume de formamida. O ADN de EMBL3 é precipitado pela adição de 2 volumes de etanol a 100%. O ADN de EMBL-3 é recuperado por microcentrifugação, lavado com etanol a 70% e ressuscitado em tampão TE (Tris-HCl 10 mM, pH 7,5; Na₂-EDTA 1 mM). O ADN é extractado várias vezes com fenol:clorofórmio:alcóol isomílico (24:24:1), precipitado com etanol, ressuscitado em tampão TE, digerido com vários enzimas de restrição e submetido a electroforese através dum gel de agarose a 1%. Após a electroforese, o ADN digerido é transferido para nitrocelulose e hibridado com sondas de rNT-4 ou hNT-4 marcado com ³²P, em condições supra descritas. A banda hibridada é subclonada no vector plasmídico pBS-KS e submetida a análise da sequência do ADN pelo processo de terminação de cadeia por didesoxi (Sanger et al., 1977, Proc. Natl. Acad. Sci. USA 74:5463-5467).

As preparações de plasmídeo λ -ZAP foram efectuadas como se segue: são combinados 200 λ de células XL1-Blue com OD₆₀₀=1,0, 200 λ de caldo de fagos de título elevado, e 1 λ do fago auxiliar R408 (1x10 minutos pfu/ml). Para controlo negativo não se adicionou caldo de fagos. Incuba-se a 37 °C durante 15 minutos. Adiciona-se 5 ml de meio 2xYT, agita-se durante 3 horas a 37 °C. Ao fim de 3 horas, o controlo negativo deve estar turvo e as amostras límpidas. As amostras são aquecidas a 65 °C durante 30 minutos, e centrifugadas a 4 000 g durante 5 minutos. O sobrenadante contém um caldo de fagemídeos. Para resgatar os fagemídeos, adiciona-se 0,5 λ do caldo a 200 λ de células XL1-Blue (OD₆₀₀=1). Incuba-se durante 15 minutos a 37 °C. Plaqueia-se 1-100 λ (preferencialmente 10 λ) em placas de LB com ampicilina. Incuba-se a 37 °C durante a noite e picam-se as colónias grandes. Depois de purificar o ADN plasmídico,

sequencia-se como acima.

As bibliotecas de ADNc de fagos são seleccionadas de acordo com processos padrão (Maniatis et al., supra) como supra descrito.

As placas positivas são purificadas, re-isoladas e submetidas a análise da sequência de ADN como supra descrito.

8.2. RESULTADOS E DISCUSSÃO

Uma região da sequência codificante de NT-4 de Xenopus foi usado como modelo para a síntese de iniciadores de oligonucleótidos degenerados. A figura 13 denota que o iniciador de oligonucleótido-5' 2Y [SEQ ID NO:53] (QYFFET) e os iniciadores de oligonucleótidos-3,' 3Y [SEQ ID NO:55] (WISECK), 3Z [SEQ ID NO:56] (CKAKQS) e 4Z [SEQ ID NO:58] (WIRIDT) eram derivados da sequência de aminoácidos de xNT-4. O iniciador de oligonucleótido-5' 2Z [SEQ ID NO:54] (QYFYET) é derivado da região homóloga de rBDNF. Todas as possíveis combinações destes oligonucleótidos degenerados foram utilizadas para amplificar ADN tanto de bibliotecas de ADN genómico de rato como de humano. Uma vez que os iniciadores representados por 3Y [SEQ ID NO:55] e 3Z [SEQ ID NO:56] de xNT-4 não são conservadas na família de genes NGF/BDNF/NT-3, e portanto não fosse provável que amplificassem NGF, BDNF ou NT-3, estes dois iniciadores foram utilizados na reamplificação ou no PCR secundário.

Os fragmentos de ADN com os tamanhos aproximados esperados foram obtidos por amplificação por PCR e reamplificação tanto de bibliotecas genómicas de rato como de humano, quando foram usadas as seguintes combinações de iniciadores:

- (1) 2Y/3Z (PCR primário): 2Y, 2Z/3Y (PCR secundário)
- (2) 2Y/3Z (PCR primário): 2Y, 2Z/32 (PCR secundário)
- (3) 2Y/4Z (PCR primário): 2Y, 2Z/32 (PCR secundário)
- (4) 2Z/4Z (PCR primário): 2Y, 2Z/3Z (PCR secundário)

Os produtos secundários de PCR com o tamanho aproximado

esperado foram submetidos a electroforese através de um gel de agarose a 2%, eluídos por técnicas padrão, digeridos com EcoRI e SalI e ligados ao ADN de pBS-KS digerido com EcoRI-XhoI. Os transformantes positivos foram seleccionados, e os fragmentos inseridos foram sujeitos a sequenciação de ADN pelo processo de terminação de cadeia didesoxi (Sanger et al., supra).

Foi deduzida uma estrutura de leitura aberta para uma porção de NT-4 de rato (Fig. 14 [SEQ ID NO:62]) e a sequência de aminoácidos que codifica NT-4 humano (Fig. 15 [SEQ ID NO:64]). A Figura 16 ilustra a região homóloga dos fragmentos rNT-4 (SEQ ID NO:62) e hNT-4 (SEQ ID NO:64) com membros representativos da família de genes NGF/BDNF/NT-3.

Uma estrutura de leitura aberta codificando uma porção de NT-4 humano maior do que a apresentada na Figura 15 está mostrada na Figura 17A (SEQ ID NO:69 e SEQ ID NO:70). A figura 17A apresenta informação adicional da sequência 3' para a região de codificação do NT-4 humano 3'. O fragmento de ácido nucleico de 192 bp foi isolado como supra descrito na Descrição das Figuras.

O tamanho actual dos produtos de PCR recuperados do procedimento de reamplificação era maior do que o previsto devido a 7 aminoácidos adicionais no fragmento de ADN de NT-4 de rato (GPGVGGG) [SEQ ID NO:101] e de NT-4 humano (GPGAGGG) [SEQ ID NO:102].

As inserções de 7 aminoácidos de rNT-4 e hNT-4 estão descritas como 'GPGXGGG' [SEQ ID NO:100], em que X=V para rNT-4 e X=A para hNT-4. A valina e a alanina possuem um grupo R não polar. Não é presentemente conhecido se a posição 4 é conservada para conter um grupo não polar em outras proteínas NT-4 de mamífero, ou se a inserção de 7 bp será por si só característica de outros genes de NT-4 de mamífero. É interessante notar que o NGF de peixe possui uma inserção de 22 aminoácidos na mesma região como revelado no presente invento.

9. EXEMPLO: ISOLAMENTO E CARACTERIZAÇÃO
DE UM CLONE GENÔMICO DE NT-4 HUMANO

Pesquisámos uma biblioteca genômica de placenta humana em EMBL3/SP6/T7 (Clontech, K802 como hospedeiro). Um total de $1,25 \times 10^6$ pfu foram plaqueados em placas NZY grandes. Foram feitos duplicados de filtros de nitrocelulose da Schleicher and Schuell, e foram hibridados com uma sonda de 120 bp (a partir do clone 17B de hNT-4, que foi obtido a partir de ADN genômico de humano utilizando os iniciadores 2Z4Z seguidos de 2Z3Z), marcados por PCR utilizando iniciadores de oligonucleótidos 2Z/3Z. Os filtros foram hibridados a 60°C com a sonda marcada radioativamente (10^6 cpm/ml) nas seguintes condições de hibridação: NaPO_4 0,5 M, BSA 1%, SDS 7%, EDTA 1 mM, e 100 $\mu\text{g/ml}$ de ADN de esperma de salmão. Os filtros foram então lavados a 60°C com $2\times\text{SSC}$ e 0,1% SDS, e submetidos a auto-radiografia. Após quatro dias de exposição, os sinais positivos foram identificados nos filtros em duplicado. Um total de 7 poços foram repicados, postos em 1 ml de tampão SM, agitadas durante 2 horas, e re-plaqueadas da forma que se segue: 1) 100 μl de uma diluição 10^{-3} (1 μl em 1 ml), misturados com 100 μl de células e plaqueados, quase sempre obtendo uma placa confluenta; 2) 200 μl de uma diluição 10^{-5} , que resulta em placas ("plaque") isoladas. Foram feitos levantamentos duplicados e pesquisaram-se como acima descrito com a sonda de 120 kb de hNT-4. Após 2 dias de exposição, muitos positivos foram identificados na placa confluenta para os poços HG2, 4 e 7. Foi identificado um positivo bem isolado tanto nas placas HG4-2 como nas HG7-2. Foi repicada uma única placa para HG4-2 e HG7-2, colocada em 500 μl de tampão SM, e agitada durante 2 horas, após as quais foi misturado 100 μl de eluente com 100 μl de células e foi plaqueado. A placa foi então inundada com 3 ml de tampão SM, e o sobrenadante foi recolhido como o primeiro caldo de elevado título. As três placas foram então plaqueadas utilizando 100 μl deste primeiro caldo misturado com 100 μl de células. As placas foram então inundadas com 3 ml de tampão SM e agitadas num aparelho rotativo durante 3 horas à temperatura ambiente. O sobrenadante foi removido, centrifugado para remoção dos resíduos celulares, seguiu-se a adição de clorofórmio e

utilizou-se o resultante como segundo caldo de elevado título. Dois μ l de caldo de elevado título de HG4-2 e HG7-2 foram aplicados num filtro de nitrocelulose da Schleicher and Schuell, e verificou-se que hibridava com a sonda de rNT-4 de 180 bp [isolado a partir do plasmídeo contendo uma inserção obtida por PCR de ADN genómico de rato utilizando iniciadores concebidos com base na sequência do nosso clone de NT-4 de rato que codifica os aminoácidos GELSVCD (SEQ ID NO:112) (iniciador degenerado) e KAESAG (SEQ ID NO:113) (iniciador exacto)]. Foram preparados lisados de placas e lisados líquidos para HG4-2, HG7-2 e HG2-1. Foi preparado ADN fágico, uma alíquota de cada um foi submetida a electroforese em gel de agarose e submetida à análise de Southern. Verificou-se que HG4-2, HG7-2 e HG2-1 hibridavam com a sonda de rNT-4 de 180 bp (hibridação NaPO_4 como acima, 65 °C), e com uma sonda de oligonucleótido de mero 45 (GGAGGGGGCTGCCGGGGAGTGGACAGGAGGCACTGGGTATCTGAG) [SEQ ID NO:114] correspondente aos aminoácidos GGGCRGVDRRHVVSE [SEQ ID NO:115] codificados pelo fragmento de PCR humano do clone 17B (6xSSC, hibridação a 45 °C). O tamanho da inserção para estes três clones genómicos é aproximadamente de 9-23 kb. Ambos contêm o exão que codifica o(s) gene(s) que está proximamente relacionado com as sondas utilizadas para a selecção, hNT-4 (120 bp) e rNT-4 (180 bp). O ADN de fago para os clones genómicos foi digerido com vários enzimas de restrição e submetido a análise de Southern. O fragmento apropriado que hibrida com a sonda rNT-4 (180 bp) pode ser subclonada no vector Bluescript. O tamanho dos fragmentos de ADN a ser subclonados é o seguinte: clone 2-1 (fragmento XhoI de 1,0 kb), clone 4-2 (fragmento XhoI de 4,0 kb) e clone 7-2 (fragmento BamHI de 5,0 kb). A sequência de codificação completa pode ser obtida e esta informação pode ser usada para identificar os limites do exão para permitir a subclonagem deste gene num vector de expressão apropriado.

Para este fim, a análise da sequência nucleotídica foi efectuada no clone fágico do genoma humano 7-2, que foi obtido por pesquisa de uma biblioteca genómica humana com um fragmento de PCR derivado de ADN genómico humano utilizando oligonucleótidos degenerados para a sequência de ADN de NT-4 de

Xenopus (ver discussão, supra). A análise da sequência revelou que o clone fágico humano 7-2 contém uma sequência idêntica à sequência do fragmento de PCR utilizado como sonda para pesquisa da biblioteca genómica. Esta sequência está contida naquilo que parece ser um exão que codifica um novo factor neurotrófico (Figura 18, SEQ ID NO:75 e SEQ ID NO:76).

O alinhamento da proteína codificada por este exão (figura 19, SEQ ID NO:77) com as neurotrofinas conhecidas revelou que partilha características encontradas em todas as neurotrofinas conhecidas (Figura 19, SEQ ID Nos: 78-92). Contém uma região prepro em que são conservados muitos dos aminoácidos idênticos conservados entre as regiões prepro de neurotrofinas previamente definidas. Além disso, esta região prepro é precedida por um local de aceitação de união localizado na mesma região como em outros genes de neurotrofina. A região prepro contém também um local de glicosilação de consenso na posição apropriada, e termina num local de clivagem que era muito semelhante aos locais de clivagem encontrados nas outras neurotrofinas (Figura 18). A região prepro de 7-2 é, no entanto, invulgar, devido ao seu curto comprimento quando comparada com a região prepro de neurotrofinas conhecidas. O decréscimo no comprimento ocorre na porção N-terminal da região prepro, que é a menos conservada das porções de prepros entre os membros da família. A região madura retém todas as 6 cisteínas encontradas em todas as neurotrofinas previamente identificadas. Muitos dos resíduos partilhados entre os diferentes membros da família das neurotrofinas são também conservados. Excluindo a extensa similaridade de sequência partilhada por um fragmento de PCR derivado de ADN genómico de rato, que pode corresponder à proteína de rato equivalente à codificada pelo clone humano 7-2, alinhamentos por computador revelaram que a neurotrofina codificada pelo clone fágico 7-2 era muito similar à NT-4 de Xenopus. Isto verificava-se tanto para as regiões prepro como para as regiões maduras. A proteína codificada pelo clone 7-2 é invulgar, quando comparada com as neurotrofinas conhecidas, devido à presença de uma inserção situada entre a segunda e a terceira cisteína na região madura.

A análise da sequência foi também efectuada em dois clones humanos adicionais isolados no mesmo procedimento de pesquisa que deu o clone 7-2 (ver discussão, supra). A sequência destes clones era similar, mas não idêntica, à obtida a partir do clone 7-2, aumentando a possibilidade de que estes codifiquem novas neurotrofinas mais proximamente relacionadas com 7-2 do que com as outras neurotrofinas. A sequência parcial de um destes clones, o clone 2-1, é apresentado na Figura 20 (SEQ ID NO:93 e SEQ ID NO:94). A sequência descrita começa na posição correspondente ao aminoácido 50 nos alinhamentos apresentados na Figura 19. A sequência parcial do outro clone, o clone 4-2, é apresentada na Figura 21 (SEQ ID NO:116 e SEQ ID NO:117).

10. EXEMPLO: EXPRESSÃO ESPECÍFICA EM TECIDOS DE NT-4 HUMANA

Um fragmento XhoI-NotI de 680 bp, contendo a região de codificação completa de clone NT-4 genómico humano, HG7-2, foi marcado radioactivamente e utilizado em análise de Northern de vários PolIA⁺ARN específicos de tecidos humanos. Os ARNm específicos de tecidos humanos foram fraccionados por electroforese através de um gel de 1% de agarose-formaldeído seguida por transferência por capilaridade para uma membrana de nylon com 10xSSC. Os ARN foram reticulados nas membranas por exposição à luz ultravioleta e hibridados a 65 °C com a sonda de NT-4 XhoI-NotI de 680 bp marcada radioactivamente na presença de NaPO₄ 0,5 M (pH 7), 1% de albumina de soro de bovino (Fraction V, Sigma), 7% de SDS, EDTA 1 mM e 100 ng/ml de ADN de esperma de salmão desnaturado e tratado com ultra-sons. O filtro foi lavado a 65 °C com 2xSSC, 0,1% de SDS e submetido a auto-radiografia durante a noite com uma câmara de intensificação e uma película de raios-X a -70 °C. A coloração do gel com brometo de etídeo demonstrou que estavam a ser pesquisados níveis equivalentes de ARN total para as diferentes amostras.

A sonda de NT-4 humana hibridou fortemente com o ARNm de músculo esquelético, próstata, timo, testículos e placenta (Figura 22). A sonda de NT-4 hibridou com um transcrito de ARNm maior no músculo esquelético do que na próstata. Estes dados

sugerem que pode estar presente uma família multigene pequena de NT-4 humana, possuindo diferentes níveis de expressão bem como dimensões dos transcritos.

A elevada expressão de NT-4 humana no tecido muscular esquelético sugere que o presente invento pode ser utilizado para tratar desordens do sistema nervoso, especificamente o vasto conjunto de desordens neurológicas que afectam os neurónios motores (ver discussão, supra). Adicionalmente, a expressão elevada de NT-4 humana no tecido da próstata sugere que o presente invento pode ser utilizado para tratar doenças da próstata, preferencialmente BPH e impotência (ver discussão, supra). Finalmente, a expressão de NT-4 humana em tecido do timo sugere que o presente invento pode ser utilizado para tratar desordens imunológicas relacionadas, do nervo e do tecido muscular, incluindo mas não se limitando a miastenia grave (ver discussão, supra).

11. CONSTRUÇÃO DE NT-4 HUMANA EM VECTORES DE EXPRESSÃO EUCARIÓTICOS E MEDIÇÃO DA ACTIVIDADE BIOLÓGICA DA NT-4 HUMANA

11.1 MATERIAIS E MÉTODOS

11. 1. 1. CONSTRUÇÃO DE VECTORES DE EXPRESSÃO EUCARIÓTICOS QUE CODIFICAM A NT-4 HUMANA

Foram construídos em pCMX (Nº de Acesso NRRL B-18790) dois vectores de expressão eucarióticos contendo a região que codifica o precursor prepro do clone genómico humano HG7-2. A primeira construção utilizou o local normal de início de tradução de pCMX (pCMX-HG7-2Q), enquanto o outro utilizou o local de início de tradução consenso de Kozak (pCMX-HG7-2M). Um fragmento genómico de 5 kb de HG7-2, contendo a região de codificação completa clonada no local Bam HI de Bluescript, foi amplificado por PCR utilizando os seguintes oligonucleótidos:

hNT4-5'XhoM:CGGTACCCTCGAGCCACCATGCTCCCTCTCCCCTCA

[SEQ ID NO:118]

hNT4-3'Not:CGGTACAAGCGGCCGCTTCTTGGGCATGGGTCTCAG

[SEQ ID NO:119]

hNT4-5'XhoQ:CGGTACCCTCGAGCCACCCAGGTGCTCCGAGAGATG

[SEQ ID NO:120]

As combinações de iniciadores de oligonucleótidos de hNT4-5'XhoM e hNT4-3'Not foram utilizados para construir pCMX-HG7-2Q, enquanto os oligo hNT4-5'XhoQ e hNT4-3'Not foram utilizados para construir o pCMX-HG7-2Q. O fragmento de PCR foi digerido com XhoI/NotI e subclonado em pCMX digerido com XhoI/NotI.

11. 1. 2. CONSTRUÇÃO DE GENES QUIMÉRICOS FUNDINDO UMA REGIÃO PREPRO DE NEUROTROFINA COM A REGIÃO DE CODIFICAÇÃO MADURA DO NT-4 HUMANO

Foram construídos dois vectores de expressão eucarióticos que codificam a porção madura de NT-4 humana. Em primeiro lugar, a região prepro de NT-4 humana foi substituída pela região prepro de NT-4 de *Xenopus* (pCMX-xNT4/hNT4). Em segundo lugar, a região prepro de NT-4 humana foi substituída pela região prepro de NT-3 humana (pCMX-hNT3/hNT4). Os seguintes oligonucleótidos foram utilizados na construção do pCMX-xNT4/hNT4 e do pCMX-hNT3/hNT4:

- (1) 5'CDM8: GAGACCGGAAGCTTCTAGAGATC [SEQ ID NO:121]
- (2) Fusão hNT3/HNT4 (oligonucleótido "US"):
TGCAGTTTCGCTCACCCCCGTTTCCGCCGTGATGT [SEQ ID NO:122]
- (3) Fusão hNT3/HNT4 (oligonucleótido "DS"):
ACATCACGGCGGAAACGGGGGGTGAGCGAAACTGCA [SEQ ID NO:123]
- (4) Fusão xNT4/hNT3 (oligonucleótido "DS"):
ACTTCCCGGCTAAAACGGGGGGTGAGCGAAACTGCA [SEQ ID NO:124]
- (5) Fusão xNT4/hNT4 (oligonucleótido "US"):
TGCAGTTTCGCTCACCCCCGTTTCCGCCGTGATGT [SEQ ID NO:109]

O vector plasmídico contendo hNT-3 (pC8-hNT3) foi amplificado por PCR com os oligonucleótidos de fusão 5'CDM8 e hNT3/hNT4 como iniciadores. O plasmídeo contendo hNT-4 (pCMX-HG7-2Q) foi amplificado por PCR com o oligonucleótido de

fusão de hNT3/hNT4 "DS" e com o oligonucleótido hNT4-3'Not1. O fragmento de PCR obtido foi excisado do gel, e reamplificado por PCR com os oligonucleótidos 5'CDM8 e hNT4-3'Not1. O produto foi então digerido com HindIII e PstI e subclonado em pCMX-HG7-2Q digerido com HindIII/PstI. Desta maneira, o plasmídeo de expressão pCMX-hNT3/hNT4 continha a região prepro de hNT3 fundida com a região que codifica NT-4 madura. Do mesmo modo, o plasmídeo de expressão de NT-4 humana (pCMX-HG7-2Q) foi amplificado por PCR com os oligonucleótidos "US" de fusão xNT4/hNT4 e 5'CDM8 como iniciadores, enquanto o pCMX-HG7-2Q foi amplificado com os oligonucleótidos "DS" de fusão xNT4/hNT4 e hNT4-3'Not1. O fragmento de PCR foi excisado gel, reamplificado com os oligonucleótidos 5'CDM8 e o hNT4-3'Not. O produto foi então digerido com HindIII e PstI e subclonado em pCMX-HG7-2Q digerido com HindIII/PstI. Assim, o plasmídeo de expressão eucariótico resultante, pCMX-xNT4/hNT4 contém a região prepro de NT-4 de Xenopus fundida com a região de codificação de NT-4 humana madura.

11. 1. 3. EXPRESSÃO DE NT-4 HUMANA RECOMBINANTE EM CÉLULAS COS

Foram desenvolvidas células COS M5 até uma densidade de $1,5 \times 10^5$ células/poço numa placa de 6 poços Costar em meios DMEM suplementados com 10% de FBS, glutamina e piruvato de Na (todos da Irvine Scientific excepto o FBS).

