

(12) **FASCÍCULO DE PATENTE DE INVENÇÃO**

(22) Data de pedido: **2001.10.06**

(30) Prioridade(s): **2000.10.06 US 238361 P**

(43) Data de publicação do pedido: **2003.07.09**

(45) Data e BPI da concessão: **2010.02.24**
104/2010

(73) Titular(es):

BIOGEN IDEC MA INC.

14 CAMBRIDGE CENTER CAMBRIDGE,

MASSACHUSETTS 02142

US

YALE UNIVERSITY

US

(72) Inventor(es):

DINAH W. Y. SAH

US

STEPHEN M. STRITTMATTER

US

RICHARD L. CATE

US

(74) Mandatário:

ANTÓNIO JOÃO COIMBRA DA CUNHA FERREIRA
RUA DAS FLORES, Nº 74, 4º AND 1249-235 LISBOA

PT

(54) Epígrafe: **HOMÓLOGOS DO RECEPTOR DE NOGO**

(57) Resumo:

DESCRIÇÃO

"Homólogos do receptor de Nogo"

CAMPO DO INVENTO

O presente invento refere-se a neurologia e biologia molecular. Mais particularmente, o presente invento refere-se a neurónios do SNC e crescimento axonal.

ANTECEDENTES DO INVENTO

Entre os mecanismos através dos quais as células de um organismo comunicam umas com as outras e obtêm informação e estímulos do seu ambiente é através de moléculas receptoras da membrana celular expressas na superfície celular. Muitos de tais receptores foram identificados, caracterizados e por vezes classificados em superfamílias principais de receptores com base em motivos estruturais e características de transdução de sinal. Os receptores são um primeiro elo essencial para a tradução de um sinal extracelular numa resposta fisiológica celular.

Os receptores dos neurónios são particularmente importantes no desenvolvimento do sistema nervoso durante a embriogénese. Os neurónios formam ligações com células alvo durante o desenvolvimento através de alongamento axonal dos neurónios para as células alvo num processo mediado por receptores. Os axónios e as dendrites têm uma região especializada das suas pontas distais conhecidas como o cone de crescimento. Os cones de crescimento permitem que o neurónio sinta o ambiente local através de um processo mediado por receptores e dirija o movimento do axónio ou da dendrite do neurónio para a célula alvo do neurónio. Este processo é conhecido como alongamento. Os cones de crescimento podem ser sensíveis a várias pistas orientadoras, por exemplo, capacidade de adesão da superfície, factores de crescimento, neurotransmissores e campos eléctricos. A orientação do crescimento no cone depende de várias classes de moléculas de adesão, sinais intercelulares, bem como factores que estimulam e inibem os cones de crescimento.

É interessante que os neurónios danificados não se alongam no sistema nervoso central (SNC) após lesão devida a traumatismo ou doença, enquanto os axónios no sistema nervoso periférico (SNP) regeneram-se facilmente. O facto de os neurónios do SNC lesionados não conseguirem alongar não é devido a uma propriedade intrínseca dos axónios do SNC, mas antes devido ao ambiente do SNC que não é permissivo ao alongamento axonal. Experiências clássicas de enxertia por Aguayo e colegas (p.ex., Richardson *et al.*, *Nature* 284: 264-265, 1980) demonstraram que os axónios do SNC podem de facto alongar-se ao longo de distâncias substanciais com enxertos de nervos periféricos e que a mielina do SNC inibe o alongamento dos axónios do SNC. Portanto, dado o ambiente apropriado, os axónios do SNC podem regenerar, implicando que a lesão axonal do SNC pode potencialmente ser abordada através da manipulação apropriada do ambiente do SNC.

A ausência de regeneração dos axónios após lesão pode ser atribuída à presença de inibidores do crescimento dos axónios. Estes inibidores estão predominantemente associados a mielina e constituem uma importante barreira à regeneração. Os inibidores do crescimento dos axónios estão presentes em mielina derivada do SNC e na membrana plasmática de oligodendrócitos que sintetizam mielina no SNC (Schwab *et al.*, *Annu. Rev. Neurosci.* 16: 565-595, 1993). Os inibidores associados à mielina parecem ser um contribuinte primário para a insuficiência da regeneração dos axónios do SNC *in vivo* após uma interrupção da continuidade axonal, enquanto outros inibidores do crescimento dos axónios não associados à mielina no SNC podem ter um papel menor. Estes inibidores bloqueiam a regeneração axonal após lesão neuronal devido a traumatismo, AVC ou infecção viral.

Numerosos inibidores do crescimento de axónios derivados de mielina foram caracterizados (ver, para revisão, David *et al.*, (1999) WO995394547,; Bandman *et al.*, Patente U.S. N.º 5858708; Schwab, *Neurochem. Res.* 21: 755-761, 1996). Vários componentes da substância branca do SNC, NI35, NI250 (Nogo) e a glicoproteína associada à Mielina (MAG), que têm actividade inibidora da extensão axonal, foram também descritos (Schwab *et al.*, (1990) WO9005191; Schwab *et al.*, (1997) Patente U.S.

N.º 5684133). Em particular, Nogo é um inibidor do crescimento dos axónios associado à mielina de 250 kDa que foi originalmente caracterizado com base nos efeitos da proteína purificada *in vitro* e de anticorpos monoclonais que neutralizam a actividade da proteína (Schwab, *Exp. Neurol.* 109: 2-5, 1990). O ADNc de Nogo foi identificado pela primeira vez através de análise aleatória do ADNc do cérebro e não foi sugerida qualquer função (Nagase *et al.*, *DNA Res.* 5: 355-364: 1998). A identificação deste ADNc de Nogo como o ADNc codificando o inibidor do crescimento de axónios associado à mielina de 250 kDa foi descoberta apenas recentemente (GrandPre *et al.*, *Nature* 403: 439-444, 2000; Chen *et al.*, *Nature* 403: 434-439, 2000; Prinjha *et al.*, *Nature* 403: 383-384, 2000).

Importantemente, mostrou-se que Nogo era o componente primário da mielina do SNC responsável pela inibição do alongamento e da regeneração axonal. A expressão selectiva de Nogo por oligodendrócitos e não por células de Schwann (as células que mielinizam os axónios do SP) é consistente com os efeitos inibidores da mielina do SNC, em contraste com a mielina do SP (GrandPre *et al.*, *Nature* 403: 434-439, 2000). Em cultura, Nogo inibe o alongamento axonal e causa colapso do cone de crescimento (Spillmann *et al.*, *J. Biol. Chem.* 272: 19283-19293, 1998). Foi mostrado que anticorpos (p.ex., IN-1) contra Nogo bloqueiam a maior parte da acção inibidora da mielina do SNC no crescimento de neurites *in vitro* (Spillmann *et al.*, *J. Biol. Chem.* 272: 19283-19293, 1998). Estas experiências indicam que Nogo é o componente principal da mielina do SNC responsável pela inibição do alongamento axonal em cultura. Para além disso, mostrou-se que *in vivo* o anticorpo IN-1 aumenta a regeneração axonal após lesão da medula espinal, resultando em recuperação de comportamentos tais como contacto, colocação e comprimento do avanço (Schnell e Schwab, *Nature* 343: 269-272, 1990; Bregman *et al.*, *Nature* 378: 498-501, 1995). Assim, há provas substanciais de que Nogo é um alvo molecular relevante de doença. Esperar-se-ia que agentes que interferem com a ligação de Nogo ao seu receptor melhorassem a regeneração axonal em estados clínicos nos quais foram danificados axónios, e melhorasse o estado do doente.

A modulação de Nogo foi descrita como um meio para tratamento de regeneração de neurónios danificados por traumatismo, enfarte e distúrbios degenerativos do SNC (Schwab *et al.*, (1994) W09417831, Tatagiba *et al.*, *Neurosurgery* 40: 541-546, 1997) bem como tumores malignos no SNC tais como glioblastoma (Schwab *et al.*, (1993) Patente U.S. N.º 5250414; Schwab *et al.*, (2000) Patente U.S. N.º 6025333).

Foi sugerido que os anticorpos que reconhecem Nogo são úteis no diagnóstico e tratamento de lesões nos nervos resultantes de traumatismo, enfarte e distúrbios degenerativos do SNC (Schnell & Schwab, *Nature* 343: 269-272, 1990; Schwab *et al.*, (1997) Patente U.S. N.º 5684133). Para axónios do SNC, existe uma correlação entre a presença de mielina e a inibição da regeneração dos axónios ao longo de distâncias longas (Savio e Schwab, *Proc. Natl. Acad. Sci.* 87: 4130-4133, 1990; Keirstead *et al.*, *Proc. Natl. Acad. Sci.* 89: 11664-11668, 1992). Após Nogo ser bloqueado por anticorpos, os neurónios podem novamente estender-se através das lesões causadas por danos nos nervos (Schnell e Schwab, *Nature* 343: 269-272, 1990).

NgR1 é um Receptor de Nogo que foi identificado através de pesquisa de moléculas que interagem com os 66 aminoácidos do domínio activo de Nogo, tal como descrito no Exemplo 7 de W00151520, um documento que faz parte do estado da arte apenas sob as condições do Artigo 54(3) EPC04-11-2009.

SUMÁRIO DO INVENTO

Foram identificados genes codificando homólogos (NgR2 e NgR3) de um receptor de Nogo (NgR1) em ratinhos e humanos. Foram identificados vários domínios nos polipéptidos codificados pelos genes de NgR2 e NgR3 que foram comparados com domínios nos polipéptidos NgR1 de ratinho e humano. Esta comparação conduziu à identificação de uma sequência de consenso (sequência de consenso NgR) que caracteriza uma família de proteínas (família NgR). Com base nestas e noutras descobertas, o presente invento apresenta moléculas e métodos para modulação do crescimento axonal em neurónios do SNC. Assim, o presente invento proporciona um polipéptido

compreendendo uma sequência de aminoácidos seleccionada de entre o grupo que consiste em:

- (a) aminoácidos 31 a 310 de SEQ ID NO:2;
- (b) aminoácidos 31 a 395 de SEQ ID NO:2;
- (c) aminoácidos 1 a 395 de SEQ ID NO:2;
- (d) aminoácidos 31 a 310 de SEQ ID NO:2 excepto 20 ou menos substituições de aminoácidos conservativas por 100 resíduos de aminoácidos;
- (e) aminoácidos 31 a 310 de SEQ ID NO:2 e até cinco aminoácidos no terminal N e/ou terminal C do referido polipéptido;
- (f) aminoácidos 31 a 395 de SEQ ID NO:2 excepto 20 ou menos substituições de aminoácidos conservativas por 100 resíduos de aminoácidos;
- (g) aminoácidos 31 a 395 de SEQ ID NO:2 e até cinco aminoácidos no terminal N e/ou terminal C do referido polipéptido; e
- (h) aminoácidos 1 a 395 de SEQ ID NO:2 excepto 20 ou menos substituições de aminoácidos conservativas por 100 resíduos de aminoácidos;

em que o referido polipéptido diminui a inibição do crescimento axonal em neurónios do SNC humano.

O presente invento proporciona também um polipéptido compreendendo uma sequência de aminoácidos seleccionada de entre o grupo que consiste em:

- (a) uma sequência de aminoácidos pelo menos 70% idêntica aos aminoácidos 1 a 420 de SEQ ID NO: 2;
- (b) aminoácidos 1 a 420 de SEQ ID NO:2
- (c) uma sequência de aminoácidos pelo menos 70% idêntica aos aminoácidos 31 a 420 de SEQ ID NO:2; e
- (d) aminoácidos 31 a 420 de SEQ ID NO:2;

em que o referido polipéptido diminui a inibição do crescimento axonal em neurónios do SNC humano.

Os polipéptidos do presente invento podem ainda compreender um polipéptido heterólogo. O polipéptido

heterólogo pode ser seleccionado de entre o grupo que consiste em Fc, Glutathione-S-transferase (GST), e uma etiqueta de Histidina (etiqueta His).

O presente invento proporciona também um polinucleótido compreendendo um ácido nucleico codificando um polipéptido do presente invento. O polinucleótido pode ainda compreender um ou mais elementos de controlo da expressão operativamente ligados ao referido ácido nucleico.

O presente invento proporciona também um vector compreendendo um polinucleótido do presente invento.

O presente invento proporciona também uma célula hospedeira compreendendo um polinucleótido do presente invento ou um vector do presente invento, em que a célula hospedeira exclui uma célula capaz de formar um ser humano. A célula hospedeira pode ser uma célula eucariótica.

O presente invento proporciona também um polipéptido receptor de NOGO de acordo com o presente invento expresso por uma célula hospedeira do presente invento.

O presente invento proporciona também um anticorpo ou fragmento de ligação ao antigénio deste que se liga especificamente a um polipéptido do presente invento. Os anticorpos ou fragmentos de ligação ao anticorpo podem ser seleccionados de entre o grupo que consiste em anticorpos policlonais, anticorpos monoclonais, anticorpos humanizados, e anticorpos humanos.

O presente invento proporciona também uma composição compreendendo um polinucleótido do presente invento, um vector do presente invento, um polipéptido do presente invento ou um anticorpo do presente invento, e um transportador.

O presente invento proporciona também um método *in vitro* de diminuição da inibição do crescimento axonal de um neurónio do SNC compreendendo o contacto de um neurónio com um polipéptido do presente invento ou um anticorpo do presente invento.

O presente invento proporciona também a utilização de um polipéptido do presente invento ou um anticorpo do presente invento para a preparação de uma composição farmacêutica para modulação da inibição do crescimento axonal.

O presente invento proporciona também a utilização de um polipéptido do presente invento ou um anticorpo do presente invento para a preparação de uma composição farmacêutica para tratamento de uma doença do sistema nervoso central (SNC), um distúrbio ou ferimento num mamífero, em que a referida doença do SNC, distúrbio ou ferimento é seleccionado de entre o grupo que consiste em ferimento cerebral, ferimento da medula espinal, AVC, ou uma doença desmielinizante. Doenças desmielinizantes que podem ser tratadas incluem esclerose múltipla, desmielinização monofásica, encefalomielite, leucoencefalopatia multifocal, panencefalite, doença de Marchiafava-Bignami, degenerescência esponjosa, doença de Alexander, doença de Canavan, leucodistrofia metacromática e doença de Krabbe.

As composições farmacêuticas que são preparadas de acordo com este aspecto do presente invento podem ser formuladas para administração através de uma via seleccionada a partir do grupo que consiste em administração parentérica, administração subcutânea, administração intravenosa, administração transmucosa, administração intradérmica e administração transdérmica.

A menos que definido de outro modo, todos os termos técnicos e científicos aqui utilizados têm o mesmo significado que o vulgarmente utilizado por um perito na arte a que o presente invento se refere. No caso de conflito, o presente pedido, incluindo as definições, dominarão.

Os materiais, métodos e exemplos apresentados abaixo são apenas ilustrativos e não se pretende que sejam limitantes. Outras características e vantagens do presente invento serão aparentes a partir da descrição detalhada e das reivindicações.

BREVE DESCRIÇÃO DAS FIGURAS

A Fig. 1A-1B mostra um alinhamento de NgR2 (SEQ ID NO:2) e NgR3 (SEQ ID NO:4) com o NgR conhecido, NgR1 (SEQ ID NO:5) e a Sequência de Consenso (SEQ ID NO:6).

Fig. 2. mNgR3 não se liga a hNogoA(1055-1120). Células COS-7 foram transfectadas com vectores codificando myc-NgR1 ou myc-NgR3, fixadas e coradas com anticorpos anti-myc ou AP-hNogoA(1055-1120).

Fig. 3. Um alinhamento das sequências de aminoácidos de NgR1 humano, NgR1 de murídeo, NgR3 de murídeo, NgR3 humano e NgR2 humano. A numeração começa com o aminoácido #1 do NgR3 de murídeo. A sequência de consenso está listada abaixo. O domínio LRR NT está indicado por uma caixa sombreada; os domínios LLR 1, LLR 3, LLR 5 e LLR 7 estão indicados por caixas vazias; LLR 2, LLR 4, LLR 6 e LLR 8 estão indicados por caixas sombreadas; e o domínio LLR CT está indicado por uma caixa sombreada. Os aminoácidos a negrito em LLR 8 indicam locais de glicosilação conservados. Um ponto indica um resíduo de cisteína conservado em LRR4. A caixa no terminal C indica sinais de GPI putativos.

DESCRIÇÃO DETALHADA DO INVENTO

A presente divulgação refere-se a moléculas e métodos para modelação do crescimento axonal em neurónios do SNC.

A divulgação proporciona um polipéptido que contém um polipéptido contendo um domínio LRRCT rico em triptofano consistindo na sequência de aminoácidos:

N X₁ W X₂ C X₃ C R A R X₄ L W X₅ W X₆ X₇ X₈ X₉ R X₁₀ S S S X₁₁ V

X₁₂ C X₁₃ X₁₄ P X₁₅ X₁₆ X₁₇ X₁₈ X₁₉ X₂₀ D L X₂₁ X₂₂ L X₂₃ X₂₄ X₂₅ D

X₂₆ X₂₇ X₂₈ C [SEQ ID NO: 19]

em que X é quaisquer aminoácido da proteína ou um intervalo, e o polipéptido não inclui a sequência de aminoácidos do resíduo 260 a 309 de SEQ ID NO:5 (NgR1 humano) ou SEQ ID NO:17 (NgR1 de ratinho).

De preferência, X17 e X23 são (independentemente) arginina ou lisina. Nalguns casos a sequência de aminoácidos do domínio LRRCT é os resíduos 261-310 de SEQ ID NO:2, ou os resíduos 261-310 de SEQ ID NO:2 com até 10 substituições de aminoácidos conservativas. Nalguns casos, o polipéptido contém a seguinte sequência de aminoácidos NTLRRCT:

**CPX₁X₂CX₃CYX₄X₅PX₆X₇TX₈SCX₉X₁₀X₁₁X₁₂X₁₃X₁₄X₁₅X₁₆P
 X₁₇X₁₈X₁₉PX₂₀X₂₁X₂₂X₂₃RX₂₄FLX₂₅X₂₆NX₂₇IX₂₈X₂₉X₃₀X₃₁X₃₂X₃₃
 X₃₄FX₃₅X₃₆X₃₇X₃₈X₃₉X₄₀X₄₁X₄₂LWX₄₃X₄₄SNX₄₅X₄₆X₄₇X₄₈IX₄₉
 X₅₀X₅₁X₅₂FX₅₃X₅₄X₅₅X₅₆X₅₇LEX₅₈LDLX₅₉DNX₆₀X₆₁LX₆₂X₆₃X₆₄
 X₆₅PX₆₆TFX₆₇GLX₆₈X₆₉LX₇₀X₇₁LX₇₂LX₇₃X₇₄CX₇₅LX₇₆X₇₇LX₇₈
 X₇₉X₈₀X₈₁FX₈₂GLX₈₃X₈₄LQYLYLQX₈₅NX₈₆X₈₇X₈₈X₈₉LX₉₀D

 X₉₁X₉₂FX₉₃DLX₉₄NLX₉₅HLFLHGNX₉₆X₉₇X₉₈X₉₉X₁₀₀X₁₀₁X₁₀₂
 X₁₀₃X₁₀₄FRGLX₁₀₅X₁₀₆LDRLLLHX₁₀₇NX₁₀₈X₁₀₉X₁₁₀X₁₁₁VHX₁₁₂
 X₁₁₃AFX₁₁₄X₁₁₅LX₁₁₆RLX₁₁₇X₁₁₈LX₁₁₉LFX₁₂₀NX₁₂₁LX₁₂₂X₁₂₃L
 X₁₂₄X₁₂₅X₁₂₆X₁₂₇LX₁₂₈X₁₂₉LX₁₃₀X₁₃₁LX₁₃₂X₁₃₃LRLNX₁₃₄NX₁₃₅W
 X₁₃₆CX₁₃₇CRX₁₃₈RX₁₃₉LWX₁₄₀WX₁₄₁X₁₄₂X₁₄₃X₁₄₄RX₁₄₅SSSX₁₄₆
 VX₁₄₇CX₁₄₈X₁₄₉PX₁₅₀X₁₅₁X₁₅₂X₁₅₃X₁₅₄X₁₅₅DLX₁₅₆X₁₅₇LX₁₅₈X₁₅₉X₁₆₀
 DX₁₆₁X₁₆₂X₁₆₃C [SEQ ID NO:18]**

em que X é qualquer resíduo de aminoácidos ou um intervalo e em que o polipéptido não é o polipéptido de SEQ ID NO:5 (NgR1 humano) ou SEQ ID NO:17 (NgR1 de ratinho). Por exemplo, X₆, X₃₇ e X₃₈ podem representar um intervalo. Exemplos específicos de polipéptidos desta divulgação são SEQ ID NO:2 (NgR2 humano), SEQ ID NO:4 (NgR3 de ratinho) e SEQ ID NO:14 (NgR3 humano). Nalguns casos, o polipéptido contém: (a) um domínio NTLRRCT, e (b) menos de um domínio CTS completo, desde que um domínio CTS parcial, se presente, consistir em não mais do que os primeiros 39 aminoácidos do domínio CTS. Embora o polipéptido possa conter um domínio GPI funcional, um domínio GPI funcional pode estar ausente, p.ex., quando é desejado um polipéptido solúvel. Um polipéptido desta divulgação inclui opcionalmente uma sequência de aminoácidos de um polipéptido heterólogo, p.ex., uma porção Fc de um anticorpo.

A divulgação proporciona também um ácido nucleico codificando um polipéptido acima descrito; um vector contendo o ácido nucleico, ácido nucleico esse que pode estar

operativamente ligado a uma sequência de controlo da expressão; e uma célula hospedeira transformada contendo o vector. É também proporcionado um método de produção de um polipéptido desta divulgação. O método inclui a introdução de um ácido nucleico codificando o polipéptido acima descrito numa célula hospedeira, a cultura da célula sob condições adequadas para expressão do polipéptido e a recuperação do polipéptido.

A divulgação proporciona também uma molécula anti-sentido cuja sequência nucleotídica é complementar a uma sequência nucleotídica codificando um polipéptido seleccionado de entre o grupo que consiste num polipéptido consistindo nos resíduos 311-395 de SEQ ID NO:2, um polipéptido consistindo nos resíduos 256-396 de SEQ ID NO:14 e um polipéptido consistindo nos resíduos 321-438 de SEQ ID NO:4, em que o ácido nucleico tem de 8 a 100 nucleótidos de comprimento, p.ex., cerca de 20, 30, 40, 50, 60, 70, 80 ou 90 nucleótidos. A divulgação proporciona também um ácido nucleico codificando uma tal molécula anti-sentido.

A divulgação proporciona também um anticorpo que se liga a um polipéptido acima descrito. Os polipéptidos ou anticorpos desta divulgação podem ser formulados em composições farmacêuticas contendo o polipéptido ou anticorpo e um transportador farmacêuticamente aceitável.

A divulgação proporciona também um método para diminuição da inibição do crescimento axonal de um neurónio do SNC. O método inclui o passo de contacto do neurónio com uma quantidade eficaz de um polipéptido ou anticorpo desta divulgação. A divulgação proporciona também um método para tratamento de uma doença, um distúrbio ou ferimento do sistema nervoso central. O método inclui a administração a um mamífero, p.ex., um humano, de uma quantidade eficaz de um polipéptido ou anticorpo desta divulgação. Doenças, distúrbios e ferimentos exemplares que podem ser tratados utilizando moléculas e métodos desta divulgação incluem, mas não se limitam a ferimento cerebral, ferimento da medula espinal, AVC, doenças desmielinizantes, p.ex., esclerose múltipla, desmielinização monofásica, encefalomielite, leucoencefalopatia multifocal, panencefalite, doença de Marchiafava-Bignami,

degenerescência esponjosa, doença de Alexander, doença de Canavan, leucodistrofia metacromática e doença de Krabbe.

A divulgação proporciona também um método para a identificação de uma molécula que se liga a um polipéptido desta divulgação. O método inclui os passos de: (a) proporcionar um polipéptido desta divulgação; (b) contacto do polipéptido com a molécula candidata; e (c) detecção da ligação da molécula candidata ao polipéptido.

Uma parte da divulgação aqui contida proporciona os polipéptidos, polinucleótidos, vectores, células hospedeiras, anticorpos e fragmentos de ligação ao antigénio, composições, métodos e utilizações do presente invento. Estes aspectos do presente invento são expostos nas reivindicações anexas.

A presente divulgação proporciona polinucleótidos purificados e isolados (p.ex., sequências de ADN e transcritos de ARN, tanto cadeias com sentido como anti-sentido complementares, tanto de cadeia simples como de cadeia dupla, incluindo variantes de processamento destas) codificando homólogos de NgR, aqui referidos como NgR. A menos que indicado de outro modo, tal como aqui se utiliza, a abreviatura em minúsculas (NgR) refere-se a um gene, ADNc, ARN ou sequência de ácido nucleico, enquanto que a versão em maiúsculas (NgR) se refere a uma proteína, polipéptido, péptido, oligopéptido ou sequência de aminoácidos. As proteínas específicas são designadas pelo número, p.ex., "NgR2" é um homólogo de NgR humano, "NgR3" é um homólogo de NgR derivado de murídeo e "NgR1" é o NgR conhecido identificado pelo Dr. Stephen Strittmatter. Os NgR conhecidos são aqui referidos como "NgR". Os polinucleótidos de ADN desta divulgação incluindo os do presente invento, incluem ADN genómico, ADNc e ADN que foi quimicamente sintetizado em todo ou em parte.

Trabalhos de referência padrão expõem os princípios gerais da tecnologia de ADN recombinante conhecidos dos peritos na arte incluem Ausubel *et al.*, "Current Protocols in Molecular Biology", John Wiley & Sons, New York (1998); Sambrook *et al.*, "Molecular Cloning: A Laboratory Manual", 2ª Ed., Cold Spring Harbor Laboratory Press, Plainview, New York

(1989); Kaufman et al., Eds., "Handbook of Molecular and Cellular Methods in Biology and Medicine", CRC Press, Boca Raton (1995); McPherson, Ed., "Directed Mutagenesis: A Practical Approach", IRL Press, Oxford (1991).

Tal como aqui se utiliza, o termo "axónio" refere-se a uma protrusão celular longa de um neurónio, através da qual são conduzidos potenciais de acção, para ou desde o corpo celular.

Tal como aqui se utiliza, o termo "crescimento axonal" refere-se a uma extensão do longo processo ou axónio, originário no corpo celular e continuado pelo cone de crescimento.

Tal como aqui se utiliza, o termo "distúrbio do sistema nervoso central" refere-se a qualquer estado patológico associado a função anormal do sistema nervoso central (SNC). O termo inclui, mas não se limita a, funcionamento alterado do SNC resultante de traumatismo físico do tecido cerebral, infecção viral, mecanismos auto-imunes e mutação genética.

Tal como aqui se utiliza, o termo "doença desmielinizante" refere-se a um distúrbio patológico caracterizado pela degradação da bainha de mielina da membrana celular do oligodendrócito.

Tal como aqui se utiliza, o termo "cone de crescimento" refere-se uma região especializada na ponta de uma neurite em crescimento que é responsável por sentir o ambiente local e mover o axónio em direcção à sua célula alvo sináptica apropriada.

Tal como aqui se utiliza, o termo "movimento do cone de crescimento" refere-se à extensão ou colapso do cone de crescimento em direcção à célula alvo do neurónio.

Tal como aqui se utiliza, o termo "neurite" refere-se a um processo que cresce para fora de um neurónio. Uma vez que é por vezes difícil distinguir uma dendrite de um axónio em cultura, o termo "neurite" é utilizado para ambos.