No dia seguinte as células foram aspiradas e alimentadas com 2 ml/poço de meios RPMI contendo 400 μ g/ml de DEAE-Dextran (Pharmacia), cloroquina 400 μ M (Sigma), glutamina 4 mM (Irvine), 1 x ITS (insulina, transferrina, selénio, Sigma). A cada poço foi adicionado 2 μ g do ADN apropriado e misturou-se por rotação. Foram utilizadas três construções em separado: pCMX-xNT4, contendo o precursor prepro de NT-4 de Xenopus; e duas construções de NT-4 humana, pCMX-HG7-2M e pCMX-HG7-2Q. Após a adição do ADN as placas foram recolocadas a 37 °C numa incubadora com 5% de CO₂ durante 3 horas e 15 minutos. A mistura meio/ADN foi então aspirada e foram adicionados 2 ml/poço de

DMSO a 10% em PBS isento de Ca^{2+} , Mg^{2+} durante 2 minutos. O DMSO/PBS foi aspirado, os poços foram lavados mais uma vez com 10% de FBS/DMEM e realimentados com 10% de FBS/DMEM. Na manhã seguinte, as placas a ser bioensaiadas foram lavadas uma vez com Meio Definido (DM) e realimentadas com 2 ml/poço de DM. Três dias após a transfecção, os sobrenadantes foram removidos das células e os resíduos celulares foram precipitados por microcentrifugação. Os sobrenadantes foram transferidos para tubos novos e ensaiados quanto à bioactividade.

11.1.4. PREPARAÇÃO DE CULTURAS ENRIQUECIDAS EM NEURÓNIOS MOTORES

Foram utilizados em todas as experiências embriões (E14) de ratos Sprague-Dawley (HSD OU Zivic-Miller). Foram sacrificadas ratas grávidas por asfixia por dióxido de carbono, e os embriões foram rapidamente removidos e colocados num meio gelado para dissecação posterior. As espinais medulas foram removidas assépticamente dos embriões de rato com 14 dias de gestação. A região caudal da espinal medula foi separada da região do bolbo (ao nível do primeiro gânglio da raiz dorsal), liberta dos gânglios sensoriais e das meninges aderentes. A medula foi então subdividida em segmentos ventrais e médiodorsais para culturas separadas. Os tecidos da espinal medula ventral foram segmentados em pequenos pedaços e incubados em 0,1% de tripsina (GIBCO) e 0,01% de desoxiribonuclease tipo I (Sigma) em PBS a 37 °C durante 20 minutos. A solução de tripsina foi então removida, lavada e substituída com meio consistindo em 45% de meio essencial mínimo de Eagle (MEM), 45% de mistura nutriente F12 de Ham(F12), 5% de soro fetal de bovino inactivado pelo calor (GIBCO), 5% de soro de cavalo inactivado pelo calor (GIBCO), glutamina (2 mM), penicilina G (0,5 U/ml), e estreptomicina (0,5 µg/ml). O tecido foi então mecanicamente dissociado por trituração suave através de uma pipeta de Pasteur, e os sobrenadantes foram combinados e filtrados através de uma fibra de nylon (Nitex, Tekto; 40 µm). A suspensão de células filtrada foi então sujeita a uma modificação do procedimento de fraccionamento descrito por Schnaar e Schnaffner (1981, J.

Neurosci. 1:204-217). Todos os passos foram efectuados a 4 °C. A metrizamida foi dissolvida em meio F12:MEM (1:1), e foi estabelecido um gradiente descontínuo que consistiu na preparação de camadas de metrizamida a 18% (0,5 ml), 3 ml de metrizamida a 17%, 3 ml de metrizamida a 12%, e 3 ml de metrizamida a 8%. A suspensão filtrada de células da espinal medula ventral (2,5 ml) obtida como acima descrito, foi acamada sobre o gradiente descontínuo, o tubo foi centrifugado a 2500 x g durante 15 minutos utilizando um rotor oscilante (Sorvall HB4). A centrifugação resultou em três camadas de células: fracção I (na interface 0-8%), fracção II (na interface 8-12%), e fracção III (na interface 12-17%). As células de cada interface foram removidas num volume pequeno (cerca de 1 ml), lavadas duas vezes com meio definido isento de soro consistindo em 50% de F12 e 50% de MEM, suplementado com glutamina (2 mM), insulina (5 µg/ml), transferrina (100 µg/ml), progesterona (20 nM), putrescina (100 µM), e selenite de sódio (30 nM) (Bottenstein e Sato, 1979, Proc. Natl. Acad. Sci. 76:514-517). A contagem das células viáveis foi obtida por contagem em hemocítmetro na presença de azul de tripano. As células ventrais da espinal medula fraccionada (enriquecidas com neurónios motores) foram então plaqueadas com uma densidade de 100 000 células/cm² em poços de 6 mm previamente revestidos com poli-L-ornitina (Sigma: 10 µg/ml) e laminina (GIBCO: 10 µg/ml). O tratamento com sobrenadantes de células COS contendo NT-4 de células COS foi efectuado no dia do plaqueamento. As culturas foram mantidas em meio definido isento de soro a 37 °C numa atmosfera de 95% de ar/5% de CO₂ com quase 100% de humidade relativa. No dia 2 (quarenta e oito horas), as células foram colhidas para medição da colina-acetiltransferase (CAT) como descrito em Fonnum, 1975, J. Neurochem. 24:407-409.

11.2 RESULTADOS

11.2.1. EXPRESSÃO EUCARIOTA DE NT-4 HUMANA RECOMBINANTE BIOLOGICAMENTE ACTIVA

O ADN plasmídico de cada uma das construções baseadas no

pCMX (pCMX-HG7-2M, pCMX-hNT3/hNT4 e pCMX-xNT4/hNT4) foi preparado e individualmente transfectado para células COS. Os sobrenadantes de COS de cada linha de células transfectadas foram utilizados de modo a avaliar a actividade biológica de cada forma recombinante respectiva de NT-4. Os volumes de sobrenadantes de COS testados foram 10, 50 e 250 μ l num volume total de 2 ml. Q1 (pCMX-HG7-2Q), N7 (pCMX-hNT3/hNT4 de fusão), e X1 (pCMX-xNT4/hNT4) possuíam actividade de promoção de neurites em explantes DRG (Figura 23). Adicionalmente, tanto Q (pCMX-HG7-2Q) como M (pCMX-HG7-2M) foram examinados quanto à sua actividade de promoção de sobrevivência em células dissociadas de DRG. Os volumes testados foram de 5-250 μ l num volume total de 2 ml. Quando adicionados às culturas de neurónios dissociados de DRG, os sobrenadantes de COS contendo hNT-4 promoveram 30% de sobrevivência neuronal comparados com 10% de sobrevivência com sobrenadantes de células COS simuladamente transfectadas. (Figura 24).

O efeito biológico da proteína humana recombinante de sobrenadantes de células COS transfectadas com pCMX-HG7-2M foi testado em culturas enriquecidas em neurónios motores preparadas como supra descrito no Exemplo da Secção 11.1.4. O tratamento das culturas enriquecidas em neurónios motores com pCMX-HG7-2M derivado de NT-4 humana diluído de 1:5, resultou num aumento de 2,9 vezes na actividade da colina-acetiltransferase (CAT) após 48 horas, quando comparado com controlos não tratados (C-NT) e simuladamente transfectados (MOC COS) (Figura 25). O aumento na actividade de CAT caiu 1,7 vezes quando foi testada uma diluição de 1:50, sugerindo que era uma resposta dependente da dose (Figura 25).

11.3 DISCUSSÃO

O presente invento proporciona a utilização de um sistema de expressão eucariótico in vitro para a expressão de NT-4 humana recombinante. O presente invento descreve várias estratégias para a expressão de uma forma biologicamente activa de NT-4 humana recombinante nas células COS. Num exemplo, a sequência de ADN que codifica o precursor prepro de NT-4 foi

amplificada utilizando duas estratégias de amplificação por PCR, para produzir plasmídeos de expressão com base no pCMX contendo o local de início de tradução pCMX(pCMXHG7-2M), ou um local de tradução consenso de Kozak (pCMX-HG7-2Q). Noutro exemplo, foram construídos para expressão nas células COS, dois genes quiméricos de neurotrofina fundindo a região prepro de NT-4 de Xenopus (pCMX-xNT4/hNT4) ou da região prepro de NT-3 humana (pCMX-hNT3/hNT4), com a região de codificação da NT-4 madura (ver Secção 5, supra, para uma discussão da utilização de construções quiméricas para a expressão de NT-4 in vitro).

A expressão de uma forma biologicamente activa de NT-4 humana num sistema eucariótico de expressão in vitro, aumenta substancialmente a facilidade com que a produção de NT-4 humana recombinante, péptidos ou seus derivados, podem ser escalonados tanto para aplicações terapêuticas como de diagnóstico, discutidas supra. Do ponto de vista do presente invento, qualquer perito na arte pode facilmente construir um plasmídeo contendo uma sequência de ADN idêntica, como descrito, ou uma sequência similar codificando uma proteína ou derivado desta, semelhante a uma NT-4 homóloga, porém distinta. O perito na arte pode também usar ou escolher entre vários vectores plasmídicos de ADN conhecidos na arte, para construir um plasmídeo de expressão para utilização num sistema de expressão eucariótico. Demostrámos que a NT-4 humana recombinante, produzida como precursor prepro completo ou por via de uma construção quimérica baseada na neurotrofina, é biologicamente activa, como demonstrado pelo efeito de estimulação de sobrenadantes de COS de NT-4 recombinante na excrescência de neurites em explantes DRG e bioactividade de culturas de neurónios motores.

12. EXEMPLO: trkB É UM RECEPTOR PARA A NEUROTROFINA-4

Sobrenadantes de células COS foram também examinados num ensaio de sobrevivência utilizando fibroblastos 3T3. Neste sistema de ensaio, os fibroblastos 3T3, que não expressam proteínas receptoras de neurotrofina, são transfectados com vectores de expressão de mamíferos que codificam trkA ou trkB. A sobrevivência de fibroblastos 3T3, é dependente da adição de um

factor neurotrófico e respectiva ligação a um receptor específico.

Células COS-M5 foram cultivadas e transfectadas com pCMX-HG7-2Q, pCMX-HG7-2M ou pCMX-HG7-2Q como descrito no Exemplo da Secção 11.1.3.

Um clone de ADNc de trkA completo de rato foi obtido do Dr. Eric Shooter da Universidade de Stanford. O ADNc de trkA de rato foi sub-clonado no vector de expressão de mamífero, pCMX para originar pCMX-trkA.

Um clone de ADNc de trkB de rato foi obtido por pesquisa de uma biblioteca de ADNc de cérebro de rato, no vector lambda Zap2 (Stratagene) com oligonucleótidos específicos para trkB de rato, correspondentes à maior parte das regiões de codificação 5' e 3' de trkB. O ADNc de trkB de rato foi subclonado em pCMX para produzir pCMX-trkB.

Fibroblastos 3T3 foram cultivados e transfectados como descrito em Glass et al., 1991, Cell 66:405-413.

Neste sistema de ensaio de sobrevivência, fibroblastos 3T3, que não expressam proteínas receptoras de neurotrofinas, foram transfectados com trkA, um proto-oncogene que codifica um receptor de tirosina quinase para NGF, ou com trkB, uma tirosina quinase que serve como uma proteína de ligação funcional para BDNF e NT-3. As células transfectadas estão dependentes da adição da neurotrofina correspondente para a sobrevivência, e assim podem ser utilizadas para o ensaio de actividade biológica das neurotrofinas. A adição de sobrenadantes de células COS contendo NT-4 neste bioensaio indicou que as células viáveis permanecem após 48 horas apenas em culturas de 3T3 trkB (Tabela 2). Assim, estes resultados demonstram que a proteína NT-4 possui actividade biológica neste sistema, e sugere que o trkB, mas não o trkA, serve como uma proteína de ligação funcional para o NT-4.

TABELA 2



Ensaio de Sobrenadantes de COS em Linhas celulares
3T3 que Expressam TrkA e TrkB.

	Diluições	Simulados	HG7- 2Q	HG7- 2M	hNT3/ hNT4
3T3	1:5	-	-	-	-
	1:10	-	-	-	-
	1:20	-	-	-	-
	1:50	-	-	-	-
3T3-trkA	1:5	-	-	-	-
	1:10	-	-	-	-
	1:20	-	-	-	-
	1:50	-	-	-	-
3T3-trkB	1:5	-	+	+	+
	1:10	-	+	+	+
	1:20	-	+	+	+
	1:50	-	+	+	+

13. DEPÓSITOS DE MICRORGANISMOS

Os seguintes bacteriófagos recombinantes contendo uma sequência do genoma humano afim com neurotrofina-4, foram depositados em 22 de Agosto, 1991 (HG4-2 e HG7-2) e em 11 de Setembro, 1991 (HG2-1) com a American Type Culture Collection, 12301 Parklawn Drive, Rockville, Maryland 20852, e marcadas com o número de acesso indicado. Adicionalmente, a construção do gene quimérico, pCMX-hNT3/hNT4, foi depositada em 30 de Outubro, 1991 com a American Type Culture Collection e marcada com o número de acesso indicado.

<u>Bacteriófago</u>	<u>Número de acesso ATCC</u>
HG4-2	75069
HG7-2	75070
HG2-1	75098
pCMX-hNT3/hNT4	75133

O presente invento não está limitado em âmbito pelos microorganismos depositados ou pelas concretizações específicas aqui descritas. De facto, várias modificações do invento em adição às aqui descritas, tornar-se-ão aparentes aos peritos na arte a partir da descrição anterior e figuras acompanhantes. Pretende-se que tais modificações caiam no âmbito das reivindicações em apêndice.

Várias publicações foram aqui citadas as quais são aqui incorporadas como referência na sua totalidade.



(1) INFORMAÇÃO PARA SEQ. ID N° 1:

- (i) Características da Sequência
 - (A) Comprimento: 130 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genômico)

(xi) Descrição da Sequência: SEQ ID N°: 1:

CAAATGTAAT CCCGCTGGTG GAACTGTGGG TGGCTGCCGG GGTGTTGATC GACGCCATTG	60
GATCTCTGAG TGCAAGGCTA AACAGTCTTA CGTGAGGGCT CTGACTATGG ATTCTGACAA	120
GATTGTTGGC	130

(1) INFORMAÇÃO PARA SEQ. ID N° 2:

- (i) Características da Sequência
 - (A) Comprimento: 130 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida

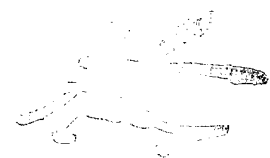
(ii) Tipo de Molécula: ADN (genômico)

(xi) Descrição da Sequência: SEQ ID N°: 2:

CAAGTGCAAT CCATCAGGCA GCACCACTAG AGGATGCCGA GGTGTAGACA AAAAGCAATG	60
GATATCTGAG TGCAAAGCAA AACAGTCTTA TGTGAGGGCT CTGACCATAG ATGCCAACAA	120
GCTTGTGGGT	130

(1) INFORMAÇÃO PARA SEQ. ID N° 3:

- (i) Características da Sequência
 - (A) Comprimento: 127 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla



(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 3:

```
CAAGTGCCGG GACCCAAATC CCGTTGACAG CGGGTGCCGG GGCATTGACT CAAAGCACTG      60
GAACTCATAT TGTACCACGA CTCACACCTT TGTCAAGGCG CTGACCATGG ATGGCAAGCA      120
GGCTGCC                                     127
```

(1) INFORMAÇÃO PARA SEQ. ID N° 4:

(i) Características da Sequência

- (A) Comprimento: 127 pares de bases
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 4:

```
CAAGTGCCGG GCCCCAAATC CTGTAGAGAG TGGATGCCGG GGCATTGACT CCAAGCACTG      60
GAACTCATAC TGCACCACGA CTCACACCTT TGTCAAGGCG TTGACAACAG ACGACAAACA      120
GGCTGCC                                     127
```

(2) INFORMAÇÃO PARA SEQ. ID N° 5:

(i) Características da Sequência

- (A) Comprimento: 127 pares de bases
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 5:

CAAGTGCAGG GACCCTAGGC CGGTGTCCAG CGGGTGCCGA GGGATCGATG CGAAGCATTG	60
GAACTCTTAC TGCACCACGA CACACACCTT CGTCAAAGCA CTGACCATGG AGGGCAAGCA	120
AGCAGCC	127

(2) INFORMAÇÃO PARA SEQ. ID N° 6:

(i) Características da Sequência

- (A) Comprimento: 127 pares de bases
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 6:

CAAGTGCAA AATCCAAGTC CAGTATCAGG TGGGTGCAGG GGCATTGATG CCAAGCATTG	60
GAATTCGTAT TGCACCACAA CAGACACATT TGTCAGGGCA TTAACCATGG AAGGCAATCA	120
GGCATCT	127

(2) INFORMAÇÃO PARA SEQ. ID N° 7:

(i) Características da Sequência:

- (A) Comprimento: 127 pares de bases
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 7:

CAAATGCAGG GACCCAAAGC TAGTTTCAAG CGGATGCCGT GGGATTGATG CAAAGCATTG	60
GAACTCTTAT TGTACCACCA CGCACACCTT TGTCAAAGCA TTAACAATGG AAGGGAAGCA	120
AGCAGCA	127

(2) INFORMAÇÃO PARA SEQ. ID N° 8:

- (i) Características da Sequência:
 - (A) Comprimento: 127 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 8:

CACGTGCCGT GCGCCCCGGG CGGGCAGCTC TGGCTGCCTG GGCATCGACG GGCGACACTG	60
GAACTCCTAC TGCACCAACT CGCACACCTT CGTGCGGGCG CTGACTTCCT TTAAGGACCT	120
GGTGGCC	127

(2) INFORMAÇÃO PARA SEQ. ID N° 9:

- (i) Características da Sequência:
 - (A) Comprimento: 130 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 9:

CAAGTGCAAT CCCATGGGTT ACACAAAAGA AGGCTGCAGG GGCATAGACA AAAGGCATTG	60
GAACTCCCAG TGCCGAAC TA CCCAGTCGTA CGTGCGGGCC CTTACCATGG ATAGCAAAAA	120
GAGAATTGGC	130

(2) INFORMAÇÃO PARA SEQ. ID N° 10:

- (i) Características da Sequência
 - (A) Comprimento: 130 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla

(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 10:

```
CAAGTGTAAT CCCATGGGTT ACACGAAGGA AGGCTGCAGG GGCATAGACA AAAGGCACTG    60
GAACTCGCAA TGCCGAACTA CCCAATCGTA TGTTCGGGCC CTTACTATGG ATAGCAAAAA    120
GAGAATTGGC                                     130
```

(2) INFORMAÇÃO PARA SEQ. ID N° 11:

(i) Características da Sequência

- (A) Comprimento: 130 pares de bases
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 11:

```
CAAATGCAAC CCCAAGGGGT ACACAAAGGA AGGCTGCAGG GGCATAGACA AGAGGCACTG    60
GAACTCACAG TGCCGAACTA CCCAGTCTTA CGTGAGAGCT CTCACCATGG ATAACAAAAA    120
GAGAGTTGGC                                     130
```

(2) INFORMAÇÃO PARA SEQ. ID N° 12:

(i) Características da Sequência

- (A) Comprimento: 130 pares de bases
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

CAAGTGCAGC	ACGAAGGGTT	ATGCAAAAGA	AGGCTGTAGA	GGCATAGACA	AGAGGTACTG	60
GAATTCCCAG	TGCCGAACTA	CTCAGTCTTA	CGTCCGCGCT	CTCACCATGG	ATAACAAAAA	120
GAGGATTGGA						130

(i) Características da Sequência

- (A) Comprimento: 130 pares de bases
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(xi) Descrição da Sequência: SEQ ID N°: 13:

CAAAATGCAAC	CCTATGGGTT	ACATGAAAGA	AGGCTGCAGA	GGCATAGACA	AAAGGTACTG	60
GAAGTCTCAG	TGCCGAAC	CTCAGTCTTA	CGTGCGGGCT	TTCACCATGG	ATAGCAGAAA	120
AAAAGTTGGT						130

(i) Características da Sequência

- (A) Comprimento: 130 pares de bases
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(xi) Descrição da Sequência: SEQ ID N°: 14:

CAAATGTAAC	CCTATGGGGT	ACACAAAGGA	GGGCTGCCGT	GGAATAGACA	AGAGGCATTA	60
TAACTCCCAA	TGCAGGACAA	CCCAGTCCTA	CGTGCGAGCG	CTCACCATGG	ATAGCAAAAA	120
GAAGATTGGC						130

(2) INFORMAÇÃO PARA SEQ. ID N° 15:

- (i) Características da Sequência
 - (A) Comprimento: 130 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 15:

CAAATGTAAC CCCAAGGGTT TCACCAACGA AGGCTGCAGA GGGATAGACA AGAAACATTG	60
GAATTCGCAG TGTAGAACCA GCCAATCCTA TGTGCGAGCT CTAACCATGG ATAGTAGGAA	120
GAAGATTGGG	130

(2) INFORMAÇÃO PARA SEQ. ID N° 16:

- (i) Características da Sequência
 - (A) Comprimento: 130 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 16:

GCGATGTAAG GAAGCCAGGC CGGTCAAAAA CGGTTGCAGG GGTATTGATG ATAAACACTG	60
GAACTCTCAG TGCAAAACAT CCCAAACCTA CGTCCGAGCA CTGACTTCAG AGAACAATAA	120
ACTCGTGGGC	130

(2) INFORMAÇÃO PARA SEQ. ID N° 17:

- (i) Características da Sequência
 - (A) Comprimento: 130 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla

(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 17:

```
GAGGTGTAAA GAAGCCAGGC CAGTCAAAAA CGGTTGCAGG GGGATTGATG ACAAACACTG    60
GAACTCTCAG TGCAAAACGT CGCAAACCTA CGTCCGAGCA CTGACTTCAG AAAACAACAA    120
ACTCGTAGGC                                     130
```

(2) INFORMAÇÃO PARA SEQ. ID N° 18:

(i) Características da Sequência

(A) Comprimento: 130 pares de bases

(B) Tipo: ácido nucleico

(C) Tipo de cadeia: dupla

(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 18:

```
AAGGTGTAAA GAAGCCAAAC CTGTTAAAAA TGGCTGCCGA GGCATTGACG ACAAGCACTG    60
GAACTCCCAG TGCAAGACAT CCCAACTTA CGTTAGAGCA TTGACTTCAG AAAACAATAA    120
ACTTGTAGGC                                     130
```

(2) INFORMAÇÃO PARA SEQ. ID N° 19:

(i) Características da Sequência

(A) Comprimento: 66 pares de bases

(B) Tipo: ácido nucleico

(C) Tipo de cadeia: dupla

(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 19:

AAGGTGTAAA GAGGCAAGAC CTGTCAAAAA TGGCTGTCTGA GGCATAGACG ACAAACACTG 60
GAATTC 66

(2) INFORMAÇÃO PARA SEQ. ID N° 20:

- (i) Características da Sequência
 - (A) Comprimento: 128 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: não conhecida
 - (D) Topologia: linear

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 20:

CAAGTGTCTGG ACTGCCAAAC CTTTAAAGAG CGGCTGTCTGC GGCATCGATG ACAAACACTG 60
GAACTCGCAG TGTAAGACCT CTCAGACGTA CGTCAGAGTC TCTGACGCAG GACCGTACCT 120
CTGTGGGC 128

(2) INFORMAÇÃO PARA SEQ. ID N° 21:

- (i) Características da Sequência
 - (A) Comprimento: 130 pares de bases
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 21:

CCGCTGCAAG GAGTCGAAGC CGGGCAAGAA CGGGTGCCGG GGCATCGACG ACAAACACTG 60
GAACTCGCAG TGCAAGACCA GCCAGACCTA TGTCCGAGCG CTGAGCAAGG AGAACAATAA 120
ATATGTGGGC 130