Tal como aqui se utiliza, o termo “oligodendrócito” refere-se a uma célula neuroglial do SNC cuja função seja mielinizar axónios do SNC.

“Sintetizado” tal como aqui se utiliza e é entendido na arte, refere-se a polinucleótidos produzidos através de métodos puramente químicos, em oposição a enzimáticos. Sequências de ADN “inteiramente” sintetizadas são portanto produzidas inteiramente por meios químicos e os ADN “parcialmente” sintetizados englobam aqueles em que apenas porções do ADN resultante foram produzidas por meios químicos. Pelo termo “região” entenda-se uma porção fisicamente contígua da estrutura primária de uma biomolécula. No caso das proteínas, uma região é definida por uma porção contígua da sequência de aminoácidos dessa proteína. O termo “domínio” é aqui definido como referindo-se a uma parte estrutural de uma biomolécula que contribui para uma função conhecida ou suspeita da biomolécula. Os domínios podem ser co-extensivos com regiões ou porções destes, os domínios podem também incorporar uma porção de uma biomolécula que seja distinta de uma região particular, para além de toda ou parte dessa região. Exemplos de domínios da proteína NgR incluem, mas não se limitam ao péptido sinal, domínio extracelular (i.e., N-terminal), e domínios de repetição ricos em leucina.

Tal como aqui se utiliza, o termo “actividade” refere-se a uma variedade de indícios mensuráveis sugerindo ou revelando ligação, directa ou indirecta; afectando uma resposta, i.e., possuindo um efeito mensurável em resposta a alguma exposição ou estímulo, incluindo, por exemplo, a afinidade de um composto para se ligar directamente a um polipéptido ou polinucleótido desta divulgação, ou, por exemplo, a medição de quantidades de proteínas a montante ou a jusante ou outras funções semelhantes após algum estímulo ou evento. Tais actividades podem ser medidas através de ensaios tais como inibição competitiva da ligação de NgR1 a ensaios de Nogo em que, por exemplo, é adicionado NgR2 solúvel não marcado a um sistema de ensaio em concentrações crescentes para inibir a ligação de Nogo a NgR1 expresso à superfície de células CHO. Como outro exemplo, pode-se avaliar a capacidade dos neurónios para se estenderem através de lesões causadas por danos nos nervos (tal como em Schnell e Schwab, *Nature* 343: 269-272,

1990) após inibição de Nogo através de várias formas de NgR2 e/ou NgR3 como indicador biológico da função de NgR.

Tal como aqui se utiliza, pretende-se que o termo "anticorpo" se refira a anticorpos completos, intactos e Fab, Fab', F(ab)2, e outros fragmentos destes. Os anticorpos completos, intactos incluem anticorpos monoclonais tais como anticorpos monoclonais de murídeo, anticorpos quiméricos, anticorpos anti-idiotípicos, anticorpos anti-anti-idiotípicos e anticorpos humanizados.

Tal como aqui se utiliza, o termo "ligação" significa a interacção física ou química entre duas proteínas ou compostos ou proteínas ou compostos associados ou combinações destes. A ligação inclui uniões iónicas, não iónicas, de hidrogénio, de Van der Waals, interacções hidrófobas, etc. A interacção física, a ligação, pode ser directa ou indirecta, sendo a indirecta através ou devida aos efeitos de outra proteína ou composto. A ligação directa refere-se a interacções que não ocorrem através ou devido ao efeito de outra proteína ou composto mas em vez disso são sem outros intermediários químicos substanciais.

Tal como aqui se utiliza, o termo "composto" significa qualquer químico ou molécula identificável, incluindo, mas não se limitando a, moléculas pequenas, péptidos, proteínas, açúcares, nucleótidos ou ácidos nucleicos, e tal composto pode ser natural ou sintético.

Tal como aqui se utiliza, o termo "complementar" refere-se a emparelhamento de Watson-Crick entre unidades de nucleótidos de uma molécula de ácido nucleico.

Tal como aqui se utiliza, o termo "contacto" significa colocar juntos, directamente ou indirectamente, um composto na vizinhança física de um polipéptido ou polinucleótido desta divulgação. O polipéptido ou polinucleótido pode estar num qualquer número de tampões, sais, soluções, etc. O contacto inclui, por exemplo, a colocação do composto num copo, numa placa de microtitulação, num balão de cultura celular ou num *microarray*, tal como um *chip* de genes, ou semelhante, que

contenha a molécula de ácido nucleico ou o polipéptido codificando o NgR ou fragmento deste.

Tal como aqui se utiliza, a frase “sequência nucleotídica homóloga” ou “sequência de aminoácidos homóloga” ou variações destas referem-se a sequências caracterizadas por uma identidade ao nível nucleotídico, ou uma homologia ao nível dos aminoácidos, de pelo menos a percentagem especificada. Sequências nucleotídicas homólogas incluem as sequências codificando para isoformas de proteínas. Tais isoformas podem ser expressas em diferentes tecidos do mesmo organismo em resultado de: por exemplo, processamento alternativo do ARN. Alternativamente, as isoformas podem ser codificadas por diferentes genes. Sequências nucleotídicas homólogas incluem sequências nucleotídicas codificando para uma proteína de uma espécie diferente da humana, incluindo, mas não se limitando a, mamíferos. Sequências nucleotídicas homólogas incluem também, mas não se limitam a, variações alélicas de ocorrência natural e mutações das sequências nucleotídicas aqui expostas. Uma sequência nucleotídica homóloga, contudo, não inclui a sequência nucleotídica codificando NgR1. Sequências de aminoácidos homólogas incluem as sequências de aminoácidos que contêm substituições de aminoácidos conservativas e cujos polipéptidos têm a mesma ligação e/ou actividade. Uma sequência de aminoácidos homóloga não inclui, no entanto, a sequência de aminoácidos codificando outros NgR conhecidos. A percentagem de homologia pode ser determinada, por exemplo, através do programa Gap (*Wisconsin Sequence Analysis Package*, Versão 8 para Unix, Genetics Computer Group, University Research Park, Madison WI), utilizando as definições por defeito, que utiliza o algoritmo de Smith e Waterman (*Adv. Appl. Math.* 2: 482-489, 1981, que é aqui incorporado por referência na sua totalidade).

Tal como aqui se utiliza, o termo molécula de ácido nucleico “isolada” refere-se a uma molécula de ácido nucleico (ADN ou ARN) que é substancialmente livre de ácidos nucleicos codificando outras proteínas com as quais está associada na natureza, i.e., um ácido nucleico que foi removido do seu ambiente nativo. Exemplos de moléculas de ácido nucleico isoladas incluem, mas não se limitam a, moléculas de ADN recombinantes contidas num vector, moléculas de ADN

recombinantes mantidas numa célula hospedeira heteróloga, moléculas de ácido nucleico parcialmente ou substancialmente purificadas e moléculas de ADN ou ADN sintéticas. De preferência, um ácido nucleico "isolado" está livre de sequências que flanqueiam naturalmente o ácido nucleico (i.e., sequências situadas nas extremidades 5' e 3' do ácido nucleico) no ADN genómico do organismo do qual o ácido nucleico é derivado. Por exemplo, em várias concretizações, a molécula de ácido nucleico de NgR isolada pode conter menos de cerca de 50 kb, 25 kb, 5 kb, 4 kb, 3 kb, 2 kb, 1 kb, 0,5 kb ou 0,1 kb de sequências nucleotídicas que flanqueiam naturalmente a molécula de ácido nucleico no ADN genómico da célula da qual o ácido nucleico é derivado. Para além disso, uma molécula de ácido nucleico "isolada", tal como uma molécula de ADNc, pode ser substancialmente livre de outros material celular ou meio de cultura quando produzida através de técnicas recombinantes, ou de precursores químicos ou outros produtos químicos quando sintetizada quimicamente.

Tal como aqui se utiliza, o termo "heterólogo" refere-se a uma sequência nucleotídica ou de aminoácidos que é uma sequência diferente ou não correspondente, ou uma sequência derivada de uma espécie diferente. Por exemplo, uma sequência nucleotídica ou de aminoácidos de NgR de ratinho é heteróloga à sequência nucleotídica ou de aminoácidos de NgR humano, e uma sequência nucleotídica ou de aminoácidos de NgR humano é heteróloga a uma sequência nucleotídica ou de aminoácidos de uma imunoglobulina humana.

Tal como aqui se utiliza, um "polipéptido NgR solúvel" é um polipéptido NgR que não se ancora por si próprio numa membrana. Tais polipéptidos solúveis incluem, por exemplo, os polipéptidos NgR2 e NgR3 que não têm uma porção suficiente do seu sinal de âncora GPI para ancorarem o polipéptido ou são modificados para que o sinal de âncora GPI não seja adequado para resultar na substituição do péptido com uma âncora GPI. Em casos preferidos, são removidos até 5, 10, 20 ou 25 aminoácidos do terminal C de NgR2 ou NgR3 para produzir as respectivas proteínas solúveis. Tal como aqui se utiliza polipéptidos NgR solúveis incluem NgR inteiros ou truncados (p.ex., com deleções internas).

Os polipéptidos NgR solúveis podem incluir a proteína NgR inteira até à putativa sequência sinal GPI (p.ex., do aminoácido 1 até cerca do aminoácido 395 de NgR2, e do aminoácido 1 até cerca do aminoácido 438 de NgR3). Noutros casos, o péptidos sinal das proteínas pode ser removido ou truncado (p.ex., pode ser removida toda ou parte da sequência sinal de NgR2, que abrange do aminoácido 1 a cerca do aminoácido 30 de SEQ ID NO:2; pode ser removida toda ou parte da sequência sinal de NgR3, que abrange do aminoácido 1 a cerca do aminoácido 40 de SEQ ID NO:4). Nalguns casos, são utilizados o NgR2 maduro (SEQ ID NO:8) e o NgR3 maduro (SEQ ID NO:9).

Os polipéptidos NgR solúveis incluem pelo menos uma das putativas porções de ligação ao ligando de NgR, incluindo a primeira região rica em cisteína (SEQ ID NO:10, a região de repetição de leucina (SEQ ID NO:12) e a segunda região rica em cisteína (SEQ ID NO:11)). Nalguns casos, os polipéptidos NgR solúveis consistem em do aminoácido 1 até cerca do aminoácido 395 de SEQ ID NO:2, ou do aminoácido 1 até cerca do aminoácido 438 de SEQ ID NO:4.

Noutros casos, os polipéptidos NgR solúveis são proteínas de fusão que contêm os aminoácidos 30 até cerca do aminoácido 395 do NgR2 maduro ou do aminoácido 40 até cerca do aminoácido 438 de NgR3, os 10 aminoácidos C-terminais de uma região charneira da IgG1 humana contendo os dois resíduos de cisteína que se pensa que participem na ligação dissulfureto intercadeias, e as regiões CH2 e CH3 de um domínio constante da cadeia pesada da IgG1 humana. Este tipo de proteína recombinante é concebido para modular a inibição do alongamento axonal através da inibição da ligação do ligando Nogo a NgR1, ou através da inibição da interacção do ligando do NgR com o NgR da superfície celular. A porção NgR da fusão liga-se ao ligando Nogo e a porção IgG1 liga-se aos receptores FcγRI (macrófagos) e FcγIII (células NK e neutrófilos).

A produção dos polipéptidos solúveis aqui divulgada incluindo os que são úteis neste invento, pode ser alcançada através de uma variedade de métodos conhecidos na arte. Por exemplo, os polipéptidos podem ser derivados de moléculas NgR transmembranares intactas através de proteólise utilizando

endopeptidases específicas em combinação com exopeptidases, degradação de Edman, ou ambas. A molécula NgR intacta pode, por sua vez, ser purificada a partir da sua fonte natural utilizando métodos convencionais. Alternativamente, o NgR intacto pode ser produzido através de técnicas de ADN recombinante utilizando ADNc, vectores de expressão e técnicas bem conhecidas para expressão génica recombinante.

De preferência, os polipéptidos solúveis aqui divulgados incluindo os que são úteis no presente invento, são produzidos directamente, eliminando assim a necessidade de um NgR inteiro como material de partida. Isto pode ser alcançado através de técnicas convencionais de síntese química ou através de técnicas de ADN recombinante bem conhecidas em que apenas as sequências de ADN que codificam os péptidos desejados são expressas em hospedeiros transformados. Por exemplo, um gene que codifica o polipéptido NgR solúvel desejado pode ser sintetizado através de meios químicos utilizando um sintetizador de oligonucleótidos. Tais oligonucleótidos são desenhados com base na sequência de aminoácidos do polipéptido NgR solúvel desejado. Sequências específicas de ADN codificando para o péptido desejado podem também ser derivadas da sequência de ADN inteira através de isolamento de fragmentos específicos por endonucleases de restrição ou através de síntese por PCR da região especificada a partir de ADNc.

Uma molécula de ácido nucleico da presente divulgação, p.ex., uma molécula de ácido nucleico possuindo a sequência nucleotídica de SEQ ID NO:1, 3 ou um complemento de uma destas sequências nucleotídicas, pode ser isolada utilizando técnicas padrão de biologia molecular e a informação da sequência aqui proporciona. Utilizando toda ou uma porção das sequências de ácido nucleico de SEQ ID NOs:1 ou 3 como sonda de hibridação, podem ser isoladas sequências de ácido nucleico de NgR utilizando técnicas padrão de hibridação e clonagem (p.ex., tal como descrito em Sambrook et al., eds., "Molecular Cloning: A Laboratory Manual" 2ª Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989; e Ausubel, et al., eds., "Current Protocols in Molecular Biology", John Wiley & Sons, New York, NY, 1993).

Um ácido nucleico desta divulgação, p.ex. um ácido nucleico do presente invento, pode ser amplificado utilizando ADNc, ARNm ou, alternativamente, ADN genómico, como molde e iniciadores oligonucleotídicos apropriados de acordo com técnicas padrão de amplificação por PCR. O ácido nucleico assim amplificado pode ser clonado num vector apropriado e caracterizado através de análise da sequência de ADN. Para além disso, oligonucleótidos correspondentes a sequências nucleotídicas de NgR podem ser preparados através de técnicas sintéticas padrão, p.ex., utilizando um sintetizador de ADN automático.

Tal como aqui se utiliza, os termos “modula” ou “modifica” significam um aumento ou diminuição na quantidade, qualidade ou efeito de uma determinada actividade ou proteína.

Tal como aqui se utiliza, o termo “oligonucleótido” refere-se a uma série de resíduos de nucleótidos ligados que tenha um número de bases suficiente para ser utilizado numa reacção em cadeia da polimerase (PCR). Esta curta sequência é baseada em (ou desenhada a partir de) uma sequência de ADNc ou genómico e é utilizada para amplificar, confirmar ou revelar a presença de um ADN ou ARN idêntico, semelhante ou complementar numa determinada célula ou tecido. Os oligonucleótidos compreendem porções de uma sequência de ADN possuindo pelo menos cerca de 10 nucleótidos e tantos como cerca de 50 nucleótidos, de preferência cerca de 15 a 30 nucleótidos. Estes são quimicamente sintetizados e podem ser utilizados como sondas.

Tal como aqui se utiliza, o termo “sonda” refere-se a sequências de ácido nucleico de comprimento variável, de preferência entre pelo menos cerca de 10 e tantos como cerca de 6000 nucleótidos, dependendo da utilização. Estas são utilizadas na detecção de sequências de ácido nucleico idênticas, semelhantes ou complementares. Sondas mais compridas são habitualmente obtidas a partir de uma fonte natural ou recombinante, são altamente específicas e muito mais lentas a hibridar que oligómeros. Estas podem ser de cadeia simples ou dupla e cuidadosamente desenhadas para terem especificidade em tecnologias de PCR, baseadas em membranas de hibridação ou semelhantes a ELISA.

O termo "prevenção" refere-se à diminuição da probabilidade de um organismo contrair ou desenvolver uma condição anormal.

O termo "tratamento" refere-se a possuir um efeito terapêutico e pelo menos parcialmente aliviar ou abolir uma condição anormal no organismo.

O termo "efeito terapêutico" refere-se aos factores de inibição ou activação que causam ou contribuem para a condição anormal. Um efeito terapêutico alivia até certo ponto um ou mais sintomas da condição anormal. Em referência ao tratamento de condições anormais, um efeito terapêutico pode referir-se a um ou mais dos seguintes: (a) um aumento na proliferação, no crescimento e/ou na diferenciação de células; (b) inibição (i.e., abrandamento ou paragem) de morte celular; (c) inibição de degenerescência; (d) alívio até certo ponto de um ou mais dos sintomas associados à condição anormal; e (e) aumento da função da população de células afectada. Compostos demonstrando eficácia contra condições anormais podem ser identificados tal como aqui descrito.

O termo "condição anormal" refere-se a uma função nas células ou tecidos de um organismo que se desvia das suas funções normais nesse organismo. Uma condição anormal pode relacionar-se com proliferação celular, diferenciação celular, sinalização celular ou sobrevivência celular. Uma condição anormal pode também incluir obesidade, complicações diabéticas tais como degenerescência da retina e irregularidades na tomada e no metabolismo da glicose e na tomada e no metabolismo de ácidos gordos.

Condições proliferativas celulares anormais, por exemplo, incluem cancros tais como distúrbios fibróticos e mesangiais, angiogénese anormal e vasculogénese, cicatrização de feridas, psoríase, diabetes *mellitus* e inflamação.

Condições de diferenciação anormais incluem, por exemplo, distúrbios neurodegenerativos, velocidades lentas de cicatrização de feridas e velocidades lentas de cicatrização de enxertos de tecido.

Condições de sinalização anormais incluem, por exemplo, distúrbios psiquiátricos envolvendo excesso de actividade de neurotransmissores.

Condições de sobrevivência celular anormais podem também referir-se a condições nas quais as vias de morte celular programada (apoptose) são activadas ou abolidas. Várias proteína-quinases estão associadas às vias de apoptose. Aberrações na função de qualquer uma das proteína-quinases podem conduzir a imortalidade celular ou morte celular prematura.

O termo "administração" refere-se a um método de incorporação de um composto em células ou tecidos de um organismo. A condição anormal pode ser prevenida ou tratada quando as células ou tecidos do organismo existem dentro do organismo ou fora do organismo. As células existentes fora do organismo podem ser mantidas ou criadas em caixas de cultura celular. Para células ancoradas dentro do organismo, existem muitas técnicas na arte para administrar compostos, incluindo (mas não se limitando a) aplicações orais, parentéricas, dérmicas, de injeção e aerossol. Para células fora do organismo, existem múltiplas técnicas na arte para administrar os compostos, incluindo (mas não se limitando a) técnicas de micro-injeção celular, técnicas de transformação e técnicas de transportadores.

A condição anormal pode também ser prevenida ou tratada através de administração de um composto a um grupo de células possuindo uma aberração numa via de transdução de sinal a um organismo. O efeito de administração de um composto na função do organismo pode ser monitorizada. O organismo é de preferência um ratinho, rato, coelho, cobaia ou cabra, de maior preferência um macaco e hominóide, e de maior preferência um humano.

Por "amplificação" entenda-se números crescentes de ADN ou ARN numa célula em comparação com células normais. "Amplificação" quando se refere a ARN pode ser a presença detectável de ARN em células, uma vez que nalgumas células normais não há expressão basal de ARN. Noutras células

normais, existe um nível basal de expressão, portanto nestes casos a amplificação é a detecção de pelo menos 1-2 vezes e de preferência mais, em comparação com o nível basal.

As sequências de aminoácidos são apresentadas na direcção amino para carboxilo, da esquerda para a direita. Os grupos amino e carboxilo não são apresentados na sequência. As sequências nucleotídicas são apresentadas apenas através da cadeia simples, na direcção 5' para 3', da esquerda para a direita. Os nucleótidos e aminoácidos são representados da maneira recomendada pela Comissão de Nomenclatura Bioquímica IUPAC-IUB ou (para os aminoácidos) através do código de três letras.

Ácidos Nucleicos

O ADN genómico desta divulgação compreende a região de codificação de proteína para um polipéptido desta divulgação e pretende-se também que inclua variantes alélicas desta. É amplamente entendido que, para muitos genes, o ADN genómico é transcrito em transcritos de ARN que sofrem um ou mais eventos de processamento em que os intrões (i.e., regiões não codificadoras) dos transcritos são removidos ou "cortados". Os transcritos de ARN que podem ser processados através de mecanismos alternativos e portanto ser sujeitos a remoção de diferentes sequências de ARN mas ainda codificar um polipéptido NgR, são referidos na arte como variantes de processamento que são englobadas por esta divulgação. As variantes de processamento compreendidas pela divulgação são portanto codificadas pelas mesmas sequências de ADN genómico originais mas surgem de transcritos de ARNm distintos. As variantes alélicas são formas modificadas de uma sequência génica de tipo selvagem, resultando a modificação de recombinação durante segregação cromossómica ou exposição a condições que dão origem a mutação genética. Variantes alélicas, como genes de tipo selvagem, são sequências de ocorrência natural (em oposição a variantes de ocorrência natural que surgem de manipulação *in vitro*).

A divulgação compreende também ADNc que é obtido através de transcrição reversa de um polinucleótido de ARN codificando NgR (convencionalmente seguido de síntese da segunda cadeia de

uma cadeia complementar para proporcionar um ADN de cadeia dupla).

As sequências de ADN preferidas codificando um polipéptido NgR humano são expostas em SEQ ID NO:1 e 13. Um ADN preferido desta divulgação compreende uma molécula de cadeia dupla compreendendo a molécula de codificação (i.e., a "cadeia de codificação") juntamente com a molécula complementar (a "cadeia não codificadora" ou "complemento") possuindo uma sequência inequivocamente dedutível a partir da cadeia de codificação de acordo com as regras de emparelhamento de bases de Watson-Crick para ADN. São também preferidos outros polinucleótidos codificando polipéptidos NgR, tal como mostrado em SEQ ID NO:3, que compreendem o homólogo de NgR de murídeo, NgR3.

São também preferidas sequências nucleotídicas que codifiquem pelo menos uma porção de um polipéptido NgR que tenha pelo menos uma função biológica de um NgR. Mais preferidas são sequências nucleotídicas que codifiquem uma porção de NgR que codifique pelo menos o NgR maduro sem o sinal GPI C-terminal hidrófobo. Também preferidas são sequências nucleotídicas que codifiquem a porção de NgR que codifique pelo menos a região de ligação ao ligando de NgR.

Esta divulgação engloba ainda homólogos de outras espécies, de preferência de mamífero, do ADN de NgR humano. Os homólogos de espécies, por vezes referidos como "ortólogos" partilham, em geral, pelo menos 35%, pelo menos 40%, pelo menos 45%, pelo menos 50%, pelo menos 60%, pelo menos 65%, pelo menos 70%, pelo menos 75%, pelo menos 80%, pelo menos 85%, pelo menos 90%, pelo menos 95%, pelo menos 98% ou pelo menos 99% de homologia com ADN humano desta divulgação. Geralmente, a percentagem de "homologia" de sequência em relação aos polinucleótidos da divulgação pode ser calculada como a percentagem de bases nucleotídicas na sequência candidata que são idênticas aos nucleótidos nas sequências NgR expostas em SEQ ID NOS:1, 3 ou 13, após alinhamento das sequências e introdução de intervalos, se necessário, para alcançar a percentagem máxima de identidade de sequência.

A informação da sequência polinucleotídica proporcionada por esta divulgação torna possível a expressão em larga escala do polipéptido codificado através de técnicas bem conhecidas e rotineiramente praticadas na arte. Os polinucleótidos desta divulgação permitem também a identificação e o isolamento de polinucleótidos codificando polipéptidos NgR aparentados, tais como variantes alélicas e homólogos de espécies, através de técnicas bem conhecidas incluindo hibridação *Southern* e/ou *Northern*, e reacção em cadeia da polimerase (PCR). Exemplos de polinucleótidos aparentados incluem sequências genómicas humanas e não humanas, incluindo variantes alélicas, bem como polinucleótidos codificando polipéptidos homólogos a NgR e polipéptidos estruturalmente aparentados partilhando uma ou mais propriedades biológicas, imunológicas e/ou físicas de NgR. Genes de espécies não humanas codificando proteínas homólogas a NgR podem também ser identificados através de análise *Southern* e/ou PCR e são úteis em modelos animais para distúrbios de NgR. O conhecimento da sequência de um ADN de NgR humano torna também possível através da utilização de hibridação *Southern* ou reacção em cadeia da polimerase (PCR) a identificação de sequências de ADN genómico codificando sequências reguladoras de controlo da expressão de NgR tais como promotores, operadores, estimuladores, repressores e semelhantes. Os polinucleótidos desta divulgação incluindo os do presente invento, são também úteis em ensaios de hibridação para detectar a capacidade das células para expressar NgR. Os polinucleótidos desta divulgação, incluindo os do presente invento, podem também proporcionar uma base para métodos de diagnóstico úteis para identificação de uma ou mais alterações genéticas num *locus* de NgR que estejam subjacentes a um estado ou estados de doença, cuja informação seja útil tanto para diagnóstico como para selecção de estratégias terapêuticas.

A divulgação aqui de um polinucleótido inteiro codificando um polipéptido NgR torna facilmente disponível ao vulgar perito na arte qualquer fragmento possível do polipéptido inteiro. Esta divulgação proporciona, portanto, fragmentos de polinucleótidos codificando NgR compreendendo pelo menos 6, e de preferência pelo menos 14, 16, 18, 20, 25, 50 ou 75 nucleótidos consecutivos de um polinucleótido codificando NgR. De preferência, os fragmentos de polinucleótidos desta divulgação compreendem sequências únicas

da sequência polinucleotídica codificando NgR, e portanto, hibridam apenas sob condições altamente rigorosas ou moderadamente rigorosas (i.e., "especificamente") com polinucleótidos codificando NgR (ou fragmentos deste). Os fragmentos polinucleotídicos das sequências genómicas desta divulgação compreendem não só sequências únicas à região de codificação, como incluem também fragmentos da sequência inteira derivados de intrões, regiões reguladoras e/ou outras sequências não traduzidas. Sequências únicas aos polipéptidos desta divulgação são reconhecíveis através de comparação de sequências com outros polinucleótidos conhecidos, e podem ser identificados através da utilização de programas de alinhamento rotineiramente utilizados na arte, p.ex., os tornados disponíveis em bases de dados de sequências públicas. Tais sequências são também reconhecíveis a partir de análises de hibridação *Southern* para determinar o número de fragmentos de ADN genómico com os quais hibridará um polinucleótido. Os polinucleótidos desta divulgação, incluindo os do presente invento, podem ser marcados de um modo que permita a sua detecção, incluindo marcação radioactiva, fluorescente e enzimática.

Os fragmentos de polinucleótidos são particularmente úteis como sondas para detecção de polinucleótidos NgR inteiros ou fragmentados. Um ou mais polinucleótidos podem ser incluídos em estojos que são utilizados para detectar a presença de um polinucleótido codificando NgR ou utilizados para detectar variações numa sequência polinucleotídica codificando NgR.

Esta divulgação engloba também ADN codificando polipéptidos NgR que hibridam sob condições moderadamente rigorosas ou altamente rigorosas com a cadeia não codificadora, ou complemento, do polipéptido em qualquer uma de SEQ ID NO:1 ou 3.

As condições rigorosas são conhecidas dos peritos na arte e podem ser verificadas em "Current Protocols in Molecular Biology", John Wiley & Sons, N.Y (1998), 6.3.1-6.3.6. De preferência, as condições são tais que sequências pelo menos cerca de 65%, 70%, 75%, 85%, 90%, 95%, 98% ou 99% homólogas uma à outra permanecem tipicamente hibridadas uma com a outra.