(2) INFORMAÇÃO PARA SEQ. ID N° 22:

(i) Características da Sequência

- (A) Comprimento: 43 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 22:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 23:

(i) Características da Sequência

- (A) Comprimento: 43 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 23:

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

(2) INFORMAÇÃO PARA SEQ. ID Nº 24:

(i) Características da Sequência

- (A) Comprimento: 42 aminoácidos
(B) Tipo: aminoácido
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 24:

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

(2) INFORMAÇÃO PARA SEQ. ID Nº 25:

(i) Características da Sequência

- (A) Comprimento: 42 aminoácidos
(B) Tipo: aminoácido
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 25:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 26:

- (i) Características da Sequência
(A) Comprimento: 42 aminoácidos
(B) Tipo: aminoácido
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 26:

Lys Cys Arg Asp Pro Arg Pro Val Ser Ser Gly Cys Arg Gly Ile Asp
1 5 10 15

Ala Lys His Trp Asn Ser Tyr Cys Thr Thr Thr His Thr Phe Val Lys
20 25 30

Ala Leu Thr Met Glu Gly Lys Gln Ala Ala
35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 27:

- (i) Características da Sequência
(A) Comprimento: 42 aminoácidos
(B) Tipo: aminoácido
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 27:

Lys Cys Lys Asn Pro Ser Pro Val Ser Gly Gly Cys Arg Gly Ile Asp
1 5 10 15

Ala Lys His Trp Asn Ser Tyr Cys Thr Thr Thr Asp Thr Phe Val Arg
20 25 30

Ala Leu Thr Met Glu Gly Asn Gln Ala Ser
35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 28:

- (i) Características da Sequência
(A) Comprimento: 42 aminoácidos
(B) Tipo: aminoácido
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 28:

Lys Cys Arg Asp Pro Lys Pro Val Ser Ser Gly Cys Arg Gly Ile Asp
1 5 10 15

Ala Lys His Trp Asn Ser Tyr Cys Thr Thr Thr His Thr Phe Val Lys
20 25 30

Ala Leu Thr Met Glu Gly Lys Gln Ala Ala
35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 29:

- (i) Características da Sequência
(A) Comprimento: 42 aminoácidos
(B) Tipo: aminoácido
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 29:

Thr Cys Arg Gly Ala Arg Ala Gly Ser Ser Gly Cys Leu Gly Ile Asp
1 5 10 15

Gly Arg His Trp Asn Ser Tyr Cys Thr Asn Ser His Thr Phe Val Arg
20 25 30

Ala Leu Thr Ser Phe Lys Asp Leu Val Ala
35 40

(2) INFORMAÇÃO PARA SEQ. ID Nº 30:

- (i) Características da Sequência
 - (A) Comprimento: 43 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID Nº: 30:

Lys Cys Asn Pro Met Gly Tyr Thr Lys Glu Gly Cys Arg Gly Ile Asp
1 5 10 15

Lys Arg His Trp Asn Ser Gln Cys Arg Thr Thr Gln Ser Tyr Val Arg
20 25 30

Ala Leu Thr Met Asp Ser Lys Lys Arg Ile Gly
35 40

(2) INFORMAÇÃO PARA SEQ. ID Nº 31:

- (i) Características da Sequência
 - (A) Comprimento: 43 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID Nº: 31:

Lys Cys Asn Pro Met Gly Tyr Thr Lys Glu Gly Cys Arg Gly Ile Asp
1 5 10 15

Lys Arg His Trp Asn Ser Gln Cys Arg Thr Thr Gln Ser Tyr Val Arg
20 25 30

Ala Leu Thr Met Asp Ser Lys Lys Arg Ile Gly
35 40

(2) INFORMAÇÃO PARA SEQ. ID Nº 32:

(i) Características da Sequência

- (A) Comprimento: 43 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID Nº: 32:

Lys Cys Asn Pro Lys Gly Tyr Thr Lys Glu Gly Cys Arg Gly Ile Asp
 1 5 10 15

Lys Arg His Trp Asn Ser Gln Cys Arg Thr Thr Gln Ser Tyr Val Arg
 20 25 30

Ala Leu Thr Met Asp Asn Lys Lys Arg Val Gly
 35 40

(2) INFORMAÇÃO PARA SEQ. ID Nº 33:

(i) Características da Sequência

- (A) Comprimento: 43 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID Nº: 33:

Lys Cys Ser Thr Lys Gly Tyr Ala Lys Glu Gly Cys Arg Gly Ile Asp
 1 5 10 15

Lys Arg Tyr Trp Asn Ser Gln Cys Arg Thr Thr Gln Ser Tyr Val Arg
 20 25 30

Ala Leu Thr Met Asp Asn Lys Lys Arg Ile Gly
 35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 34:

(i) Características da Sequência

- (A) Comprimento: 43 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 34:

Lys Cys Asn Pro Met Gly Tyr Met Lys Glu Gly Cys Arg Gly Ile Asp
 1 5 10 15

Lys Arg Tyr Trp Asn Ser Gln Cys Arg Thr Thr Gln Ser Tyr Val Arg
 20 25 30

Ala Phe Thr Met Asp Ser Arg Lys Lys Val Gly
 35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 35:

(i) Características da Sequência

- (A) Comprimento: 43 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 35:

Lys Cys Asn Pro Met Gly Tyr Thr Lys Glu Gly Cys Arg Gly Ile Asp
 1 5 10 15

Lys Arg His Tyr Asn Ser Gln Cys Arg Thr Thr Gln Ser Tyr Val Arg
 20 25 30

Ala Leu Thr Met Asp Ser Lys Lys Lys Ile Gly
 35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 36:

(i) Características da Sequência

- (A) Comprimento: 43 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 36:

Lys Cys Asn Pro Lys Gly Phe Thr Asn Glu Gly Cys Arg Gly Ile Asp
1 5 10 15

Lys Lys His Trp Asn Ser Gln Cys Arg Thr Ser Gln Ser Tyr Val Arg
20 25 30

Ala Leu Thr Met Asp Ser Arg Lys Lys Ile Gly
35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 37:

(i) Características da Sequência

- (A) Comprimento: 43 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 37:

Arg Cys Lys Glu Ala Arg Pro Val Lys Asn Gly Cys Arg Gly Ile Asp
1 5 10 15

Asp Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr Val Arg
20 25 30

Ala Leu Thr Ser Glu Asn Asn Lys Leu Val Gly
35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 38:

- (i) Características da Sequência
 - (A) Comprimento: 43 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 38:

Arg Cys Lys Glu Ala Arg Pro Val Lys Asn Gly Cys Arg Gly Ile Asp
 1 5 10 15

Asp Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr Val Arg
 20 25 30

Ala Leu Thr Ser Glu Asn Asn Lys Leu Val Gly
 35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 39:

- (i) Características da Sequência
 - (A) Comprimento: 43 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 39:

Arg Cys Lys Glu Ala Lys Pro Val Lys Asn Gly Cys Arg Gly Ile Asp
 1 5 10 15

Asp Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr Val Arg
 20 25 30

Ala Leu Thr Ser Glu Asn Asn Lys Leu Val Gly
 35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 40:

- (i) Características da Sequência
(A) Comprimento: 21 aminoácidos
(B) Tipo: aminoácido
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 40:

Arg Cys Lys Glu Ala Arg Pro Val Lys Asn Gly Cys Arg Gly Ile Asp
1 5 10 15

Asp Lys His Trp Asn
20

(2) INFORMAÇÃO PARA SEQ. ID N° 41:

- (i) Características da Sequência
(A) Comprimento: 42 aminoácidos
(B) Tipo: aminoácido
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 41:

Lys Cys Arg Thr Ala Lys Pro Phe Lys Ser Gly Cys Arg Gly Ile Asp
1 5 10 15

Asp Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr Val Arg
20 25 30

Ala Leu Thr Gln Asp Arg Thr Ser Val Gly
35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 42:

- (i) Características da Sequência
 - (A) Comprimento: 43 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 42:

Arg Cys Lys Glu Ser Lys Pro Gly Lys Asn Gly Cys Arg Gly Ile Asp
1 5 10 15

Asp Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr Val Arg
 20 25 30

Ala Leu Ser Lys Glu Asn Asn Lys Tyr Val Gly
 35 40

(2) INFORMAÇÃO PARA SEQ. ID N° 43:

- (i) Características da Sequência
 - (A) Comprimento: 1302 pares de bases
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

- (ix) Estrutura
 - (A) Nome/Código: CDS
 - (B) Localização: ligação (529..534, 538..1248)

(xi) Descrição da Sequência: SEQ. ID N°: 43

(Segue Sequência)

-101-

CAATCATACT TATGAACAGC AGGGGGAGCC CTCGCCTTAC TTCCCAGCCA TGCAGAACTC	60
AAGCAGCTTT GTTTATGCCG ATCCCTAAGC AGCCCAGACC ACACTGAGCA TGTGCACAGT	120
CTTAGTCTTG CAAAGATGTT TAACAAAGTT ACAAGATGGT GACCCCTGT AGCCAACTTT	180
GAAAGCATAA ATCATTTGTT TGATTAGGCT TGTGGTGCAG TAAGTTCATG TTTATATTTA	240
GCATACAAAA TACAGCATTT CTAGCCTTAT TCTATTTTAG ACTTTACCCT TTAATGCCCA	300
GTTCTGCCCC TTGCCTTATA GATGTTAAAG TCCCAATATC ACATTGGCAT CCTCGGCTGT	360
TTACAACAAA CATTAAACT TGTACTTATA TTTAACATTC TGTGTCTCTT CCAATATTCC	420
ATCACACTTA GACCCTAAAA GAATTATATG TATATAATTT GCATAAATTA TATAATGGCA	480
GCCGTATTCT AATTCTGTTT TTTTTTTTTT TTTTGCAGT GGTCTGAG GTG GAT	534
Val Asp	
1	
TAA GTA ATG ATC CTC CGC CTT TAT GCC ATG GTG ATC TCA TAC TGT TGT	582
Val Met Ile Leu Arg Leu Tyr Ala Met Val Ile Ser Tyr Cys Cys	
5 10 15	
GCC ATC TGC GCT GCC CCC TTC CAG AGC CGG ACC ACA GAT TTG GAT TAT	630
Ala Ile Cys Ala Ala Pro Phe Gln Ser Arg Thr Thr Asp Leu Asp Tyr	
20 25 30	
GGC CCC GAT AAA ACA TCA GAA GCC TCA GAC CGG CAA TCA GTT CCC AAC	678
Gly Pro Asp Lys Thr Ser Glu Ala Ser Asp Arg Gln Ser Val Pro Asn	
35 40 45	
AAC TTC AGT CAT GTC CTG CAA AAT GGG TTC TTT CCA GAT TTG TCA TCC	726
Asn Phe Ser His Val Leu Gln Asn Gly Phe Phe Pro Asp Leu Ser Ser	
50 55 60 65	
ACC TAT TCC AGC ATG GCT GGT AAG GAC TGG AAC CTA TAC TCA CCT AGA	774
Thr Tyr Ser Ser Met Ala Gly Lys Asp Trp Asn Leu Tyr Ser Pro Arg	
70 75 80	
GTG ACT CTT TCA AGT GAG GAG CCT TCT GGA CCT CCA CTA CTT TTC TTG	822
Val Thr Leu Ser Ser Glu Glu Pro Ser Gly Pro Pro Leu Leu Phe Leu	
85 90 95	
TCA GAG GAG ACA GTG GTA CAT CCA GAA CCA GCC AAC AAG ACT TCC CGG	870
Ser Glu Glu Thr Val Val His Pro Glu Pro Ala Asn Lys Thr Ser Arg	
100 105 110	
CTA AAA CGG GCA TCA GGA TCT GAT TCG GTC AGC TTG TCC CGT CGG GGA	918
Leu Lys Arg Ala Ser Gly Ser Asp Ser Val Ser Leu Ser Arg Arg Gly	
115 120 125	
GAG CTC TCT GTG TGT GAC AGT GTC AAC GTC TGG GTT ACC GAT AAA CGT	966
Glu Leu Ser Val Cys Asp Ser Val Asn Val Trp Val Thr Asp Lys Arg	
130 135 140 145	
ACA GCC GTG GAT GAT CGG GGT AAA ATA GTG ACT GTC ATG TCT GAG ATT	1014
Thr Ala Val Asp Asp Arg Gly Lys Ile Val Thr Val Met Ser Glu Ile	
150 155 160	
CAG ACT CTA ACA GGA CCA CTG AAG CAA TAC TTC TTT GAG ACC AAG TGC	1062
Gln Thr Leu Thr Gly Pro Leu Lys Gln Tyr Phe Phe Glu Thr Lys Cys	
165 170 175	
AAT CCA TCA GGC AGC ACC ACT AGA GGA TGC CGA GGT GTA GAC AAA AAG	1110
Asn Pro Ser Gly Ser Thr Thr Arg Gly Cys Arg Gly Val Asp Lys Lys	
180 185 190	
CAA TGG ATA TCT GAG TGC AAA GCA AAA CAG TCT TAT GTG AGG GCT CTG	1158
Gln Trp Ile Ser Glu Cys Lys Ala Lys Gln Ser Tyr Val Arg Ala Leu	
195 200 205	

ACC ATA GAT GCC AAC AAG CTT GTG GGT TGG CGT TGG ATC CGT ATT GAC 1206
Thr Ile Asp Ala Asn Lys Leu Val Gly Trp Arg Trp Ile Arg Ile Asp
210 215 220 225

ACA GCG TGT GTC TGT ACC TTG TTG AGT CGG ACA GGA AGG ACG 1248
Thr Ala Cys Val Cys Thr Leu Leu Ser Arg Thr Gly Arg Thr
230 235

TAAAAGACGA GGGTTAGCAA AATAGAGAGA AGAGGTTGAT CCGTTGACCT GCAG 1302

(2) INFORMAÇÃO PARA SEQ. ID N° 44:

(i) Características da Sequência

- (A) Comprimento: 239 aminoácidos
- (B) Tipo: aminoácido
- (D) Topologia: linear

(ii) Tipo de Molécula: Proteína

(xi) Descrição da Sequência: SEQ ID N°: 44:

Val Asp Val Met Ile Leu Arg Leu Tyr Ala Met Val Ile Ser Tyr Cys
1 5 10 15

Cys Ala Ile Cys Ala Ala Pro Phe Gln Ser Arg Thr Thr Asp Leu Asp
20 25 30

Tyr Gly Pro Asp Lys Thr Ser Glu Ala Ser Asp Arg Gln Ser Val Pro
35 40 45

Asn Asn Phe Ser His Val Leu Gln Asn Gly Phe Phe Pro Asp Leu Ser
50 55 60

Ser Thr Tyr Ser Ser Met Ala Gly Lys Asp Trp Asn Leu Tyr Ser Pro
65 70 75 80

Arg Val Thr Leu Ser Ser Glu Glu Pro Ser Gly Pro Pro Leu Leu Phe
85 90 95

Leu Ser Glu Glu Thr Val Val His Pro Glu Pro Ala Asn Lys Thr Ser
100 105 110

Arg Leu Lys Arg Ala Ser Gly Ser Asp Ser Val Ser Leu Ser Arg Arg
115 120 125

Gly Glu Leu Ser Val Cys Asp Ser Val Asn Val Trp Val Thr Asp Lys
130 135 140

Arg Thr Ala Val Asp Asp Arg Gly Lys Ile Val Thr Val Met Ser Glu
145 150 155 160

Ile Gln Thr Leu Thr Gly Pro Leu Lys Gln Tyr Phe Phe Glu Thr Lys
165 170 175

Cys Asn Pro Ser Gly Ser Thr Thr Arg Gly Cys Arg Gly Val Asp Lys
180 185 190

Lys Gln Trp Ile Ser Glu Cys Lys Ala Lys Gln Ser Tyr Val Arg Ala
 195 200 205
 Leu Thr Ile Asp Ala Asn Lys Leu Val Gly Trp Arg Trp Ile Arg Ile
 210 215 220
 Asp Thr Ala Cys Val Cys Thr Leu Leu Ser Arg Thr Gly Arg Thr
 225 230 235

(2) INFORMAÇÃO PARA SEQ. ID N° 45:

(i) Características da Sequência

- (A) Comprimento: 123 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 45:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 46:

- (i) Características da Sequência
 - (A) Comprimento: 118 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 46:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 47:

- (i) Características da Sequência
 - (A) Comprimento: 119 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 47:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 48:

(i) Características da Sequência

- (A) Comprimento: 119 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 48:

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

Asn Gly Cys Arg Gly Ile Asp Asp Lys His Trp Asn Ser Gln Cys Lys
65 70 75 80
Thr Ser Gln Thr Tyr Val Arg Ala Leu Thr Ser Glu Asn Asn Lys Leu
85 90 95
Val Gly Trp Arg Trp Ile Arg Ile Asp Thr Ser Cys Val Cys Ala Leu
100 105 110
Ser Arg Lys Ile Gly Arg Thr
115

(2) INFORMAÇÃO PARA SEQ. ID N° 49:

(i) Características da Sequência

- (A) Comprimento: 1313 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(ix) Estrutura:

- (A) Nome/Código: CDS
- (B) Localização: 552..1259

(xi) Descrição da Sequência: SEQ ID N°: 49:

CTGCAGGGAA ACAATCATAC TTATGAACAG CAGGGGGAGC CCTCGCCTTA CTCCCAGCC	60
ATGCAGAACT CAAGCAGCTT TGTTTATGCC GATCCCTAAG CAGCCCAGAC CACTGTGAGC	120
ATGTGCACAG TCTTAGTCTT GCAAAGATGT TTAACAAAGT TACAAGATGG TGACCCCTG	180
TAGCCAACTT TGAAAGCATA AATCATTTGT TTGATTAGGC TTGTGGTGCA GTAAGTTCAT	240
GTTTATATTT AGCATACAAA ATACAGCATT TCTAGCCTTA TTCTATTTTA GACTTTACCC	300
TTTAATGCCC AGTTCTGCCC ATTGCCTTAT AGATGTAAAG GTCCCAATAT CACATTGGCA	360
TCCTCGGCTG TTTACAACAA ACATTAAAC TTGTACTTAT ATTTAACATT CTGTTGTTCT	420
TCCAATATTC CATCACAATT AGACCCTAAA AGAATTATAT GTATATAATT TGCATAAATT	480
ATATAATGGC AGCCGTATTC TAATTCTGTT TTTTTTTTTT TTTTTCGAG TGGTCTGAGG	540
TGGATTAACT A ATG ATC CTC CGC CTT TAT GCC ATG GTG ATC TCA TAC TGT	590
Met Ile Leu Arg Leu Tyr Ala Met Val Ile Ser Tyr Cys	
1 5 10	
TGT GCC ATC TGC GCT GCC CCC TTC CAG AGC CGG ACC ACA GAT TTG GAT	638
Cys Ala Ile Cys Ala Ala Pro Phe Gln Ser Arg Thr Thr Asp Leu Asp	
15 20 25	

-107-

TAT GGC CCC GAT AAA ACA TCA GAA GCC TCA GAC CGG CAA TCA GTT CCC Tyr Gly Pro Asp Lys Thr Ser Glu Ala Ser Asp Arg Gln Ser Val Pro 30 35 40 45	686
AAC AAC TTC AGT CAT GTC CTG CAA AAT GGG TTC TTT CCA GAT TTG TCA Asn Asn Phe Ser His Val Leu Gln Asn Gly Phe Phe Pro Asp Leu Ser 50 55 60	734
TCC ACC TAT TCC AGC ATG GCT GGT AAG GAC TGG AAC CTA TAC TCA CCT Ser Thr Tyr Ser Ser Met Ala Gly Lys Asp Trp Asn Leu Tyr Ser Pro 65 70 75	782
AGA GTG ACT CTT TCA AGT GAG GAG CCT TCT GGA CCT CCA CTA CTT TTC Arg Val Thr Leu Ser Ser Glu Glu Pro Ser Gly Pro Pro Leu Leu Phe 80 85 90	830
TTG TCA GAG GAG ACA GTG GTA CAT CCA GAA CCA GCC AAC AAG ACT TCC Leu Ser Glu Glu Thr Val Val His Pro Glu Pro Ala Asn Lys Thr Ser 95 100 105	878
CGG CTA AAA CGG GCA TCA GGA TCT GAT TCG GTC AGC TTG TCC CGT CGG Arg Leu Lys Arg Ala Ser Gly Ser Asp Ser Val Ser Leu Ser Arg Arg 110 115 120 125	926
GGA GAG CTC TCT GTG TGT GAC AGT GTC AAC GTC TGG GTT ACC GAT AAA Gly Glu Leu Ser Val Cys Asp Ser Val Asn Val Trp Val Thr Asp Lys 130 135 140	974
CGT ACA GCC GTG GAT GAT CGG GGT AAA ATA GTG ACT GTC ATG TCT GAG Arg Thr Ala Val Asp Asp Arg Gly Lys Ile Val Thr Val Met Ser Glu 145 150 155	1022
ATT CAG ACT CTA ACA GGA CCA CTG AAG CAA TAC TTC TTT GAG ACC AAG Ile Gln Thr Leu Thr Gly Pro Leu Lys Gln Tyr Phe Phe Glu Thr Lys 160 165 170	1070
TGC AAT CCA TCA GCC AGC ACC ACT AGA GGA TGC CGA GGT GTA GAC AAA Cys Asn Pro Ser Gly Ser Thr Thr Arg Gly Cys Arg Gly Val Asp Lys 175 180 185	1118
AAG CAA TGG ATA TCT GAG TGC AAA GCA AAA CAG TCT TAT GTG AGG GCT Lys Gln Trp Ile Ser Glu Cys Lys Ala Lys Gln Ser Tyr Val Arg Ala 190 195 200 205	1166
CTG ACC ATA GAT GCC AAC AAG CTT GTG GGT TGG CGT TGG ATC CGT ATT Leu Thr Ile Asp Ala Asn Lys Leu Val Gly Trp Arg Trp Ile Arg Ile 210 215 220	1214
GAC ACA GCG TGT GTC TGT ACC TTG TTG AGT CGG ACA GGA AGG ACG Asp Thr Ala Cys Val Cys Thr Leu Leu Ser Arg Thr Gly Arg Thr 225 230 235	1259
TAAAAGACGA GGGTTAGCAA AATAGAGAGA AGAGGTTGAT CCGTTGACCT GCAG	1313

(2) INFORMAÇÃO PARA SEQ. ID Nº 50:

(i) Características da Sequência

(A) Comprimento: 236 aminoácidos

(B) Tipo: aminoácido

(D) Topologia: linear

(ii) Tipo de Molécula: Proteína

(xi) Descrição da Sequência: SEQ ID N°: 50:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 51:

(i) Características da Sequência

- (A) Comprimento: 57 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 51:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 52:

(i) Características da Sequência

- (A) Comprimento: 6 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 52:

Gln	Tyr	Phe	Tyr	Glu	Thr
1				5	

(2) INFORMAÇÃO PARA SEQ. ID N° 53:

(i) Características da Sequência

- (A) Comprimento: 17 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 53:

CARTAYTTYT TYGARAC

17

(2) INFORMAÇÃO PARA SEQ. ID N° 54:

(i) Características da Sequência

- (A) Comprimento: 17 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 54:

CARTAYTTYT AYGARAC

(2) INFORMAÇÃO PARA SEQ. ID N° 55:

(i) Características da Sequência

- (A) Comprimento: 17 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(ix) Estrutura:

- (A) Nome/Código; base modificada
- (B) Localização: 9
- (C) Outra informação: /marcação = N
/nota= "N = I"

(ix) Estrutura:

- (A) Nome/Código; base modificada

- (B) Localização: 12
(C) Outra informação: /marcação = N
/nota= "N = I"

(xi) Descrição da Sequência: SEQ ID N°: 55:

TTRCAYTCNS WNATCCA

(2) INFORMAÇÃO PARA SEQ. ID N° 56:

- (i) Características da Sequência
(A) Comprimento: 17 pares de base
(B) Tipo: ácido nucleico
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida
- (ii) Tipo de Molécula: ADN (genómico)
- (ix) Estrutura:
(A) Nome/Código; base modificada
(B) Localização: 9
(C) Outra informação: /marcação = N
/nota= "N = I"

(xi) Descrição da Sequência: SEQ ID N°: 56:

GAYTGYTTNG CYTTRCA

(2) INFORMAÇÃO PARA SEQ. ID N° 57:

- (i) Características da Sequência
(A) Comprimento: 17 pares de base
(B) Tipo: ácido nucleico
(C) Tipo de cadeia: simples
(D) Topologia: não conhecida
- (ii) Tipo de Molécula: ADN (genómico)

(ix) Estrutura:

- (A) Nome/Código; base modificada
- (B) Localização: 9
- (C) Outra informação: /marcação = N
/nota= "N = I"

(xi) Descrição da Sequência: SEQ ID N°: 57:

CTYTGYYTNG CYTTRCA

(2) INFORMAÇÃO PARA SEQ. ID N° 58:

(i) Características da Sequência

- (A) Comprimento: 17 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(ix) Estrutura:

- (A) Nome/Código; base modificada
- (B) Localização: 6
- (C) Outra informação: /marcação = n
/nota= "N = I"

(ix) Estrutura:

- (A) Nome/Código; base modificada
- (B) Localização: 9
- (C) Outra informação: /marcação = N
/nota= "N = I"

(ix) Estrutura:

- (A) Nome/Código; base modificada
- (B) Localização: 12
- (C) Outra informação: /marcação = N
/nota= "N = I"

(xi) Descrição da Sequência: SEQ ID N°: 58:

GTRTCNATNC KNATCCA

(2) INFORMAÇÃO PARA SEQ. ID N° 59:

(i) Características da Sequência

(A) Comprimento: 17 pares de base

(B) Tipo: ácido nucleico

(C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 59:

CCAAGCTTCT AGAATTC

(2) INFORMAÇÃO PARA SEQ. ID N° 60:

(i) Características da Sequência

(A) Comprimento: 17 pares de base

(B) Tipo: ácido nucleico

(C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 60:

GACTCGAGTC GACATCG

(2) INFORMAÇÃO PARA SEQ. ID N° 61:

(i) Características da Sequência

(A) Comprimento: 126 pares de base

(B) Tipo: ácido nucleico

(C) Tipo de cadeia: dupla

(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genômico)

(ix) Estrutura:

(A) Nome/Código: CDS

(B) Localização: 1..126

(xi) Descrição da Sequência: SEQ ID N°: 61:

CAG TAT TTT TAC GAG ACG CGC TGC AAG GCC GAA AGC GCT GGG GAA GGT	48
Gln Tyr Phe Tyr Glu Thr Arg Cys Lys Ala Glu Ser Ala Gly Glu Gly	
1 5 10 15	
GGC CCA GGT GTG GGC GGA GGG GGC TGT CGC GGC GTG GAT CGG AGG CAC	96
Gly Pro Gly Val Gly Gly Gly Gly Cys Arg Gly Val Asp Arg Arg His	
20 25 30	
TGG CTC TCA GAA TGT AAA GCC AAA CAA TCG	126
Trp Leu Ser Glu Cys Lys Ala Lys Gln Ser	
35 40	

(2) INFORMAÇÃO PARA SEQ. ID N° 62:

(i) Características da Sequência

(A) Comprimento: 42 aminoácidos

(B) Tipo: aminoácido

(D) Topologia: linear

(ii) Tipo de Molécula: Proteína

(xi) Descrição da Sequência: SEQ ID N°: 62:

Gln Tyr Phe Tyr Glu Thr Arg Cys Lys Ala Glu Ser Ala Gly Glu Gly	
1 5 10 15	
Gly Pro Gly Val Gly Gly Gly Gly Cys Arg Gly Val Asp Arg Arg His	
20 25 30	
Trp Leu Ser Glu Cys Lys Ala Lys Gln Ser	
35 40	

(2) INFORMAÇÃO PARA SEQ. ID N° 63:

(i) Características da Sequência

- (A) Comprimento: 126 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genômico)

(ix) Estrutura:

- (A) Nome/Código; CDS
- (B) Localização: 1..126

(xi) Descrição da Sequência: SEQ ID N°: 63:

CAG TAT TTT TAC GAA ACC CGC TGC AAG GCT GAT AAC GCT GAG GAA GGT	48
Gln Tyr Phe Tyr Glu Thr Arg Cys Lys Ala Asp Asn Ala Glu Glu Gly	
1 5 10 15	
GGC CCG GGG GCA GGT GGA GGG GGC TGC CGG GGA GTG GAC AGG AGG CAC	96
Gly Pro Gly Ala Gly Gly Gly Gly Cys Arg Gly Val Asp Arg Arg His	
20 25 30	
TGG GTA TCT GAG TGT AAA GCC AAA CAA TCG	126
Trp Val Ser Glu Cys Lys Ala Lys Gln Ser	
35 40	

(2) INFORMAÇÃO PARA SEQ. ID N° 64:

(i) Características da Sequência

- (A) Comprimento: 42 aminoácidos
- (B) Tipo: aminoácido
- (D) Topologia: linear

(ii) Tipo de Molécula: Proteína

(xi) Descrição da Sequência: SEQ ID N°: 64:

Gln Tyr Phe Tyr Glu Thr Arg Cys Lys Ala Asp Asn Ala Glu Glu Gly	
1 5 10 15	
Gly Pro Gly Ala Gly Gly Gly Gly Cys Arg Gly Val Asp Arg Arg His	
20 25 30	
Trp Val Ser Glu Cys Lys Ala Lys Gln Ser	
35 40	

(i) Características da Sequência

- (A) Comprimento: 35 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(xi) Descrição da Sequência: SEQ ID N°: 65:

Gln Tyr Phe Phe Glu Thr Lys Cys Asn Pro Ser Gly Ser Thr Thr Arg
1 5 10 15
Gly Cys Arg Gly Val Asp Lys Lys Gln Trp Ile Ser Glu Cys Lys Ala
20 25 30
Lys Gln Ser
35

(2) INFORMAÇÃO PARA SEQ. ID N° 66:

(i) Características da Sequência

- (A) Comprimento: 35 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 66:

Gln Tyr Phe Phe Glu Thr Lys Cys Arg Ala Pro Asn Pro Val Glu Ser
1 5 10 15
Gly Cys Arg Gly Ile Asp Ser Lys His Trp Asn Ser Tyr Cys Thr Thr
20 25 30
Thr His Thr
35

(i) Características da Sequência

- (A) Comprimento: 35 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 67:

Gln Tyr Phe Tyr Glu Thr Lys Cys Asn Pro Met Gly Tyr Thr Lys Glu
1 5 10 15
Gly Cys Arg Gly Ile Asp Lys Arg His Trp Asn Ser Gln Cys Arg Thr
20 25 30
Thr Gln Ser
35

(2) INFORMAÇÃO PARA SEQ. ID N° 68:

(i) Características da Sequência

- (A) Comprimento: 35 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 68:

Gln Tyr Phe Tyr Glu Thr Arg Cys Lys Glu Ala Arg Pro Val Lys Asn
1 5 10 15
Gly Cys Arg Gly Ile Asp Asp Lys His Trp Asn Ser Gln Cys Lys Thr
20 25 30
Ser Gln Thr
35

(2) INFORMAÇÃO PARA SEQ. ID N° 69:

(i) Características da Sequência

- (A) Comprimento: 192 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(ix) Estrutura:

- (A) Nome/Código: CDS
- (B) Localização: 1..192

(xi) Descrição da Sequência: SEQ ID N°: 69:

CAA TAT TTT TTC GAG ACC CGC TGC AAG GCT GAT AAC GCT GAG GAA GGT	48
Gln Tyr Phe Phe Glu Thr Arg Cys Lys Ala Asp Asn Ala Glu Glu Gly	
1 5 10 15	
GGT CCG GGG GCA GGT GGA GGG GGC TGC CGG GGA GTG GAC AGG AGG CAC	96
Gly Pro Gly Ala Gly Gly Gly Gly Cys Arg Gly Val Asp Arg Arg His	
20 25 30	
TGG CTA TCT GAG TGC AAG GCC AAG CAG TCC TAT GTG CGG GCA TTG ACC	144
Trp Val Ser Glu Cys Lys Ala Lys Gln Ser Tyr Val Arg Ala Leu Thr	
35 40 45	
GCT GAT GCC CAG GGC CGT GTG GGC TGG CGA TGG ATC CGC ATC GAT ACG	192
Ala Asp Ala Gln Gly Arg Val Gly Trp Arg Trp Ile Arg Ile Asp Thr	
50 55 60	

(2) INFORMAÇÃO PARA SEQ. ID N° 70:

(i) Características da Sequência

- (A) Comprimento: 64 aminoácidos
- (B) Tipo: aminoácido
- (D) Topologia: linear

(ii) Tipo de Molécula: Proteína

(xi) Descrição da Sequência: SEQ ID N°: 70:

Gln Tyr Phe Phe Glu Thr Arg Cys Lys Ala Asp Asn Ala Glu Glu Gly	
1 5 10 15	
Gly Pro Gly Ala Gly Gly Gly Gly Cys Arg Gly Val Asp Arg Arg His	
20 25 30	

Trp Val Ser Glu Cys Lys Ala Lys Gln Ser Tyr Val Arg Ala Leu Thr
 35 40 45

Ala Asp Ala Gln Gly Arg Val Gly Trp Arg Trp Ile Arg Ile Asp Thr
 50 55 60

(2) INFORMAÇÃO PARA SEQ. ID Nº 71:

- (i) Características da Sequência
 - (A) Comprimento: 35 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida
- (ii) Tipo de Molécula: ADN (genómico)
- (ix) Estrutura:
 - (A) Nome/Código: CDS
 - (B) Localização: 18..35
- (xi) Descrição da Sequência: SEQ ID Nº: 71:

GACTCGAGTC GACATCG GAA ACC CGC TGC AAG GCT
 Glu Thr Arg Cys Lys Ala
 1 5

(2) INFORMAÇÃO PARA SEQ. ID Nº 72:

- (i) Características da Sequência
 - (A) Comprimento: 6 aminoácidos
 - (B) Tipo: aminoácido
 - (D) Topologia: linear
- (ii) Tipo de Molécula: Proteína
- (xi) Descrição da Sequência: SEQ ID Nº: 72:

Glu Thr Arg Cys Lys Ala
 1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 73:

- (i) Características da Sequência
 - (A) Comprimento: 35 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: dupla
 - (D) Topologia: não conhecida
- (ii) Tipo de Molécula: ADN (genómico)
- (ix) Estrutura:
 - (A) Nome/Código: CDS
 - (B) Localização: 18..35
- (xi) Descrição da Sequência: SEQ ID N°: 73:

GACTCGAGTC GACATCG GAT AAC GCT GAG GAA GGT
Asp Asn Ala Glu Glu Gly
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 74:

- (i) Características da Sequência
 - (A) Comprimento: 6 aminoácidos
 - (B) Tipo: aminoácido
 - (D) Topologia: linear
- (ii) Tipo de Molécula: Proteína
- (xi) Descrição da Sequência: SEQ ID N°: 74:

Asp Asn Ala Glu Glu Gly
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 75:

- (i) Características da Sequência
 - (A) Comprimento: 1404 pares de base
 - (B) Tipo: ácido nucleico

- (C) Tipo de cadeia: dupla
(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(ix) Estrutura:

- (A) Nome/Código: CDS
(B) Localização: 460..1104

(xi) Descrição da Sequência: SEQ ID N°: 75:

CTGTGACCCC AGGTGGCACC CGAGTGGTGC ACTCTCTGCT CACTGCAACC TCGGCCTCCT	60
GGGTTTCGAGT GATTCTCCTA CCTCAGCCTA CTGAGTAGCT GGGATTACAG GCGTGCAGCA	120
CTATGCCCCGG TTAATTTTGG TATTTTGGT AGAGATGAGG TTTCACAATG TTGACCAGCT	180
GCTCTGGAAC TCCTGACCTC AAGTCATCCA CCTGCCTCAG CCTCCCAGAG TGCTGGGATT	240
AGAGGTGTGG GGCACAGTGC CTGGCCTGTA GTAGTTGAAT ATTTATTATT AATCTACAAG	300
TTGCGCATTA CGCAAGCCCT AGATATAGGG TCCCCCAAAC TTCTAGAACA AGGGCTTCCC	360
CACAATCCTG GCAGGCAAGC CTCCCCTGGG GTTCCCAACT TCTTTCCCCA CTGAAGTTTT	420
TACCCCTTC TCTAATCCCA GCCTCCCTCT TTCTGTCTC CAG GTG CTC CGA GAG	474
	Gln Val Leu Arg Glu 1 5
ATG CTC CCT CTC CCC TCA TGC TCC CTC CCC ATC CTC CTC CTT TTC CTC	522
Met Leu Pro Leu Pro Ser Cys Ser Leu Pro Ile Leu Leu Leu Phe Leu	10 15 20
CTC CCC AGT GTG CCA ATT GAG TCC CAA CCC CCA CCC TCA ACA TTG CCC	570
Leu Pro Ser Val Pro Ile Glu Ser Gln Pro Pro Pro Ser Thr Leu Pro	25 30 35
CCT TTT CTG GCC CCT GAG TGG GAC CTT CTC TCC CCC CGA GTA GTC CTG	618
Pro Phe Leu Ala Pro Glu Trp Asp Leu Leu Ser Pro Arg Val Val Leu	40 45 50
TCT AGG GGT GCC CCT GCT GGG CCC CCT CTG CTC TTC CTG CTG GAG GCT	666
Ser Arg Gly Ala Pro Ala Gly Pro Pro Leu Leu Phe Leu Leu Glu Ala	55 60 65
GGG GCC TTT CGG GAG TCA GCA GGT GCC CCG GCC AAC CGC AGC CGG CGT	714
Gly Ala Phe Arg Glu Ser Ala Gly Ala Pro Ala Asn Arg Ser Arg Arg	70 75 80 85
GGG GTG AGC GAA ACT GCA CCA GCG AGT CGT CGG GGT GAG CTG GCT GTG	762
Gly Val Ser Glu Thr Ala Pro Ala Ser Arg Arg Gly Glu Leu Ala Val	90 95 100
TGC GAT GCA GTC AGT GGC TGG GTG ACA GAC CGC CGG ACC GCT GTG GAC	810
Cys Asp Ala Val Ser Gly Trp Val Thr Asp Arg Arg Thr Ala Val Asp	105 110 115

TTG CGT GGG CGC GAG GTG GAG GTG TTG GGC GAG GTG CCT GCA GCT GGC	858
Leu Arg Gly Arg Glu Val Glu Val Leu Gly Glu Val Pro Ala Ala Gly	
120 125 130	
GGC AGT CCC CTC CGC CAG TAC TTC TTT GAA ACC CGC TGC AAG GCT GAT	906
Gly Ser Pro Leu Arg Gln Tyr Phe Phe Glu Thr Arg Cys Lys Ala Asp	
135 140 145	
AAC GCT GAG GAA GGT GGT CCG GGG GCA GGT GGA GGG GGC TGC CGG GGA	954
Asn Ala Glu Glu Gly Gly Pro Gly Ala Gly Gly Gly Gly Cys Arg Gly	
150 155 160 165	
GTG GAC AGG AGG CAC TGG GTA TCT GAG TGC AAG GCC AAG CAG TCC TAT	1002
Val Asp Arg Arg His Trp Val Ser Glu Cys Lys Ala Lys Gln Ser Tyr	
170 175 180	
GTG CGG GCA TTG ACC GCT GAT GCC CAG GGC CGT GTG GGC TGG CGA TGG	1050
Val Arg Ala Leu Thr Ala Asp Ala Gln Gly Arg Val Gly Trp Arg Trp	
185 190 195	
ATT CGA ATT GAC ACT GCC TGC GTC TGC ACA CTC CTC AGC CGG ACT GGC	1098
Ile Arg Ile Asp Thr Ala Cys Val Cys Thr Leu Leu Ser Arg Thr Gly	
200 205 210	
CGG GCC TGAGACCCAT GCCCAGGAAA ATAACAGAGC TGGATGCTGA GAGACCTCAG	1154
Arg Ala	
215	
GGATGGCCCA GCTGATCTAA GGACCCCACT TTGGGAACTC ATCAAATAAT CACAAAATCA	1214
CAATTCTCTG ATTTGGAGCT CAATCTCTGC AGGATGGGTG AAACCACATG GGGTTTGGGA	1274
GGTTGAATAG GAGTTCTCCT GGAGCAACTT GAGGGTAATA ATGATGATGA TATAATAATA	1334
ATAGCCACTA TTTACTGAGT GTTTACTGTT TCTTATCCCT AATACATAAC TCCTCAGATC	1394
AACTCTCATG	1404

(2) INFORMAÇÃO PARA SEQ. ID Nº 76:

(i) Características da Sequência

- (A) Comprimento: 215 aminoácidos
- (B) Tipo: aminoácido
- (D) Topologia: linear

(ii) Tipo de Molécula: Proteína

(xi) Descrição da Sequência: SEQ ID Nº: 76:

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

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

(2) INFORMAÇÃO PARA SEQ. ID N° 77:

(i) Características da Sequência

- (A) Comprimento: 214 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 77:

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

Thr Ala Val Asp Leu Arg Gly Arg Glu Val Glu Val Leu Gly Glu Val
115 120 125

Pro Ala Ala Gly Gly Ser Pro Leu Arg Gln Tyr Phe Phe Glu Thr Arg
130 135 140

Cys Lys Ala Asp Asn Ala Glu Glu Gly Gly Pro Gly Ala Gly Gly Gly
145 150 155 160

Gly Cys Arg Gly Val Asp Arg Arg His Trp Val Ser Glu Cys Lys Ala
165 170 175

Lys Gln Ser Tyr Val Arg Ala Leu Thr Ala Asp Ala Gln Gly Arg Val
180 185 190

Gly Trp Arg Trp Ile Arg Ile Asp Thr Ala Cys Val Cys Thr Leu Leu
195 200 205

Ser Thr Arg Gly Arg Ala
210

(2) INFORMAÇÃO PARA SEQ. ID N° 78:

(i) Características da Sequência

- (A) Comprimento: 176 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 78:

Ser Ser Thr Tyr Ser Ser Met Ala Gly Lys Asp Trp Asn Leu Tyr Ser
1 5 10 15

Pro Arg Val Thr Leu Ser Ser Glu Glu Pro Ser Gly Pro Pro Leu Leu
20 25 30

Phe Leu Ser Glu Glu Thr Val Val His Pro Glu Pro Ala Asn Lys Thr
35 40 45

Ser Arg Leu Lys Arg Ala Ser Gly Ser Asp Ser Val Ser Leu Ser Arg
50 55 60

Arg Gly Glu Leu Ser Val Cys Asp Ser Val Asn Val Trp Val Thr Asp
65 70 75 80

Lys Arg Thr Ala Val Asp Asp Arg Gly Lys Ile Val Thr Val Met Ser
85 90 95

Glu Ile Gln Thr Leu Thr Gly Pro Leu Lys Gln Tyr Phe Phe Glu Thr
100 105 110

Lys Cys Asn Pro Ser Gly Ser Thr Thr Arg Gly Cys Arg Gly Val Asp
115 120 125

Lys Lys Gln Trp Ile Ser Glu Cys Lys Ala Lys Gln Ser Tyr Val Arg
130 135 140

Ala Leu Thr Ile Asp Ala Asn Lys Leu Val Gly Trp Arg Trp Ile Arg
145 150 155 160

Ile Asp Thr Ala Cys Val Cys Thr Leu Leu Ser Arg Thr Gly Arg Thr
165 170 175

(2) INFORMAÇÃO PARA SEQ. ID N° 79:

(i) Características da Sequência

- (A) Comprimento: 177 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 79:

Glu Phe Gln Pro Met Ile Ala Thr Asp Thr Glu Leu Leu Arg Gln Gln
1 5 10 15
Arg Arg Tyr Asn Ser Pro Arg Val Leu Leu Ser Asp Ser Thr Pro Leu
20 25 30
Glu Pro Pro Pro Leu Tyr Leu Met Glu Asp Tyr Val Gly Asn Pro Val
35 40 45
Val Ala Asn Arg Thr Ser Pro Arg Arg Lys Tyr Ala Glu His Lys Ser
50 55 60
His Arg Gly Glu Tyr Ser Val Cys Asp Ser Glu Ser Leu Trp Val Thr
65 70 75 80
Asp Lys Ser Ser Ala Ile Asp Ile Arg Gly His Gln Val Thr Val Leu
85 90 95
Gly Glu Ile Lys Thr Gly Asn Ser Pro Val Lys Gln Tyr Phe Tyr Glu
100 105 110
Thr Arg Cys Lys Glu Ala Arg Pro Val Lys Asn Gly Cys Arg Gly Ile
115 120 125
Asp Asp Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr Val
130 135 140
Arg Ala Leu Thr Ser Glu Asn Asn Lys Leu Val Gly Trp Arg Trp Ile
145 150 155 160
Arg Ile Asp Thr Ser Cys Val Cys Ala Leu Ser Arg Lys Ile Gly Arg
165 170 175
Thr

(2) INFORMAÇÃO PARA SEQ. ID N° 80:

(i) Características da Sequência

- (A) Comprimento: 175 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 80:

Gln Pro Val Ile Ala Met Asp Thr Glu Leu Leu Arg Gln Gln Arg Arg
1 5 10 15
Tyr Asn Ser Pro Arg Val Leu Leu Ser Asp Thr Thr Pro Leu Glu Pro
20 25 30
Pro Pro Leu Tyr Leu Met Glu Asp Tyr Val Gly Ser Pro Val Val Ala
35 40 45
Asn Arg Thr Ser Arg Arg Lys Arg Tyr Ala Glu His Lys Ser His Arg
50 55 60
Gly Glu Tyr Ser Val Cys Asp Ser Glu Ser Leu Trp Val Thr Asp Lys
65 70 75 80
Ser Ser Ala Ile Asp Ile Arg Gly His Gln Val Thr Val Leu Gly Glu
85 90 95
Ile Lys Thr Gly Asn Ser Pro Val Lys Gln Tyr Phe Tyr Glu Thr Arg
100 105 110
Cys Lys Glu Ala Arg Pro Val Lys Asn Gly Gly Arg Gly Ile Asp Asp
115 120 125
Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr Val Arg Ala
130 135 140
Leu Thr Ser Glu Asn Asn Lys Leu Val Gly Trp Arg Trp Ile Arg Ile
145 150 155 160
Asp Thr Ser Cys Val Cys Ala Leu Ser Arg Lys Ile Gly Arg Thr
165 170 175