Um exemplo não limitante de condições de hibridação rigorosas é a hibridação num tampão de elevada salinidade compreendendo SSC 6x, Tris-HCl 50 mM (pH 7,5), EDTA 1 mM, PVP a 0,02%, Ficoll a 0,02%, BSA a 0,02% e 500 mg/ml de ADN de esperma de salmão desnaturado a 65°C. Esta hibridação é seguida de uma ou mais lavagens em SSC 0,2x, BSA a 0,01% a 50°C. Uma molécula de ácido nucleico isolada desta divulgação que hibride sob condições rigorosas com a sequência de SEQ ID NO:1 ou 3 corresponde a uma molécula de ácido nucleico de ocorrência natural. Tal como aqui se utiliza, uma molécula de ácido nucleico de "ocorrência natural" refere-se a uma molécula de ARN ou ADN possuindo uma sequência nucleotídica que ocorre na natureza (p.ex., codifica uma proteína natural). Tal como aqui se utiliza, "condições de hibridação rigorosas" significa: 42°C numa solução de hibridação compreendendo formamida a 50%, SDS a 1%, NaCl 1 M, sulfato de dextrano a 10% (p/vol), e lavagem duas vezes durante 30 minutos a 60°C numa solução de lavagem compreendendo SSC 0,1x e SDS a 1%.

Vectores

Outro aspecto da presente divulgação é dirigido a vectores ou vectores de expressão recombinante, compreendendo qualquer uma das moléculas de ácido nucleico descritas acima, p.ex. vectores compreendendo os polinucleótidos do presente invento. Os vectores são aqui utilizados para amplificar ADN ou ARN codificando NgR e/ou para expressar ADN que codifique NgR. Tal como aqui se utiliza, o termo "vector" refere-se a uma molécula de ácido nucleico capaz de transportar outro ácido nucleico ao qual foi ligada. Um tipo de vector é um "plasmídeo", que se refere a uma volta de ADN de cadeia dupla circular na qual podem ser ligados segmentos de ADN adicionais. Outro tipo de vector é um vector viral, em que segmentos de ADN adicionais podem ser ligados no genoma viral. Certos vectores são capazes de replicação autónoma numa célula hospedeira na qual são introduzidos (p.ex., vectores bacterianos possuindo uma origem de replicação bacteriana e vectores de mamífero epissómicos). Outros vectores (p.ex., vectores de mamífero não epissómicos) são integrados no genoma de uma célula hospedeira após introdução na célula hospedeira, e são deste modo replicados juntamente com o genoma do hospedeiro. Para além disso, certos vectores são capazes de dirigir a expressão de genes aos quais estão operativamente

ligados. Tais vectores são aqui referidos como "vectores de expressão". Em geral, os vectores de expressão de utilidade em técnicas de ADN recombinante estão frequentemente na forma de plasmídeos. No presente fascículo, "plasmídeo" e "vector" podem ser utilizados indiferentemente uma vez que o plasmídeo é a forma de vector mais vulgarmente utilizada. No entanto, pretende-se que esta divulgação e o presente invento incluam tais outras formas de vectores de expressão, tais como vectores virais (p.ex., retrovírus deficientes na replicação, adenovírus e vírus adeno-associados), que têm funções equivalentes.

A expressão de proteínas em procariotas é muito frequentemente efectuada em *E. coli* com vectores contendo promotores constitutivos ou indutíveis dirigindo a expressão de proteínas de fusão ou sem serem de fusão. Os vectores de fusão adicionam vários aminoácidos a uma proteína aí codificada, habitualmente ao terminal amino da proteína recombinante. Tais vectores de fusão têm tipicamente três objectivos: (1) aumentar a expressão da proteína recombinante; (2) aumentar a solubilidade da proteína recombinante; e (3) ajudar na purificação da proteína recombinante actuando como um ligando em purificação por afinidade. Frequentemente, em vectores de expressão de fusão, é introduzido um local de clivagem proteolítica na junção da porção de fusão e da proteína recombinante para permitir a separação da proteína recombinante da porção de fusão subsequente à purificação da proteína de fusão. Tais enzimas e suas sequências de reconhecimento cognatas incluem Factor Xa, trombina e enteroquinase. Vectores de expressão de fusão típicos incluem pGEX (Pharmacia Biotech Inc; Smith e Johnson, *Gene* 67: 31-40, 1988), pMAL (New England Biolabs, Beverly, Mass.) e pRIT5 (Pharmacia, Piscataway, N.J.) que fundem glutathione-S-transferase (GST), proteína de ligação a maltose E, ou proteína A, respectivamente, com a proteína recombinante alvo.

Exemplos de vectores de expressão de *E. coli* sem serem de fusão indutíveis adequados incluem pTrc (Amrann *et al.*, *Gene* 69: 301-315, 1988) e pET11d (Studier *et al.*, "Gene Expression Technology: Methods in Enzymology" 185, Academic Press, San Diego, CA. (1990) 60-89).

Uma estratégia para maximizar a expressão de proteínas recombinantes em *E. coli* é expressar a proteína em bactérias hospedeiras com uma capacidade diminuída para clivar proteoliticamente a proteína recombinante. Ver, Gottesman, "Gene Expression Technology: Methods in Enzymology" 185, Academic Press, San Diego, CA. (1990) 119-128. Outra estratégia é alterar a sequência de ácido nucleico do ácido nucleico a inserir num vector de expressão para que os codões individuais para cada aminoácido sejam utilizados preferencialmente em *E. coli* (Wada *et al.*, *Nucleic Acids Res.* 20: 2111-2118, 1992). Tal alteração das sequências de ácido nucleico desta divulgação pode ser realizada através de técnicas padrão de síntese de ADN.

Noutros casos, o vector de expressão de NgR é um vector de expressão de levedura. Exemplos de vectores para expressão em levedura *S. cerevisiae* incluem pYepSec1 (Baldari, *et al.*, *EMBO J.* 6: 229-234, 1987), pMFa (Kurjan e Herskowitz, *Cell* 30: 933-943, 1982), pJRY88 (Schultz *et al.*, *Gene* 54: 113-123, 1987), pYES2 (Invitrogen Corporation, San Diego, CA), e picZ (Invitrogen Corp, San Diego, CA).

Alternativamente, NgR pode ser expresso em células de insecto utilizando vectores de expressão de baculovírus. Vectores de baculovírus disponíveis para expressão de proteínas em células de insecto em cultura (p.ex., células SF9) incluem as séries pAc (Smith *et al.*, *Mol. Cell. Biol.* 3: 2156-2165, 1983) e a série pVL (Lucklow e Summers, *Virology* 170: 31-39, 1989).

Ainda noutros casos, um ácido nucleico desta divulgação, p.ex. um ácido nucleico do presente invento, é expresso em células de mamífero utilizando um vector de expressão de mamífero. Exemplos de vectores de expressão de mamífero incluem pCDM8 (Seed, *Nature* 329: 840, 1987) e pMT2PC (Kaufman *et al.*, *EMBO J.* 6: 187-195, 1987). Quando utilizado em células de mamífero, as funções de controlo do vector de expressão são frequentemente proporcionadas por elementos reguladores virais. Por exemplo, os promotores vulgarmente utilizados são derivados de políoma, adenovírus 2, citomegalovírus e Vírus Símio 40. Para outros sistemas de expressão adequados para células procarióticas e eucarióticas, ver, p.ex., Capítulos 16

e 17 de Sambrook et al., (Eds.) "Molecular Cloning: A Laboratory Manual. 2ª Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

Noutros casos, o vector de expressão de mamífero recombinante é capaz de dirigir a expressão do ácido nucleico preferencialmente num determinado tipo celular (p.ex., são utilizados elementos reguladores específicos de tecido para expressar o ácido nucleico). Elementos reguladores específicos de tecido são conhecidos na arte. Exemplos não limitantes de promotores específicos de tecido adequados incluem o promotor da albumina (específico do fígado; Pinkert et al. *Genes Dev.* 1: 268-277, 1987), promotores específicos linfóides (Calame e Eaton, *Adv. Immunol.* 43: 235-275, 1988), em particular promotores de receptores de células T (Winoto e Baltimore, *EMBO J.* 8: 729-733, 1989) e imunoglobulinas (Banerji et al., *Cell* 33: 729-740, 1983; Queen e Baltimore, *Cell* 33: 741-748, 1983), promotores específicos de neurónios (p.ex., o promotor do neurofilamento; Byrne e Ruddle, *Proc. Natl. Acad. Sci. USA* 86: 5473-5477, 1989), promotores específicos do pâncreas (Edlund et al., *Science* 230: 912-916, 1985), e promotores específicos da glândula mamária (p.ex., promotor do soro do leite; Pat. U.S. N.º 4873316 e Publicação de Pedido Europeu N.º 264166). Promotores regulados no desenvolvimento estão também englobados, p.ex., os promotores *hox* de murídeo (Kessel e Gruss, *Science* 249: 374-379, 1990) e o promotor da α -fetoproteína (Campes e Tilghman, *Genes Dev.* 3: 537-546, 1989).

A divulgação proporciona ainda um vector de expressão recombinante compreendendo uma molécula de ADN desta divulgação clonada no vector de expressão numa orientação anti-sentido. Isto é, a molécula de ADN é operativamente ligada a uma sequência reguladora de um modo que permita a expressão (através de transcrição da molécula de ADN) de uma molécula de ARN que seja ARNm anti-sentido de NgR. Podem ser escolhidas sequências reguladoras operativamente ligadas a um ácido nucleico clonado na orientação anti-sentido que dirijam a expressão contínua da molécula de ARN anti-sentido numa variedade de tipos celulares, por exemplo podem ser escolhidos promotores e/ou estimuladores virais ou sequências reguladoras que dirijam a expressão directa constitutiva, específica de um

tecido ou específica de um tipo celular do ARN anti-sentido. O vector de expressão anti-sentido pode estar na forma de um plasmídeo recombinante, fagemídeo ou vírus atenuado no qual ácidos nucleicos anti-sentido sejam produzidos sob o controlo de uma região reguladora de elevada eficiência, cuja actividade possa ser determinada pelo tipo celular no qual o vector é introduzido. Para uma discussão da regulação da expressão génica utilizando genes anti-sentido ver Weintraub *et al.*, "Antisense RNA as a molecular tool for genetic analysis", "REVIEWS-TRENDS IN GENETICS", Vol. 1(1) 1986.

Vectores preferidos incluem, mas não se limitam a, plasmídeos, fagos, cosmídeos, episomas, partículas virais ou vírus e fragmentos de ADN integráveis (i.e., fragmentos integráveis no genoma hospedeiro através de recombinação homóloga). Partículas virais incluem, mas não se limitam a, adenovírus, baculovírus, parvovírus, herpesvírus, poxvírus, vírus adeno-associados, vírus de Semliki Forest, vírus vacínia e retrovírus. Vectores de expressão preferidos incluem, mas não se limitam a, pcDNA3 (Invitrogen) e pSVL (Pharmacia Biotech). Outros vectores de expressão incluem, mas não se limitam a, vectores pSPORTTM, vectores pGEMTM (Promega), vectores pPROEXTM (LTI, Bethesda, MD), vectores BluescriptTM (Stratagene), vectores pQETM (Qiagen), pSE420TM (Invitrogen) e pYES2TM (Invitrogen).

Vectores de expressão preferidos são construções de ADN replicáveis nas quais uma sequência de ADN codificando NgR está operativamente ligada ou unida a sequências de controlo adequadas capazes de efectuar a expressão do NgR num hospedeiro adequado. As regiões de ADN estão operativamente ligadas ou unidas quando estão funcionalmente relacionadas umas com as outras. Por exemplo, um promotor está operativamente ligado ou unido a uma sequência de codificação se controlar a transcrição da sequência. Os vectores de amplificação não requerem domínios de controlo da expressão, mas em vez disso necessitam apenas da capacidade de se replicarem num hospedeiro, habitualmente conferida por uma origem de replicação e um gene de selecção para facilitar o reconhecimento de transformantes. A necessidade de sequências de controlo no vector de expressão variará dependendo do hospedeiro seleccionado e do método de transformação

escolhido. Geralmente, as sequências de controlo incluem, mas não se limitam a, um promotor da transcrição, estimuladores e uma sequência operadora opcional para controlar a transcrição, sinais de poliadenilação, uma sequência codificando uma ligação adequada ao ribossoma do ARNm e sequências que controlam a terminação da transcrição e da tradução. Tais sequências reguladoras são descritas, por exemplo, em Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Sequências reguladoras incluem as que dirigem a expressão constitutiva de uma sequência nucleotídica em muito tipos de células hospedeiras e as que dirigem a expressão da sequência nucleotídica apenas em certas células hospedeiras (p.ex., sequências reguladoras específicas de tecido). Será apreciado pelos peritos na arte que o desenho do vector de expressão pode depender de factores tais como a escolha da célula hospedeira a transformar, o nível de expressão da proteína desejada, etc. Os vectores de expressão desta divulgação podem ser introduzidos em células hospedeiras para produzir deste modo proteínas ou péptidos, incluindo proteínas ou péptidos de fusão, codificados por ácidos nucleicos tal como aqui descrito (p.ex., proteínas NgR, formas mutantes de NgR, proteínas de fusão, etc.).

Os vectores preferidos contêm de preferência um promotor que é reconhecido pelo organismo hospedeiro. As sequências promotoras podem ser procarióticas, eucarióticas ou virais. Exemplos de sequências procarióticas adequadas incluem os promotores PR e PL do bacteriófago lambda ("The Bacteriophage Lambda, Hershey, A.D. (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1973), que é aqui incorporado por referência na sua totalidade; "Lambda II", Hendrix, R.W. (Ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1980), que é aqui incorporado por referência na sua totalidade); os promotores *trp*, *recA*, de choque térmico e *lacZ* de *E. coli* e o promotor inicial de SV40 (Benoist et al., *Nature* 290: 304-310, 1981). Promotores adicionais incluem, mas não se limitam aos do vírus do tumor mamário de ratinho, da repetição terminal longa do vírus da imunodeficiência humana, do vírus de maloney, promotor inicial imediato de citomegalovírus, do vírus de Epstein-Barr, do vírus do sarcoma de Rous, da actina humana, da miosina humana, da hemoglobina

humana, da creatina muscular humana e da metalotioneína humana.

Sequências reguladoras adicionais podem também ser incluídas nos vectores preferidos. Exemplos preferidos de sequências reguladoras adequadas são representados pela sequência de Shine-Dalgarno do gene da replicase do fago MS-2 e do gene cII do bacteriófago lambda. A sequência de Shine-Dalgarno pode ser directamente seguida por ADN codificando NgR e resultar na expressão da proteína NgR madura.

Para além disso, vectores de expressão adequados podem incluir um marcador apropriado que permite a pesquisa das células hospedeiras transformadas. A transformação do hospedeiro seleccionado é efectuada utilizando qualquer uma das várias técnicas bem conhecidas dos peritos na arte e descritas em Sambrook *et al.*, *supra*.

Uma origem de replicação pode também ser proporcionada através da construção do vector para incluir uma origem exógena ou pode ser proporcionada através do mecanismo de replicação cromossómica da célula hospedeira. Se o vector for integrado no cromossoma da célula hospedeira, o último pode ser suficiente. Alternativamente, em vez de utilizar vectores que contêm origens virais de replicação, um perito na arte pode transformar células de mamífero através do método de co-transformação com um marcador seleccionável e ADN de NgR. Um exemplo de um marcador adequado é di-hidrofolato-redutase (DHFR) ou timidina-quinase (ver, Patente U.S. N.º 4399216).

As sequências nucleotídicas codificando NgR podem ser recombinadas com o ADN vector de acordo com técnicas convencionais, incluindo terminais tornados cegos ou tornados desfasados para ligação, digestão com enzimas de restrição para proporcionar terminais apropriados, preenchimento de extremidades coesivas conforme apropriado, tratamento com fosfatase alcalina para evitar união indesejável e ligação com ligases apropriadas. Técnicas para tal manipulação são divulgadas por Sambrook *et al.*, *supra* e são bem conhecidas na arte. Métodos para construção de vectores de expressão de mamífero são divulgados, por exemplo, em Okayama *et al.*, *Mol. Cell. Biol.* 3: 280, 1983, Cosman *et al.* *Mol. Immunol.* 23: 935,

1986, Cosman *et al.*, *Nature* 312: 768, 1984, EP-A-0367566 e WO 91/18982.

Células Hospedeiras e Células Hospedeiras Transformadas

De acordo com outro aspecto desta divulgação e do presente invento, são proporcionadas células hospedeiras, incluindo células procarióticas e eucarióticas, compreendendo um polinucleótido ou um vector desta divulgação, p.ex. um polinucleótido do presente invento ou um vector do presente invento de um modo que permita a expressão do polipéptido NgR codificado. De preferência, a célula produz pouco ou nenhum polipéptido NgR endógeno. Os polinucleótidos desta divulgação, incluindo os do presente invento, podem ser introduzidos na célula hospedeira como parte de um plasmídeo circular, ou como ADN linear compreendendo uma região de codificação de uma proteína isolada ou um vector viral. Métodos para introdução de ADN na célula hospedeira que são bem conhecidos e rotineiramente praticados na arte incluem transformação, transfecção, electroporação, injeção nuclear ou fusão com transportadores tais como lipossomas, micelas, células fantasma e protoplastos. Sistemas de expressão desta divulgação, incluindo os do presente invento, incluem sistemas celulares bacterianos, de levedura, fúngicos, vegetais, de insecto, de invertebrado, de vertebrado e de mamífero.

As células hospedeiras desta divulgação, incluindo as do presente invento, são uma fonte valiosa de imunogénio para o desenvolvimento de anticorpos especificamente imunorreactivos com NgR. As células hospedeiras desta divulgação, incluindo as do presente invento, são também úteis em métodos para a produção em larga escala de polipéptidos NgR em que as células são criadas num meio de cultura adequado e os produtos de polipéptido desejados são isolados a partir das células ou do meio em que as células são criadas, através de métodos de purificação conhecidos na arte, p.ex., métodos cromatográficos convencionais incluindo cromatografia de imunoafinidade, cromatografia de afinidade de receptores, cromatografia de interacção hidrófoba, cromatografia de afinidade de lectina, filtração por exclusão de tamanho, cromatografia de permuta cationica ou aniónica, cromatografia líquida de alta pressão (HPLC), HPLC de fase inversa e semelhantes. Ainda outros métodos de purificação incluem os métodos em que a proteína

desejada é expressa e purificada como uma proteína de fusão possuindo uma porção etiqueta, marcadora ou quelante específica que seja reconhecida por um parceiro ou agente de ligação específica. A proteína purificada pode ser clivada para produzir a proteína desejada ou pode ser deixada como uma proteína de fusão intacta. A clivagem da componente de fusão pode produzir uma forma da proteína desejada possuindo resíduos de aminoácidos adicionais em resultado do processo de clivagem.

O conhecimento das sequências de ADN de NgR permite a modificação de células para permitir ou aumentar a expressão de NgR endógena. As células podem ser modificadas (p.ex., através de recombinação homóloga) para proporcionar maior expressão através da substituição, no todo ou em parte, do promotor de NgR de ocorrência natural com todo ou parte de um promotor heterólogo para que as células expressem NgR a níveis mais elevados. O promotor heterólogo é inserido de um modo tal que fica operativamente ligado a sequências endógenas codificando NgR. (Ver, por exemplo, Publicação Internacional PCT N.º WO 94/12650, Publicação Internacional PCT N.º WO 92/20808, e Publicação Internacional PCT N.º WO 91/09955). Está também contemplado que, para além do ADN promotor heterólogo, ADN marcador amplificável (p.ex., *ada*, *dhfr* e o gene *CAD* multifuncional que codifica carbamoílfosfato-sintetase, aspartato-transcarbamilase e di-hidro-orotase) e/ou ADN de intrões podem ser inseridos juntamente com o ADN promotor heterólogo. Se ligados à sequência de codificação de NgR, a amplificação do ADN marcador através de métodos padrão de selecção resulta em co-amplificação das sequências de codificação de NgR nas células.

A informação da sequência de ADN proporcionada pela presente divulgação torna também possível o desenvolvimento (p.ex., através de recombinação homóloga ou estratégias "*knock-out*"; ver Capecchi, *Science* 244: 1288-1292, 1989) de animais que não conseguem expressar NgR funcional ou que expressam uma variante de NgR. Tais animais (especialmente pequenos animais de laboratório tais como ratos, coelhos e ratinhos) são úteis como modelos para estudo das actividades *in vivo* de NgR e moduladores de NgR.

Células hospedeiras adequadas para expressão dos polipéptidos desta divulgação, incluindo os polipéptidos do presente invento, incluem, mas não se limitam a, procariotas, levedura e eucariotas. Se for empregue um vector de expressão procariótico, então a célula hospedeira apropriada será qualquer célula procariótica capaz de expressar as sequências clonadas. Células procarióticas adequadas incluem, mas não se limitam a, bactérias do género *Escherichia*, *Bacillus*, *Salmonella*, *Pseudomonas*, *Streptomyces* e *Staphylococcus*.

Se for empregue um vector de expressão eucariótico, então a célula hospedeira apropriada será qualquer célula eucariótica capaz de expressar a sequência clonada. De preferência, células eucarióticas são células de eucariotas superiores. Células eucarióticas adequadas incluem, mas não se limitam a, células de cultura de tecidos de mamífero não humano e células de cultura de tecidos humanos. As células hospedeiras preferidas incluem, mas não se limitam a, células de insecto, células HeLa, células de ovário de hamster chinês (CHO), células de rim de macaco verde africano (células COS), células 293 humanas, e fibroblastos 3T3 de murídeo. A propagação de tais células em cultura celular tornou-se um procedimento de rotina (ver, "Tissue Culture", Academic Press, Kruse e Patterson, Eds. (1973)).

Adicionalmente, pode ser empregue uma célula de levedura como célula hospedeira. As células de levedura preferidas incluem, mas não se limitam aos géneros *Saccharomyces*, *Pichia* e *Kluveromyces*. Os hospedeiros de levedura preferidos são *S. cerevisiae* e *P. pastoris*. Os vectores de levedura preferidos podem conter uma sequência de origem de replicação de um plasmídeo de levedura 2T, uma sequência de replicação autónoma (ARS), uma região promotora, sequências para poliadenilação, sequências para terminação da transcrição e um gene de um marcador seleccionável. Vectores vaivém para replicação tanto em levedura como em *E. coli* são também incluídos aqui.

Alternativamente, podem ser utilizadas células de insecto como células hospedeiras. Num caso preferido, os polipéptidos desta divulgação, incluindo os do presente invento, são expressos utilizando um sistema de expressão de baculovírus (ver, Luckow et al., *Bio/Technology* 6: 47, 1988; "Baculovirus

Expression Vectors: A Laboratory Manual", O'Rielly et al. (Eds.), W.H. Freeman and Company, New York, 1992; e Patente U.S. N.º 4879236). Adicionalmente, o sistema de expressão de baculovírus completo MAXBAC™ (Invitrogen) pode, por exemplo, ser utilizado para produção em células de insecto.

Células hospedeiras adequadas são melhor discutidas em Goeddel, "Gene Expression Technology: Methods In Enzymology" 185, Academic Press, San Diego, CA (1990). Alternativamente, o vector de expressão recombinante pode ser transcrito e traduzido *in vitro*, por exemplo utilizando sequências reguladoras promotoras T7 e polimerase T7.

O ADN vector pode ser introduzido em células procarióticas ou eucarióticas através de técnicas convencionais de transformação ou transfecção. Tal como aqui se utilizam, pretende-se que os termos "transformação" e "transfecção" refiram uma variedade de técnicas reconhecidas na arte para introdução do ácido nucleico estranho (p.ex., ADN) numa célula hospedeira, incluindo co-precipitação em fosfato de cálcio ou cloreto de cálcio, transfecção mediada por DEAE-dextrano, lipofecção ou electroporação. Métodos adequados para transformação ou transfecção de células hospedeiras podem ser encontrados em Sambrook, et al. (Molecular Cloning: A Laboratory Manual. 2ª ed., Cold Spring Harbor Laboratory, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1989), e outros manuais de laboratório.

Para transfecção estável de células de mamífero, sabe-se que, dependendo do vector de expressão e da técnica de transfecção utilizados, apenas uma pequena fracção de células pode integrar o ADN estranho no seu genoma. Para identificar e seleccionar estes integrantes, um gene que codifica um marcador seleccionável (p.ex., resistência a antibióticos) é geralmente introduzido nas células hospedeiras juntamente com o gene de interesse. Vários marcadores seleccionáveis incluem os que conferem resistência a fármacos, tais como G418, higromicina, di-hidrofolato-redutase (DHFR) e metotrexato. Ácido nucleico codificando um marcador seleccionável pode ser introduzido numa célula hospedeira no mesmo vector que o que codifica NgR ou pode ser introduzido num vector separado. As células estavelmente transfectadas com o ácido nucleico

introduzido podem ser identificadas através de selecção com fármacos (p.ex., as células que incorporaram o gene do marcador seleccionável sobreviverão, enquanto as outras células morrem).

Num caso preferido, os polipéptidos desta divulgação, incluindo formas de NgR2 e NgR3, formas solúveis de NgR, polipéptidos NgR quiméricos, fusões NgR/Ig e fragmentos e variações de cada um dos de cima são expressos em células de ovário de hamster chinês (CHO).

Para introduzir o fragmento de ADN codificando para a proteína ou o polipéptido NgR numa célula CHO para expressar a proteína ou o polipéptido NgR recombinante, é necessário construir o vector de expressão.

Os vectores para expressão em CHO incluem, mas não se limitam a, pAl-11, pXT1, pRc/CMV, pRc/RSV e pcDNA1Neo. O promotor não está especificamente limitado desde que promova eficazmente a expressão em células CHO. Exemplos de promotores adequados são: SR α , SV40, LTR, CMV e HSV-TK. Destes, são preferidos os promotores CMV e SR α .

Para além dos promotores acima mencionados, os vectores de expressão podem conter estimuladores, sinais de processamento, sinais de poliadenilação, marcadores seleccionáveis e uma origem de replicação de SV40. Marcadores seleccionáveis adequados incluem, mas não se limitam ao gene da di-hidrofolato-redutase (DHFR) que proporciona resistência a metotrexato (MTX), o gene de resistência a ampicilina e o gene de resistência a neomicina.

Exemplos de vectores de expressão contendo cada um o ADN codificando para NgR, porções, fragmentos e construções solúveis deste, incluem o vector (tal como um descrito acima), no qual o promotor está operativamente ligado (de preferência a montante) à sequência nucleotídica codificando a construção de NgR desejada; um sinal de poliadenilação a jusante da sequência nucleotídica codificando a construção de NgR; e, de preferência, o vector inclui um gene de DHFR operável. De preferência, o gene de resistência à ampicilina está também operativamente contido no vector.

A célula CHO sem o gene DHFR (Urlaub, G. et al., *Proc. Natl. Acad. Sci. USA* 77: 4216-4220, 1980) e CHO-K1 (*Proc. Natl. Acad. Sci. USA* 60: 1275, 1968) são adequadas para utilização.

Os vectores de expressão de NgR preparados tal como acima são introduzidos em células CHO através de qualquer método conhecido, incluindo, mas não se limitando ao método de fosfato de cálcio (Graham e van der Eb, *Virology* 52: 456-467, 1973) e electroporação (Niemann et al., *EMBO J.* 1: 841-845, 1982).