(2) INFORMAÇÃO PARA SEQ. ID N° 81:

(i) Características da Sequência

- (A) Comprimento: 178 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 81:

Glu Phe Gln Pro Met Ile Ala Thr Asp Thr Glu Leu Leu Arg Gln Gln
1 5 10 15
Arg Arg Tyr Asn Ser Pro Arg Val Leu Ser Asp Ser Thr Pro Leu
20 25 30
Glu Pro Pro Pro Leu Tyr Leu Met Glu Asp Tyr Val Gly Asn Pro Val
35 40 45
Val Thr Asn Arg Thr Ser Pro Arg Arg Lys Arg Tyr Ala Glu His Lys
50 55 60
Ser His Arg Gly Glu Tyr Ser Val Cys Asp Ser Glu Ser Leu Trp Val
65 70 75 80
Thr Asp Lys Ser Ser Ala Ile Asp Ile Arg Gly His Gln Val Thr Val
85 90 95
Leu Gly Glu Ile Lys Thr Gly Asn Ser Pro Val Lys Gln Tyr Phe Tyr
100 105 110
Glu Thr Arg Cys Lys Glu Ala Arg Pro Val Lys Asn Gly Gly Arg Gly
115 120 125
Ile Asp Asp Lys His Trp Asn Ser Gln Cys Lys Thr Ser Gln Thr Tyr
130 135 140
Val Arg Ala Leu Thr Ser Glu Asn Asn Lys Leu Val Gly Trp Arg Trp
145 150 155 160
Ile Arg Ile Asp Thr Ser Cys Val Cys Ala Leu Ser Arg Lys Ile Gly
165 170 175
Arg Thr

(2) INFORMAÇÃO PARA SEQ. ID N° 82:

(i) Características da Sequência

- (A) Comprimento: 160 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 82:

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

(2) INFORMAÇÃO PARA SEQ. ID Nº 83:

(i) Características da Sequência

- (A) Comprimento: 160 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID Nº: 83:

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

				85						90					95				
Thr	Lys	Cys	Asn	Pro	Met	Gly	Tyr	Thr	Lys	Glu	Gly	Gly	Arg	Asp	Ile				
			100					105					110						
Asp	Lys	Arg	His	Trp	Asn	Ser	Gln	Cys	Arg	Thr	Thr	Gln	Ser	Tyr	Val				
		115					120					125							
Arg	Ala	Leu	Thr	Met	Asp	Ser	Lys	Lys	Arg	Ile	Gly	Trp	Arg	Phe	Ile				
		130				135					140								
Arg	Ile	Asp	Thr	Ser	Cys	Val	Cys	Thr	Leu	Thr	Ile	Lys	Arg	Gly	Arg				
145					150					155					160				

(2) INFORMAÇÃO PARA SEQ. ID N° 84:

(i) Características da Sequência

- (A) Comprimento: 160 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 84:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 85:

(i) Características da Sequência

- (A) Comprimento: 160 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 85:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 86:

(i) Características da Sequência

- (A) Comprimento: 180 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 86:

Ala Ala Arg Val Thr Gly Gln Thr Arg Asn Ile Thr Val Asp Pro Arg
1 5 10 15
Leu Phe Lys Lys Arg Arg Leu His Ser Pro Arg Val Leu Phe Ser Thr
20 25 30
Gln Pro Pro Pro Thr Ser Ser Asp Thr Leu Asp Leu Asp Phe Gln Ala
35 40 45
His Gly Thr Ile Pro Phe Asn Arg Thr His Arg Ser Lys Arg Ser Ser
50 55 60
Ser His Pro Ile Phe His Arg Gly Glu Phe Ser Val Cys Asp Ser Val
65 70 75 80
Ser Val Trp Val Gly Asp Lys Thr Thr Ala Thr Asp Ile Lys Gly Lys
85 90 95
Glu Val Met Val Leu Gly Glu Val Asn Ile Asn Asn Ser Val Phe Lys
100 105 110
Gln Tyr Phe Phe Glu Thr Lys Cys Arg Asp Pro Asn Pro Val Asp Ser
115 120 125
Gly Gly Arg Asp Ile Asp Ser Lys His Trp Asn Ser Tyr Cys Thr Thr
130 135 140
Thr His Thr Phe Val Lys Ala Leu Thr Thr Asp Glu Lys Gln Ala Ala
145 150 155 160
Trp Arg Phe Ile Arg Ile Asp Thr Ser Cys Val Cys Val Leu Ser Arg
165 170 175
Lys Ala Thr Arg
180

(2) INFORMAÇÃO PARA SEQ. ID N° 87:

- (i) Características da Sequência
- (A) Comprimento: 180 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 87:

Ala Ala Arg Val Thr Gly Gln Thr Arg Asn Ile Thr Val Asp Pro Lys
 1 5 10 15
 Leu Phe Lys Lys Arg Arg Leu Arg Ser Pro Arg Val Leu Phe Ser Thr
 20 25 30
 Gln Pro Pro Pro Thr Ser Ser Asp Thr Leu Asp Leu Asp Phe Gln Ala
 35 40 45
 His Gly Thr Ile Ser Phe Asn Arg Thr His Arg Ser Lys Arg Ser Ser
 50 55 60
 Thr His Pro Val Phe His Met Gly Glu Phe Ser Val Cys Asp Ser Val
 65 70 75 80
 Ser Val Trp Val Gly Asp Lys Thr Thr Ala Thr Asp Ile Lys Gly Lys
 85 90 95
 Glu Val Thr Val Leu Gly Glu Val Asn Ile Asn Asn Ser Val Phe Lys
 100 105 110
 Gln Tyr Phe Phe Glu Thr Lys Cys Arg Ala Pro Asn Pro Val Glu Ser
 115 120 125
 Gly Gly Arg Asp Ile Asp Ser Lys His Trp Asn Ser Tyr Cys Thr Thr
 130 135 140
 Thr His Thr Phe Val Lys Ala Leu Thr Thr Asp Asp Lys Gln Ala Ala
 145 150 155 160
 Trp Arg Phe Ile Arg Ile Asp Thr Ser Cys Val Cys Val Leu Ser Arg
 165 170 175
 Lys Ala Ala Arg
 180

(2) INFORMAÇÃO PARA SEQ. ID N° 88:

(i) Características da Sequência

- (A) Comprimento: 179 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 88:

Ala Ala Arg Val Ala Gly Gln Thr Arg Asn Ile Thr Val Asp Pro Arg
1 5 10 15
Phe Lys Lys Arg Arg Leu Arg Ser Pro Arg Val Leu Phe Ser Thr Gln
20 25 30
Pro Pro Arg Glu Ala Asp Thr Thr Gln Asp Leu Asp Phe Glu Val Gly
35 40 45
Gly Ala Ala Pro Phe Asn Arg Thr His Arg Ser Lys Arg Ser Ser Ser
50 55 60
His Pro Ile Phe His Arg Gly Glu Phe Ser Val Cys Asp Ser Val Ser
65 70 75 80
Val Trp Val Gly Asp Lys Thr Thr Ala Thr Asp Ile Lys Gly Lys Glu
85 90 95
Val Met Val Leu Gly Glu Val Asn Ile Asn Asn Ser Val Phe Lys Gln
100 105 110
Tyr Phe Phe Glu Thr Lys Cys Arg Asp Pro Asn Pro Val Asp Ser Gly
115 120 125
Gly Arg Asp Ile Asp Ser Lys His Trp Asn Ser Tyr Cys Thr Thr Thr
130 135 140
His Thr Phe Val Lys Ala Leu Thr Met Asp Gly Lys Gln Ala Ala Trp
145 150 155 160
Arg Phe Ile Arg Ile Asp Thr Ser Cys Val Cys Val Leu Ser Arg Lys
165 170 175
Ala Val Arg

(2) INFORMAÇÃO PARA SEQ. ID N° 89:

(i) Características da Sequência

- (A) Comprimento: 163 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 89:

Phe Phe Lys Lys Lys Arg Phe Arg Ser Ser Arg Val Leu Phe Ser Thr
1 5 10 15
Gln Pro Pro Pro Glu Ser Arg Lys Gly Gln Ser Thr Gly Phe Leu Ser
20 25 30

Ser Ala Val Ser Leu Asn Arg Thr Ala Arg Thr Lys Arg Thr Ala His
35 40 45

Pro Val Leu His Arg Gly Glu Phe Ser Val Cys Asp Ser Val Ser Met
50 55 60

Trp Val Gly Asp Lys Thr Thr Ala Thr Asp Ile Lys Gly Lys Glu Val
65 70 75 80

Thr Val Leu Gly Glu Val Asn Ile Asn Asn Asn Val Phe Lys Gln Tyr
85 90 95

Phe Phe Glu Thr Lys Cys Arg Asp Pro Arg Pro Val Ser Ser Gly Gly
100 105 110

Arg Asp Ile Asp Ala Lys His Trp Asn Ser Tyr Cys Thr Thr Thr His
115 120 125

Thr Phe Val Lys Ala Leu Thr Met Glu Gly Lys Gln Ala Ala Trp Arg
130 135 140

Phe Ile Arg Ile Asp Thr Ser Cys Val Cys Val Leu Ser Arg Lys Ser
145 150 155 160

Gly Arg Pro

(2) INFORMAÇÃO PARA SEQ. ID N° 90:

(i) Características da Sequência

- (A) Comprimento: 124 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 90:

His Arg Ser Lys Arg Ser Ser Glu Ser His Pro Val Phe His Arg Gly
1 5 10 15

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

Thr Ala Thr Asp Ile Lys Gly Lys Glu Val Met Val Leu Gly Glu Val
35 40 45

Asn Ile Asn Asn Ser Val Phe Lys Gln Tyr Phe Phe Glu Thr Lys Cys
50 55 60

Arg Asp Pro Asn Pro Val Asp Ser Gly Gly Arg Asp Ile Asp Ala Lys
65 70 75 80

His Trp Asn Ser Tyr Cys Thr Thr Thr His Thr Phe Val Lys Ala Leu
85 90 95

Thr Met Asp Gly Lys Gln Ala Ala Trp Arg Phe Ile Arg Ile Asp Thr
100 105 110

Ser Cys Val Cys Val Leu Ser Arg Lys Thr Gly Gln
115 120

(2) INFORMAÇÃO PARA SEQ. ID N° 91:

(i) Características da Sequência

- (A) Comprimento: 164 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 91:

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

(2) INFORMAÇÃO PARA SEQ. ID N° 92:

(i) Características da Sequência

- (A) Comprimento: 196 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 92:

```

Pro Ala Gly Ser Ser Pro Asp Pro Ser Ser Pro Val Val Asp Pro Lys
1          5          10          15
Leu Phe Ser Lys Arg His Tyr Pro Ser Pro Arg Val Val Phe Ser Glu
20          25          30
Val Ile Pro Ser His Asp Val Leu Asp Gly Glu Gly Tyr Asp Phe Glu
35          40          45
Arg Val Arg Gly Leu Arg Val Arg Arg Lys Ala Val Ser His Thr Met
50          55          60
His Arg Gly Glu Tyr Ser Val Cys Asp Ser Ile Asn Thr Trp Val Asn
65          70          75          80
Thr Lys Arg Ala Thr Asp Met Ser Gly Asn Glu Val Thr Val Leu Ser
85          90          95
His Val Thr Val Asn Asn Lys Val Lys Lys Gln Leu Phe Tyr Glu Thr
100         105         110
Thr Cys Arg Ser Pro Thr His Arg Ser Ser Gly Ile Val Ile Gly Gly
115         120         125
Arg Ser Gly Gly Arg Gly Gly Ser Gln Gly Ser Lys Thr Gly Asn Ser
130         135         140
Gly Gly Arg Asp Ile Asp Ser Arg Tyr Trp Asn Ser His Cys Thr Asn
145         150         155         160
Thr Asp Ile Tyr Val Ser Ala Leu Thr Val Phe Lys Glu Gln Thr Ala
165         170         175
Trp Arg Phe Ile Arg Ile Asn Ala Ser Cys Val Cys Val Ser Arg Thr
180         185         190
Asn Ser Trp Ser
195

```

(2) INFORMAÇÃO PARA SEQ. ID N° 93:

(i) Características da Sequência

- (A) Comprimento: 225 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(ix) Estrutura:

- (A) Nome/Código: CDS
- (B) Localização: 1..225

(xi) Descrição da Sequência: SEQ ID N°: 93:

CGA	GTG	GTC	CTG	TCT	AGG	GGT	GCC	GCT	GCC	GGG	CCC	CCT	CTG	GTC	TTC	48
Arg	Val	Val	Leu	Ser	Arg	Gly	Ala	Ala	Ala	Gly	Pro	Pro	Leu	Val	Phe	
1				5					10					15		
CTG	CTG	GAG	ACT	GGA	GCC	TTT	CGG	GAG	TCA	GCA	GGC	GCC	CGG	GCC	AAC	96
Leu	Leu	Glu	Thr	Gly	Ala	Phe	Arg	Glu	Ser	Ala	Gly	Ala	Arg	Ala	Asn	
			20					25					30			
CGC	AGC	CAG	CGA	GGG	GTG	AGC	GAT	ACT	TCA	CCG	GCG	AGT	CAT	CAG	GGT	144
Arg	Ser	Gln	Arg	Gly	Val	Ser	Asp	Thr	Ser	Pro	Ala	Ser	His	Gln	Gly	
		35					40					45				
GAG	CTG	GCC	CTG	TGC	GAT	GCA	GTC	AGT	GTC	TGG	GTG	ACA	GAC	CCC	TGG	192
Glu	Leu	Ala	Leu	Cys	Asp	Ala	Val	Ser	Val	Trp	Val	Thr	Asp	Pro	Trp	
	50					55					60					
ACT	GCT	GTG	GAC	TTG	GGT	GTG	CTC	GAG	GTC	GAG						225
Thr	Ala	Val	Asp	Leu	Gly	Val	Leu	Glu	Val	Glu						
65					70					75						

(2) INFORMAÇÃO PARA SEQ. ID N° 94:

(i) Características da Sequência

(A) Comprimento: 75 aminoácidos

(B) Tipo: aminoácido

(D) Topologia: linear

(ii) Tipo de Molécula: Proteína

(xi) Descrição da Sequência: SEQ ID N°: 94:

Arg	Val	Val	Leu	Ser	Arg	Gly	Ala	Ala	Ala	Gly	Pro	Pro	Leu	Val	Phe
1				5					10					15	
Leu	Leu	Glu	Thr	Gly	Ala	Phe	Arg	Glu	Ser	Ala	Gly	Ala	Arg	Ala	Asn
			20					25					30		
Arg	Ser	Gln	Arg	Gly	Val	Ser	Asp	Thr	Ser	Pro	Ala	Ser	His	Gln	Gly
		35					40					45			
Glu	Leu	Ala	Leu	Cys	Asp	Ala	Val	Ser	Val	Trp	Val	Thr	Asp	Pro	Trp
	50					55					60				
Thr	Ala	Val	Asp	Leu	Gly	Val	Leu	Glu	Val	Glu					
65					70					75					

(2) INFORMAÇÃO PARA SEQ. ID N°95:

(i) Características da Sequência

- (A) Comprimento: 53 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADNc

(xi) Descrição da Sequência: SEQ ID N°: 95:

CCCACAAGCT TGTGGCATC TATGGTCAGA GCCCTCACAT AAGACTGTTT TGC
53

(2) INFORMAÇÃO PARA SEQ. ID N° 96:

- (i) Características da Sequência
 - (A) Comprimento: 10 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 96:

Lys	Cys	Asn	Pro	Ser	Gly	Ser	Thr	Thr	Arg
1			5						10

(2) INFORMAÇÃO PARA SEQ. ID N° 97:

- (i) Características da Sequência
 - (A) Comprimento: 7 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 97:

Arg Gly Cys Arg Gly Val Asp
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 98:

- (i) Características da Sequência
 - (A) Comprimento: 5 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 98:

Lys Gln Trp Ile Ser
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 99:

- (i) Características da Sequência
 - (A) Comprimento: 6 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 99:

Lys Gln Ser Tyr Val Arg
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 100:

- (i) Características da Sequência
 - (A) Comprimento: 7 aminoácidos

- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 100:

Gly Pro Gly Xaa Gly Gly Gly
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 101:

- (i) Características da Sequência
 - (A) Comprimento: 7 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 101:

Gly Pro Gly Val Gly Gly Gly
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 102:

- (i) Características da Sequência
 - (A) Comprimento: 7 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 102:

Gly Pro Gly Ala Gly Gly Gly
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 103:

(i) Características da Sequência

(A) Comprimento: 5 aminoácidos

(B) Tipo: aminoácido

(C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 103:

Glu Ser Ala Gly Glu
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 104:

(i) Características da Sequência

(A) Comprimento: 5 aminoácidos

(B) Tipo: aminoácido

(C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 104:

Asp Asn Ala Glu Glu
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 105:

(i) Características da Sequência

- (A) Comprimento: 18 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 105:

CAGTATTTTT ACGAAACC

(2) INFORMAÇÃO PARA SEQ. ID N° 106:

- (i) Características da Sequência
 - (A) Comprimento: 18 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADNc

(xi) Descrição da Sequência: SEQ ID N°: 106:

ACATTTCGGT TTGTTCTG

(2) INFORMAÇÃO PARA SEQ. ID N° 107:

- (i) Características da Sequência
 - (A) Comprimento: 18 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 107:

CAGTATTTTT ACGAGACG

(2) INFORMAÇÃO PARA SEQ. ID N° 108:

(i) Características da Sequência

- (A) Comprimento: 18 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADNc

(xi) Descrição da Sequência: SEQ ID N°: 108:

ACATTTCGGT TTGTTAGC

(2) INFORMAÇÃO PARA SEQ. ID N° 109:

(i) Características da Sequência

- (A) Comprimento: 36 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: simples
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 109:

TGCAGTTTCG CTCACCCCCC GTTTTAGCCG GGAAGT

(2) INFORMAÇÃO PARA SEQ. ID N° 110:

(i) Características da Sequência

- (A) Comprimento: 7 aminoácidos
- (B) Tipo: aminoácido
- (C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 110:

Lys Gln Tyr Phe Tyr Glu Thr
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 111:

(i) Características da Sequência

(A) Comprimento: 7 aminoácidos

(B) Tipo: aminoácido

(C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 111:

Trp Arg Phe Ile Arg Ile Asp
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 112:

(i) Características da Sequência

(A) Comprimento: 7 aminoácidos

(B) Tipo: aminoácido

(C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 112:

Gly Glu Leu Ser Val Cys Asp
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 113:

- (i) Características da Sequência
 - (A) Comprimento: 6 aminoácidos
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 113:

Lys Ala Glu Ser Ala Gly
1 5

(2) INFORMAÇÃO PARA SEQ. ID N° 114:

- (i) Características da Sequência
 - (A) Comprimento: 45 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 114:

GGAGGGGGCT GCCGGGGAGT GGACAGGAGG CACTGGGTAT CTGAG

(2) INFORMAÇÃO PARA SEQ. ID N° 115:

- (i) Características da Sequência
 - (A) Comprimento: 15 aminoácidos.
 - (B) Tipo: aminoácido
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: Péptido

(xi) Descrição da Sequência: SEQ ID N°: 115:

Gly	Gly	Gly	Cys	Arg	Gly	Val	Asp	Arg	Arg	His	Trp	Val	Ser	Glu
1				5				10					15	

(2) INFORMAÇÃO PARA SEQ. ID N° 116:

(i) Características da Sequência

- (A) Comprimento: 582 pares de base
- (B) Tipo: ácido nucleico
- (C) Tipo de cadeia: dupla
- (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(ix) Estrutura

- (A) Nome/Código: CDS
- (B) Localização: 1..396

(xi) Descrição da Sequência: SEQ ID N°:116:

GTA	GTT	TGT	CCA	ATT	ATG	TCA	CAC	CAC	AGA	AGT	AAG	GTT	CCT	TCA	CAA
48															
Val	Val	Cys	Pro	Ile	Met	Ser	His	His	Arg	Ser	Lys	Val	Pro	Ser	Gln
1				5					10					15	
AGA	TCC	TCT	AGA	GTC	GCG	CCC	GCG	ACC	TGC	AGG	CGC	AGA	ACT	GGT	AGG
96															
Arg	Ser	Ser	Arg	Val	Ala	Pro	Ala	Thr	Cys	Arg	Arg	Arg	Thr	Gly	Arg
			20					25					30		
TAT	GGA	AGA	TCC	CTC	GAG	GTG	GAG	GTG	TTG	GGC	GAG	GTG	CCT	CCA	GCT
144															
Tyr	Gly	Arg	Ser	Leu	Glu	Val	Glu	Val	Leu	Gly	Glu	Val	Pro	Pro	Ala
	35				40							45			
GTC	GGC	AGT	TCC	CTC	CGC	CAG	CAC	TTC	TTT	GTT	GCC	CGC	TTC	GAG	GCC
192															
Val	Gly	Ser	Ser	Leu	Arg	Gln	His	Phe	Phe	Val	Ala	Arg	Phe	Glu	Ala
	50				55						60				
GAT	AAA	TCT	GAG	GAA	GGT	GGC	CCG	GGG	GTA	GGT	GGA	GGG	GCT	GCC	GCC
240															
Asp	Lys	Ser	Glu	Glu	Gly	Gly	Pro	Gly	Val	Gly	Gly	Gly	Ala	Ala	Ala
65					70					75				80	
GGG	GTG	TGG	ACC	GGG	GGG	CAC	TGG	GTG	TCT	GAG	TGC	AAG	GCC	AAG	CAG
288															
Gly	Val	Trp	Thr	Gly	Gly	His	Trp	Val	Ser	Glu	Cys	Lys	Ala	Lys	Gln
				85					90					95	

TCC TAT GTG CGG GCA TTG ACC GCT GAT GCC CAG GGC CGT GTG GAC TGG
336
Ser Tyr Val Arg Ala Leu Thr Ala Asp Ala Gln Gly Arg Val Asp Trp
100 105 110

CGA TGG ATT CAA ACT GGC ACA GCC TGT GTC TGC ACA CTC CTC AGC CGG
384
Arg Trp Ile Gln Thr Gly Thr Ala Cys Val Cys Thr Leu Leu Ser Arg
115 120 125