Os transformantes possuindo os vectores de expressão são seleccionados com base nos marcadores seleccionáveis acima mencionados. A selecção clonal repetida dos transformantes utilizando os marcadores seleccionáveis permite a selecção de linhas celulares estáveis possuindo elevada expressão das construções de NgR. Concentrações crescentes de MTX no meio de selecção permitem amplificação génica e maior expressão da proteína desejada. A célula CHO contendo o NgR recombinante pode ser produzida através de cultura das células CHO contendo os vectores de expressão de NgR expressando constitutivamente as construções de NgR.

Os meios utilizados na cultura de células CHO incluem meio DMEM suplementado com cerca de 0,5 a 20% de soro fetal de vitelo, meio DMEM e meio RPMI 1640. O pH do meio é de preferência de cerca de 6 a 8. A cultura é de preferência a cerca de 30 a 40°C durante cerca de 15 a 72 horas com arejamento.

Uma célula hospedeira desta divulgação, p.ex., uma célula hospedeira do presente invento, tal como uma célula hospedeira procariótica ou eucariótica em cultura, pode ser utilizada para produzir (i.e., expressar) a proteína NgR. Assim, a divulgação proporciona ainda métodos para a produção da proteína NgR utilizando as células hospedeiras desta divulgação. Num caso, o método compreende a cultura da célula hospedeira desta divulgação (na qual foi introduzido um vector de expressão recombinante codificando NgR), p.ex. uma célula hospedeira do presente invento, num meio adequado para que

seja produzida a proteína NgR. Noutro caso, o método compreende ainda o isolamento de NgR a partir do meio ou da célula hospedeira.

Em situações onde o polipéptido NgR se verificará principalmente intracelularmente, o material intracelular (incluindo corpos de inclusão para bactérias Gram-negativas) pode ser extraído da célula hospedeira utilizando qualquer técnica padrão conhecida de um vulgar perito na arte. Tais métodos englobarão, por exemplo e não como limitação, a lise das células hospedeiras para libertar os conteúdos do periplasma/citoplasma através de prensa francesa, homogeneização e/ou ultra-sons seguido de centrifugação.

Se o polipéptido NgR tiver formado corpos de inclusão no citosol, tais corpos de inclusão podem ligar-se frequentemente às membranas celulares interna e/ou externa. Após centrifugação, os corpos de inclusão encontrar-se-ão principalmente no material sedimentado. O material sedimentado pode então ser tratado a pH extremos ou com um ou mais agentes caotrópicos tais como um detergente, guanidina, derivados de guanidina, ureia ou derivados de ureia na presença de um agente redutor tal como ditiotreitol a pH alcalino ou tris-carboxietilfosfina a pH ácido para libertar, quebrar e solubilizar os corpos de inclusão. O polipéptido NgR uma vez solubilizado pode ser analisado utilizando em electroforese em gel, imunoprecipitação ou semelhantes. Vários métodos de isolamento do polipéptido NgR serão aparentes a um vulgar perito na arte, por exemplo, o isolamento pode ser alcançado utilizando métodos padrão tais como os expostos abaixo e em Marston et al., *Meth. Enzymol.* 182: 264-275, 1990.

Se o polipéptido NgR isolado não for biologicamente activo após o procedimento de isolamento empregue, vários métodos para "redobragem" ou conversão do polipéptido na sua estrutura terciária e geração de ligações dissulfureto, podem ser utilizados para restaurar a actividade biológica. Os métodos conhecidos de um vulgar perito na arte incluem o ajustamento do pH do polipéptido solubilizado para um pH habitualmente acima de 7 e na presença de uma concentração particular de um caotrope. A selecção do caotrope é muito semelhante às escolhas utilizadas para solubilização de corpos

de inclusão mas habitualmente a uma concentração inferior e não é necessariamente o mesmo caotrope que o utilizado para a solubilização. Pode ser necessário empregar um agente redutor ou o agente redutor mais a sua forma oxidada numa proporção específica, para gerar um determinado potencial redox permitindo que ocorra baralhamento de dissulfuretos na formação da ponte ou pontes de cisteína da proteína. Alguns dos pares redox vulgarmente utilizados incluem cisteína/cistamina, glutathione (GSH)/ditiobis GSH, cloreto de cobre, ditiotreitól (DTT)/ditiano DTT, 2-mercaptoetanol (bME)/ditio-b(ME). Para aumentar a eficiência da redobragem, pode ser necessário empregar um co-solvente, tal como glicerol, polietilenoglicol de vários pesos moleculares e arginina.

Animais Transgénicos

As células hospedeiras desta divulgação, incluindo as do presente invento, podem também ser utilizadas para produzir animais transgénicos não humanos. Por exemplo, num caso, uma célula hospedeira desta divulgação é um oócito fertilizado ou uma célula estaminal embrionária na qual foram introduzidas sequências de codificação de NgR. Tais células hospedeiras podem então ser utilizadas para criar animais transgénicos não humanos nos quais foram introduzidas no seu genoma sequências exógenas de NgR ou animais recombinantes homólogos nos quais foram alteradas sequências endógenas de NgR. Tais animais são úteis para estudar a função e/ou actividade de NgR e para identificação e/ou avaliação de moduladores da actividade de NgR. Tal como aqui se utiliza, um "animal transgénico" é um animal não humano, de preferência um mamífero, de maior preferência um roedor tal como um rato ou ratinho, no qual uma ou mais das células do animal incluem um transgene. Outros exemplos de animais transgénicos incluem primatas não humanos, ovelhas, cães, vacas, cabras, galinhas, anfíbios, etc. Um transgene é ADN exógeno que é integrado no genoma de uma célula a partir da qual um animal transgénico se desenvolve e que permanece no genoma do animal maduro, dirigindo deste modo a expressão de um produto génico codificado num ou mais tipos celulares ou tecidos do animal transgénico. Tal como aqui se utiliza, um "animal recombinante homólogo" é um animal não humano, de preferência um mamífero, de maior preferência um ratinho, no qual um gene de NgR endógeno foi alterado através

de recombinação homóloga entre o gene endógeno e uma molécula de ADN exógena introduzida numa célula do animal, p.ex., uma célula embrionária do animal, antes do desenvolvimento do animal.

Um animal transgénico desta divulgação pode ser criado através da introdução de ácido nucleico codificando NgR no pró-núcleo masculino de um oócito fertilizado, p.ex., através de micro-injecção, infecção retroviral e permitindo que o oócito se desenvolva numa fêmea adoptiva pseudográvida. A sequência de ADN de NgR humano de SEQ ID NOs:1 ou 3 pode ser introduzida como um transgene no genoma de um animal não humano. Alternativamente, um homólogo não humano do gene humano de NgR, tal como um gene de ratinho de NgR, pode ser isolado com base em hibridação com o ADNC do NgR humano (descrito mais acima) e utilizado como transgene. Sequências intrónicas e sinais de poliadenilação podem também ser incluídos no transgene para aumentar a eficiência da expressão do transgene. Uma ou mais sequências reguladoras específicas de tecido podem ser operativamente ligadas ao transgene de NgR para dirigir a expressão da proteína NgR para determinadas células. Os métodos para criação de animais transgénicos através de manipulação embrionária e micro-injecção, particularmente animais tais como ratinhos, tornaram-se convencionais na arte e estão descritos, por exemplo, nas Pat. U.S. N^{os}. 4736866; 4870009; e 4873191; e Hogan 1986, em *Manipulating the Mouse Embryo*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY. Métodos semelhantes são utilizados para produção de outros animais transgénicos. Um animal transgénico fundador pode ser identificado com base na presença do transgene de NgR no seu genoma e/ou expressão do ARNm de NgR em tecidos ou células dos animais. Um animal transgénico fundador pode então ser utilizado para criar mais animais possuindo o transgene. Para além disso, animais transgénicos possuindo um transgene codificando NgR podem ainda ser criados para gerar outros animais transgénicos possuindo outros transgenes.

Para criar um animal recombinante homólogo, é preparado um vector que contenha pelo menos uma porção de um gene de NgR no qual uma deleção, adição ou substituição foi introduzida para deste modo alterar, p.ex., destruir funcionalmente, o

gene de NgR. O gene de NgR pode ser um gene humano (p.ex., SEQ ID NO:1 ou 13), mas de preferência, é um homólogo não humano de um gene de NgR humano. Por exemplo, um homólogo de ratinho do gene de NgR humano de SEQ ID NO:1 ou 13 pode ser utilizado para construir um vector de recombinação homóloga adequado para alterar um gene endógeno de NgR no genoma de ratinho. Num caso, o vector é desenhado para que, após recombinação homóloga, o gene endógeno de NgR seja funcionalmente destruído (i.e., deixe de codificar uma proteína funcional; também referido como um vector "knock out").

Alternativamente, o vector pode ser desenhado para que, após recombinação homóloga, o gene endógeno de NgR seja mutado ou de outro modo alterado mas codificando ainda a proteína funcional (p.ex., a região reguladora a montante pode ser alterada para deste modo alterar a expressão da proteína NgR endógena). No vector de recombinação homóloga, a porção alterada do gene de NgR é flanqueada nas suas extremidades 5' e 3' por ácido nucleico adicional do gene de NgR para permitir que ocorra recombinação homóloga entre o gene exógeno de NgR transportado pelo vector e um gene endógeno de NgR numa célula estaminal embrionária. O ácido nucleico adicional de NgR flanqueador é de comprimento suficiente para recombinação homóloga com sucesso com o gene endógeno. Tipicamente, estão incluídas no vector várias quilobases de ADN flanqueador (tanto nas extremidades 5' como 3'). Ver p.ex., Thomas et al., *Cell* 51: 503, 1987 para uma descrição de vectores de recombinação homóloga. O vector é introduzido numa linha de células estaminais embrionárias (p.ex., através de electroporação) e as células em que o gene de NgR introduzido se recombinau de forma homóloga com o gene de NgR endógeno são seleccionadas (ver p.ex., Li et al., *Cell* 69: 915, 1992).

As células seleccionadas são então injectadas num blastocisto de um animal (p.ex., um ratinho) para formar quimeras de agregação. Ver p.ex., Bradley 1987, Em: "Teratocarcinomas and Embryonic Stem Cells: A Practical Approach", Robertson, ed. IRL, Oxford, págs. 113-152. Um embrião quimérico pode então ser implantado numa fêmea adoptiva pseudográvida adequada e o embrião levado até termo. A descendência ancorando o ADN recombinado de forma homóloga nas suas células germinais pode ser utilizada para criar

animais nos quais todas as células do animal contenham o ADN recombinado de forma homóloga através de transmissão pela linha germinal do transgene. Métodos para construção de vectores de recombinação homóloga e animais recombinantes homólogos são descritos em Bradley, *Curr. Opin. Biotechnol.* 2: 823-829, 1991; Publicação Internacional PCT N^{os}.: WO 90/11354; WO 91/01140; WO 92/0968; e WO 93/04169.

Noutro caso, podem ser produzidos animais transgénicos não humanos que contenham sistemas seleccionados que permitam a expressão regulada do transgene. Um exemplo de um tal sistema é o sistema de recombinase cre/loxP do bacteriófago P1. Para uma descrição do sistema de recombinase cre/loxP, ver, p.ex., Lakso et al., *Proc. Natl. Acad. Sci. USA* 89: 6232-6236, 1992. Outro exemplo de um sistema de recombinase é o sistema de recombinase FLP de *Saccharomyces cerevisiae* (O'Gorman et al., *Science* 251: 1351-1355, 1991. Se for utilizado um sistema de recombinase cre/loxP para regular a expressão do transgene, são necessários animais contendo transgenes codificando tanto a recombinase Cre como uma proteína seleccionada. Tais animais podem ser proporcionados através da construção de animais "duplos" transgénicos, p.ex., através do acasalamento de dois animais transgénicos, um contendo um transgene codificando uma proteína seleccionada e o outro contendo um transgene codificando uma recombinase.

Os clones dos animais transgénicos não humanos aqui descritos podem também ser produzidos de acordo com os métodos descritos em Wilmut et al., *Nature* 385: 810-813, 1997. Em resumo, uma célula, p.ex., uma célula somática, do animal transgénico pode ser isolada e induzida a sair do ciclo de crescimento e entrar na fase G₀. A célula quiescente pode então ser fundida, p.ex., através da utilização de pulsos eléctricos, com um óócito enucleado de um animal da mesma espécie da qual é isolada a célula quiescente. O óócito reconstruído é então cultivado para se desenvolver até mórula ou blastocisto e então transferido para uma fêmea adoptiva pseudográvida. A descendência nascida desta fêmea adoptiva será um clone do animal do qual a célula, p.ex., a célula somática, é isolada.

Anti-sentido

São também proporcionados por esta divulgação polinucleótidos anti-sentido que reconhecem e hibridam com polinucleótidos NgR. São proporcionados polinucleótidos anti-sentido inteiros e fragmentados. As moléculas anti-sentido fragmentadas desta divulgação incluem (i) as que reconhecem e hibridam especificamente com ARN de NgR (tal como determinado através de comparação de sequências de ADN codificando NgR com ADN codificando outras moléculas conhecidas). A identificação de sequências únicas de polinucleótidos codificando NgR pode ser deduzida através da utilização de qualquer base de dados de sequências publicamente disponível e/ou através da utilização de programas de comparação de sequências comercialmente disponíveis. Após identificação das sequências desejadas, pode ser efectuado o isolamento através de digestão de restrição ou amplificação utilizando qualquer uma das várias técnicas de reacção em cadeia da polimerase bem conhecidas na arte. Os polinucleótidos anti-sentido são particularmente relevantes para regulação da expressão de NgR pelas células expressando ARNm de NgR.

Oligonucleótidos anti-sentido ou fragmentos de uma sequência nucleotídica exposta em SEQ ID NO:1, 3, 13 ou sequências complementares ou homólogas destas, derivados das sequências nucleotídicas da presente divulgação codificando NgR são úteis como ferramentas de diagnóstico para sondar a expressão génica em vários tecidos. Por exemplo, o tecido pode ser sondado *in situ* com sondas oligonucleotídicas possuindo grupos detectáveis através de técnicas convencionais de autorradiografia para investigar a expressão nativa desta enzima ou condições patológicas relacionadas com ela. Em aspectos específicos, são proporcionadas moléculas de ácido nucleico anti-sentido que compreendem uma sequência complementar a pelo menos cerca de 10, 25, 50, 100, 250 ou 500 nucleótidos de uma cadeia de codificação de NgR inteira, ou a apenas uma parte desta. São adicionalmente proporcionadas moléculas de ácido nucleico codificando fragmentos, homólogos, derivados e análogos de uma proteína NgR de SEQ ID NO:2, 4 ou 14 ou ácidos nucleicos anti-sentido complementares a uma sequência de ácido nucleico de NgR de SEQ ID NOs: 1, 3 ou 13.

Num caso, uma molécula de ácido nucleico anti-sentido é anti-sentido a uma "região de codificação" da cadeia de codificação de uma sequência nucleotídica codificando NgR. O termo "região de codificação" refere-se à região da sequência nucleotídica compreendendo codões que são traduzidos em resíduos de aminoácidos (p.ex., a região de codificação de proteína do NgR humano corresponde à região de codificação de SEQ ID NO:1, 3 ou 13). Noutro caso, a molécula de ácido nucleico anti-sentido é anti-sentido a uma "região não codificadora" da cadeia de codificação de uma sequência nucleotídica codificando NgR. O termo "região não codificadora" refere-se a sequências 5' e 3' que flanqueiam a região de codificação que não são traduzidas em aminoácidos (i.e., também referidas como regiões 5' e 3' não traduzidas).

Os oligonucleótidos anti-sentido são de preferência dirigidos a regiões reguladoras de uma sequência nucleotídica de SEQ ID NO:1, 3, 13 ou ARNm correspondendo a estas, incluindo, mas não se limitando ao codão de iniciação, caixa TATA, sequências estimuladoras e semelhantes. Dadas as sequências da cadeia de codificação codificando NgR aqui divulgadas (p.ex., SEQ ID NO:1, 3 ou 13), os ácidos nucleicos anti-sentido desta divulgação podem ser desenhados de acordo com as regras de emparelhamento de bases de Watson e Crick ou de Hoogsteen. A molécula de ácido nucleico anti-sentido pode ser complementar à região de codificação inteira do ARNm de NgR, mas é de preferência um oligonucleótido que é anti-sentido a apenas uma porção da região de codificação ou não codificadora do ARNm de NgR. Por exemplo, o oligonucleótido anti-sentido pode ser complementar à região em volta do local de início da tradução do ARNm de NgR. Um oligonucleótido anti-sentido pode ter, por exemplo, cerca de 5, 10, 15, 20, 25, 30, 35, 40, 45 ou 50 nucleótidos de comprimento. Um ácido nucleico anti-sentido desta divulgação pode ser construído utilizando síntese química ou reacções de ligação enzimática utilizando procedimentos conhecidos na arte. Por exemplo, um ácido nucleico anti-sentido (p.ex., um oligonucleótido anti-sentido) pode ser quimicamente sintetizado utilizando nucleótidos de ocorrência natural ou nucleótidos modificados de vários modos desenhados para aumentar a estabilidade biológica das moléculas ou para aumentar a estabilidade física do dúplex formado entre os ácidos nucleicos anti-sentido e com sentido,

p.ex., podem ser utilizados derivados de fosforotioato e nucleótidos substituídos com acridina.

Exemplos de nucleótidos modificados que podem ser utilizados para gerar o ácido nucleico anti-sentido incluem: 5-fluorouracilo, 5-bromouracilo, 5-clorouracilo, 5-iodouracilo, hipoxantina, xantina, 4-acetilcitosina, 5-(carboxi-hidroxiimetil)uracilo, 5-carboximetilaminometil-2-tiouridina, 5-carboximetilaminometiluracilo, di-hidrouracilo, beta-D-galactosilqueosina, inosina, N6-isopenteniladenina, 1-metilguanina, 1-metilinosina, 2,2-dimetilguanina, 2-metiladenina, 2-metilguanina, 3-metilcitosina, 5-metilcitosina, N6-adenina, 7-metilguanina, 5-metilaminometiluracilo, 5-metoxiaminometil-2-tiouracilo, beta-D-manosilqueosina, 5'-metoxicarboximetiluracilo, 5-metoxiuracilo, 2-metiltio-N6-isopenteniladenina, ácido uracil-5-oxiacético (v), wybutoxosina, pseudouracilo, queosina, 2-tiocitosina, 5-metil-2-tiouracilo, 2-tiouracilo, 4-tiouracilo, 5-metiluracilo, éster metílico de ácido uracil-5-oxiacético, ácido uracil-5-oxiacético (v), 5-metil-2-tiouracilo, 3-(3-amino-3-N-2-carboxipropil)uracilo, (acp3)w, e 2,6-diaminopurina. Alternativamente, o ácido nucleico anti-sentido pode ser produzido biologicamente utilizando um vector de expressão no qual um ácido nucleico foi subclonado numa orientação anti-sentido (i.e., o ARN transcrito a partir do ácido nucleico inserido será de uma orientação anti-sentido a um ácido nucleico alvo de interesse, melhor descrito na seguinte subsecção).

As moléculas de ácido nucleico anti-sentido desta divulgação (de preferência oligonucleótidos de 10 a 20 nucleótidos de comprimento) são tipicamente administradas a um sujeito ou geradas *in situ* para que hibridem ou se liguem a ARNm celular e/ou ADN genómico codificando uma proteína NgR para inibir deste modo a expressão da proteína, p.ex., através da inibição da transcrição e/ou da tradução. A supressão da expressão de NgR ao nível da transcrição ou da tradução é útil para gerar modelos celulares ou animais para doenças/condições caracterizadas por expressão de NgR aberrante. A hibridação pode ser através de complementaridade de nucleótidos convencional para formar um dúplex estável ou, por exemplo, no caso de uma molécula de ácido nucleico anti-sentido que se

ligue a cópias de ADN, através de interacções específicas no sulco principal da dupla hélice.

Os oligonucleótidos anti-sentido de fosforotioato e metilfosfonato estão especificamente contemplados para utilização terapêutica. Os oligonucleótidos anti-sentido podem ainda ser modificados através da adição de poli-L-lisina, polilisina transferrina ou porções colesterol na sua extremidade 5'.

Um exemplo de uma via de administração de moléculas de ácido nucleico anti-sentido da divulgação inclui injeção directa num local de tecido. Alternativamente, as moléculas de ácido nucleico anti-sentido podem ser modificadas para atingir células seleccionadas e depois administradas sistemicamente. Por exemplo, para administração sistémica, as moléculas anti-sentido podem ser modificadas para se ligarem especificamente a receptores ou antigénios expressos numa superfície celular seleccionada, p.ex., através de ligação das moléculas de ácido nucleico anti-sentido a péptidos ou anticorpos que se liguem a receptores ou antigénios da superfície celular. As moléculas de ácido nucleico anti-sentido podem também ser distribuídas a células utilizando os vectores aqui descritos. Para alcançar concentrações intracelulares suficientes de moléculas anti-sentido, são preferidas as construções de vector nas quais a molécula de ácido nucleico anti-sentido é colocada sob o controlo de um promotor forte pol II ou pol III.

Ainda noutro caso, a molécula de ácido nucleico anti-sentido desta divulgação é uma molécula de ácido nucleico α -anomérica. Uma molécula de ácido nucleico α -anomérica forma híbridos de cadeia dupla específicos com ARN complementar nos quais, ao contrário das unidades β habituais, as cadeias correm paralelas umas às outras (Gaultier *et al.*, *Nucleic Acids Res.* 15: 6625-6641, 1987). A molécula de ácido nucleico anti-sentido pode também compreender um 2'-O-metilribonucleótido (Inoue *et al.*, *Nucleic Acids Res.* 15: 6131-6148, 1987) ou um análogo de ARN-ADN quimérico (Inoue *et al.*, *FEBS Lett.* 215: 327-330, 1987).

As sequências de NgR ensinadas na presente divulgação, incluindo as do presente invento, facilitam o desenho de novos

factores de transcrição para modulação da expressão de NgR em células nativas e animais, e células transformadas ou transfectadas com polinucleótidos NgR. Por exemplo, foi mostrado que proteínas com dedos de zinco Cys₂-His₂, que se ligam ao ADN através dos seus domínios dedos de zinco, são passíveis de alterações estruturais que conduzam ao reconhecimento de diferentes sequências alvo. Estas proteínas de dedos de zinco artificiais reconhecem locais alvo específicos com elevada afinidade e baixas constantes de dissociação, e são capazes de actuar como interruptores génicos para modular a expressão génica. O conhecimento da sequência alvo de NgR particular da presente divulgação facilita o desenho de proteínas de dedos de zinco específicas para a sequência alvo utilizando métodos conhecidos tais como uma combinação de modelação com base na estrutura e pesquisa de bibliotecas de apresentação fágica (Segal *et al.*, *Proc. Natl. Acad. Sci. USA* 96: 2758-2763, 1999; Liu *et al.*, *Proc. Natl. Acad. Sci. USA* 94: 5525-5530, 1997; Greisman *et al.*, *Science* 275: 657-661, 1997; Choo *et al.*, *J. Mol. Biol.* 273: 525-532, 1997). Cada domínio dedo de zinco reconhece habitualmente três ou mais pares de bases. Uma vez que uma sequência de reconhecimento de 18 pares de bases tem geralmente um comprimento suficiente para a tornar única em qualquer genoma conhecido, esperar-se-ia que uma proteína de dedos de zinco consistindo em 6 repetições em cadeia de dedos de zinco assegurasse especificidade para uma determinada sequência (Segal *et al.*, (1999), acima). As repetições de dedos de zinco artificiais, desenhadas com base no promotor de sequências de NgR, são fundidas com domínios de activação ou repressão para promover ou suprimir a expressão de NgR (Liu *et al.*, (1997), acima). O promotor de NgR pode ser obtido através de métodos padrão conhecidos do vulgar perito na arte com a divulgação aqui contida e o conhecimento da sequência de NgR. Alternativamente, os domínios dedo de zinco podem ser fundidos com o factor de ligação à caixa TATA (TBP) com comprimentos variáveis da região ligante entre o péptido dedo de zinco e o TBP para criar activadores ou repressores da transcrição (Kim *et al.*, *Proc. Natl. Acad. Sci. USA* 94: 3616-3620, 1997. Tais proteínas e polinucleótidos que os codificam, têm utilidade para modulação da expressão de NgR *in vivo* em células nativas, animais e humanos; e/ou células transfectadas com sequências codificando NgR. O novo factor de transcrição pode ser

distribuído às células alvo através de construções de transfecção que expressem o factor de transcrição (terapia génica), ou através da introdução da proteína. Proteínas de dedos de zinco modificadas podem também ser desenhadas para se ligarem a sequências de ARN para utilizar em terapêuticos como alternativas aos métodos de ARN anti-sentido ou catalítico (McColl *et al.*, *Proc. Natl. Acad. Sci. USA* 96: 9521-9526, 1997; Wu *et al.*, *Proc. Natl. Acad. Sci. USA* 92: 344-348, 1995). A presente divulgação contempla métodos de desenho de tais factores de transcrição com base na sequência génica da divulgação, bem como proteínas de dedos de zinco feitas à medida, que sejam úteis para modular a expressão de NgR em células (nativas ou transformadas) cujo complemento genético incluía estas sequências.

Ribozimas e porções PNA

Ainda noutro caso, um ácido nucleico anti-sentido desta divulgação é uma ribozima. As ribozimas são moléculas de ARN catalíticas com actividade ribonucleásica que são capazes de clivar um ácido nucleico de cadeia simples, tal como um ARNm, para o qual têm uma região complementar. Assim, as ribozimas (p.ex., ribozimas cabeça-de-martelo, descritas em Haselhoff e Gerlach, *Nature* 334: 585-591, 1988) podem ser utilizadas para clivar cataliticamente transcritos de ARNm de NgR para deste modo inibir a tradução do ARNm de NgR. A ribozima possuindo especificidade para um ácido nucleico codificando NgR pode ser desenhada com base na sequência nucleotídica de um ADN de NgR aqui divulgado (i.e., SEQ ID NO:1, 3 ou 13). Por exemplo, pode ser construído um derivado de ARN L-19 IVS de *Tetrahymena* no qual a sequência nucleotídica do local activo é complementar com a sequência nucleotídica a clivar num ARNm codificando NgR. Ver, p.ex., Cech *et al.*, Patente U.S. N.º 4987071; e Cech *et al.*, Patente U.S. N.º 5116742. Alternativamente, o ARNm de NgR pode ser utilizado para seleccionar um ARN catalítico possuindo uma actividade ribonucleásica específica a partir de um banco de moléculas de ARN. Ver, p.ex., Bartel *et al.*, *Science* 261: 1411-1418, 1993.

Alternativamente, a expressão do gene de NgR pode ser inibida através do direccionamento de sequências nucleotídicas complementares à região reguladora do NgR (p.ex., o promotor e/ou estimuladores de NgR) para formar estruturas helicoidais

triplas que impedem a transcrição do gene de NgR em células alvo. Ver geralmente, Helene, *Anticancer Drug Des.* 6: 569-584, 1991; Helene. *et al.*, *Ann. N.Y. Acad Sci.* 660: 27-36, 1992; e Maher, *BioEssays* 14: 807-815, 1992.

Em vários casos, os ácidos nucleicos de NgR podem ser modificados na porção base, porção açúcar ou estrutura de fosfato para melhorar, p.ex., a estabilidade, hibridação ou solubilidade da molécula. Por exemplo, a estrutura de desoxirribose-fosfato dos ácidos nucleicos pode ser modificada para gerar ácidos nucleicos peptídicos (ver Hyrup *et al.*, *Bioorg. Med. Chem. Lett.* 4: 5-23, 1996). Tal como aqui se utilizam, os termos "ácidos nucleicos peptídicos" ou "PNAs" referem-se a imitadores de ácidos nucleicos, p.ex., imitadores de ADN, nos quais a estrutura de desoxirribose-fosfato é substituída por uma estrutura pseudopeptídica e são apenas mantidas as quatro nucleobases naturais. Mostrou-se que a estrutura neutra dos PNA permite a hibridação específica a ADN e ARN sob condições de baixa força iónica. A síntese de oligómeros PNA pode ser efectuada utilizando protocolos de síntese peptídica de fase sólida padrão tal como descrito em Hyrup *et al.*, (1996) acima; Perry-O'Keefe *et al.*, *Proc. Natl. Acad. Sci. USA* 93: 14670-14675, 1996.

Os PNA de NgR podem ser utilizados em aplicações terapêuticas e de diagnóstico. Por exemplo, os PNA podem ser utilizados como agentes anti-sentido ou anti-gene para modulação da expressão génica específica da sequência através, p.ex., de indução da paragem da transcrição ou da tradução ou inibição da replicação. Os PNA de NgR podem também ser utilizados, p.ex., na análise de mutações de um único par de bases num gene através, por exemplo, de fixação de PCR dirigida por PNA; como enzimas de restrição artificiais quando utilizadas em combinação com outras enzimas, p.ex., nucleases S1 (Hyrup (1996), acima); ou como sondas ou iniciadores para a sequência de ADN e hibridação (Hyrup *et al.*, (1996), acima; Perry-O'Keefe (1996), acima).

Noutro caso, os PNA de NgR podem ser modificados, p.ex., para aumentar a sua estabilidade ou tomada celular, através de união lipófila ou outros grupos ajudantes a PNA, através da formação de quimeras PNA-ADN ou através da utilização de

lipossomas ou outras técnicas de distribuição de fármacos conhecidas na arte. Por exemplo, podem ser geradas quimeras PNA-ADN de NgR que podem combinar as propriedades vantajosas de PNA e ADN. Tais quimeras permitem que enzimas de reconhecimento de ADN, p.ex., ARNase H e ADN-polimerases, interajam com a porção ADN enquanto a porção PNA proporcionaria elevada afinidade de ligação e especificidade. As quimeras PNA-ADN podem ser ligadas utilizando ligantes de comprimentos apropriados seleccionados em termos do empacotamento de bases, do número de ligações entre as nucleobases e da orientação (Hyrup (1996), acima). A síntese de quimeras PNA-ADN pode ser efectuada tal como descrito em Hyrup (1996), acima e Finn *et al.*, *Nucleic Acids Res.* 24: 3357-3363, 1996. Por exemplo, uma cadeia de ADN pode ser sintetizada num suporte sólido utilizando a química de conjugação de fosforamidito padrão, e análogos de nucleósidos modificados, p.ex., pode ser utilizado 5'-(4-metoxitritil)-amino-5'-desoxi-timidina-fosforamidite entre o PNA e a extremidade 5' do ADN (Mag *et al.*, *Nucleic Acids Res.* 17: 973-988, 1989). Os monómeros de PNA são então unidos de um modo passo-a-passo para produzir uma molécula quimérica com um segmento PNA a 5' e um segmento ADN a 3' (Finn *et al.*, (1996), acima). Alternativamente, podem ser sintetizadas moléculas quiméricas com um segmento ADN a 5' e um segmento PNA a 3'. Ver, Petersen *et al.*, *Bioorg. Med. Chem. Lett.* 5: 1119-1124, 1975.

Noutros casos, o oligonucleótido pode incluir outros grupos anexos tais como péptidos (p.ex., para se dirigir a receptores da célula hospedeira *in vivo*), ou agentes facilitadores do transporte através da membrana celular (ver Letsinger *et al.*, *Proc. Natl. Acad. Sci. USA* 86: 6553-6556, 1989; Lemaitre *et al.*, *Proc. Natl. Acad. Sci. USA* 84: 648-652, 1987; Publicação PCT N.º WO 88/09810) ou da barreira hemato-encefálica (ver, p.ex., Publicação PCT N.º WO 89/10134). Adicionalmente, os oligonucleótidos podem ser modificados com agentes de clivagem desencadeados por hibridação (ver, p.ex., Krol *et al.*, *Biotechniques* 6: 958-976, 1988) ou agentes intercalantes (ver, p.ex., Zon, *Pharm. Res.* 5: 539-549, 1988). Para este fim, o oligonucleótido pode ser conjugado com outra molécula, p.ex., um péptido, um agente de ligação cruzada

desencadeado por hibridação, um agente de transporte, um agente de clivagem desencadeado por hibridação, etc.

Podem ser utilizados métodos de sequenciação automáticos para obter ou verificar a sequência nucleotídica de NgR. Crê-se que as sequências nucleotídicas de NgR desta divulgação, incluindo as do presente invento, sejam 100% exactas. No entanto, tal como é conhecido na arte, a sequência nucleotídica obtida através de métodos automáticos pode conter alguns erros. As sequências nucleotídicas determinadas através de automação são tipicamente pelo menos cerca de 90%, mais tipicamente pelo menos de cerca de 95% a pelo menos cerca 99,9% idênticas à sequência nucleotídica real de uma dada molécula de ácido nucleico. A sequência real pode ser determinada mais precisamente utilizando métodos de sequenciação manuais, que são bem conhecidos na arte. Um erro numa sequência que resulte numa inserção ou deleção de um ou mais nucleótidos pode resultar numa mudança de enquadramento na tradução de tal modo que a sequência de aminoácidos prevista diferirá da que seria prevista a partir da sequência nucleotídica real da molécula de ácido nucleico, a começar no ponto da mutação.

Polipéptidos

Esta divulgação proporciona também polipéptidos NgR de mamífero purificados e isolados codificados por um polinucleótido desta divulgação. É actualmente preferido um polipéptido NgR humano compreendendo a sequência de aminoácidos exposta em SEQ ID NO:2 ou SEQ ID NO:14. Outro caso preferido é um polipéptido NgR de ratinho compreendendo a sequência de aminoácidos de NgR3, tal como exposta em SEQ ID NO:4.

Os polipéptidos desta divulgação que proporcionam os polipéptidos do presente invento são expostos nas reivindicações anexas.

Um aspecto desta divulgação refere-se a proteínas NgR isoladas, e porções biologicamente activas destas, ou derivados, fragmentos, análogos ou homólogos destas. São também proporcionados fragmentos polipeptídicos adequados para utilizar como imunogénios para criar anticorpos anti-NgR. De

preferência, os fragmentos de proteínas NgR compreendem pelo menos uma actividade biológica de NgR. Num caso, proteínas NgR nativas podem ser isoladas a partir de fontes de células ou tecidos através de um esquema de purificação apropriado utilizando técnicas padrão de purificação de proteínas. Noutro caso, são produzidas proteínas NgR através de técnicas de ADN recombinante. Alternativamente à expressão recombinante, pode ser sintetizada quimicamente uma proteína ou polipéptido NgR utilizando técnicas padrão de síntese peptídica.

A divulgação engloba também polipéptidos que têm pelo menos 99%, pelo menos 95%, pelo menos 90%, pelo menos 85%, pelo menos 80%, pelo menos 75%, pelo menos 70%, pelo menos 65%, pelo menos 60%, pelo menos 55%, pelo menos 50% ou pelo menos 45% de identidade e/ou homologia com o polipéptido preferido da divulgação. Adicionalmente, a divulgação engloba polipéptidos possuindo a sequência de consenso mostrada em SEQ ID NO:6 (mostrada na Tabela 5) excluindo o NgR caracterizado anteriormente ("NgR1"), e polipéptidos compreendendo pelo menos cerca de 90% da sequência de consenso.

O termo "percentagem de identidade de sequência" é calculado através da comparação de duas sequências alinhadas de modo óptimo ao longo da região de comparação, determinação do número de posições nas quais ocorre a base de ácido nucleico idêntica (p.ex., A, T, C, G, U ou I, no caso de ácidos nucleicos) em ambas as sequências para dar o número de posições correspondentes, dividindo o número de posições correspondentes pelo número total de posições na região de comparação (i.e., o tamanho da janela) e multiplicando o resultado por 100 para dar a percentagem de identidade de sequência. O termo "identidade substancial" tal como aqui se utiliza denota uma característica de uma sequência polinucleotídica, em que o polinucleótido compreende uma sequência que tem pelo menos 80 por cento de identidade, de preferência pelo menos 85 por cento de identidade e frequentemente de 90 a 95 por cento de identidade de sequência, mais habitualmente pelo menos 99 por cento de identidade de sequência em comparação com uma sequência de referência ao longo de uma região de comparação.

Num aspecto, a percentagem de homologia é calculada como a percentagem de resíduos de aminoácidos na menor de duas sequências que alinham com um resíduo de aminoácido idêntico na sequência a comparar, onde podem ser introduzidos quatro intervalos num troço de 100 aminoácidos para maximizar o alinhamento (Dayhoff, em "Atlas of Protein Sequence and Structure", Vol. 5, pág. 124, National Biochemical Research Foundation, Washington, D.C. (1972)).

Uma determinação da homologia ou identidade é tipicamente feita através de um programa de computador de homologias conhecido na arte. Um programa exemplar é o programa Gap (*Wisconsin Sequence Analysis Package*, Versão 8 para UNIX, Genetics Computer Group, University Research Park, Madison, WI) utilizando as definições por defeito, que utiliza o algoritmo de Smith e Waterman (*Adv. Appl. Math.* 2: 482-489, 1981). Empregando o suporte lógico GAP no pacote de programas de GCG, (ver Needleman e Wunsch, *J. Mol. Biol.* 48: 443-453, 1970) podem ser utilizadas as seguintes definições para comparação de sequências de ácidos nucleicos: penalidade de criação de GAP de 5.0 e penalidade de extensão de GAP de 0.3, a região de codificação das sequências de ácido nucleico análogas referidas acima exhibe um grau de identidade de preferência de pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 98% ou 99%, com a parte CDS (de codificação) da sequência de ADN mostrada em SEQ ID NO:1, 3 ou 13. BestFit foi originalmente escrito para a Versão 1.0 por Paul Haeberli a partir de uma cuidadosa leitura dos artigos de Needleman e Wunsch (1970), acima, e Smith e Waterman (1981), acima. Podem ser utilizadas as seguintes definições de Bestfit para comparação de sequências de ácido nucleico: penalidade de criação de GAP de 8.0 e penalidade de extensão de GAP de 2, a região de codificação das sequências de ácido nucleico análogas referidas acima exhibe um grau de identidade de preferência de pelo menos 70%, 75%, 80%, 85%, 90%, 95%, 98% ou 99%, com a parte CDS (de codificação) da sequência de aminoácidos mostrada em SEQ ID NO:2, 4 ou 14.

Alternativamente, a homologia pode ser determinada através de análise de hibridação em que uma sequência de ácido nucleico hibrida com o complemento de uma sequência codificando as proteínas acima mencionadas sob condições

rigorosas, moderadamente rigorosas ou de baixo rigor. Ver *p.ex.*, Ausubel, *et al.*, (Eds.) "Current Protocols in Molecular Biology", John Wiley & Sons, New York, NY, 1993, e abaixo.

Os polipéptidos desta divulgação, incluindo os do presente invento, podem ser isolados a partir de fontes celulares naturais ou podem ser quimicamente sintetizados, mas são de preferência produzidos através de procedimentos recombinantes envolvendo células hospedeiras desta divulgação, *p.ex.*, células hospedeiras do presente invento.

Uma proteína "isolada" ou "purificada" ou uma porção biologicamente activa desta é substancialmente livre de material celular ou outras proteínas contaminantes da fonte celular ou de tecido da qual a proteína NgR é derivada, ou substancialmente livre de precursores químicos ou outros químicos quando quimicamente sintetizada. A expressão "substancialmente livre de material celular" inclui preparações da proteína NgR em que a proteína é separada de componentes celulares das células das quais é isolada ou produzida de forma recombinante. Num caso, a expressão "substancialmente livre de material celular" inclui preparações de proteína NgR possuindo menos de cerca de 30% (por peso seco) de proteína não NgR (também aqui referida como "proteína contaminante"), de maior preferência menos de cerca de 20% de proteína não NgR, ainda de preferência menos de cerca de 10% de proteína não NgR e ainda de maior preferência menos de cerca de 5% de proteína não NgR. Quando a proteína NgR ou uma porção biologicamente activa desta é produzida de forma recombinante, é também de preferência substancialmente livre de meio de cultura, *i.e.*, o meio de cultura representa menos de cerca de 20%, de preferência menos de cerca de 10% e de maior preferência menos de cerca de 5% do volume da preparação proteica.

A expressão "substancialmente livre de precursores químicos ou outros químicos" inclui preparações de proteína NgR na qual a proteína é separada de precursores químicos ou outros químicos que estejam envolvidos na síntese da proteína. Num caso, a expressão "substancialmente livre de precursores químicos ou outros químicos" inclui preparações de proteína NgR possuindo menos de cerca de 30% (por peso seco) de

precursores químicos ou químicos não NgR, de preferência menos de cerca de 20% de precursores químicos ou químicos não NgR, ainda de preferência menos de cerca de 10% de precursores químicos ou químicos não NgR e de maior preferência menos de cerca de 5% de precursores químicos ou químicos não NgR.

Porções biologicamente activas de uma proteína NgR incluem péptidos compreendendo sequências de aminoácidos suficientemente homólogas ou derivadas da sequência de aminoácidos da proteína NgR, p.ex., a sequência de aminoácidos mostrada em SEQ ID NO:2, 4 ou 14 que inclui menos aminoácidos que as proteínas NgR inteiras, e exibem pelo menos uma actividade de uma proteína NgR. Tipicamente, as porções biologicamente activas compreendem um domínio ou motivo com pelo menos uma actividade da proteína NgR. Uma porção biologicamente activa de uma proteína NgR pode ser um polipéptido que tenha, por exemplo, 10, 25, 50, 100 ou mais aminoácidos de comprimento.

Uma porção biologicamente activa de uma proteína NgR desta divulgação, p.ex. uma proteína NgR do presente invento, pode conter pelo menos uma das características que é conservada entre as proteínas NgR (p.ex., uma cisteína conservada como o terminal N da proteína madura, quatro cisteínas conservadas no terminal N antes de uma região rica em leucina, quatro cisteínas conservadas C-terminais em relação a uma região de repetição de leucinas, oito repetições ricas em leucina e um terminal C hidrófobo). Uma porção biologicamente activa alternativa de uma proteína NgR pode conter pelo menos dois dos domínios acima identificados. Outra porção biologicamente activa de uma proteína NgR pode conter pelo menos três dos domínios acima identificados. Ainda outra porção biologicamente activa de uma proteína NgR desta divulgação, p.ex. uma proteína NgR do presente invento, pode conter pelo menos quatro dos domínios acima identificados.

Para além disso, outras porções biologicamente activas, nas quais são suprimidas outras regiões da proteína, podem ser preparadas através de técnicas recombinantes e avaliadas quanto a uma ou mais das actividades funcionais de uma proteína NgR nativa.

Num caso, a proteína NgR tem uma sequência de aminoácidos mostrada em SEQ ID NO:2, 4 ou 14. Noutros casos, a proteína NgR é substancialmente homóloga a SEQ ID NO:2, 4 ou 14 e mantém a actividade funcional da proteína de SEQ ID NO:2, 4 ou 14, contudo difere na sequência de aminoácidos devido a variação alélica ou mutagénese, tal como descrito em detalhe abaixo.

Assim, noutro caso, a proteína NgR é uma proteína que compreende uma sequência de aminoácidos pelo menos cerca de 45% homóloga à sequência de aminoácidos de SEQ ID NO:2 ou SEQ ID NO:4 ou SEQ ID NO:14 e mantém a actividade funcional das proteínas NgR de SEQ ID NO:2, 4 ou 14.

Espera-se que a utilização de células hospedeiras de mamífero proporcione tais modificações pós-tradução (p.ex., glicosilação, truncagem, lipidação e fosforilação) conforme seja necessário para conferir uma actividade biológica óptima a produtos de expressão recombinantes desta divulgação, incluindo os do presente invento. Formas glicosiladas e não glicosiladas de polipéptidos NgR estão englobadas por esta divulgação e pelo presente invento.

A divulgação engloba também polipéptidos NgR variantes (ou análogos). Num exemplo, são proporcionadas variantes de inserção em que um ou mais resíduos de aminoácidos suplementam uma sequência de aminoácidos de NgR. As inserções podem ser localizadas num ou em ambos os terminais da proteína, ou podem ser posicionados dentro de regiões internas da sequência de aminoácidos de NgR. As variantes de inserção com resíduos adicionais num ou em ambos os terminais podem incluir, por exemplo, proteínas de fusão e proteínas incluindo etiquetas de aminoácidos ou marcadores.

As variantes de inserção incluem polipéptidos NgR em que um ou mais resíduos de aminoácidos são adicionados a uma sequência de aminoácidos de NgR ou a um fragmento biologicamente activo desta.

Os produtos variantes desta divulgação incluem também produtos NgR maduros, i.e., produtos NgR em que a sequência líder ou sinal é removida, com resíduos amino-terminais

adicionais. Os resíduos amino-terminais adicionais podem ser derivados de outra proteína ou podem incluir um ou mais resíduos que não sejam identificáveis como sendo derivados de proteínas específicas. Produtos NgR com um resíduo de metionina adicional na posição -1 ($\text{Met}^{-1}\text{-NgR}$) estão contemplados, como variantes com resíduos adicionais de metionina e lisina nas posições -2 e -1 ($\text{Met}^{-2}\text{-Lys}^{-1}\text{-NgR}$). As variantes de NgR com resíduos Met, Met-Lys, Lys adicionais (ou um ou mais resíduos básicos em geral) são particularmente úteis para maior produção de proteína recombinante em células hospedeiras bacterianas.

Variantes de Polipéptidos

A divulgação engloba também variantes de NgR possuindo resíduos de aminoácidos adicionais que resultam da utilização de sistemas de expressão específicos.

Tal como aqui se utiliza, uma “proteína quimérica” NgR ou “proteína de fusão” compreende um polipéptido NgR operativamente ligado a um polipéptido não NgR. As proteínas de fusão de NgR do presente invento são expostas nas reivindicações anexas. Um “polipéptido NgR” refere-se a um polipéptido possuindo uma sequência de aminoácidos correspondente a NgR, enquanto um “polipéptido não NgR” refere-se a um polipéptido possuindo uma sequência de aminoácidos correspondendo a uma proteína que não é homóloga à proteína NgR, p.ex., uma proteína que seja diferente da proteína NgR e que seja derivada do mesmo organismo ou de um diferente. Dentro de uma proteína de fusão de NgR o polipéptido NgR pode corresponder a toda ou a uma porção da uma proteína NgR. Num caso, uma proteína de fusão de NgR compreende pelo menos uma porção biologicamente activa de uma proteína NgR. Noutro caso, uma proteína de fusão de NgR compreende pelo menos duas porções biologicamente activas de uma proteína NgR. Ainda noutro caso, uma proteína de fusão de NgR compreende pelo menos três porções biologicamente activas de uma proteína NgR. Dentro da proteína de fusão, pretende-se que o termo “operativamente ligado” indique que o polipéptido NgR e o polipéptido não NgR estão fundidos enquadrados um com o outro. O polipéptido não NgR pode ser fundido com o terminal N ou o terminal C do polipéptido NgR.

Por exemplo, num caso a proteína de fusão de NgR compreende um domínio NgR operativamente ligado ao domínio extracelular de uma segunda proteína. Tais proteínas de fusão podem ainda ser utilizadas em ensaios de pesquisa de compostos que modulem a actividade de NgR (tais ensaios são descritos em detalhe abaixo).

Por exemplo, a utilização de vectores comercialmente disponíveis que expressam um polipéptido desejado como parte de um produto de fusão com glutathione-S-transferase (GST) proporciona o polipéptido desejado possuindo um resíduo de glicina adicional na posição -1 após clivagem do componente GST a partir do polipéptido desejado.

Noutro caso, a proteína de fusão é uma proteína NgR contendo uma sequência sinal heteróloga no seu terminal N. Por exemplo, a sequência sinal nativa de NgR (i.e., aminoácidos 1-30 de SEQ ID NO:2 e aminoácidos 1-40 de SEQ ID NO:4) pode ser removida e substituída por uma sequência sinal de outra proteína. Em certas células hospedeiras (p.ex., células hospedeiras de mamífero), a expressão e/ou secreção de NgR pode ser aumentada através da utilização de uma sequência sinal heteróloga.

Ainda noutro caso, a proteína de fusão é uma proteína de fusão NgR-imunoglobulina na qual sequências de NgR compreendendo um ou mais domínios são fundidas com sequências derivadas de um membro da família de proteínas imunoglobulinas. As proteínas de fusão NgR-imunoglobulina desta divulgação, incluindo as do presente invento, podem ser incorporadas em composições farmacêuticas e administradas a um sujeito para inibir uma interacção entre o ligando de NgR e uma proteína NgR à superfície de uma célula, para deste modo suprimir a transdução de sinal mediada por NgR *in vivo*. As proteínas de fusão NgR-imunoglobulina podem ser utilizadas para afectar a biodisponibilidade de um ligando cognato de NgR. A inibição da interacção ligando de NgR/NgR pode ser terapêuticamente útil tanto para o tratamento de distúrbios proliferativos e de diferenciação, bem como para modulação (p.ex., promoção ou inibição) da sobrevivência celular. Para além disso, as proteínas de fusão NgR-imunoglobulina desta divulgação, incluindo as do presente invento, podem ser

utilizadas como imunogénios para produzir anticorpos anti-NgR num sujeito, para purificar ligandos de NgR, e em ensaios de pesquisa para identificar moléculas que inibam a interacção de NgR com o ligando de NgR.

Uma proteína quimérica ou de fusão de NgR desta divulgação, p.ex., uma proteína de fusão de NgR do presente invento, pode ser produzida através de técnicas padrão de ADN recombinante. Por exemplo, fragmentos de ADN codificando para as diferentes sequências polipeptídicas são ligados uns com os outros enquadrados de acordo com técnicas convencionais, p.ex., através do emprego de terminais tornados cegos ou tornados desfasados para ligação, digestão com enzimas de restrição para proporcionar terminais apropriados, preenchimento de extremidades coesivas conforme apropriado, tratamento com fosfatase alcalina para evitar união indesejável e ligação enzimática. Noutra concretização, o gene de fusão pode ser sintetizado através de técnicas convencionais incluindo sintetizadores automáticos de ADN. Alternativamente, pode ser realizada amplificação por PCR de fragmentos génicos utilizando iniciadores âncora que dão origem a pontas salientes complementares entre dois fragmentos génicos consecutivos que podem ser subsequentemente ligados e re-amplificados para gerar uma sequência génica quimérica (ver, por exemplo, Ausubel *et al.*, (Eds.) "Current Protocols in Molecular Biology", John Wiley & Sons, 1992). Para além disso, estão comercialmente disponíveis muitos vectores de expressão que codificam já uma porção da fusão (p.ex., um polipéptido GST). Um ácido nucleico codificando NgR pode ser clonado num tal vector de expressão para que a porção de fusão seja ligada enquadrada à proteína NgR.

Variantes resultantes da expressão noutros sistemas de vectores estão também contempladas.

As variantes de inserção incluem também proteínas de fusão em que o terminal amino e/ou o terminal carboxilo de NgR é/são fundidos com outro polipéptido.

Noutro aspecto, a divulgação proporciona variantes de deleção em que é removido um ou mais resíduos de aminoácidos de um polipéptido NgR. As deleções podem ser efectuadas num ou

em ambos os terminais do polipéptido NgR, ou com remoção de um ou mais resíduos de aminoácidos não terminais de NgR. As variantes de deleção incluem, portanto, todos os fragmentos de um polipéptido NgR.

A divulgação engloba também fragmentos polipeptídicos da sequência exposta em SEQ ID NO:2, 4 ou 14 em que os fragmentos mantêm propriedades biológicas (p.ex., ligação ao ligando e/ou sinalização intracelular) imunológicas de um polipéptido NgR. Fragmentos compreendendo pelo menos 4, 5, 10, 15, 20, 25, 30, 35 ou 40 aminoácidos consecutivos de SEQ ID NO:2, 4 ou 14 são contemplados por esta divulgação. Os fragmentos polipeptídicos preferidos apresentam propriedades antigénicas únicas ou específicas do NgR humano e seus homólogos alélicos e específicos. Fragmentos dos polipéptidos desta divulgação possuindo as propriedades biológicas e imunológicas desejadas podem ser preparados através de qualquer um dos métodos bem conhecidos e rotineiramente praticados na arte.

Ainda noutro aspecto, a divulgação proporciona variantes de substituição dos polipéptidos NgR. As variantes de substituição incluem os polipéptidos em que um ou mais dos resíduos de aminoácidos de um polipéptido NgR são removidos e substituídos por resíduos alternativos. Num aspecto, as substituições são de natureza conservativa; no entanto, a divulgação engloba também substituições que são não conservativas. Substituições conservativas para este fim podem ser definidas tal como exposto nas Tabelas 2, 3 ou 4 abaixo.