ACT GGC TGG GCC TGAGACTTAT ACCCAGGAAC TGGTCAGGCA GAAAAAGAAC 436
Thr Gly Trp Ala
130

AGAGCTGGAT GCTGAGAGAC CTCAGGGTTG GCCCAGCTGC TCTACGGACG GACCCCAGTT 496

GGGGAACCTCA TGAAATCATC ACAAATCAC AACTCTCTGA ATTTGAGCTC AATCTCTGCA 556

GGATGGGTGC CACCACATGT GGT TTT 582

(2) INFORMAÇÃO PARA SEQ. ID N° 117:

(i) Características da Sequência

- (A) Comprimento: 132 aminoácidos
- (B) Tipo: aminoácido
- (D) Topologia: linear

(ii) Tipo de Molécula: Proteína

(xi) Descrição da Sequência: SEQ ID N°: 117:

Val Val Cys Pro Ile Met Ser His His Arg Ser Lys Val Pro Ser Gln
1 5 10 15

Arg Ser Ser Arg Val Ala Pro Ala Thr Cys Arg Arg Arg Thr Gly Arg
20 25 30

Tyr Gly Arg Ser Leu Glu Val Glu Val Leu Gly Glu Val Pro Pro Ala
35 40 45

Val Gly Ser Ser Leu Arg Gln His Phe Phe Val Ala Arg Phe Glu Ala
50 55 60

Asp Lys Ser Glu Glu Gly Gly Pro Gly Val Gly Gly Gly Ala Ala Ala
65 70 75 80

Gly Val Trp Thr Gly Gly His Trp Val Ser Glu Cys Lys Ala Lys Gln
85 90 95

Ser Tyr Val Arg Ala Leu Thr Ala Asp Ala Gln Gly Arg Val Asp Trp
100 105 110

Arg Trp Ile Gln Thr Gly Thr Ala Cys Val Cys Thr Leu Leu Ser Arg
115 120 125

Thr Gly Trp Ala
130

(2) INFORMAÇÃO PARA SEQ. ID N° 118:

- (i) Características da Sequência
 - (A) Comprimento: 36 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 118:
CGGTACCCTC GAGCCACCAT GCTCCCTCTC CCCTCA

(2) INFORMAÇÃO PARA SEQ. ID N° 119:

- (i) Características da Sequência
 - (A) Comprimento: 36 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 119:

CGGTACAAGC GGCCGCTTCT TGGGCATGGG TCTCAG

(2) INFORMAÇÃO PARA SEQ. ID N° 120:

- (i) Características da Sequência
 - (A) Comprimento: 36 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 120:

CGGTACCCTC GAGCCACCCA GGTGCTCCGA GAGATG

(2) INFORMAÇÃO PARA SEQ. ID N° 121:

- (i) Características da Sequência
 - (A) Comprimento: 23 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 121:

GAGACCGGAA GCTTCTAGAG ATC

(2) INFORMAÇÃO PARA SEQ. ID N° 122:

- (i) Características da Sequência
 - (A) Comprimento: 36 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: simples
 - (D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 122:

TGCAGTTTCG CTCACCCCCC GTTCCGCCG TGATGT

(2) INFORMAÇÃO PARA SEQ. ID N° 123:

- (i) Características da Sequência
 - (A) Comprimento: 36 pares de base
 - (B) Tipo: ácido nucleico
 - (C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 123:

ACATCACGGC GGAAACGGGG GGTGAGCGAA ACTGCA

(2) INFORMAÇÃO PARA SEQ. ID N° 124:

(i) Características da Sequência

(A) Comprimento: 36 pares de base

(B) Tipo: ácido nucleico

(C) Tipo de cadeia: simples

(D) Topologia: não conhecida

(ii) Tipo de Molécula: ADN (genómico)

(xi) Descrição da Sequência: SEQ ID N°: 124:

ACTTCCCGGC TAAACGGGG GGTGAGCGAA ACTGCA

REIVINDICAÇÕES

1 - Uso de NT-4 ou de um péptido afim com NT-4 para o fabrico de um medicamento para o tratamento de uma desordem de infertilidade envolvendo oócitos.

2 - Uso de acordo com a reivindicação 1, caracterizado por a proteína ou péptido serem derivados de NT-4 humano.

3 - Uso de uma proteína NT-4, péptido de NT-4 ou derivado capazes de suportar a sobrevivência, crescimento e/ou diferenciação de neurónios motores, para o fabrico de um medicamento para o tratamento de uma desordem dos neurónios motores relacionada com NT-4.

4 - Uso de acordo com a reivindicação 3, caracterizado por a desordem ser seleccionada do grupo consistindo em esclerose amiotrófica lateral, atrofia muscular espinal progressiva, atrofia muscular infantil, poliomielite, síndrome pós-polio, doença de Charlot-Marie Tooth, trauma dos nervos ou lesão dos nervos.

5 - Uso de acordo com a reivindicação 3, caracterizado por a proteína NT-4 ser codificada por uma sequência de ácido nucleico recombinante, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

6 - Processo para promover a sobrevivência, crescimento e/ou diferenciação de neurónios motores, caracterizado por se exporem os neurónios motores a uma concentração terapeuticamente eficaz de uma proteína NT-4, péptido de NT-4 ou derivado que são capazes de promover a sobrevivência, crescimento e/ou diferenciação de neurónios motores.

7 - Processo de acordo com a reivindicação 6, caracterizado por a proteína NT-4 ser codificada por uma molécula de ácido nucleico recombinante compreendendo a sequência de ADN, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o

número de acesso 75070.

8 - Processo de acordo com a reivindicação 6, caracterizado por a proteína NT-4, péptido de NT-4 ou derivado serem codificados por uma molécula de ácido nucleico recombinante que tem uma homologia de pelo menos cerca de 70% com a correspondente sequência de ADN, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

9 - Processo para diagnosticar uma desordem de neurónios motores num paciente caracterizado por compreender os passos de:

- (a) determinar uma concentração de NT-4 numa amostra retirada de um paciente, por meio de um ensaio imunométrico utilizando um anticorpo anti-NT-4;
- (b) calcular a diferença entre a concentração de NT-4 da amostra retirada do paciente e a concentração média de NT-4 de amostras análogas de indivíduos normais; e
- (c) diagnosticar uma desordem de neurónios motores quando a diferença entre a concentração de NT-4 da amostra retirada do paciente e a concentração de NT-4 de uma amostra análoga de pelo menos um indivíduo normal se correlaciona positivamente com a desordem de neurónios motores.

10 - Processo de acordo com a reivindicação 9, caracterizado por o anticorpo anti-NT-4 ser capaz de se ligar a uma proteína NT-4 codificada por uma molécula de ácido nucleico recombinante compreendendo a sequência de ADN afim com NT-4, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

11 - Processo de acordo com a reivindicação 9, caracterizado por o anticorpo anti-NT-4 ser capaz de se ligar a uma proteína NT-4 codificada por uma molécula de ácido nucleico recombinante que tem uma homologia de pelo menos 70% com a correspondente sequência de ADN, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de

acesso 75070.

12 - Uso de um oligonucleótido que:

(a) consiste em pelo menos seis nucleótidos;

(b) compreende uma sequência complementar a pelo menos uma porção de um transcrito de ARN do gene NT-4; e

(c) é hibridável com o transcrito de ARN, para o fabrico de um medicamento para o tratamento de uma doença localizada na próstata caracterizada por transcrição aumentada de um gene NT-4 em tecido da próstata, relativamente aos níveis de transcrição em tecido da próstata de pacientes normais.

13 - Uso de acordo com a reivindicação 12, caracterizado por a doença localizada na próstata ser hipertrofia prostática benigna.

14 - Uso de acordo com a reivindicação 13, caracterizado por o oligonucleótido compreender uma sequência que é 100% complementar à porção correspondente do transcrito de ARN codificado pela sequência de ADN, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

15 - Uso de acordo com a reivindicação 14, caracterizado por o oligonucleótido compreender uma sequência que é pelo menos 70% complementar à porção correspondente do transcrito de ARN codificado pela sequência de ADN, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

16 - Uso de uma proteína NT-4 para o fabrico de um medicamento para o tratamento de uma desordem de impotência envolvendo a próstata humana.

17 - Uso de acordo com a reivindicação 16, caracterizado por a proteína NT-4 ser codificada por uma molécula de ácido nucleico recombinante compreendendo a sequência de ADN afim com NT-4, tal como contida no bacteriófago HG7-2, tal como

depositado no ATCC com o número de acesso 75070.

18 - Uso de acordo com a reivindicação 16, caracterizado por a proteína NT-4 ser codificada por uma molécula de ácido nucleico recombinante que tem uma homologia de pelo menos cerca de 70% com a correspondente sequência de ADN, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

19 - Conjunto de diagnóstico para detectar o transporte retrógrado num neurónio motor, caracterizado por compreender num recipiente adequado uma proteína NT-4, derivado ou fragmento peptídico marcados detectavelmente.

20 - Conjunto de diagnóstico de acordo com a reivindicação 19, caracterizado por a proteína NT-4, derivado ou fragmento peptídico marcados detectavelmente serem codificados por uma molécula de ácido nucleico recombinante compreendendo a sequência de ADN afim com NT-4, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

21 - Conjunto de diagnóstico de acordo com a reivindicação 19, caracterizado por a proteína NT-4, derivado ou fragmento peptídico marcados detectavelmente serem codificados por um fragmento de ácido nucleico que tem uma homologia de pelo menos cerca de 70% com a correspondente sequência de ADN, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

22 - Proteína prepro quimérica, caracterizada por compreender:

- (a) uma sequência de aminoácidos de uma região prepro de uma neurotrofina diferente de NT-4;
- (b) uma sequência de aminoácidos de uma região prepro de uma forma madura de uma proteína afim com NT-4.

23 - Proteína prepro quimérica de acordo com a reivindicação 22, caracterizada por a sequência de aminoácidos da forma

madura de uma proteína afim co com NT-4 ser codificada por uma molécula de ácido nucleico recombinante compreendendo pelo menos uma porção da sequência de ADN que codifica NT-4, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

24 - Proteína prepro quimérica de acordo com a reivindicação 22, caracterizada por a neurotrofina diferente de NT-4 ser factor de crescimento de nervo, factor neurotrófico derivado de cérebro ou neurotrofina-3.

25 - Proteína prepro quimérica de acordo com a reivindicação 22, caracterizada por a região prepro da neurotrofina diferente de NT-4 ser derivada de xenopus NT-4.

26 - Proteína prepro quimérica de acordo com a reivindicação 22, caracterizada por a região prepro da neurotrofina diferente de NT-4 ser derivada de NT-3 humana.

27 - Molécula de ácido nucleico, caracterizada por compreender uma sequência de nucleótidos que codifica a proteína prepro quimérica da reivindicação 22.

28 - Vector de ADN, caracterizado por ser pCMX-hNT3/hNT4 depositado no ATCC com o número de acesso 75133.

29 - Vector de ADN, caracterizado por ser pCMX-xNT4/hNT4.

30 - Proteína NT-4 recombinante substancialmente pura, caracterizada por ser codificada pelo vector de ADN da reivindicação 28.

31 - Proteína NT-4 recombinante substancialmente pura, caracterizada por ser codificada pelo vector de ADN da reivindicação 29.

32 - Uso de uma proteína NT-4 para o fabrico de um medicamento para o tratamento de uma desordem imunológica, relacionada

com NT-4, que afecta a transmissão neuromuscular.

33 - Uso de acordo com a reivindicação 32, caracterizado por a desordem ser miastenia grave.

34 - Uso de acordo com a reivindicação 32, caracterizado por a proteína NT-4 ser codificada por uma molécula de ácido nucleico recombinante compreendendo a sequência de ADN afim com NT-4, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

35 - Uso de acordo com a reivindicação 32, caracterizado por a proteína NT-4 ser codificada por uma molécula de ácido nucleico recombinante que tem uma homologia de pelo menos cerca de 70% com a correspondente sequência de ADN, tal como contida no bacteriófago HG7-2, tal como depositado no ATCC com o número de acesso 75070.

36 - Processo de produção de NT-4, caracterizado por compreender o crescimento de uma célula recombinante contendo a molécula de ácido nucleico da reivindicação 26, sob condições tais que a proteína prepro quimérica seja expressa e processada pela célula de modo a produzir a forma madura de NT-4.

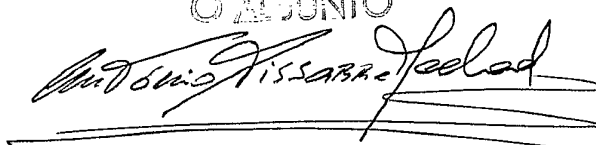
37 - Processo para detectar ou medir a actividade de NT-4, caracterizado por compreender (a) a exposição de uma célula que expressa trkB a um agente de teste; e (b) detectar ou medir a ligação específica do agente de teste a trkB, correlacionando-se a ligação específica a trkB positivamente com a actividade de NT-4 no agente de teste.

21. MAI 1992
Lisboa,

Por REGENERON PHARMACEUTICALS, INC., FINN HALLBOOK,
CARLOS FERNANDO IBANEZ MOLINER e HAKAN BENGT PERSSON

=O AGENTE OFICIAL=

© ADJUNTO



[illegible]

17

FIGURA 2

	R69	S67	D63	A68
NGF				
HUMANO	67	MTBDPNPVDSCGRGIDSLDMSYCTTITFEVKALTMGIDMA	00	
RATO		...A...E.....T.D...		
GALINHA		...R..S.....A.....E...		
VÍBORA		...K..S..SG...A.....D...B...E.N..S		
XENOPUS		...K..S.../...A.....E...		
SALMÃO		T..GRVGS...L...GB.....NS...B...SKDLV.		
BDNF				
HUMANO	65	KCNPMGYTIDGCRGIDITMANSQCTIQSIVVWLTMSIG	100	
RATO		...K.....N..V.		
GALINHA		...K.....N..V.		
VÍBORA		...STK..A.....Y.....F...R..KV.		
XENOPUS		...M.....Y.....F...R..KV.		
SALMÃO		...K.F.N.....K.....S.....R..K..		
RAIA				
HDNF/NT-3				
HUMANO	68	BCKEAPVINGCRGIDDMANSQCTISQTYVRALTSENKLVG	00	
RATO		...K.....		
GALINHA		...K.....		
XENOPUS		...K.E.T..K.F.S.....QRT-S..		
SALMÃO		...SK.G.....SK...Y..		
RAIA				
NT-4				
VÍBORA		KCNPAAGGTVGGRGVDRWVISECKAKQSYVRALTMSKLVG		
XENOPUS		...S.S.7R.....KQO.....I..AN..L..		

FIGURA 3

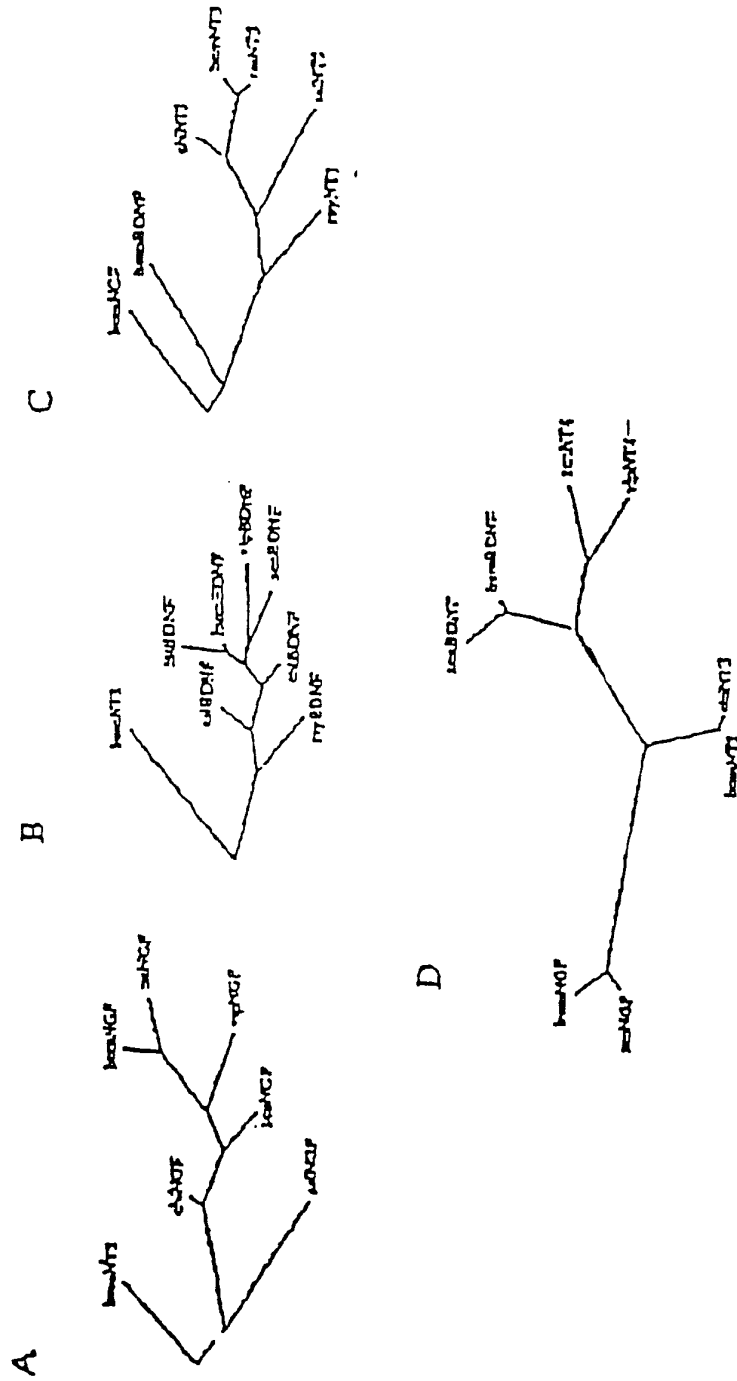


FIGURA 4

A

1 CATTGATCTTATGAAAGCAGCGGGGAGCCCTCGCCTTACTTCCCAGCCATGCAGAACTC
 61 AAGCAGCTTTGTTTAIGCCGATCCCTAAGCAGCCGACACACAGCAGATGTGCACAGT
 121 CTTAGTCTTGCAGAGAGTGTAAACAAAGTTACAAAGATGGTGACCCCTGTAGCCAACTTT
 181 GAAAGCATAAATCATTTCTTGAITAGGCTTGTGGTGCAGTAAAGTTCAITGTTTATATTTA
 241 GCATACAAATACAGCAITTCATGCTTATTTCTATTTTAGACTTTACCCCTTTAATGCCCA
 301 GTTCTGCCCATTTGCCITATAGATGTAAAGTCCCAATATCAGATTGGCATCCTCGGCTGT
 361 TTACAAACAAATTAAGCTTGTACTTATATTTAAATCTGTGTTGTTCTTCCATATTC
 421 ATCAGCTTAGACCCTAAGAAATATATGTATATATTTTGCATAAATATATATATGGCA

 401 GCCGTATTCTAATTCGTGTTTTTTTTTTTTTTTTTTCAGTGGTCTGAGTGGGATTAAGTA
 INTRÃO V D FIM V
 1 M I L R L Y A H V I S Y C C A I C A A P SC
 541 ATGATCCCTCCGCTTATATGCCATGGTGATCTCATCTGTGTGCCATCTCGCGTCCCCC
 21 F Q S R T T D L D Y G P D K T S E A S D
 601 TTCCAGAGCCGGACCAAGATTTGGATTATGGCCCCGATAAACATCAGAAAGCTCAGAC
 41 R Q S V F H H F S H V L Q H G F F P D L
 661 CGGCAATCAGTTCCTAACAACTTCAGTCAITGCTCGCAAAATGGGTTCTTCCAGATTG
 61 S S T Y S S H A G K D W H L Y S P R V T
 721 TCATCCACCTATTCAGCATGGCTGGTAAAGGACTGGAACTATATCTACCTAGAGTGACT
 81 L S S E E P S G P P L L F L S E E T V V
 781 CTTTCAGTGAGGAGCCTTCTGGACCTCCACTACTTTTCTGTGAGGAGACAGTGGTA
 101 H P E P A H K T S R L K R A S G S D S V
 841 CATCCAGAACCCAGCCAAAGACTTCCCGGCTAAACGGGCATCAGGATCTGATTCGGTC
 121 S L S R R G E L S V C D S V N V W V T D
 901 AGCTTGTCCTCGTGGGAGAGCTCTCTGTGTGTGACAGTGTCAACGTCTGGGTTACCGAT
 141 K R T A V D D R G K I V T V H S E I Q T
 961 AAACGTAACGCCGTGGATGATCGGGTAAATAGTGAAGTGTCTGAGATTACAGACT
 161 L T G P L K Q Y F F E T K C N P S G S T
 1021 CTAAAGGACCACTGAAAGCAATCTTCTTTGAGAACAGTGCATTCATCAGGAGCACC
 181 T R G C R G V D K K Q W I S E C K A K Q
 1081 ACTAGAGGATGCCGAGGTGTAGACAAAGCAATGCATATCTGAGTGCAAAGCAAGACAG
 201 S Y V R A L T I D A N K L V G W R W I R
 1141 TCTTATGTGAGGGCTCTGACCATAGATGCCAACAGCTTGTGGGTTGGCGTTGGATCCGT
 221 I D T A C V C T L L S R T G, R T End 236
 1201 ATTGACACAGCGGTGTGTCTGTACCTTGTTCAGTCCGACAGGAAGGACGTAAAGACGAGG
 1261 GTTAGCAAAATAGAGAGAGAGGTTGATCCGTTGACCTGCAG

B

NT-4 ASGSDSVLSRRGELSVCDSVNWVT--DQRTAVDDRGIKIVTMSEIQTLTGPLKQYFFET
 NGF .STIIPVFIH..F.....SV..G--..T..T.IK..E...LA.VNINNSVFR.....
 BDNF HSDPAR.....ISE...M..K....MS.GT...LEKVPVSK.Q.....Y..
 NT-3 YAEIHK.H...Y.....ESL....--..SS.I.I..HQ...LG..K.GNS.V....Y..