Tabela 1

X_{aa}# (com base num domínio NTLRRCT)	Coluna I (R1, R2, R3)	Coluna II (apenas R2+R3)
X ₁	G, R, M	
X ₂	A, D, C	
X ₃	V, T	
X ₄	N, P, S	
X ₅	E, A, S	
X ₆	nada, K	nada
X ₇	V, M, P	
X ₈	T, V	V

X_{aa}# (com base num domínio NTLRRCT)	Coluna I (R1, R2, R3)	Coluna II (apenas R2+R3)
X ₉	Q, P	Q
X ₁₀	Q, A	Q
X ₁₁	Q, H, N	
X ₁₂	G, N	N
X ₁₃	L, F	F
X ₁₄	Q, A, S	
X ₁₅	A, S	
X ₁₆	V, I	
X ₁₇	V, T, E, L	
X ₁₈	S, G	
X ₁₉	L, I	
X ₂₀	A, E, V, P	
X ₂₁	A, S, D	
X ₂₂	S, T	
X ₂₃	Q, E	
X ₂₄	IVL	
X ₂₅	Q, H	Q
X ₂₆	N, G	N
X ₂₁	R, L	
X ₂₈	T, G, R, S	
X ₂₉	F, L, T, H	
X ₃₀	L, V	L
X ₃₁	Q, R, P	
X ₃₂	Q, P, A	P
X ₃₃	G, A	G
X ₃₄	H, T, S	
X ₃₅	S, G, R	
X ₃₆	P, S, A	
X ₃₇	C, nada	nada
X ₃₉	R, nada	nada
X ₃₉	A, N	
X ₄₀	M, L	
X ₄₁	V, L, T	

X_{aa}# (com base num domínio NTLRRCT)	Coluna I (R1, R2, R3)	Coluna II (apenas R2+R3)
X ₄₂	T, I	T
X ₄₃	L, I	
X ₄₄	Y, F, H	
X ₄₅	N, V	N
X ₄₆	I, L	
X ₄₇	T, S, A	
X ₄₈	F, Y, T, R	
X ₄₉	A, H, Y, D	
X ₅₀	P, A	P
X ₅₁	N, S, G, A	
X ₅₂	T, A	T
X ₅₃	E, R, T	
X ₅₄	G, H	
X ₅₅	F, L	
A ₅₆	V, Q, H	
X ₅₇	H, A, L	
X ₅₈	E, Q	E
X ₅₉	G, S	G
X ₆₀	R, A	R
X ₆₁	Q, H	Q
X ₆₂	R, H	H
X ₆₃	T, S	
X ₆₄	L, V	L
X ₆₅	A, E, D	
X ₆₆	E, D, A	
X ₆₇	Q, H	Q
X ₆₈	V, E, G	
X ₆₉	K, R	
X ₇₀	H, Q	
X ₇₁	A, S, T	
X ₇₂	Y, H	
X ₇₃	Y, D	Y
X ₇₄	K, R	

X_{aa}# (com base num domínio NTLRRCT)	Coluna I (R1, R2, R3)	Coluna II (apenas R2+R3)
X ₇₅	G, Q	
X ₇₆	S, Q	S
X ₇₇	A, S, E	
X ₇₈	P, G	P
X ₇₉	A, G, P	
X ₈₀	G, N	
X ₈₁	I, V, L	
X ₈₂	G, R	
X ₈₃	H, V, A	
X ₈₄	S, A	S
X ₈₅	D, E	
X ₈₆	H, S, A	
X ₈₇	I, L	
X ₈₈	E, L, Q	
X ₈₉	Y, H, A	
X ₉₀	Q, P	Q
X ₉₁	D, N	
X ₉₂	I, L, T	
X ₉₃	V, A, R	
X ₉₄	V, A, G	
X ₉₅	S, T	S
X ₉₆	K, R	
X ₉₇	L, I	L
X ₉₈	W, R, S	
X ₉₉	S, L	
X ₁₀₀	L, V	L
X ₁₀₁	G, T, P	
X ₁₀₂	Q, P, E	
X ₁₀₃	G, H, R	
X ₁₀₄	I, T, V, A	
X ₁₀₅	V, G, H	
X ₁₀₆	N, S	
X ₁₀₇	E, G, Q	

X_{aa}# (com base num domínio NTLRRCT)	Coluna I (R1, R2, R3)	Coluna II (apenas R2+R3)
X ₁₀₈	Q, R	
X ₁₀₉	L, V	
X ₁₁₀	Q, A	
X ₁₁₁	W, G, H	
X ₁₁₂	H, R, P	
X ₁₁₃	K, A, H	
X ₁₁₄	H, R	
X ₁₁₅	D, G	
X ₁₁₆	H, R, S, G	
X ₁₁₇	T, M	
X ₁₁₉	T, I	
X ₁₁₉	F, Y	
X ₁₂₀	N, A	
X ₁₂₁	S, N	
X ₁₂₂	T, A, S	
X ₁₂₃	E, S, A	
X ₁₂₄	Q, P	
X ₁₂₅	G, T	
X ₁₂₆	D, E	D
X ₁₂₇	C, A	
X ₁₂₈	P, D	
X ₁₂₉	V, G, P, R	
X ₁₃₀	A, S	
X ₁₃₁	E, Q	Q
X ₁₃₂	F, Y	F
X ₁₃₃	G, A, D	
X ₁₃₄	A, P	
X ₁₃₅	D, A, V	
X ₁₃₆	G, D	
X ₁₃₇	A, E	
X ₁₃₈	S, P	
X ₁₃₉	E, A	
X ₁₄₀	L, F	

X_{aa}# (com base num domínio NTLRRCT)	Coluna I (R1, R2, R3)	Coluna II (apenas R2+R3)
X ₁₄₁	R, Q	
X ₁₄₂	R, K	R
X ₁₄₃	R, K	R
X ₁₄₄	F, A	
X ₁₄₅	G, V	
X ₁₄₆	A, D, E	
X ₁₄₇	T, P	
X ₁₄₈	A, V, S	
X ₁₄₉	T, S, L	
X ₁₅₀	E, G, P, Q	
X ₁₅₁	L, E, R	
X ₁₅₂	R, L	R
X ₁₅₃	G, D	
X ₁₅₄	Q, H, A	
X ₁₅₅	Q, R	
X ₁₅₆	K, R	
X ₁₅₇	L, A, R	
X ₁₅₈	R, A	R
X ₁₅₉	V, A, E	
X ₁₆₀	E, A, N	
X ₁₆₁	F, L	F
X ₁₆₂	R, Q	
X ₁₆₃	N, A, G	

Polipéptidos variantes incluem aqueles em que substituições conservativas foram introduzidas através de modificação de polinucleótidos codificando polipéptidos do presente invento. Os aminoácidos podem ser classificados de acordo com as propriedades físicas e a contribuição para a estrutura secundária e terciária da proteína. Uma substituição conservativa é reconhecida na arte como uma substituição de um aminoácido por outro aminoácido que tenha propriedades semelhantes. Substituições conservativas exemplares são expostas na Tabela 2 (de WO 97/09433, página 10, publicada a

13 de Março de 1997 (PCT/GB96/02197, apresentado a 9/6/96), imediatamente abaixo).

Tabela 2
Substituições Conservativas I

CARACTERÍSTICA DA CADEIA LATERAL	AMINOÁCIDO
Alifática	
Não polar	G A P I L V
Polar sem carga	C S T M N Q
Polar com carga	D E K R
Aromática	H F W Y
Outras	N Q D E

Alternativamente, os aminoácidos conservativos podem ser agrupados tal como descrito em Lehninger (Biochemistry, Segunda Edição; Worth Publishers, Inc. NY, NY (1975), pág. 71-77) tal como exposto na Tabela 3, imediatamente abaixo.

Tabela 3
Substituições Conservativas II

CARACTERÍSTICA DA CADEIA LATERAL	AMINOÁCIDO
Não polar (hidrófoba)	
A. Alifática:	A L I V P
B. Aromática:	F W
C. Contendo Enxofre:	M
D. Limiar:	G
Polar sem carga	
A. Hidroxilo:	S T Y
B. Amidas:	N Q
C. Sulfidrilo:	C
D. Limiar:	G
Carregada Positivamente (Básico):	K R H
Carregada Negativamente (Ácido):	D E

Ainda como outra alternativa, substituições conservativas exemplificativas são expostas na Tabela 4, abaixo.

Tabela 4

Substituições Conservativas III

Resíduo Original	Substituição Exemplificativa
Ala (A)	Val, Leu, Ile
Arg (R)	Lys, Gln, Asn
Asn (N)	Gln, His, Lys, Arg
Asp (D)	Glu
Cys (C)	Ser
Gln (Q)	Asn
Glu (E)	Asp
His (H)	Asn, Gln, Lys, Arg
Ile (I)	Leu, Val, Met, Ala, Phe,
Leu (L)	Ile, Val, Met, Ala, Phe
Lys (K)	Arg, Gln, Asn
Met (M)	Leu, Phe, Ile
Phe (F)	Leu, Val, Ile, Ala
Pro (P)	Gly
Ser (S)	Thr
Thr (T)	Ser
Trp (W)	Tyr
Tyr (Y)	Trp, Phe, Thr, Ser
Val (V)	Ile, Leu, Met, Phe, Ala

Adicionalmente, prevê-se também que os resíduos de aminoácidos que são conservados entre os membros da família das proteínas NgR da presente divulgação, tal como indicado pelo alinhamento aqui apresentado, não sejam particularmente passíveis de alteração. Por exemplo, as proteínas NgR da presente divulgação podem conter pelo menos um domínio que seja uma região tipicamente conservada em NgR. Exemplos destes

domínios conservados incluem, p.ex., domínio de repetição rico em leucina. Resíduos de aminoácidos que não são conservados ou são apenas semi-conservados entre membros das proteínas NgR podem ser facilmente passíveis de alteração.

Os NgR inteiros têm uma região LRR caracterizada pela sequência de aminoácidos de consenso mostrada em SEQ ID NO:19. Pelo menos alguns NgR inteiros incluem também um domínio de sinalização CT (CTS) e um domínio GPI.

As designações dos domínios NgR aqui utilizadas são definidas como se segue:

Domínio	hNgR1 SEQ ID:5	mNgR1 SEQ ID NO:17	NgR2 SEQ ID:2	hNgR3 SEQ ID:14	mNgR3 SEQ ID:4
Seq. de sinal	1-26	1-26	1-30	-	1-40
LRRNT	27-56	27-56	31-59	-	41-69
LRR1	57-81	57-81	60-82	5-27	70-92
LRR2	82-105	82-105	83-106	28-51	93-106
LRR3	106-130	106-130	107-131	52-76	106-141
LRR4	131-154	131-154	132-155	77-100	142-165
LRR5	155-178	155-178	156-179	101-124	166-189
LRR6	179-202	179-202	180-203	125-148	190-213
LRR7	203-226	203-226	204-227	149-172	214-237
LRR8	227-250	227-250	228-251	173-196	238-261
LRRCT	260-309	260-309	261-310	206-255	271-320
CTS (sinalização CT)	310-445	310-445	311-395	256-396	321-438
GPI	446-473	456-473	396-420	370-392	439-462

Nalguns casos, os domínios acima são modificados. A modificação pode ser de um modo que conserve a funcionalidade do domínio. A modificação pode incluir adição, deleção ou substituição de certos aminoácidos. Modificações exemplares incluem substituições de aminoácidos conservativas. De preferência, tais substituições totalizam 20 ou menos por 100 resíduos. De maior preferência, tais substituições totalizam 10 ou menos 100 resíduos. Outras modificações

exemplares incluem adição de sequências flanqueadoras de até cinco aminoácidos no terminal N e/ou terminal C de um ou mais dos domínios.

Nalguns casos, a molécula de ácido nucleico isolada codifica um polipéptido pelo menos cerca de 70%, 80%, 90%, 95%, 98% e de preferência pelo menos cerca de 99% homólogo a SEQ ID NO:2, 4 ou 14.

Podem ser introduzidas mutações em SEQ ID NO:1, 3 ou 13 através de técnicas padrão, p.ex., mutagénese dirigida ao local e mutagénese mediada por PCR. Podem ser feitas substituições de aminoácidos conservativas num ou mais resíduos de aminoácidos que se preveja serem não essenciais. Alternativamente, podem ser introduzidas mutações aleatoriamente ao longo de uma sequência de codificação de NgR. Isto pode ser alcançado, p.ex., através de mutagénese de saturação. Os mutantes resultantes podem ser pesquisados quanto a actividade biológica de NgR. As actividades biológicas de NgR podem incluir mas não se limitam a: (1) interacções proteína:proteína, p.ex., com outros NgR ou outras proteínas da superfície celular envolvidas em sinalização relacionada com Nogo; (2) formação de complexos com um ligando de NgR; (3) ligação a um anticorpo anti-NgR.

Deve entender-se que a definição de polipéptidos da divulgação pretende incluir polipéptidos possuindo modificações para além de inserção, deleção ou substituição de resíduos de aminoácidos. Para fins de exemplo, as modificações podem ser de natureza covalente e incluir por exemplo, ligação química com polímeros, lípidos, outras porções orgânicas e inorgânicas. Tais derivados podem ser preparados para aumentar a semivida em circulação de um polipéptido ou podem ser desenhados para melhorar a capacidade de direccionamento do polipéptido para células, tecidos ou órgãos desejados. Semelhantemente, a divulgação engloba ainda polipéptidos NgR que foram modificados covalentemente para incluir um ou mais polímeros solúveis em água tais como polietilenoglicol, polioxietilenoglicol ou polipropilenoglicol. As variantes que apresentem propriedades de ligação ao ligando do NgR nativo e são expressas a níveis mais elevados, bem como as variantes que proporcionem receptores constitutivamente activos, são

particularmente úteis em ensaios da divulgação; as variantes são também úteis a proporcionar modelos celulares, tecidulares e animais de doenças/condições caracterizadas por actividade de NgR aberrante.

Composições de um polipéptido NgR quimicamente modificado nas quais o polipéptido NgR está ligado a um polímero estão incluídas dentro do âmbito da presente divulgação e invento. O polímero pode ser solúvel em água para prevenir a precipitação da proteína num ambiente aquoso, tal como um ambiente fisiológico. Polímeros solúveis em água podem ser seleccionados de entre o grupo que consiste em, por exemplo, polietilenoglicol (PEG), monometoxipolietilenoglicol, dextrano, celulose ou outros polímeros baseados em hidratos de carbono, poli(N-vinilpirrolidona)polietilenoglicol, homopolímeros de polipropilenoglicol, um copolímero de óxido de polipropileno/óxido de etileno, polióis polioxietilados (p.ex. glicerol) e poli(álcool vinílico). O polímero seleccionado é habitualmente modificado para ter um único grupo reactivo, tal como um éster activo para acilação ou um aldeído para alquilação, para que o grau de polimerização possa ser controlado. Os polímeros podem ser de qualquer peso molecular e podem ser ramificados ou não ramificados, e podem também ser utilizadas misturas de tais polímeros. Quando o polímero NgR quimicamente modificado é destinado a utilização terapêutica, serão seleccionados para utilização polímeros farmacêuticamente aceitáveis.

Quando o polímero é para ser modificado através de uma reacção de acilação, o polímero deve ter um único grupo éster reactivo. Alternativamente, se o polímero é para ser modificado através de alquilação redutora, o polímero deve ter um único grupo aldeído reactivo. Um aldeído reactivo preferido é polietilenoglicol-propionaldeído, que é estável em água ou mono-Cl-ClO-alcoxi ou derivados ariloxi deste (ver Patente U.S. N.º 5252714).

A peguilação de polipéptidos NgR pode ser realizada através de qualquer uma das reacções de peguilação conhecidas na arte, tal como descrito, por exemplo, nas seguintes referências: *Focus on Growth Factors* 3: 4-10, 1992; EP 0154316; e EP 0401384. De preferência, a peguilação é

realizada através de uma reacção de acilação ou uma reacção de alquilação com uma molécula de polietilenoglicol reactiva (ou um polímero análogo reactivo solúvel em água). Um polímero solúvel em água preferido para peguilação de polipéptidos tal como NgR é polietilenoglicol (PEG). Tal como aqui se utiliza, pretende-se que "polietilenoglicol" englobe qualquer uma das formas de PEG que foram utilizadas para derivar outras proteínas, tais como mono(Cl-ClO)alcoxi- ou ariloxi-polietilenoglicol.

A derivação química de polipéptidos NgR pode ser efectuada sob quaisquer condições adequadas utilizadas para fazer reagir uma substância biologicamente activa com uma molécula polimérica activada. Os métodos para preparação de polipéptidos NgR peguilados compreenderão geralmente os passos de (a) reacção do polipéptido com polietilenoglicol, tal como um éster reactivo ou derivado aldeído de PEG, sob condições através das quais o polipéptido NgR fica unido a um ou mais grupos PEG, e (b) obtenção dos produtos de reacção. Será aparente a um vulgar perito na arte seleccionar as condições reaccionais óptimas ou as reacções de acilação baseadas em parâmetros conhecidos e o resultado desejado.

Os polipéptidos polímero:NgR peguilados e outros podem ser geralmente utilizados para tratar condições que podem ser aliviadas ou moduladas através de administração dos polipéptidos NgR aqui descritos. No entanto, as moléculas de polipéptidos polímero:NgR quimicamente derivadas aqui divulgadas podem ter actividades adicionais, actividade biológica aumentada ou reduzida ou outras características tais como semivida aumentada ou diminuída em comparação com as moléculas não derivadas. Os polipéptidos NgR, fragmentos destes, variantes e derivados, podem ser empregues sozinhos, juntos ou em combinação com outras composições farmacêuticas. As citocinas, factores de crescimento, antibióticos, anti-inflamatórios e/ou agentes quimioterapêuticos conforme apropriado para a indicação a tratar.

A presente divulgação proporciona composições compreendendo polipéptidos purificados desta divulgação, p.ex. polipéptidos purificados do presente invento. Composições farmacêuticas preferidas compreendem, para além do polipéptido

desta divulgação, um diluente líquido, semi-sólido ou sólido farmacêuticamente aceitável (i.e., estéril e não tóxico) que serve como veículo, excipiente ou meio farmacêutico. Pode ser utilizado qualquer diluente conhecido na arte. Diluentes exemplares incluem, mas não se limitam a água, soluções salinas, monolaurato de polioxietileno-sorbitano, estearato de magnésio, metil- e propil-hidroxibenzoato, talco, alginatos, amidos, lactose, sacarose, dextrose, sorbitol, manitol, glicerol, fosfato de cálcio, óleo mineral e manteiga de cacau.

As variantes que apresentam propriedades de ligação ao ligando do NgR nativo e são expressas a níveis mais elevados, bem como as variantes que proporcionam receptores constitutivamente activos, são particularmente úteis em ensaios desta divulgação; as variantes são também úteis em ensaios da divulgação e a proporcionar modelos celulares, tecidulares e animais de doenças/condições caracterizadas por actividade de NgR aberrante.

Com o conhecimento da informação da sequência nucleotídica aqui divulgada, um perito na arte pode identificar e obter sequências nucleotídicas que codifiquem NgR de diferentes fontes (i.e., diferentes tecidos ou diferentes organismos) através de uma variedade de meios bem conhecidos dos peritos na arte e divulgados, por exemplo, por Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Segunda Edição, Cold Spring Harbor Press, Cold Spring Harbor, NY (1989).

Por exemplo, o ADN que codifica NgR pode ser obtido através de pesquisa de ARNm, ADNc ou ADN genómico com sondas oligonucleotídicas geradas a partir da informação da sequência génica de NgR aqui proporcionada. As sondas podem ser marcadas com um grupo detectável, tal como um grupo fluorescente, um átomo radioactivo ou um grupo quimioluminescente de acordo com procedimentos conhecidos dos peritos na arte e utilizados em ensaios de hibridação convencionais, tal como descritos, por exemplo, por Sambrook et al., (1989) acima.

Uma molécula de ácido nucleico compreendendo qualquer uma das sequências nucleotídicas de NgR descritas acima pode ser alternativamente sintetizada através da utilização do

procedimento de reacção em cadeia da polimerase (PCR), com os iniciadores oligonucleotídicos de PCR produzidos a partir das sequências nucleotídicas aqui proporcionadas. Ver Patentes U.S. N^{os}. 4683195 para Mullis *et al.* e 4683202 para Mullis. A reacção de PCR proporciona um método para aumentar selectivamente a concentração de uma determinada sequência de ácido nucleico mesmo quando essa sequência não foi previamente purificada e está presente apenas numa única cópia numa determinada amostra. O método pode ser utilizado para amplificar ADN de cadeia simples ou de cadeia dupla. A essência do método envolve a utilização de duas sondas oligonucleotídicas para actuarem como iniciadores para a replicação dependente do molde, mediada pela polimerase, de uma molécula de ácido nucleico desejada.

Uma ampla variedade de metodologias alternativas de clonagem e amplificação *in vitro* é bem conhecida dos peritos na arte. Exemplos destas técnicas encontram-se, por exemplo, em Berger *et al.*, "Guide to Molecular Cloning Techniques", Methods in Enzymology, 152 Academic Press, San Diego, CA.

As moléculas de ácido nucleico da presente divulgação, incluindo as do presente invento, e fragmentos derivados destas, são úteis para pesquisa de polimorfismo de comprimento de fragmentos de restrição (RFLP) associado a certos distúrbios, bem como para mapeamento genético.

Anticorpos

São também compreendidos pela presente divulgação anticorpos (p.ex., anticorpos monoclonais e policlonais, anticorpos de cadeia simples, anticorpos quiméricos, anticorpos bifuncionais/biespecíficos, anticorpos humanizados, anticorpos humanos e anticorpos com região determinante de complementaridade (CDR) enxertada, incluindo compostos que incluam sequências de CDR que reconheçam especificamente um polipéptido desta divulgação) específicos para NgR ou fragmentos deste. Os anticorpos e fragmentos de ligação ao antigénio que proporcionam um aspecto do presente invento são expostos nas reivindicações anexas. Os anticorpos preferidos desta divulgação e do presente invento são anticorpos humanos que são produzidos e identificados de acordo com métodos descritos em W093/11236 publicado a 20 de Junho de 1993.

Fragmentos de anticorpo, incluindo Fab, Fab', F(ab')₂, e F_v, são também proporcionados por esta divulgação e pelo presente invento. O termo "específico para", quando utilizado para descrever os anticorpos desta divulgação, incluindo os anticorpos do presente invento, indica que as regiões variáveis dos anticorpos reconhecem e ligam-se exclusivamente a polipéptidos NgR (i.e., são capazes de distinguir polipéptidos NgR de outros polipéptidos NgR conhecidos em virtude de diferenças mensuráveis na afinidade de ligação, apesar da possível existência de identidade, homologia ou semelhança de sequência localizada, entre NgR e tais polipéptidos).

O péptido antigénico de NgR compreende pelo menos 8 resíduos de aminoácidos da sequência de aminoácidos em SEQ ID NO:2, 4 ou 14 e engloba um epítipo de NgR para que um anticorpo criado contra o péptido forme um complexo imunitário específico com NgR. De preferência, o péptido antigénico compreende pelo menos 10 resíduos de aminoácidos, de preferência pelo menos 15 resíduos de aminoácidos, ainda de preferência pelo menos 20 resíduos de aminoácidos, e de maior preferência pelo menos 30 resíduos de aminoácidos. Epítipos preferidos englobados pelo péptido antigénico são regiões de NgR que se localizam à superfície da proteína, p.ex., regiões hidrófilas.

Entender-se-á que anticorpos específicos podem também interagir com outras proteínas (por exemplo, proteína A de *S. aureus* ou outros anticorpos em técnicas de ELISA) através de interações com sequências fora da região variável dos anticorpos, e, em particular, na região constante da molécula. Ensaio de pesquisa para determinar a especificidade de ligação de um anticorpo desta divulgação, p.ex. um anticorpo do presente invento, são bem conhecidos e rotineiramente praticados na arte. Para uma discussão compreensiva de tais ensaios, ver Harlow et al., em "Antibodies: A Laboratory Manual", Cold Spring Harbor Laboratory Press; Cold Spring Harbor, NY (1988), Capítulo 6. Anticorpos que reconheçam e se liguem a fragmentos dos polipéptidos NgR desta divulgação, p.ex. os polipéptidos do presente invento, estão também contemplados, desde que os anticorpos sejam específicos para polipéptidos NgR. Os anticorpos desta divulgação, incluindo os

anticorpos do presente invento, podem ser produzidos utilizando qualquer método bem conhecido e rotineiramente praticado na arte.

Para a produção de anticorpos policlonais, vários animais hospedeiros adequados (p.ex., coelho, cabra, ratinho ou outro mamífero) podem ser imunizados através de injeção com a proteína nativa ou uma variante sintética desta ou um derivado da anterior. Uma preparação imunogénica apropriada pode conter, por exemplo, proteína NgR expressa de forma recombinante ou um polipéptido NgR quimicamente sintetizado. A preparação pode ainda incluir um adjuvante. Vários adjuvantes utilizados para aumentar a resposta imunológica incluem, mas não se limitam a, adjuvante de Freund (completo e incompleto), géis minerais (p.ex., hidróxido de alumínio), substâncias de superfície activa (p.ex., lisolecitina, polióis plurónicos, polianiões, péptidos, emulsões em óleo, dinitrofenol, etc.), adjuvantes humanos tais como *Bacille Calmette-Guerin* e *Corynebacterium parvum* ou agentes imunoestimuladores semelhantes. Se desejado, as moléculas de anticorpo dirigidas contra NgR podem ser isoladas do mamífero (p.ex., do sangue) e ainda purificadas através de técnicas bem conhecidas, tais como cromatografia de proteína A para obter a fracção de IgG.

O termo "anticorpo monoclonal" ou "composição de anticorpo monoclonal", tal como aqui se utiliza, refere-se a uma população de moléculas de anticorpo que contém apenas uma espécie de um local de ligação ao antigénio capaz de imuno-reagir com um determinado epítopo de NgR. Uma composição de anticorpo monoclonal apresenta assim tipicamente uma afinidade de ligação para uma determinada proteína NgR com a qual imuno-reage. Para preparação de anticorpos monoclonais dirigidos contra uma determinada NgR, ou derivados, fragmentos, análogos ou homólogos desta, pode ser utilizada qualquer técnica que proporcione a produção de moléculas de anticorpo através de cultura contínua de uma linha celular. Tais técnicas incluem, mas não se limitam à técnica de hibridoma (ver Kohler e Milstein, *Nature* 256: 495-497, 1975); a técnica de trioma; a técnica de hibridoma de células B humanas (ver Kozbor *et al.*, *Immunol. Today* 4: 72, 1983) e a técnica de hibridoma de EBV para produzir anticorpos monoclonais humanos (ver Cole *et al.*, (1985) em *Monoclonal Antibodies and Cancer Therapy*, Alan R

Liss, Inc., págs. 77-96). Os anticorpos monoclonais humanos podem ser utilizados na prática desta divulgação incluindo a prática do presente invento e podem ser produzidos através da utilização de hibridomas humanos (ver Cote et al., *Proc. Natl. Acad. Sci. USA* 80: 2026-2030, 1983) ou através de transformação de células B humanas com Vírus de Epstein-Barr *in vitro* (ver Cole et al., (1985), acima).

De acordo com esta divulgação, podem ser adaptadas técnicas para a produção de anticorpos de cadeia simples específicos para uma proteína NgR (ver p.ex., Patente U.S. N.º 4946778). Adicionalmente, podem ser adaptados métodos para a construção de bibliotecas de expressão de Fab (ver p.ex., Huse et al., *Science* 246: 1275-1281, 1989) para permitir uma identificação rápida e eficaz de fragmentos F_{ab} monoclonais com a especificidade desejada para uma proteína NgR ou derivados, fragmentos, análogos ou homólogos desta. Anticorpos não humanos podem ser "humanizados" através de técnicas bem conhecidas na arte. Ver, p.ex., Patente U.S. N.º 5225539. Num método, as CDRs não humanas são inseridas num anticorpo humano ou sequência estrutural de anticorpo de consenso. Mais alterações podem então ser introduzidas na estrutura do anticorpo para modular a afinidade ou imunogenicidade. Fragmentos de anticorpo que contêm os idiotipos para uma proteína NgR podem ser produzidos através de técnicas conhecidas na arte incluindo, mas não se limitando a: (i) um fragmento F(ab')₂ produzido através de digestão com pepsina de uma molécula de anticorpo; (ii) um fragmento Fab gerado através de redução das pontes dissulfureto de um fragmento F(ab')₂; (iii) um fragmento Fab gerado através do tratamento da molécula de anticorpo com papaína e um agente de redução; e (iv) fragmentos F_v.

Adicionalmente, anticorpos anti-NgR recombinantes, tais como anticorpos monoclonais quiméricos e humanizados, compreendendo tanto porções humanas como não humanas, que podem ser produzidos utilizando técnicas padrão de ADN recombinante, estão dentro do âmbito desta divulgação e do presente invento. Tais anticorpos monoclonais quiméricos e humanizados podem ser produzidos através de técnicas de ADN recombinante conhecidas na arte, por exemplo utilizando métodos descritos no Pedido Internacional PCT N.º

PCT/US86/02269; Pedido de Patente Europeia N.º 184187; Pedido de Patente Europeia N.º 171496; Pedido de Patente Europeia N.º 173494; Publicação Internacional PCT N.º WO 86/01533; Pat. U.S. N.º 4816567; Pedido de Patente Europeia N.º 125023; Better *et al.*, *Science* 240: 1041-1043, 1988; Liu *et al.*, *Proc. Natl. Acad. Sci. USA* 84: 3439-3443, 1987; Liu *et al.*, *J. Immunol.* 139: 3521-3526, 1987; Sun *et al.*, *Proc. Natl. Acad. Sci. USA* 84: 214-218, 1987; Nishimura *et al.*, *Cancer Res.* 47: 999-1005, 1987; Wood *et al.*, *Nature* 314: 446-449, 1985; Shaw *et al.*, *J. Natl. Cancer Inst.* 80: 1553-1559, 1988; Morrison, *Science* 229: 1202-1207, 1985; Oi *et al.*, *BioTechniques* 4: 214, 1986; Patente U.S. N.º 5225539; Jones *et al.*, *Nature* 321: 552-525, 1986; Verhoeyan *et al.*, *Science* 239: 1534, 1988; e Beidler *et al.*, *J. Immunol.* 141: 4053-4060, 1988.

De acordo com uma parte preferida desta divulgação, uma porção de um NgR é unida a uma porção Fc de um anticorpo para formar uma proteína de fusão NgR/Fc. De preferência, a proteína de fusão de Ig é solúvel. A proteína de fusão NgR/Fc pode ser formada através de técnicas recombinantes tal como descrito acima. Num caso, uma porção de um NgR incluindo a sequência de aminoácidos inteira de NgR excepto a região hidrófoba C-terminal é fundida com uma porção Fc de um anticorpo. Em casos preferidos, o NgR é um NgR humano e a Fc também é humana. De maior preferência, a porção Fc humana é derivada de um anticorpo IgG. Noutros casos, a sequência sinal N-terminal é omitida. Tais anticorpos são úteis na ligação a Nogo para impedir a sinalização de Nogo através do NgR.

Num caso, os métodos para a pesquisa de anticorpos que possuem a especificidade desejada incluem, mas não se limitam a, ensaio imunossorvente enzimático (ELISA) e outras técnicas imunologicamente mediadas conhecidas na arte. Num caso específico, a selecção de anticorpos que são específicos para um determinado domínio de uma proteína NgR é facilitada através da geração de hibridomas que se liguem ao fragmento de uma proteína NgR possuindo um tal domínio. Os anticorpos que são específicos para um ou mais domínios dentro de uma proteína NgR, p.ex., domínios abrangendo as regiões conservadas acima identificadas de NgR ou derivados, fragmentos, análogos ou homólogos destas, são também aqui proporcionados.

Os anticorpos anti-NgR podem ser utilizados em métodos conhecidos na arte referentes à localização e/ou quantificação de uma proteína NgR (p.ex., para utilizar em medição de níveis da proteína NgR dentro de amostras fisiológicas apropriadas, para utilizar em métodos de diagnóstico, para utilizar em imagiologia da proteína e semelhantes). Num dado caso, anticorpos para proteínas NgR, ou derivados, fragmentos, análogos ou homólogos destas, que contenham o domínio de ligação derivado do anticorpo, são utilizados como compostos farmacologicamente activos (daqui em diante "Produtos terapêuticos").

Um anticorpo anti-NgR (p.ex., anticorpo monoclonal) pode ser utilizado para isolar NgR através de técnicas padrão, tais como cromatografia de afinidade ou imunoprecipitação. Um anticorpo anti-NgR pode facilitar a purificação de NgR natural a partir de células e de NgR produzido e forma recombinante expresso em células hospedeiras. Para além disso, um anticorpo anti-NgR pode ser utilizado para detectar proteína NgR (p.ex., num lisado celular ou sobrenadante celular) para avaliar a abundância e o padrão de expressão da proteína NgR. Os anticorpos anti-NgR podem ser utilizados como diagnóstico para monitorizar níveis da proteína em tecido como parte de um procedimento de testes clínicos, p.ex., para determinar a eficácia de um dado regime de tratamento. A detecção pode ser facilitada através de ligação (i.e., ligação física) do anticorpo a uma substância detectável. Exemplos de substâncias detectáveis incluem várias enzimas, grupos prostéticos, materiais fluorescentes, materiais luminescentes, materiais bioluminescentes e materiais radioactivos. Exemplos de enzimas adequadas incluem peroxidase de rábano, fosfatase alcalina, β -galactosidase ou acetilcolinesterase; exemplos de complexos de grupos protéticos adequados incluem estreptavidina/biotina e avidina/biotina; exemplos de materiais fluorescentes adequados incluem umbeliferona, fluoresceína, isotiocianato de fluoresceína, rodamina, diclorotriazinilamina fluoresceína, cloreto de dansilo ou ficoeritrina; um exemplo de um material luminescente inclui luminol; exemplos de materiais bioluminescentes incluem luciferase, luciferina e aequorina, e exemplos de um material radioactivo adequado incluem ^{125}I , ^{131}I , ^{35}S ou ^3H .

Outro aspecto da presente divulgação é dirigido a métodos de indução de uma resposta imunitária num mamífero contra um polipéptido desta divulgação através de administração ao mamífero de uma quantidade do polipéptido suficiente para induzir uma resposta imunitária. A quantidade será dependente da espécie animal, do tamanho do animal, e semelhantes mas pode ser determinada pelos peritos na arte.

Outro aspecto desta divulgação é dirigido a anticorpos anti-idiotípicos e anticorpos anti-anti-idiotípicos. Um anticorpo anti-idiotípico é um anticorpo que reconhece determinantes de outro anticorpo (um anticorpo alvo). Geralmente, o anticorpo anti-idiotípico reconhece determinantes do local de ligação ao antigénio do anticorpo alvo. Tipicamente, o anticorpo alvo é um anticorpo monoclonal. Um anticorpo anti-idiotípico é geralmente preparado através de imunização de um animal (particularmente, ratinhos) da mesma espécie e tipo genético como fonte do anticorpo monoclonal alvo, com o anticorpo monoclonal alvo. O animal imunizado monta uma resposta imunitária para os determinantes idiotípicos do anticorpo monoclonal alvo e produz anticorpos contra os determinantes idiotípicos do anticorpo monoclonal alvo. Células produtoras de anticorpos, tais como células esplénicas, do animal imunizado podem ser utilizadas para gerar anticorpos monoclonais anti-idiotípicos. Para além disso, um anticorpo anti-idiotípico pode também ser utilizado para imunizar animais para produzir anticorpos anti-anti-idiotípicos. Estes animais imunizados podem ser utilizados para gerar anticorpos monoclonais anti-anti-idiotípicos utilizando técnicas padrão. Os anticorpos anti-anti-idiotípicos podem ligar-se ao mesmo epítipo que o anticorpo monoclonal alvo original utilizado para preparar o anticorpo anti-idiotípico. Os anticorpos anti-anti-idiotípicos representam outros anticorpos monoclonais com a mesma especificidade de antigénio que o anticorpo monoclonal alvo original.

Se a ligação do anticorpo anti-idiotípico com o anticorpo alvo for inibida pelo antigénio relevante do anticorpo alvo e se o anticorpo anti-idiotípico induzir uma resposta de anticorpos com a mesma especificidade que o anticorpo alvo,

este imita o antígeno do anticorpo alvo. Um tal anticorpo anti-idiotípico é um "anti-idiotipo de imagem interna" e é capaz de induzir uma resposta de anticorpos como se fosse o antígeno original. (Bona e Kohler (1984) *Anti-idiotípico Antibodies and Internal Image*, in *Monoclonal and Anti-idiotípico Antibodies: Probes for Receptor Structure and Function*", Venter J.C. et al., (Eds), Alan, R. Liss, New York, NY, págs. 141-149, 1984). Foi mostrado que vacinas incorporando anticorpos anti-idiotipo de imagem interna induzem respostas protectoras contra vírus, bactérias, e parasitas (Kennedy et al., (1986) 232: 220-223, 1047; McNamara et al., *Science* 226: 1325-1326, 1985). Foi também mostrado que os anticorpos anti-idiotípicos de imagem interna induzem imunidade a antígenos relacionados com tumores (Raychauhuri et al., *J. Immunol.* 137: 1743-1749, 1986; Raychauhuri et al., *J. Immunol.* 139: 3902-3910, 1987; Bhattacharya-Chatterjee et al., *J. Immunol.* 139: 1354-1360, 1987; Bhattacharya-Chatterjee et al., *J. Immunol.* 141: 1398-1403, 1988; Herlyn et al., *Intern. Rev. Immunol.* 4: 347-357, 1989; Chen et al., *Cell Imm. Immunother. Cancer* 351-359, 1990; Herlyn et al., *In vivo* 5: 615-624, 1991; Furuya et al., *AntiCancer Res.* 12: 27-32, 1992; Mittelman, A. et al., *Proc. Natl. Acad. Sci. USA* 89: 466-470, 1992; Durrant et al., *Cancer Res.* 54: 4837-4840, 1994; Mittelman et al., *Cancer Res.* 54: 415-421, 1994; Schmitt et al., *Hybridoma* 13: 389-396, 1994; Chakrobarty et al., *J. Immunother.* 18: 95-103, 1995; Chakrobarty et al., *Cancer Res.* 55: 1525-1530, 1995; Foon, K. A. et al., *Clin. Cancer Res.* 1: 1205-1294, 1995; Herlyn et al., *Hybridoma* 14: 159-166, 1995; Sclebusch et al., *Hybridoma* 14: 167-174, 1995; Herlyn et al., *Cancer Immunol Immunother.* 43: 65-76, 1996).

Os anticorpos anti-idiotípicos para NgR podem ser preparados, por exemplo, através da imunização de um animal, tal como um ratinho, com uma quantidade imunogénica de uma composição compreendendo NgR2 (SEQ ID NO:2), NgR3 (SEQ ID NOs:4 ou 14), ou uma porção imunogénica destas, contendo pelo menos um epítipo antigénico de NgR. A composição pode também conter um adjuvante adequado, e qualquer transportador necessário para proporcionar imunogenicidade. Os anticorpos monoclonais que reconhecem NgR podem ser preparados a partir das células do animal imunizado tal como descrito acima. Um anticorpo monoclonal reconhecendo um epítipo de NgR é então

seleccionado e utilizado para preparar uma composição compreendendo uma quantidade imunogénica do anticorpo monoclonal anti-NgR. Tipicamente, uma dose de 25 a 200 µg de monoclonal anti-NgR purificado seria suficiente num adjuvante adequado.

Os animais podem ser imunizados 2-6 vezes em intervalos de 14 a 30 dias entre doses. Tipicamente, os animais são imunizados através de qualquer via de administração adequada, tal como intraperitoneal, subcutânea, intravenosa ou uma combinação destas. A produção de anticorpo anti-idiotípico pode ser monitorizada durante o período de imunização utilizando métodos de imunoensaio padrão. Os animais com títulos de anticorpo adequados reactivos com os anticorpos monoclonais alvo podem ser re-imunizados com o anticorpo monoclonal utilizado como imunogénio três dias antes da colheita das células produtoras de anticorpo. De preferência, são utilizadas células do baço, embora possam ser seleccionadas outras células produtoras de anticorpos. As células produtoras de anticorpos são colhidas e fundidas com células de mieloma para produzir *Hibridomas*, tal como descrito acima, e são seleccionadas as células produtoras de anticorpos anti-idiotípicos adequadas.

Os anticorpos anti-anti-idiotípicos são produzidos através de outro ciclo de imunização e produção de *Hibridoma* através da utilização do anticorpo monoclonal anti-idiotípico como imunogénio.

Os anticorpos desta divulgação, incluindo anticorpos do presente invento, são úteis para, p.ex., fins terapêuticos (através de modulação da actividade de NgR), fins de diagnóstico para detectar ou quantificar NgR, e purificação de NgR. Portanto, estojos compreendendo um anticorpo desta divulgação, p.ex., um anticorpo do presente invento, para qualquer um dos fins aqui descritos estão também compreendidos.

Estojos

A presente divulgação é também dirigida a estojos, incluindo estojos farmacêuticos. Os estojos podem compreender qualquer uma das moléculas de ácido nucleico descritas acima,

qualquer um dos polipéptidos descritos acima, ou qualquer anticorpo que se ligue a um polipéptido da divulgação tal como descrito acima, bem como controlos apropriados, tais como controlos positivos e/ou negativos. O estojo compreende de preferência componentes adicionais, tais como, por exemplo, instruções, suporte sólido, reagentes úteis para quantificação e semelhantes. Por exemplo, o estojo pode compreender: um composto ou agente marcado capaz de detectar proteína ou ARNm de NgR numa amostra biológica; meios para determinação da quantidade de NgR na amostra; e meios para comparação da quantidade de NgR na amostra com um padrão. O composto ou agente pode ser embalado num recipiente adequado.

Ensaio de Pesquisa

A informação da sequência de ADN e de aminoácidos proporcionada pela presente divulgação torna também possível a identificação de compostos parceiros de ligação com os quais um polipéptido ou polinucleótido de NgR interagirá. Métodos para identificar compostos parceiros de ligação incluem ensaios em solução, ensaios *in vitro* em que os polipéptidos NgR são imobilizados e ensaios baseados em células. A identificação de compostos parceiros de ligação de polipéptidos NgR proporciona candidatos para intervenção terapêutica ou profilática em patologias associadas a actividade biológica normal e aberrante de NgR.

A divulgação proporciona também um método (também aqui referido como um "ensaio de pesquisa") para identificação de moduladores, i.e., compostos ou agentes candidatos ou de teste (p.ex., péptidos, peptidomiméticos, moléculas pequenas (p.ex., moléculas de menos de cerca de 1000 Daltons) ou outros fármacos) que se ligam a proteínas NgR ou têm um efeito estimulador ou inibidor, por exemplo, na expressão de NgR ou actividade de NgR.

Num caso, a divulgação proporciona ensaios para pesquisa de compostos candidatos ou de teste que se liguem a ou modulem a actividade de uma proteína ou polipéptido NgR ou porção biologicamente activa deste. Os compostos de teste da presente divulgação podem ser obtidos utilizando qualquer uma das numerosas abordagens em métodos de bibliotecas combinatórias conhecidos na arte, incluindo: bibliotecas biológicas;

bibliotecas de fase sólida ou fase em solução paralelas espacialmente abordáveis; métodos de bibliotecas sintéticas requerendo desconvolução; o método de bibliotecas “uma conta um composto”; e métodos de bibliotecas sintéticas utilizando selecção por cromatografia de afinidade. A abordagem de biblioteca biológica está limitada a bibliotecas peptídicas, enquanto as outras quatro abordagens são aplicáveis a bibliotecas de compostos peptídicos, oligómeros não peptídicos ou de molécula pequena (Lam *Anticancer Drug Des.* 12: 145, 1997).

Exemplos de métodos para a síntese de bibliotecas moleculares podem ser verificados na arte, por exemplo, em: DeWitt *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 6909, 1993; Erb *et al.*, *Proc. Natl. Acad. Sci. USA* 91: 11422, 1994; Zuckermann *et al.*, *J. Med. Chem.* 37: 2678, 1994; Cho *et al.*, *Science* 261: 1303, 1993; Carrell *et al.*, *Angew Chem. Int. Ed. Engl.* 33: 2059, 1994; Carell *et al.*, *Angew Chem. Int. Ed. Engl.* 33: 2061, 1994; e Gallop *et al.*, *J. Med. Chem.* 37: 1233, 1994.

As bibliotecas de compostos podem ser apresentadas em solução (p.ex., Houghten, *BioTechniques* 13: 412-421, 1992), ou em contas (Lam, *Nature* 354: 82-84, 1991), em chips (Fodor, *Nature* 364: 555-556, 1993), bactérias (Ladner, Patente U.S. N.º 5223409), esporos (Ladner, acima), plasmídeos (Cull *et al.*, *Proc. Natl. Acad. Sci. USA* 89: 1865-1869, 1992) ou em fagos (Scott e Smith, *Science* 249: 386-390, 1990; Devlin, *Science* 249: 404-406, 1990; Cwirla *et al.*, *Proc. Natl. Acad. Sci. USA* 87: 6378-6382, 1990; Felici, *J. Mol. Biol.* 222: 301-310, 1991; Ladner, acima).

1. Ensaaios Baseados em Células

A divulgação proporciona também ensaios baseados em células para identificar compostos parceiros de ligação de um polipéptido NgR. Num caso, a divulgação proporciona um método compreendendo os passos de pôr em contacto um polipéptido NgR expresso à superfície de uma célula com um composto parceiro de ligação candidato e de detecção da ligação do composto parceiro de ligação candidato ao polipéptido NgR. Noutro caso, um ensaio é um ensaio baseado em células compreendendo pôr em contacto uma célula expressando uma forma ligada à membrana da proteína NgR, ou uma porção biologicamente activa desta, à

superfície da célula com um composto de teste e a determinação da capacidade do composto de teste para modular (p.ex., estimular ou inibir) a actividade da proteína NgR ou porção biologicamente activa desta.

Num caso, um ensaio é um ensaio baseado em células em que uma célula que expressa uma forma ligada à membrana da proteína NgR, ou uma porção biologicamente activa desta, à superfície celular é posta em contacto com um composto de teste e a capacidade do composto de teste para se ligar a uma proteína NgR é determinada. A célula pode, por exemplo, ser de origem em mamífero ou uma célula de levedura. A determinação da capacidade do composto de teste para se ligar à proteína NgR pode ser alcançada, por exemplo, através da ligação do composto de teste a um marcador radioisótopo ou enzimático para que a ligação do composto de teste à proteína NgR ou porção biologicamente activa desta possa ser determinada através da detecção do composto marcado num complexo. Por exemplo, os compostos de teste podem ser marcados com ^{125}I , ^{35}S , ^{14}C ou ^3H , directamente ou indirectamente, e o radioisótopo detectado através de contagem directa da radioemissão ou através de contagem de cintilações. Alternativamente, os compostos de teste podem ser marcados enzimaticamente com, por exemplo, peroxidase de rábano, fosfatase alcalina ou luciferase, e o marcador enzimático detectado através de determinação da conversão de um substrato apropriado em produto. Num caso, o ensaio compreende pôr em contacto uma célula que expressa uma forma ligada à membrana da proteína NgR ou uma porção biologicamente activa desta, à superfície celular com um composto conhecido que se ligue a NgR para formar uma mistura de ensaio, pôr em contacto a mistura de ensaio com um composto de teste e determinação da capacidade do composto de teste para interagir com uma proteína NgR, em que a determinação da capacidade do composto de teste para interagir com uma proteína NgR compreende a determinação da capacidade do composto de teste para se ligar preferencialmente a NgR ou a uma porção biologicamente activa deste em comparação com o composto conhecido.

A determinação da capacidade do composto de teste para modular a actividade de NgR ou uma porção biologicamente activa desta pode ser alcançada, por exemplo, através da

determinação da capacidade da proteína NgR para se ligar ou interagir com uma molécula alvo de NgR. Tal como aqui se utiliza, uma "molécula alvo" é uma molécula com a qual uma proteína NgR se liga ou interage na natureza, por exemplo, uma molécula à superfície de uma célula que expressa uma proteína NgR, uma molécula à superfície de uma segunda célula, uma molécula no meio extracelular, uma molécula associada à superfície interna de uma membrana celular ou uma molécula citoplasmática. Uma molécula alvo de NgR pode ser uma molécula não NgR ou uma proteína ou polipéptido NgR da presente divulgação, p.ex., do presente invento. Num caso, uma molécula alvo de NgR é uma componente de uma via de transdução de sinal que facilita a transdução de um sinal extracelular (p.ex., um sinal gerado através da ligação de um composto a uma molécula NgR ligada à membrana) através da membrana celular e para dentro da célula. O alvo, por exemplo, pode ser uma segunda proteína intercelular que tenha actividade catalítica ou uma proteína que facilite a associação de moléculas de sinalização a jusante com NgR. Num caso preferido, a detecção compreende detecção de um fluxo de cálcio ou outro evento fisiológico na célula causado pela ligação da molécula.

Moléculas de ligação específicas, incluindo ligandos naturais e compostos sintéticos, podem ser identificadas ou desenvolvidas utilizando produtos NgR isolados ou recombinantes, variantes de NgR ou, de preferência, células expressando tais produtos. Parceiros de ligação são úteis para purificação de produtos NgR e detecção ou quantificação de produtos NgR em amostras de fluidos e tecidos utilizando procedimentos imunológicos conhecidos. As moléculas de ligação são também manifestamente úteis em modulação (i.e., bloqueio, inibição ou estimulação) de actividades biológicas de NgR, especialmente as actividades envolvidas em transdução de sinal.

2. Ensaio Livres de Células

(a) Ligação directa:

A divulgação inclui vários sistemas de ensaio para identificação de parceiros de ligação de NgR. Em ensaios em solução, os métodos da divulgação compreendem os passos de (a) pôr em contacto um polipéptido NgR com um ou mais compostos

parceiros de ligação candidatos e (b) identificação dos compostos que se ligam ao polipéptido NgR. A identificação dos compostos que se ligam ao polipéptido NgR pode ser alcançada através de isolamento do complexo polipéptido NgR/parceiro de ligação e separação do composto parceiro de ligação do polipéptido NgR. Um passo adicional de caracterização das propriedades físicas, biológicas e/ou bioquímicas do composto parceiro de ligação está também compreendido noutra parte da divulgação. Num aspecto, o complexo polipéptido NgR/parceiro de ligação é isolado utilizando um anticorpo imuno específico para o polipéptido NgR ou o composto parceiro de ligação candidato.

Ainda noutros casos, o polipéptido NgR ou o composto parceiro de ligação candidato compreende um marcador ou etiqueta que facilite o seu isolamento, e os métodos desta divulgação para identificar compostos parceiros de ligação incluem um passo de isolamento do complexo polipéptido NgR/parceiro de ligação através da interacção com o marcador ou etiqueta. Uma etiqueta exemplar deste tipo é uma sequência poli-histidina, geralmente em torno de seis resíduos de histidina, que permite o isolamento de um composto assim marcado utilizando quelacção com níquel. Outros marcadores e etiquetas, tais como a etiqueta FLAG[®] (Eastman Kodak, Rochester, NY), bem conhecidos e rotineiramente utilizados na arte, estão englobados pela divulgação.

(b) NgR imobilizado

Numa variação de um ensaio *in vitro*, a divulgação proporciona um método compreendendo os passos de (a) pôr em contacto um polipéptido NgR imobilizado, ou um fragmento biologicamente activo deste, com um composto parceiro de ligação candidato e (b) detecção da ligação do composto candidato ao polipéptido NgR. Numa variação alternativa, o composto parceiro de ligação candidato é imobilizado e é detectada a ligação a NgR. A imobilização é alcançada utilizando qualquer um dos métodos bem conhecidos na arte, incluindo ligação covalente a um suporte, uma conta ou uma resina cromatográfica, bem como interacções não covalentes, de elevada afinidade tais como ligação de anticorpo, ou utilização de ligação estreptavidina/biotina em que o composto imobilizado inclui uma porção biotina. A ligação de um

composto de teste a NgR, ou a interacção de NgR com uma molécula alvo na presença e ausência de um composto candidato, pode ser alcançada em qualquer vaso adequado para conter os reagentes. Exemplos de tais vasos incluem placas de microtitulação, tubos de ensaio e tubos de micro-centrífuga. Num caso, pode ser proporcionada uma proteína de fusão que adiciona um domínio que permite que uma ou ambas as proteínas sejam ligadas a uma matriz. Por exemplo, e não para fins de limitação, proteínas de fusão GST-NgR ou proteínas de fusão GST-alvo podem ser adsorvidas em contas de glutathione-Sepharose (Sigma Chemical, St. Louis, MO) ou placas de microtitulação com glutathione derivada, que são então combinadas com o composto de teste ou com a proteína alvo não adsorvida ou proteína NgR, e a mistura é incubada sob condições que conduzam à formação do complexo (p.ex., a condições fisiológicas para sal e pH). Após incubação, as contas ou poços das placas de microtitulação são lavados para remover quaisquer componentes não ligados, a matriz é imobilizada no caso de contas e os complexos são determinados directamente ou indirectamente, por exemplo, tal como descrito acima. Alternativamente, os complexos podem ser dissociados da matriz e o nível de ligação de NgR ou a actividade determinados utilizando técnicas padrão.

Outras técnicas para imobilização de proteínas em matrizes podem também ser utilizadas nos ensaios de pesquisa da divulgação. Por exemplo, o NgR ou a sua molécula alvo podem ser imobilizados utilizando conjugação de biotina e estreptavidina. O NgR ou moléculas alvo biotinilados podem ser preparados a partir de biotina-NHS (N-hidroxissuccinimida) utilizando técnicas bem conhecidas na arte (p.ex., estojo de biotinilação, Pierce Chemicals, Rockford, IL), e imobilizados nos poços de placas de 96 poços revestidas com estreptavidina (Pierce Chemical). Alternativamente, anticorpos reactivos com NgR ou moléculas alvo, mas que não interfiram com a ligação da proteína NgR à sua molécula alvo, podem ser derivados para os poços da placa, e o alvo não ligado ou o NgR aprisionado nos poços através de conjugação com o anticorpo. Os métodos para detecção de tais complexos, para além dos descritos acima para os complexos imobilizados com GST, incluem imunodetecção de complexos utilizando anticorpos reactivos com o NgR ou a molécula alvo, bem como ensaios ligados a enzimas que se

baseiam na detecção de uma actividade enzimática associada à NgR ou à molécula alvo.

A detecção de ligação pode ser alcançada (i) utilizando um marcador radioactivo no composto que não está imobilizado, (ii) utilizando um marcador fluorescente no composto não imobilizado, (iii) utilizando um anticorpo imuno específico para o composto não imobilizado, (iv) utilizando um marcador no composto não imobilizado que excita um suporte fluorescente ao qual o composto imobilizado está unido, (v) determinando a actividade do NgR, bem como outras técnicas bem conhecidas e rotineiramente praticadas na arte.

A determinação da actividade da molécula alvo pode, por exemplo, ser alcançada através de detecção da indução de um segundo mensageiro celular do alvo (i.e., Ca^{2+} intracelular, diacilglicerol, IP_3 , etc.), detecção de actividade catalítica/enzimática do alvo num substrato apropriado, detecção da indução de um gene repórter (compreendendo um elemento regulador que responda a NgR operativamente ligado a um ácido nucleico codificando um marcador detectável, p.ex., luciferase), ou detecção de uma resposta celular, por exemplo, sobrevivência celular, diferenciação celular ou proliferação celular.

(c) Experiências de competição

Ainda noutro caso, o ensaio compreende pôr em contacto a proteína NgR ou uma porção biologicamente activa desta com um composto conhecido que se ligue a NgR para formar uma mistura de ensaio, pôr em contacto a mistura de ensaio com um composto de teste, e determinação da capacidade do composto de teste para interagir com uma proteína NgR, em que a determinação da capacidade do composto de teste para interagir com uma proteína NgR compreende a determinação da capacidade do composto de teste para se ligar preferencialmente a NgR ou a uma porção biologicamente activa desta em comparação com o composto conhecido.

Ainda noutro caso, o ensaio livre de células compreende pôr em contacto a proteína NgR ou uma porção biologicamente activa desta com um composto conhecido que se ligue a NgR para formar uma mistura de ensaio, pôr em contacto a mistura de

ensaio com um composto de teste, e determinação da capacidade do composto de teste para interagir com uma proteína NgR, em que a determinação da capacidade do composto de teste para interagir com uma proteína NgR compreende a determinação da capacidade da proteína NgR para modular a actividade de uma molécula alvo de NgR.

Os ensaios livres de células da presente divulgação são passíveis de utilização tanto na forma solúvel como na forma ligada à membrana de NgR. No caso de ensaios livres de células compreendendo a forma ligada à membrana de NgR, pode ser desejável utilizar um agente solubilizante para que a forma ligada à membrana de NgR seja mantida em solução. Exemplos de tais agentes solubilizantes incluem detergentes não iónicos tais como n-octilglucósido, n-dodecilglucósido, n-dodecilmaltósido, octanoil-N-metilglucamida, decanoil-N-metilglucamida, Triton® X-100, Triton® X-114, Thesit®, Isotridecipoли(éter etilenoglicólico)_n, sulfonato de 3-(3-colamidopropil)dimetilaminiol-1-propano (CHAPS), sulfonato de 3-(3-colamidopropil)dimetilaminiol-2-hidroxi-1-propano (CHAPSO) ou sulfonato de N-dodecil-N,N-dimetil-3-amónio-1-propano.