KGNPSGSTITGCRGVDIKQKQISECKAKQSYVTALTIDANKLVGWRWIRIDTACVCTLLSRIGRT 123
 ..RASNPFVES...I.S.H.N.Y.TTTITF.K...T.E-.QAA..F.....V.SRKAT. 118
H.Y.KE...I..RI.N.Q.RIT.....M.SK.RI...F.....S...T.IKT... 119
 R.KENRFVKN....I.D.H.N.Q..TS.T.....SEN.....S...A.SRKI... 119

FIGURA 5

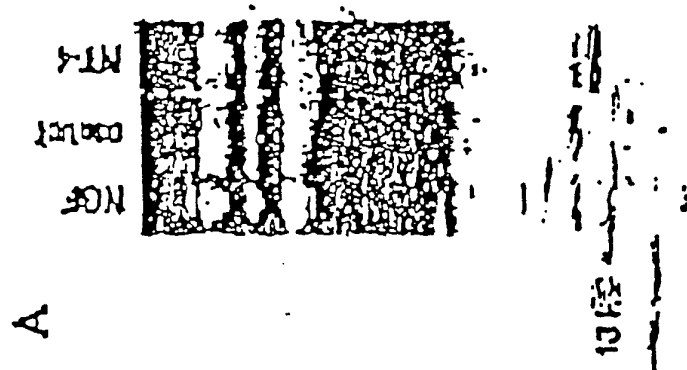
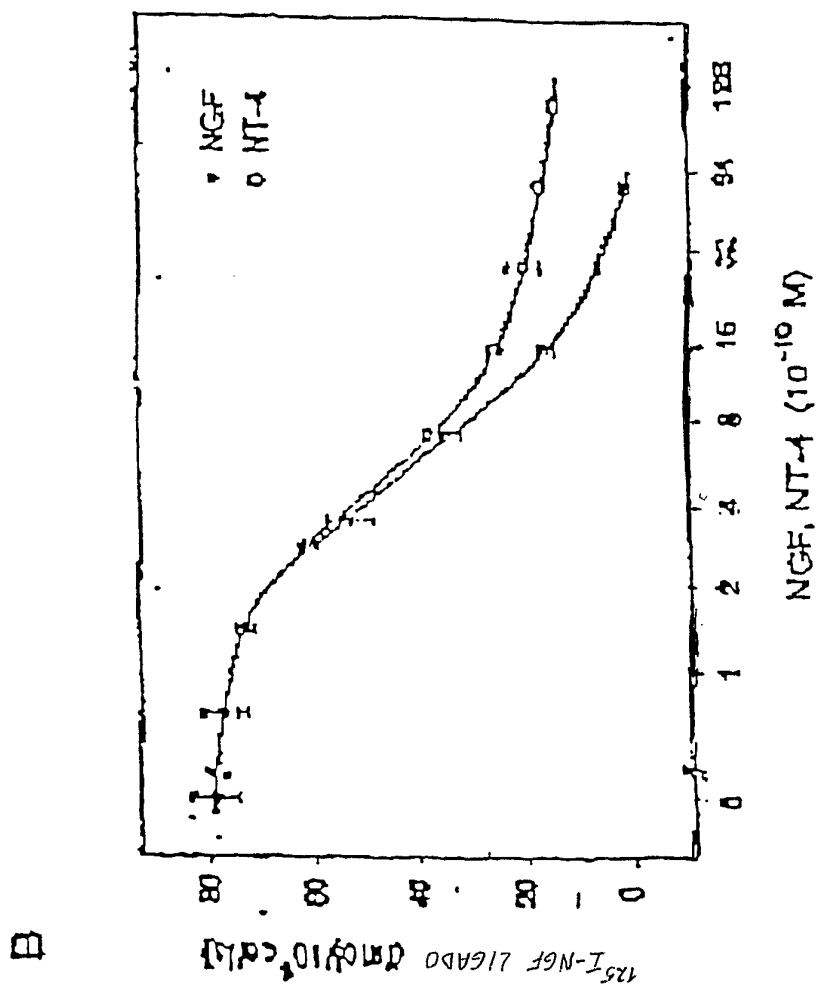


FIGURA 6

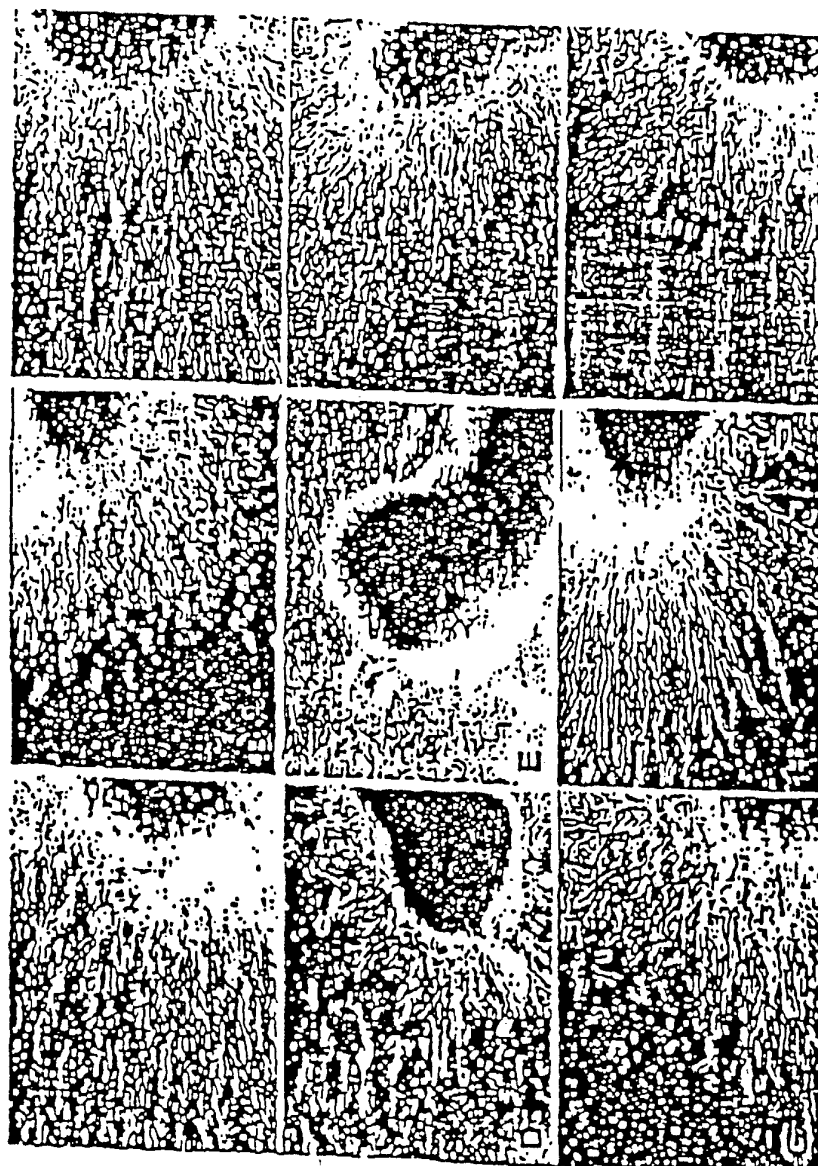


FIGURA 7



FIGURA 8

				D	
		E		s	
		c		p	
		.o		HC DID	
S	r	P		II Iv n2s	II
f	s	I		u al n8p	n
e	t	S		o 1J 16C	1
I	I	I		I VI III	I

//

CTCCAGGGAAACAAATCAIACITTAATGAAACAGGAGGGGAGCCCTCCCTTACTTCCCAAGCC
 ----- 60
 GACGTCCCTTTTCTTACTATGAAIACITTCCTCCCTCCGAGCCGAAACAAAGGCTCCG

				T	
		t			
		h	S	D	F
II	F	l	a	p	n C
1	n C	1	a	p	n C
a	u Av	ANl	u D	u D	u v
I	4 II	1b1	J p	1d	4 1
I	II uJ	uvI	Λ n	0e	II J
1	1 II	111	1 1	1 1	1 1

/

ATCCAGAACTCAAGCAGCTTTCITTAATGCCGATCCCTAAGCAGCCCAAGCCAGCTGACC
 ----- 120
 TACGTCTTCAGTTCCTCCAAACAAATACCCCTACCGATTCGTCCGCTCTCCTCTCACTCC

				D	
				s	
		N	P		
DAI	1II			II	BMT
spall	2q	D		a	sas
pals	01	d		e	tep
CLIp	6A	e		1	114
1111	11	1		1	115
				1	111

// /

ATCTCCACAGCTCTTACGTCTTCCAAAGATCTTTAAACAAAGTATACAGATCGTCAACCCCTC
 ----- 180
 TACAGCTCTCAGAAACAGAACTTCTACAAATTCITTCATCTTCTACCACTCGGGGAC

				T	
				t	
				h	
				1	
C				1	
II v				1	
p 1				1	
h J				1	
1 1				1	

/

TACCCAACTTTCAAGCATAAATCATTTCTTTGATACGCTTCTCGTCCACTAAGTTCAT
 ----- 240
 ATCCGTTGAAACTTCTGTATTIACIAAACAACTAATCCGAACACCAGCTCATTCAGTA

C

C

C

3

C

781 CACACATCCTCTTTCAAGTCAGGACGCTTCTGGACCTCCACTACTTTTCTGTGACAGG

GATCTCACTCAGAAAGTTCAGTCTCTGGGAAGACCTGGACCTGATGAAAGACACGCTTCC 810

FIGURA 8 (CONTINUAÇÃO)

	S								T
	t			D	C				a S
	y	R		r	v		SC		qDa A
	S	s		J	I		Hllcv		IsuDI
	J	a		I	J		sccl		ltJpW
	I	I		I	I		pIJJ		-IAnH
							IIII		IIIII
							//		/
041	AGACACCTGGTACATCCAGAACCAAGCCAAACAAAGCACTTCCCGCGTAAAAACGGCGATCAGGAT								
	TCGTGCACCAATGTAGGCTCTTGGTCGGCTTCTCTGAAAGCGCGCGATTTTTCGGCGTAGTCCTA								
	T V V H P E F A H K T S R L K R A S C S -								
							D		
							s		
							P		
							H T		
							t		
							H		
							CDIH		
							a s		
							h		
							I H		
							Avn2qS		
							e p		
							I 4		
							1		
							c e		
							1 5		
							1		
							1 1		
							IIIIII		
							1 1		
							1 1		
							//		
							/		
							/		
							/		
901	CTGATTCGGCTACGCTTCTGTCGGCTGGGAGAGGCTCTCTCTCTCTGACAGCTCTCAAGCTCT								
	GACTAAGCCAGTCCAAACAGGCCAGCCCCCTCTCCAGAGAGACACAGCTCTCAAGTTCAGCA								
	D S V S L S H R C E L S V C D S V H V W -								
							S		
							a		
							H T		
							a s		
							N		
							I		
							CD		
							v s D		
							v D		
							F		
							e p		
							D a D		
							o		
							1 4		
							r		
							1 d		
							k		
							1 5		
							d		
							1 e		
							1 1		
							1 1		
							1 1		
							//		
							/		
							/		
							/		
961	GGCTTACCGATAAACCTAACCGCTTGGATGATCGCGGTAAAAATAGTGACTGTCTATCTCTC								
	CCCAATCGCTATTTGCAATCTCGGCACTACTAGCCCCATTTTATCACTGACAGTACAGAC								
	V T D K R T A V D D R C K I V T V H S E -								
							S		
							Aa		
							D		
							Ec		
							IvU		
							Bs		
							co		
							sm		
							o5		
							aA		
							p7		
							II		
							II		
							//		
							/		
							/		
							/		
1021	AGATTACAGACTCTAACAGGACCACTCAAGCAATACTTCTTTGACAGCCAACTGCAAATCCAT								
	TCTAAGTCTGAGATTCTCCCTCGTCACTTCCGTTATCAAGAAACTCTCGTTACGTTAGCTA								
	I Q T L T C P L K O Y F F E T K C H P S -								
							F		
							S		
							B		
							s		
							AF		
							co		
							ck		
							II		
							//		
							/		
							/		
							/		
1001	CAGGCAGCAGCACTACAGGATCCCGAGCTGTACACAAAAAGCAATCGATATCTGACTCCA								
	GTCCGTCGTCGTCATCTCTTACGGCTCCACATCTGTTTTTTCGTTACCTATAGACTACCT								

FIGURA 8 (CONTINUAÇÃO)

```

      D
      s
      H
      P
      C DLS      R n C
      v a l l      l H d Avx      A
      l n O s      e m l l l e      l
      J 16 H      A e l u J m      w
      l l l l      l l l l l l      l
      //
AACC AAAACAGCTCTTATCTGAGCGCTCTGACCATACATCCCAACAACTTCTGGGTTCGC
1141 ----- 1200
TTCGTTTTCAGAAATACACTCCCGAGACTGCTATCTACGGTTCTTCCAAACACCCAAACCC

C      A K Q S T V R A L T I D A H K L V G H R -

      S      n
      BD Na      r
      asDlu A      a H      n      l      RH      a H
      mtpaJ l      l m      s      n      ll      e n
      HynIA w      l e      a      f      ey      l l
      l l l v l l      l l      l      l      ll      l l
      / //
CTTGGATCCGTTATGACACAGCGCTCTCTCTCTACCTTCTTGAGTCCGACAGGAAGGACCGT
1201 ----- 1260
CAACCTAGCCATAACTCTCTCCGACACACACATCGAACAACTCAGCCCTCTCTCTCTCCCA

C      H I R I D T A C V C T L L S R T C R T -

      S      H
      a H      i
      EH      A      u D b      n      S      P
      en      l      J p o      c      f      s
      rl      w      A n l      l      e      t
      ll      l      l l l      l      l      l
      /
AAAAACAGCAGCGTTACCAAAAATAGACAGAAAGCGTTGATCCGTTGACCTTCGCAG
1261 ----- 1313
TTTTCTGCTCCCAATCGTTTATCTCTCTCTCCAACTAGCCAACTGGACGTC

```

C

Enzymes that do cut:

AclI	AluI	AlwI	AlwNI	ApdLI	AvdII	BamHI	BanII
BbvI	BceII	Dpu10I	DsaI	BsaJI	BsmAI	Dsp1286I	BspCI
BspCI	BstI	BstEII	BstXI	BstYI	Cfr10I	CviJI	DdeI
DpnI	DraIII	DrdI	DrdII	DsaI	EaeI	EciI	Eco57I
EcoDXXI	EcoPI	EcoPI5I	EcoRV	FlnI	Fnu4HI	FokI	HaeIII
HqI	HhaI	HincII	HindIII	HinII	HphI	IaeI	MaeII
MaeIII	MboII	MlyI	MmeI	HnlI	HsuI	HspI	HwoI
NciI	NcoI	NlaIII	NlaIV	HspI	PleI	PstI	RleAI
RsaI	Sau96I	SauJAI	SbfI	SfoNI	SceI	SplI	SspI
SstI	StyI	StyLTI	StySJI	TaqII-1	TfiI	Tsp45I	TspCI
Tth1111	Tth11111	XcmI					

Enzymes that do not cut:

AacII	AflIII	AflIII	AgeI	AhaII	ApaI	AsuI	AvdI
AvrII	BalI	BanI	BbvII	BclI	BclI	BglI	BqlII
BsaAI	BsaBI	BsiI	BsmI	BspII	BspMI	BspMII	BssHII
BstBI	Bsu36I	CfrAI	ClaI	DraI	EaeI	Eco1111I	EcoAI
EcoBI	EcoDI	EcoEI	EcoKI	EcoNI	EcoO109I	EcoRI	EcoRII
EcoRI24I	EcoRI24/JI	EspI	EspII	FauI	FseI	FspI	GdIII
GsuI	HaeI	HaeII	HgaI	HqI	HlnfIII	HpaI	KpnI
McrI	MfeI	HluI	HncI	MarI	NdeI	NheI	NotI
MfuI	NsiI	HspDII	PfIM	PmlI	PpuMI	PshAI	PvuI
PvuII	RsrII	SacII	SalI	ScaI	SfiI	SmaI	SnaI

FIGURA 9

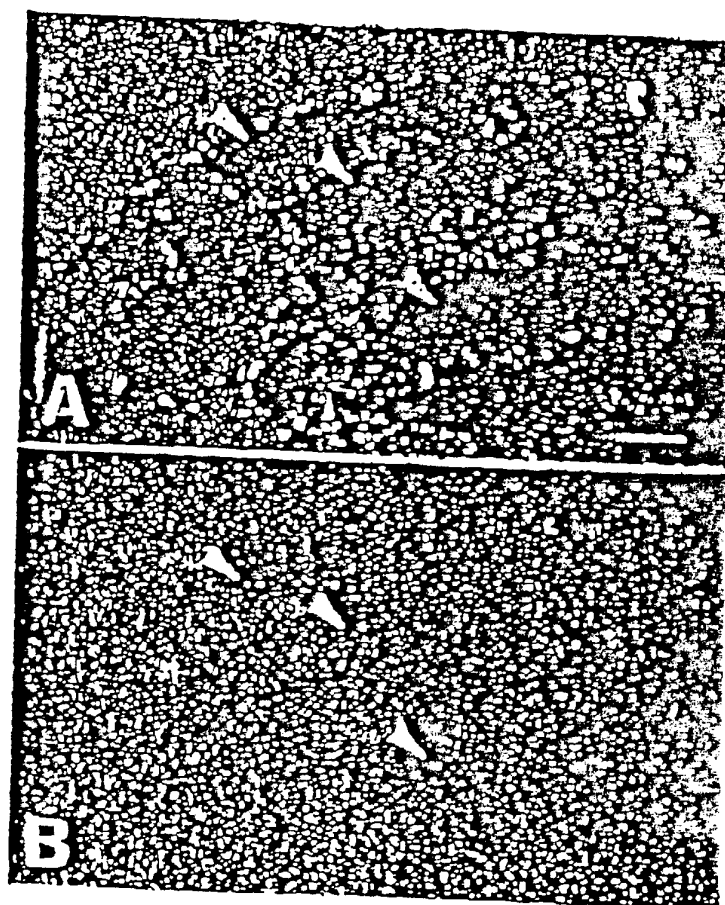


FIGURA 10

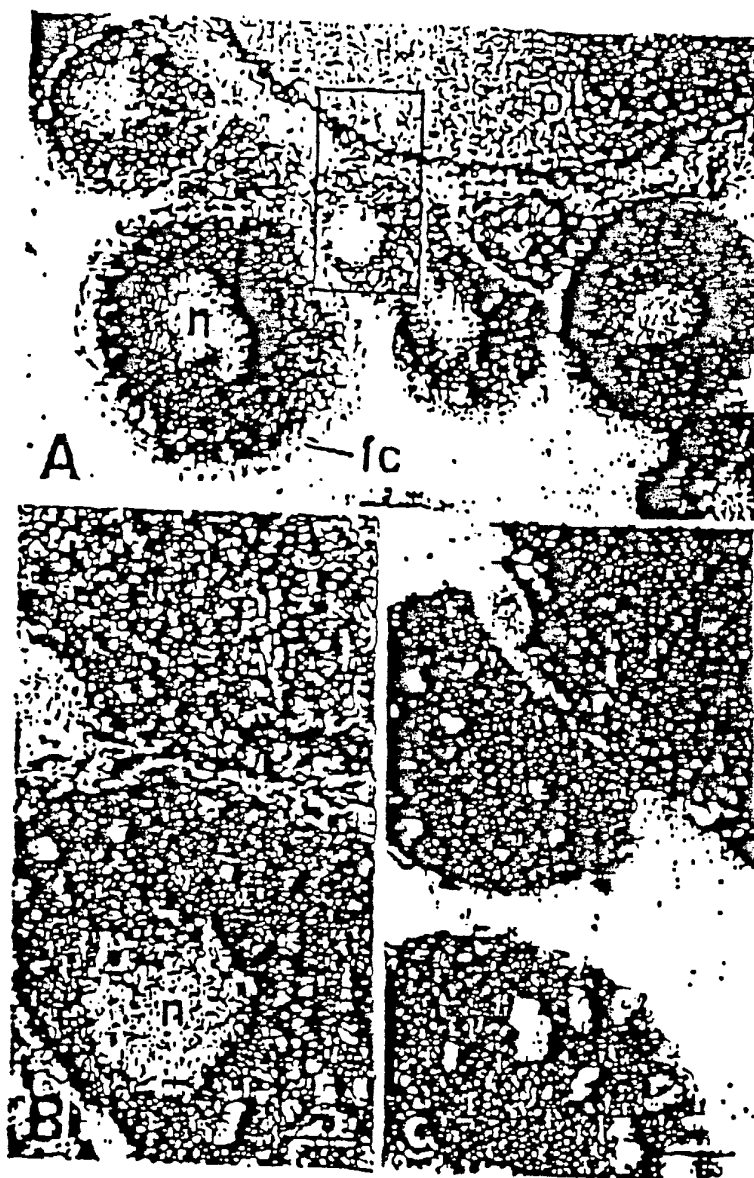


FIGURA 11

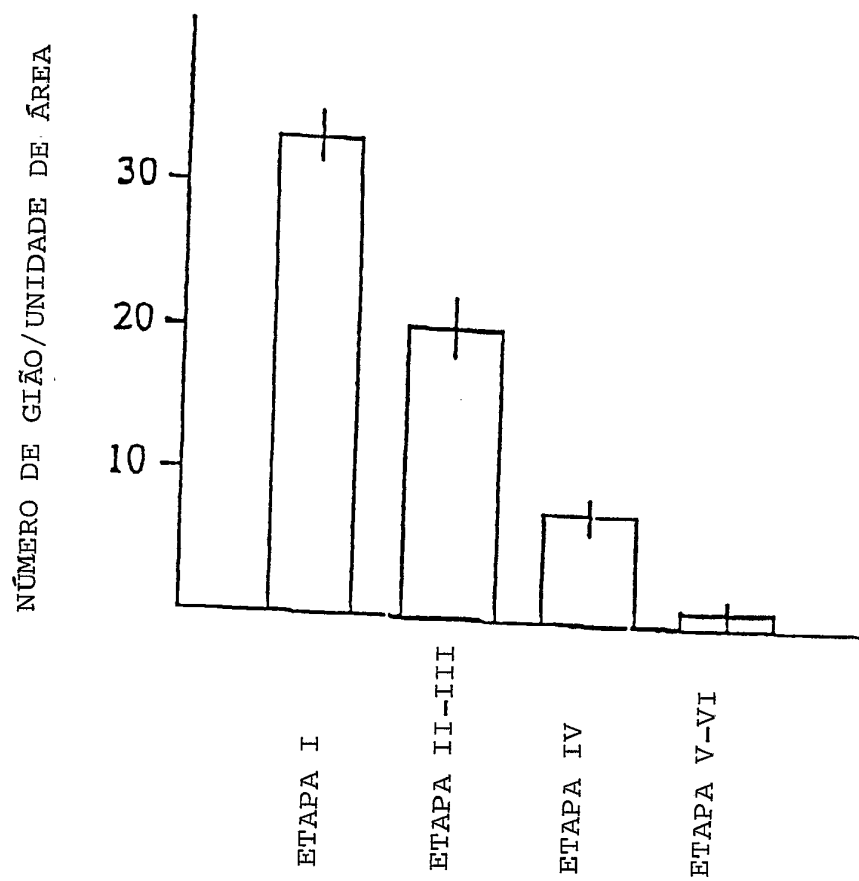


FIGURA 12

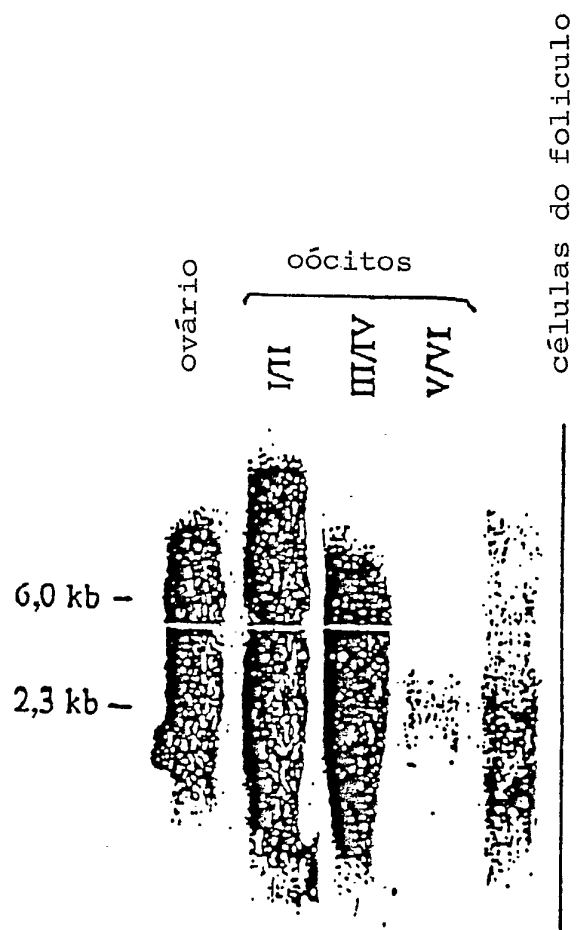


FIGURA 13

A.

$\xrightarrow{2Y}$
 $\xleftarrow{3Y} \xleftarrow{3Z}$
 $\xleftarrow{4Z}$
QYFFETKCHPSGSTTRGCRGVDIKKQWSECKAKQSYVFTALTIDANIKLVGWRWIRIDT...
 $\xrightarrow{2Z}$
 QYFFET

B.

2Y CA(G,A) TA(T,C) TT(T,C) TT(T,C) GA(G,A) AC
 2Z CA(G,A) TA(T,C) TT(T,C) TA(T,C) GA(G,A) AC
 3Y TT (A,G)CA (C,T)TC I(G,C)(A,T) IAT CCA
 3Z GA (C,T)TG (C,T)TT IGC (C,T)TT (A,G)CA
 3Z' CT (C,T)TG (C,T)TT IGC (C,T)TT, (A,G)CA
 4Z GT (G,A)TC IAT IC(T,G) IAT CCA

C.

3' (3Y, 3Z, 3Z', 4Z)
 $\begin{array}{c} \text{Hind III} \quad \text{XbaI} \quad \text{EcoRI} \\ \hline \text{CCA AGC TTC TAG AAT TC} \end{array}$

5' (2Y, 2Z)
 $\begin{array}{c} \text{XhoI} \quad \text{SalI} \\ \hline \text{GAC TCG AGT CGA CAT CG} \end{array}$

FIGURA 14

```

      20      40      60
      .      .      .
[CAG TAT TTT TAC GAG AGT] CCC TCC AAC CCC GGA ACC GGT GGG GAA GGT GGC CCA GGT GTG
  Q  Y  F  Y  L  T  R  C  K  A  L  S  A  G  E  G  G  P  G  V

      80      100      120
      .      .      .
CCC GGA GGG GGC TGT CCC GCG GTG GAT CCC AGG CAC TCC CTC TCA GAA [TGT AAA GCG AAA
  G  G  G  G  C  R  G  V  D  R  R  K  K  L  S  L  C  K  A  K

CAA TCG]
  Q  S

```



2000-00-00
00000000

FIGURA 15

20 40 60
[CAG TAT TTT TAC GAA AGC CGC TGC AAG GGT GAT AAC GGT GAG GAA GGT GGC CCG GGG GCA
Q Y F Y E F R C K A D H A E E G C P G A

80 100 120
GGT GGA GGG GGC TGC CGG GGA GTG GAC AGC AGG CAC TGG GTA TGT GAG [TGT AAA GGC AAA
G G G G C R G V D R R H W V S E C K A K

CAA TCG
Q S]

FIGURA 16

INT-1	[QYFYET]RCKAESAGEGGFGVGGGGCGVDRIHWLSE[CKAKQ3]
hNT-4	DN.E.....A.....V.....
XNT-1	...F..K.NPSGSTTR-----...KKQ.I.....
INGF	...F..K.RAPNPVES-----...I.SK..N.Y.TTTNT
IBDNF	K.NPMGYTKE-----...I.K...N.Q.RTT..
INT-3	KEARPVKH-----...I.DK..N.Q:.TS.T

FIGURA 17 A-C

A.

```

      20      40      60
[EAA TAT TTT TTT GAG ACC] CGC TGC AAG GCT GAT AAC GCT GAG GAA GGT GGT CCG GGG GCA GGT
  Q   Y   F   F   E   T   R   C   K   A   D   N   A   E   E   G   G   P   G   A   G>

      80      100      120
GGA GGG GGC TGC CCG GGA GTG GAC AGG AGG CAC TGG GTA TCT GAG TGC AAG GCC AAG CAG TCC
  G   G   G   C   R   G   V   D   R   R   H   W   V   S   E   C   K   A   K   Q   S>

      140      160      180
TAT GTG CCG GCA TTG ACC GCT GAT GCC CAG GGC CGT GTG GGC TGG CGA TGG ATC CGC ATC GAT
  Y   V   R   A   L   T   A   D   A   Q   G   R   V   G   W   R   W   I   R   I   D>
[ACG]
  T>

```

B.

```

GACTCGAGTCGACATCG.GAA.ACC.CGC.TGC.AAG.GCT
      E T R C K A

```

C.

```

GACTCGAGTCGACATCG.GAT.AAC.GCT.GAG.GAA.GGT
      D N A E E G

```

FIGURA 18

10 20 30 40 50 60 70 80
 CTTCACCCAGGTGGCAGCCGAGTGGTGCACCTCTGTGCTCACTGCAACCTCGGCTCTGGGTTGCACTGATTCTCTACCTCA
 90 100 110 120 130 140 150 160 170
 GCCTACTGASTAGCTGGGATTACACGGCTGCAGCAGTATGCCCGGTTAATTTTGGTATTTTGGTAGAGATGAGGTTTACCATTG
 180 190 200 210 220 230 240 250
 TTGACCAGCTGCTCTGAACTCCTCAGCTCAAGTCATCCAGCTGCTCAGCTGCTGAGTGGGATTAGAGGTGTGGGGCAG
 260 270 280 290 300 310 320 330 340
 AGTGGCTGGCTGTAGTAGTGAATATTTTATTAATCTCAAGTTGCGCATACGCAAGCCCTAGATATAGGGTCCCCAAC
 350 360 370 380 390 400 410 420
 TTCTAGAACAAAGGGCTTCCCCACAATCCTGGCAGGCAAGCCCTCGGCTTCCCAACTTCTTCCCCACTGAAGTTTATCC
 430 440 450 460 470 480 490
 CTTCTCTAATCCAGCCCTCTCTTCTGTCTC CAG GTG CTC CGA GAG ATG CTC CCT CTC CCC TCA TGC
 O V L R E M L P L P S C>
 500 510 520 530 540 550
 TCC CTC CCC ATC CTC CTC CTT TTC CTC CTC CCC AGT CTG CCA ATT GAG TCC CAA CCC CCA CCC
 S L P I L L L F L L P S V P I E S O P P>
 560 570 580 590 600 610 620
 TCA ACA TTG CCC CCT TTT CTC CCC CTT GAG TGG GAC CTT CTC TCC CCC CGA GTA GTC CTG TCT
 S T L P P F L A P E W D L L S P R V V L S>
 630 640 650 660 670 680
 ACC CGT GCC CCT GCT GGG CCC CTT CTG CTC TTC CTG CTC GAG GCT GGG GCC TTT CGG GAG TCA
 R G A P A G P P L I F L L E A G A F R E S>
 690 700 710 720 730 740
 GCA GGT GCC CGC GCC AAC CGC AGC CGG CGT GGG GTG ACC GAA ACT GCA CCA GCG AGT CGT CGG
 A G A P A H R S R R G V S E T A P A S R R>
 750 760 770 780 790 800 810
 GGT GAG CTG GCT GTG TGC GAT GCA GTC AGT GCC TGG CTG ACA GAC CGC CGG ACC GCT GTG GAC
 G E L A V C D A V S G W V T D R R T A V D>
 820 830 840 850 860 870
 TTG CGT GGG CGC GAG GTG GAG GTG TTC GCG GAG GTG CCT GCA GCT GGC GGC AGT CCC CTC CGC
 L R G R E V E V L G E V P A A G G S P L R>
 880 890 900 910 920 930
 CAG TAC TTC TTT GAA ACC CGC TGC AAG GCT GAT AAC GCT GAG GAA GGT GCC CCG GGG GCA GGT
 Q Y F F E T R C K A D N A E E G G P G A G>
 940 950 960 970 980 990
 GGA GGG CCC TGC CGG CCA GTG GAC AGG AGG CAC TGG GTA TCT GAG TGC AAG CCC AAG CAG TCC
 G G G C R G V D R R H W V S E C K A K Q S>
 1000 1010 1020 1030 1040 1050 1060
 TAT GTG CGG GCA TTG ACC GCT GAT GCC CAG GGC CGT CTC GCG TGG CGA TGC ATT CGA ATT GAC
 Y V R A L T A D A Q G R V G W R W I R I D>
 1070 1080 1090 1100 1110 1120 1130
 ACT GCC TGC GTC TGC ACA CTC CTC ACC CCG ACT GCG CCC TGAGACCCATGCCAGGAAATAACAGAG
 T A C V C T L L S R T G R A>
 1140 1150 1160 1170 1180 1190 1200 1210
 CTGGATGCTCAGACACCTCAGGCATGCCCGAGCTCATTAAGGACCCAGTTTGGCACTCATATAATACAAAAATCAAAAT
 1220 1230 1240 1250 1260 1270 1280 1290 1300
 TCTCTGATTTTGAAGCTCAATCTCTGAGGATCGGTGAAACCATGCGGTTTGGAGGTGAATAGGAGTTCTCTGGGCAACT
 1310 1320 1330 1340 1350 1360 1370 1380
 TCAGCGTAATAATGATGATATAATAATAATAGCCACTATTACTGACTGTTTACTGTTTCTTATCCCTAATACATAACTCCT
 1390 1400
 CAGATCAACTCTCATG

[illegible]

FIGURA 20

```

      20      40      60
      *      *      *
CGA GTG GTC CTG TCT AGG GGT GCC GCT GCC GGG CCC CCT CTG GTC TTC CTG CTG GAG ACT GGA
GCT CAC CAG GAC AGA TCC CCA CGG CGA CGG CCC GGG GGA GAC CAG AAG GAC GAC CTC TGA CCT
R   V   V   L   S   R   G   A   A   A   G   P   P   L   V   F   L   L   E   T   G>

      80      100      120
      *      *      *
GCC TTT CGG GAG TCA GCA GGC GCC CGG GCC AAC CGC AGC CAG CGA GGG GTG AGC GAT ACT TCA
CGG AAA GCC CTC AGT CGT CCG CGG GCC CGG TTG GCG TCG GTC GCT CCC CAC TCG CTA TGA AGT
A   F   R   E   S   A   G   A   R   A   N   R   S   Q   R   G   V   S   D   T   S>

      140      160      180
      *      *      *
CCG GCG AGT CAT CAG GGT GAG CTG GCC GTG TGC GAT GCA GTC AGT GTC TGG GTG ACA GAC CCC
GGC CGC TCA GTA GTC CCA CTC GAC CGG CAC ACG CTA CGT CAG TCA CAG ACC CAC TGT CTG GGG
P   A   S   H   Q   G   E   L   A   V   C   D   A   V   S   V   W   V   T   D   P>

      200      220
      *      *
TGG ACT GCT GTG GAC TTG GGT GTG CTC GAG GTC GAG
ACC TGA CGA CAC CTG AAC CCA CAC GAG CTC CAG CTC
W   T   A   V   D   L   G   V   L   E   V   E>

```

FIGURA 21

```

      20      40      60
→ GTA GTT TGT CCA ATT ATG TCA CAC CAC AGA AGT AAG GTT CCT TCA CAA AGA TCC TCT AGA GTC
   CAT CAA ACA GGT TAA TAC AGT GTG GTG TCT TCA TTC CAA GGA AGT GTT TCT AGG AGA TCT CAG
     V   V   C   P   I   M   S   H   H   R   S   K   V   P   S   Q   R   S   S   R   R   V>

      80      100      120
GCG CCC GCG ACC TGC AGG CGC AGA ACT GGT AGG TAT GGA AGA TCC CTC GAG GTG GAG GTG TTG
CGC GGG CGC TGG ACG TCC GCG TCT TGA CCA TCC ATA CCT TCT AGG GAG CTC CAC CTC CAC AAC
   A   P   A   T   C   R   R   R   T   G   R   Y   G   R   S   L   E   V   E   V   L>

      140      160      180
GCG GAG GTG CCT CCA GCT GTC GGC AGT TCC CTC CGC CAG CAC TTC TTT GTT GCC CGC TTC GAG
CGC CTC CAC GGA GGT CGA CAG CCG TCA AGG GAG GCG GTC GTG AAG AAA CAA CGG CGC AAG CTC
   G   E   V   P   P   A   V   G   S   S   L   R   Q   H   F   F   V   A   R   F   E>

      200      220      240
GCC GAT AAA TCT GAG GAA GGT GGC CCG GGG GTA GGT GGA GGG GCT GCC GCC GGG GTG TGG ACC
CGC CTA TTT AGA CTC CTT CCA CCG GGC CCC CAT CCA CCT CCC CGA CGG CGG CCC CAC ACC TGG
   A   D   K   S   E   E   G   G   P   G   V   G   G   G   A   A   A   A   G   V   W   T>

      260      280      300
GGG GGG CAC TGG GTG TCT GAG TGC AAG GCC AAG CAG TCC TAT GTG CGG GCA TTG ACC GCT GAT
CCC CCC GTG ACC CAC AGA CTC ACG TGC AAG CAG TCC TAT GTG CGG GCA TTG ACC GCT GAT
   G   G   H   W   V   S   E   C   K   A   K   Q   S   Y   V   R   A   L   T   A   D>

      320      340      360
GCC CAG GGC CGT GTG GAC TGG CGA TGG ATT CAA ACT GGC ACA GCC TGT GTC TGC ACA CTC CTC
CGG GTC CCG GCA CAC CTG ACC GCT ACC TAA GTT TGA CCG TGT CGG ACA CAG ACG TGT GAG GAG
   A   Q   G   R   V   D   N   R   W   I   Q   T   G   T   A   C   V   C   T   L   L>

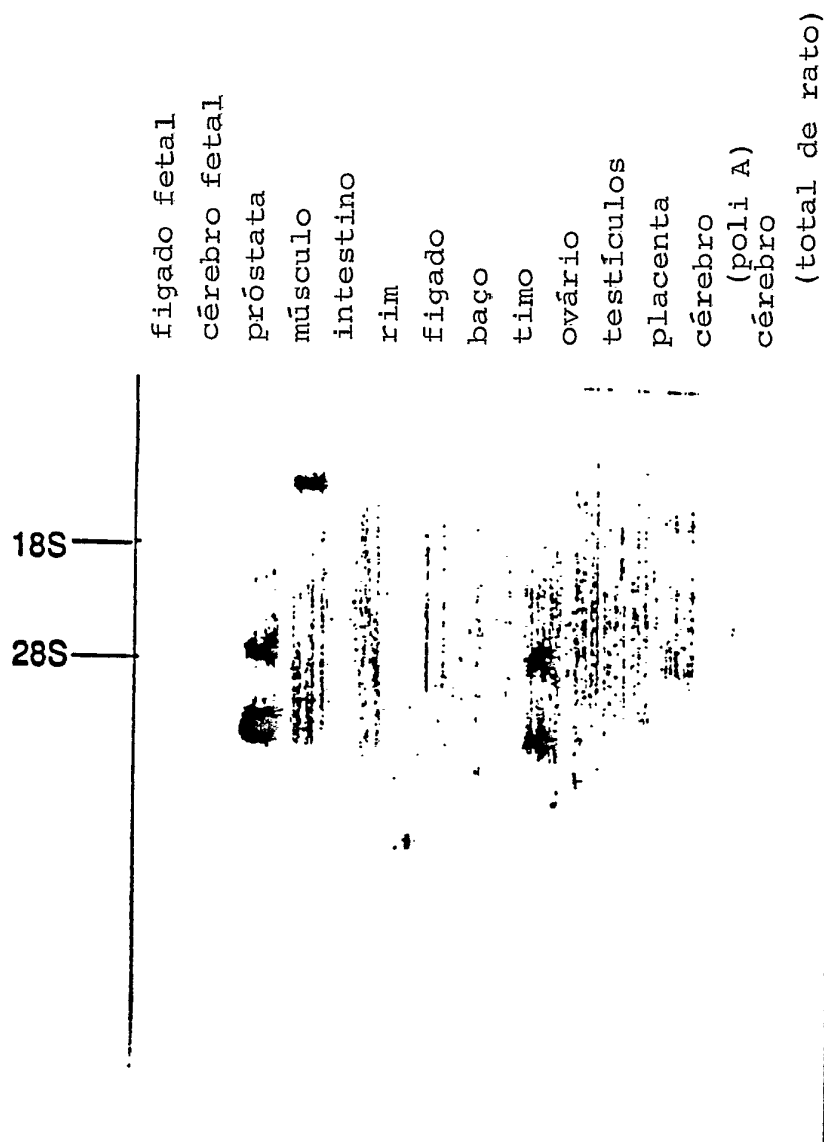
      380      400      420      440
AGC CGG ACT GGC TGG GCC TGAG ACTTATACCCAGGAAGTGGT CAGGCAGAAAAAGAACAGAG
TCG GCC TGA CCG ACC CGG ACTC TGAATATGGGTCTCTGACCA GTCGCTCTTTTCTTGCTCTC
   S   R   T   G   W   A>

      460      480      500      520
CTGGATGCTGAGAGACCTCA GGGTTGGCCAGCTGCTCTA CGGACGGACCCAGTTGGGG AACTCATGAAATCATCACAA
GACCTACGACTCTCTGGAGT CCCAACCAGGTCGACGAGAT GCCTGCTGGGGTCAACCCC TTGAGTACTTATAGTAGTGT

      540      560      580
AATCACAACCTCTCTGAATTT GAGCTCAATCTCTGCAGGAT GGGTGCCACCACATGTGGTT TT
TTAGTGTTGAGAGACTTAAA CTCGAGTTAGAGACGTCTA CCCACGGTGGTGTACACCAA AA

```

FIGURA 22



Handwritten signature

FIGURA 23

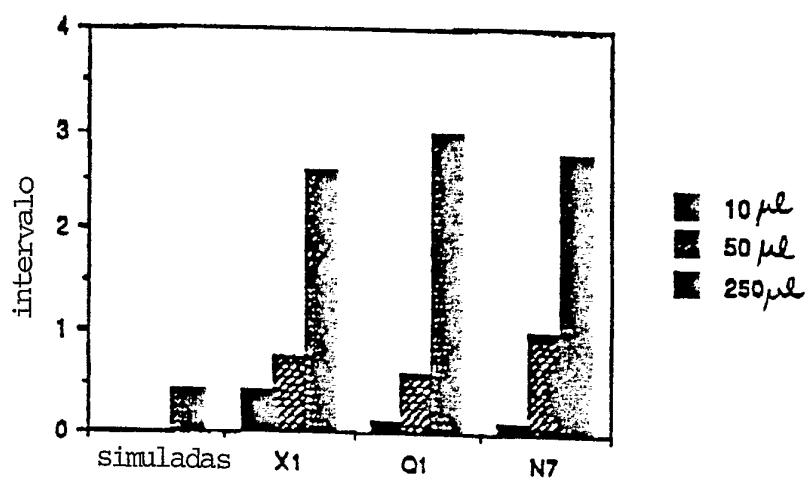


FIGURA 24

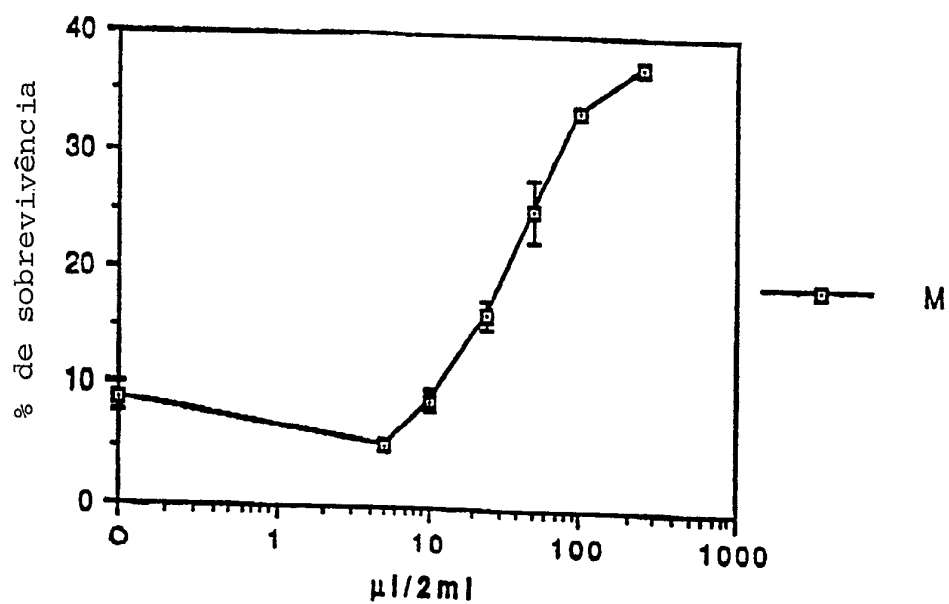
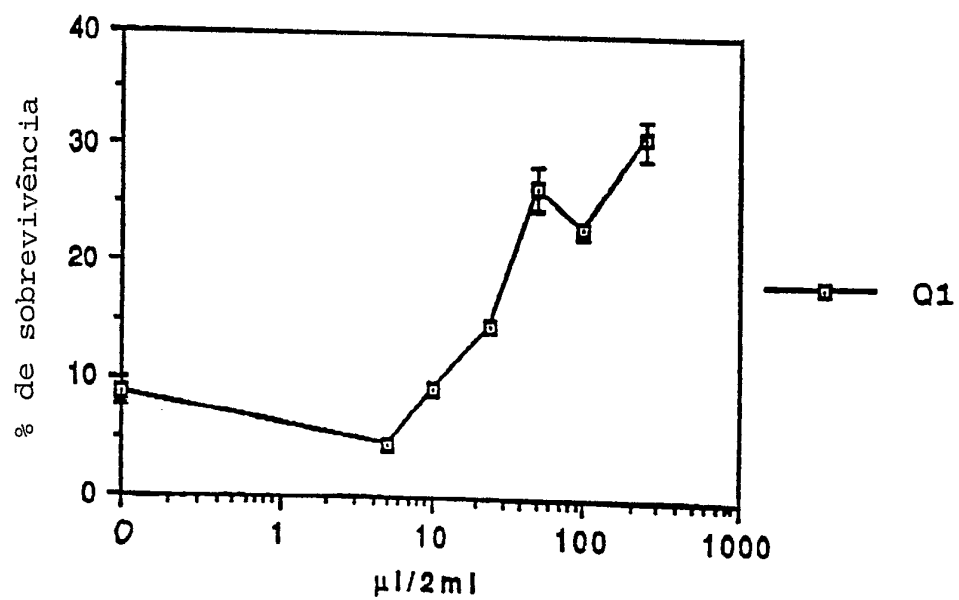


FIGURA 25