Moduladores

Os agentes que modulam (i.e., aumentam, diminuem ou bloqueiam) a actividade ou expressão de NgR podem ser identificados através da incubação de um putativo modulador com uma célula contendo um polipéptido ou polinucleótido de NgR e determinação do efeito do putativo modulador na actividade ou expressão de NgR. A selectividade de um composto que module a actividade de NgR pode ser avaliada através da comparação dos seus efeitos em NgR com o seu efeito noutros compostos de NgR. Moduladores selectivos podem incluir, por exemplo, anticorpos e outras proteínas, péptidos ou moléculas orgânicas que se liguem especificamente a um polipéptido NgR ou um ácido nucleico codificando NgR. Os moduladores da actividade de NgR serão terapeuticamente úteis no tratamento de doenças e condições fisiológicas nas quais está envolvida actividade normal ou aberrante de NgR. Os polinucleótidos, polipéptidos e moduladores de NgR podem ser utilizados no tratamento de doenças e condições associadas a desmielinização. Os polinucleótidos e polipéptidos NgR, bem

como os moduladores de NgR, podem também ser utilizados em ensaios de diagnóstico para tais doenças ou condições.

Métodos da divulgação para identificar moduladores incluem variações de qualquer um dos métodos descritos acima para identificar compostos parceiros de ligação, incluindo as variações técnicas em que um composto parceiro de ligação foi identificado e o ensaio de ligação é realizado na presença e ausência de um modulador candidato. Um modulador é identificado nos casos onde a ligação entre o polipéptido NgR e o composto parceiro de ligação se altera na presença do modulador candidato em comparação com ligação na ausência do composto modulador candidato. Um modulador que aumente a ligação entre o polipéptido NgR e o composto parceiro de ligação é descrito como um estimulador ou activador, e um modulador que diminua a ligação entre o polipéptido NgR e o composto parceiro de ligação é descrito como um inibidor.

Noutro caso, os moduladores da expressão de NgR podem ser identificados num método em que uma célula é posta em contacto com um composto candidato e é determinada a expressão de ARNm de NgR ou da proteína na célula. O nível de expressão de ARNm de NgR ou da proteína na presença do composto candidato é comparado ao nível de expressão de ARNm de NgR ou da proteína na ausência do composto candidato. O composto candidato pode então ser identificado como um modulador da expressão de NgR com base nesta comparação. Por exemplo, quando a expressão de ARNm de NgR ou da proteína é superior (estatisticamente significativamente maior) na presença do composto candidato que na sua ausência, o composto candidato é identificado como um estimulador do ARNm de NgR ou da expressão da proteína. Alternativamente, quando a expressão de ARNm de NgR ou da proteína é inferior (estatisticamente significativamente menor) na presença do composto candidato que na sua ausência, o composto candidato é identificado como um inibidor da expressão de ARNm de NgR ou da proteína. O nível de expressão de ARNm de NgR ou de proteína nas células pode ser determinado através de métodos aqui descritos para detecção de ARNm de NgR ou da proteína.

Pesquisa de Elevada Tiragem

A divulgação compreende também ensaios de pesquisa de elevada tiragem (HTS) para identificar compostos que interajam com ou inibam a actividade biológica (i.e., afectam a actividade enzimática, actividade de ligação, etc.) de um polipéptido NgR. Os ensaios HTS permitem pesquisa de um grande número de compostos de um modo eficiente. Os sistemas HTS baseados em células estão contemplados para investigar a interacção receptor de NgR-ligando. Os ensaios HTS são desenhados para identificar "acertos" ou "compostos guia" possuindo a propriedade desejada, a partir dos quais podem ser desenhadas modificações para melhorar a propriedade desejada. A modificação química do "acerto" ou "composto guia" é frequentemente baseada numa relação estrutura/actividade identificável entre o "acerto" e o polipéptido NgR.

Outro aspecto da presente divulgação é dirigido a métodos de identificação de compostos que se ligam a NgR ou moléculas de ácido nucleico codificando NgR, compreendendo pôr em contacto NgR ou uma molécula de ácido nucleico codificando a mesma com um composto e determinação de se o composto se liga a NgR ou a uma molécula de ácido nucleico codificando a mesma. A ligação pode ser determinada através de ensaios de ligação que são bem conhecidos dos peritos na arte, incluindo, mas não se limitando a, ensaios de mudança em gel, *Western blots*, ensaio de competição radiomarcado, clonagem de expressão baseada em fagos, co-fraccionamento através de cromatografia, co-precipitação, ligação cruzada, análise de aprisionamento por interacção/dois híbridos, análise *southwestern*, ELISA e semelhantes, que são descritos em, por exemplo, Ausubel *et al.*, (Eds.), "Current Protocols in Molecular Biology", 1999, John Wiley & Sons, NY. As proteínas NgR podem, por exemplo, ser utilizadas como "proteínas isco" num ensaio de dois híbridos ou ensaio de três híbridos (ver, p.ex., Patente U.S. N.º 5283317; Zervos *et al.*, *Cell* 72: 223-232, 1993; Madura *et al.*, *J. Biol. Chem.* 268: 12046-12054, 1993; Bartel *et al.*, *BioTechniques* 14: 920-924, 1993; Iwabuchi *et al.*, *Oncogene* 8: 1693-1696, 1993; e Brent WO 94/10300), para identificar outras proteínas que se ligam ou interagem com NgR ("proteínas de ligação a NgR" ou "NgR-bp") e modulam a actividade de NgR. É também provável que tais proteínas de ligação a NgR estejam envolvidas na propagação de sinais através das proteínas NgR

tal como, por exemplo, elementos a montante ou a jusante da via de NgR.

Podem ser utilizados outros ensaios para identificar ligandos específicos de um receptor NgR, incluindo ensaios que identificam ligandos da proteína alvo através de medição da ligação directa de ligandos de teste à proteína alvo, bem como ensaios que identificam ligandos das proteínas alvo através de ultrafiltração de afinidade com métodos de espectroscopia de massa de vapor iónico/HPLC ou outros métodos físicos e analíticos. Alternativamente, tais interacções de ligação são avaliadas indirectamente utilizando o sistema de dois híbridos de levedura descritos em Fields *et al.*, *Nature* 340: 245-246, 1989, e Fields *et al.*, *Trends Genet.* 10: 286-292, 1994. O sistema de dois híbridos é um ensaio genético baseado na natureza modular da maioria dos factores de transcrição utilizados para detecção de interacções entre duas proteínas ou polipéptidos. Pode ser utilizado para identificar proteínas que se ligam a uma proteína conhecida de interesse ou para delinear domínios ou resíduos críticos para uma interacção. Variações nesta metodologia foram desenvolvidas para clonar genes que codificam proteínas de ligação ao ADN, para identificar péptidos que se ligam a uma proteína e para pesquisar fármacos. O sistema de dois híbridos explora a capacidade de um par de proteínas que interagem para pôr um domínio de activação da transcrição em íntima proximidade com um domínio de ligação ao ADN que se ligue a uma sequência de activação a montante (UAS) de um gene repórter, e é geralmente efectuado em levedura. O ensaio requer a construção de genes de dois híbridos codificando (1) um domínio de ligação ao ADN que é fundido com uma primeira proteína e (2) um domínio de activação fundido com uma segunda proteína. O domínio de ligação ao ADN direcciona a primeira proteína híbrida para o UAS do gene repórter; no entanto, como a maioria das proteínas não possui um domínio de activação, esta proteína híbrida de ligação ao ADN não activa a transcrição do gene repórter. A segunda proteína híbrida, que contém o domínio de activação, não pode activar por si própria a expressão do gene repórter porque não se liga ao UAS. No entanto, quando ambas as proteínas híbridas estão presentes, a interacção não covalente da primeira e segunda proteínas prende o domínio de activação ao UAS, activando a transcrição do gene repórter. Por exemplo,

quando a primeira proteína é um produto do gene de NgR ou um fragmento deste, que se saiba interagir com outra proteína ou ácido nucleico, este ensaio pode ser utilizado para detectar agentes que interfiram com a interacção de ligação. A expressão do gene repórter é monitorizada à medida que diferentes agentes de teste são adicionados ao sistema. A presença de um agente inibidor resulta na falta de um sinal repórter. Os compostos a pesquisar incluem (o que pode incluir compostos que são suspeitos de se ligarem a NgR, ou uma molécula de ácido nucleico codificando a mesma), mas não se limitam aos de origem extracelular, intracelular, biológica ou química.

A função do produto do gene de NgR não é certa e não se encontraram ainda ligandos que se liguem ao produto do gene. O ensaio de dois híbridos de levedura é útil para identificar proteínas que se ligam ao produto do gene. Num ensaio para identificar proteínas que se ligam a um receptor NgR, ou fragmento deste, pode ser utilizado um polinucleótido de fusão codificando tanto um receptor NgR (ou fragmento) como um domínio de ligação a UAS (i.e., uma primeira proteína). Adicionalmente, um grande número de genes híbridos codificando cada um uma segunda proteína diferente fundida com um domínio de activação é produzido e pesquisado no ensaio. Tipicamente, a segunda proteína é codificada por um ou mais membros de uma biblioteca de fusão de ADNc total ou ADN genómico, com cada uma das regiões codificando a segunda proteína fundida com o domínio de activação. Este sistema é aplicável a uma ampla variedade de proteínas e nem é necessário conhecer a identidade ou função da segunda proteína de ligação. O sistema é altamente sensível e pode detectar interacções não reveladas por outros métodos; mesmo interacções transientes podem desencadear transcrição para produzir um ARNm estável que pode ser repetidamente traduzido para produzir a proteína repórter.

Outros ensaios podem ser utilizados para pesquisar agentes que se ligam à proteína alvo. Um tal método de pesquisa para identificar a ligação directa de ligandos de teste a uma proteína alvo é descrito na Patente U.S. N.º 5585277. Este método baseia-se no princípio de que as proteínas existem geralmente como uma mistura de estados dobrado e não dobrado e alternam continuamente entre os dois

estados. Quando um ligando de teste se liga à forma dobrada de uma proteína alvo (i.e., quando o ligando de teste é um ligando da proteína alvo), a molécula da proteína alvo ligada ao ligando permanece no seu estado dobrado. Assim, a proteína alvo dobrada está presente numa maior extensão na presença de um ligando de teste que se liga à proteína alvo, que na ausência de um ligando. A ligação do ligando à proteína alvo pode ser determinada através de qualquer método que distinga entre os estados dobrado e não dobrado da proteína alvo. A função da proteína alvo não necessita ser conhecida para que este ensaio seja realizado. Virtualmente qualquer agente pode ser avaliado através deste método como ligando de teste, incluindo, mas não limitado a metais, polipéptidos, proteínas, lípidos, polissacáridos, polinucleótidos e moléculas orgânicas pequenas.

Outro método para identificação de ligandos de uma proteína alvo é descrito em Wieboldt *et al.*, *Anal. Chem.* 69: 1683-1691, 1997. Esta técnica pesquisa bibliotecas combinatórias de 20-30 agentes de cada vez em fase em solução para ligação à proteína alvo. Os agentes que se ligam à proteína alvo são separados de outros componentes da biblioteca através de simples lavagem da membrana. As moléculas especificamente seleccionadas que são retidas no filtro são subsequentemente libertadas da proteína alvo e analisadas através de HPLC e espectroscopia de massa de ionização por electrovaporização pneumáticamente assistida (*spray* iónico). Este procedimento selecciona componentes da biblioteca com a maior afinidade para a proteína alvo e é particularmente útil para bibliotecas de moléculas pequenas.

Os métodos desta divulgação englobam também ligandos, especialmente neuropéptidos, que são ligados a um marcador, tal como um radiomarcador (p.ex., ^{125}I , ^{35}S , ^{32}P , ^{33}P , ^3H), um marcador fluorescente, um marcador quimioluminescente, um marcador enzimático e um marcador imunogénico. Moduladores que recaem dentro do âmbito da divulgação incluem, mas não se limitam a, moléculas não peptídicas tais como imitadores não peptídicos, efectores alostéricos não peptídicos e péptidos. O polipéptido ou polinucleótido de NgR empregue num tal teste pode estar livre em solução, ligado a um suporte sólido, originário na superfície celular ou localizado

intracelularmente ou associado a uma porção de uma célula. Um perito na arte pode, por exemplo, medir a formação de complexos entre NgR e o composto a testar. Alternativamente, um perito na arte pode examinar a diminuição na formação de complexos entre NgR e o seu substrato causada pelo composto a testar.

Outro aspecto da presente divulgação é dirigido a métodos de identificação de compostos que modulam (i.e., aumentam ou diminuem) a actividade de NgR compreendendo pôr em contacto NgR com um composto e determinando se o composto modifica a actividade de NgR. A actividade na presença do composto de teste é comparada com a actividade na ausência do composto de teste. Quando a actividade da amostra contendo o composto de teste é superior à actividade na amostra sem o composto de teste, o composto terá maior actividade. Semelhantemente, quando a actividade da amostra contendo o composto de teste é inferior à actividade da amostra sem o composto de teste, o composto terá actividade inibida.

A presente divulgação é particularmente útil para pesquisa de compostos através de utilização de NgR em qualquer uma de uma variedade de técnicas de pesquisa de fármacos. Os compostos a pesquisar incluem (o que pode incluir compostos que são suspeitos de modular a actividade de NgR), mas não se limitam aos de origem extracelular, intracelular, biológica ou química. O polipéptido NgR empregue num tal teste pode estar em qualquer forma, de preferência, livre em solução, ligado a um suporte sólido, originário na superfície celular ou localizado intracelularmente. Um perito na arte pode, por exemplo, medir a formação de complexos entre NgR e o composto a testar. Alternativamente, um perito na arte pode examinar a diminuição na formação de complexos entre Nogo-R e o seu substrato causada pelo composto a testar.

A actividade de polipéptidos NgR do presente invento pode ser determinada através, por exemplo, do exame da capacidade para se ligarem ou serem activados por ligandos peptídicos quimicamente sintetizados. Alternativamente, a actividade do NgR pode ser ensaiada através de exame da sua capacidade para se ligar a iões de cálcio, hormonas, quimocinas, neuropeptídeos, neurotransmissores, nucleótidos, lípidos,

odorantes e fotões. Alternativamente, a actividade do NgR pode ser determinada através de exame da actividade de moléculas efectoras incluindo, mas não se limitando a, adenilato-ciclase, fosfolipases e canais iónicos. Assim, moduladores da actividade de NgR podem alterar uma função do receptor NgR, tal como uma propriedade de ligação de um receptor ou uma actividade. Em várias concretizações do método, o ensaio pode tomar a forma de um ensaio de fluxo iónico, um ensaio de crescimento de levedura, um ensaio de GTP não hidrolisável tal como um ensaio de [³⁵S]-GTP S, um ensaio de AMPc, um ensaio de inositol-trifosfato, um ensaio de diacilglicerol, um ensaio de Aequorina, um ensaio de Luciferase, um ensaio FLIPR para concentração intracelular de Ca²⁺, um ensaio de mitogénese, um ensaio da actividade de MAP-quinase, um ensaio de libertação de ácido araquidónico (p.ex., utilizando [³H]-ácido araquidónico) e um ensaio para taxas de acidificação extracelular, bem como outros ensaios baseados na ligação ou função da actividade de NgR que são geralmente conhecidos na arte. A actividade de NgR pode ser determinada através de metodologias que são utilizadas para ensaiar a actividade de FaRP, que é bem conhecida dos peritos na arte. As actividades biológicas dos receptores NgR de acordo com o presente invento incluem, mas não se limitam à ligação de um ligando natural ou não natural, bem como qualquer uma das actividades funcionais de NgR conhecidas na arte. Exemplos não limitantes das actividades de NgR incluem sinalização transmembranar de várias formas, que podem envolver associação de fosfatidilinositol (PI) e/ou o exercer de uma influência sobre PI; outra actividade exemplar dos NgR é a ligação de proteínas ou polipéptidos acessórios que diferem das proteínas GPI conhecidas.

Os moduladores da divulgação exibem uma variedade de estruturas químicas, que podem ser geralmente agrupadas em imitadores não peptídicos de ligandos do receptor NgR natural, efectores alostéricos peptídicos ou não peptídicos de receptores NgR e péptidos que possam funcionar como activadores ou inibidores (competitivos, incompetitivos e não competitivos) (p.ex., produtos de anticorpos) de receptores NgR. A divulgação não restringe as fontes para os moduladores adequados, que podem ser obtidos a partir de fontes naturais tais como extractos vegetais, animais ou minerais, ou fontes

não naturais tais como bibliotecas de moléculas pequenas, incluindo os produtos de abordagens combinatórias químicas à construção de bibliotecas, e bibliotecas peptídicas.

Podem ser utilizados outros ensaios para examinar a actividade enzimática incluindo, mas não se limitando aos fotométricos, radiométricos, HPLC, electromecânicos e semelhantes, que são descritos, por exemplo, em *Enzyme Assays: A Practical Approach*, Eisenthal e Danson (Eds.), 1992, Oxford University Press.

A utilização de ADNC em programas de descoberta de fármacos é bem conhecida; ensaios capazes de testar milhares de compostos desconhecidos por dia em pesquisas de elevada tiragem (HTSs) são minuciosamente documentados. A literatura está repleta de exemplos da utilização de ligandos radiomarcados em ensaios de ligação HTS para descoberta de fármacos (ver Williams, *Med. Res. Rev.* 11: 147-184, 1991; Sweetnam *et al.*, *J. Nat. Prod.* 56: 441-455, 1993 para revisão). São preferidos receptores recombinantes para ensaio de ligação HTS porque permitem uma melhor especificidade (maior pureza relativa), proporcionam a capacidade de gerar grandes quantidades de material receptor e podem ser utilizados numa ampla variedade de formatos (ver Hodgson, *Bio/Technology* 10: 973-980, 1992).

Está disponível uma variedade de sistemas heterólogos para expressão funcional de receptores recombinantes que são bem conhecidos dos peritos na arte. Tais sistema incluem bactérias (Strosberg *et al.*, *Trends Pharmacol. Sci.* 13: 95-98, 1992), levedura (Pausch, *Trends Biotechnol.* 15: 487-494, 1997), vários tipos de células de insecto (Vanden Broeck, *Int. Rev. Cytol.* 164: 189-268, 1996), células de anfíbio (Jayawickreme *et al.*, *Curr. Opin. Biotechnol.* 8: 629-634, 1997) e várias linhas celulares de mamífero (CHO, HEK293, COS, etc.; ver Gerhardt *et al.*, *Eur. J. Pharmacol.* 334: 1-23, 1997). Estes exemplos não excluem a utilização de outros sistemas de expressão celulares possíveis, incluindo linhas celulares obtidas a partir de nemátodes (pedido PCT WO 98/37177).

Em casos preferidos, os métodos de pesquisa de compostos que modulam a actividade de NgR compreendem pôr em contacto compostos de teste com NgR e ensaiar quanto à presença de um complexo entre o composto e NgR. Em tais ensaios, o ligando está tipicamente marcado. Após incubação adequada, o ligando livre é separado do presente na forma ligada e a quantidade de marcador livre ou não complexado é uma medida da capacidade do composto particular para se ligar a NgR.

Noutro caso, é empregue pesquisa de elevada tiragem para compostos possuindo afinidade de ligação adequada a NgR. Resumidamente, é sintetizado um grande número de diferentes compostos de teste peptídicos pequenos num substrato sólido. Os compostos de teste peptídicos são postos em contacto com NgR e lavados. O NgR ligado é então detectado através de métodos bem conhecidos na arte. Os polipéptidos purificados desta divulgação, p.ex. polipéptidos purificados do presente invento, podem também ser revestidos directamente em placas para utilização nas técnicas de pesquisa de fármacos acima mencionadas. Adicionalmente, podem ser utilizados anticorpos não neutralizantes para capturar a proteína e imobilizá-la no suporte sólido.

Geralmente, um NgR expresso pode ser utilizado para ensaios de ligação HTS em conjunto com o seu ligado definido. O péptido identificado é marcado com um radioisótopo adequado, incluindo, mas não se limitando a, ^{125}I , ^3H , ^{35}S ou ^{32}P , através de métodos que são bem conhecidos dos peritos na arte. Alternativamente, os péptidos podem ser marcados através de métodos bem conhecidos com um derivado fluorescente adequado (Baindur *et al.*, *Drug Dev. Res.* 33: 373-398, 1994; Rogers *Drug Discov. Today* 2: 156-160, 1997). O ligando radioactivo especificamente ligado ao receptor em preparações de membrana feitas a partir da linha celular que expressa a proteína recombinante pode ser detectado em ensaios HTS num de vários modos padrão, incluindo filtração do complexo receptor-ligando para separar o ligando ligado do ligando não ligado (Williams *Med. Res. Rev.* 11: 147-184, 1991; Sweetnam *et al.*, *J. Nat. Prod.* 56: 441-455, 1993). Métodos alternativos incluem um ensaio de proximidade de cintilação (SPA) ou um formato FlashPlate no qual tal separação é desnecessária (Nakayama *Curr. Opin. Drug Disc. Dev.* 1: 85-91, 1998, Bossé *et al.*, *J.*

Biomol. Screening 3: 285-292, 1998). A ligação de ligandos fluorescentes pode ser detectada de vários modos, incluindo transferência de energia de fluorescência (FRET), análise espectrofotofluorométrica directa de ligando ligado ou polarização de fluorescência (Rogers *Drug Discov. Today* 2: 156-160, 1997; Hill *Curr. Opin. Drug Disc. Dev.* 1: 92-97, 1998).

Exemplos de tais respostas biológicas incluem, mas não se limitam aos seguintes: a capacidade para sobreviver na ausência de um nutriente limitante em células de levedura especificamente modificadas (Pausch *Trends in Biotechnol.* 15: 487-494, 1997); alterações na concentração intracelular de Ca^{2+} medida através de corantes fluorescentes (Murphy *et al.*, *Curr. Opin. Drug Disc. Dev.* 1: 192-199, 1998). As alterações na fluorescência podem também ser utilizadas para monitorizar alterações induzidas pelo ligando no potencial de membrana ou pH intracelular; um sistema automático adequado para HTS foi descrito para estes fins (Schroeder *et al.*, *J. Biomol. Screening* 1: 75-80, 1996). Melanóforos preparados a partir de *Xenopus laevis* mostram uma alteração dependente do ligando na organização do pigmento em resposta à activação heteróloga de NgR; esta resposta é adaptável a formatos HTS (Jayawickreme *et al.*, *Curr. Opin. Biotechnol.* 8, 629-634, 1997). Estão também disponíveis ensaios para a medição de segundos mensageiros vulgares, incluindo AMPc, fosfoinosítidos e ácido araquidónico, mas estes não são geralmente preferidos para HTS.

Métodos preferidos de HTS empregando estes receptores incluem células CHO permanentemente transfectadas, nas quais podem ser identificados agonistas e antagonistas através da capacidade para transduzir o sinal para a ligação de Nogo em membranas preparadas a partir destas células através da putativa âncora GPI. Noutro caso, podem ser utilizadas células CHO permanentemente transfectadas para a preparação de membranas que contenham quantidades significativas das proteínas receptoras recombinantes; estas preparações de membrana seriam então utilizadas em ensaios de ligação do receptor, empregando o ligando radiomarcado específico para o receptor particular. Alternativamente, um ensaio funcional, tal como monitorização fluorescente de alterações induzidas

pelo ligando na concentração interna de Ca^{2+} ou potencial de membrana em células CHO permanentemente transfectadas contendo cada um destes receptores individualmente ou em combinação seria preferido para HTS. Igualmente preferido seria um tipo alternativo de célula de mamífero, tal como células HEK293 ou COS, em formatos semelhantes. Mais preferidas seriam linhas celulares de insecto permanentemente transfectadas, tais como células S2 de *Drosophila*. Ainda mais preferido seriam células de levedura recombinantes expressando os receptores de *Drosophila melanogaster* em formatos HTS bem conhecidos dos peritos na arte (p.ex., Pausch (1997), acima).

A divulgação contempla uma variedade de ensaios para pesquisar e identificar inibidores da ligação do ligando a receptores NgR. Num exemplo, o receptor NgR é imobilizado e a interacção com um parceiro de ligação é avaliada na presença e ausência de um modulador candidato tal como um composto inibidor. Noutro exemplo, a interacção entre o receptor NgR e o seu parceiro de ligação é avaliado num ensaio em solução, tanto na presença como na ausência de um composto inibidor candidato. Em qualquer um dos ensaios, um inibidor é identificado como um composto que diminui a ligação entre o receptor NgR e o seu parceiro de ligação. Outro ensaio contemplado envolve uma variação do ensaio de dois híbridos em que um inibidor de interacções proteína/proteína é identificado através da detecção de um sinal positivo numa célula hospedeira transformada ou transfectada, tal como descrito na publicação PCT número WO 95/20652, publicada a 3 de Agosto de 1995.

Moduladores candidatos contemplados pela divulgação incluem compostos seleccionados a partir de bibliotecas de potenciais activadores ou potenciais inibidores. Existem várias bibliotecas diferentes utilizadas para a identificação de moduladores de molécula pequena, incluindo: (1) bibliotecas químicas, (2) bibliotecas de produtos naturais, e (3) bibliotecas combinatórias constituídas por péptidos, oligonucleótidos ou moléculas orgânicas aleatórios. As bibliotecas químicas consistem em estruturas químicas aleatórias, algumas das quais são análogos de compostos conhecidos ou análogos de compostos que foram identificados como "acertos" ou "guias" noutras pesquisas de descoberta de

fármacos, alguns dos quais são derivados de produtos naturais e alguns dos quais surgem de química orgânica sintética não dirigida. As bibliotecas de produtos naturais são colecções de microrganismos, animais, plantas ou organismos marinhos que são utilizados para criar misturas para pesquisa através de: (1) fermentação e extracção de caldos de microrganismos do solo, vegetais ou marinhos ou (2) extracção de organismos vegetais ou marinhos. As bibliotecas de produtos naturais incluem policétidos, péptidos não ribossómicos e variantes (de ocorrência não natural) destes. Para uma revisão, ver Cane *et al.*, *Science* 282: 63-68, 1998. As bibliotecas combinatórias são compostas por um grande número de péptidos, oligonucleótidos ou compostos orgânicos como uma mistura. Estas bibliotecas são relativamente fáceis de preparar através de métodos de síntese automática tradicionais, PCR, clonagem ou métodos sintéticos sujeitos a direito de propriedade. De particular interesse são bibliotecas combinatórias não peptídicas. Ainda outras bibliotecas de interesse incluem bibliotecas de péptidos, proteínas, peptidomiméticos, de colecções multiparalelas sintéticas, recombinatórias e de polipéptidos. Para uma revisão da química combinatória e de bibliotecas criadas a partir desta, ver Myers *Curr. Opin. Biotechnol.* 8: 701-707, 1997. A identificação de moduladores através da utilização das várias bibliotecas aqui descritas permite a modificação do "acerto" (ou "guia") candidato para optimizar a capacidade do "acerto" para modular a actividade.

Ainda outros inibidores candidatos contemplados pela divulgação podem ser desenhados e incluem formas solúveis de parceiros de ligação, bem como parceiros de ligação como proteínas quiméricas ou de fusão. Um "parceiro de ligação" tal como aqui se utiliza engloba amplamente moduladores não peptídicos, bem como péptidos moduladores tais como neuropéptidos para além dos ligandos naturais, anticorpos, fragmentos de anticorpo e compostos modificados compreendendo domínios de anticorpo que sejam imuno-específicos para o produto de expressão do gene de NgR identificado.

Outros casos compreendem a utilização de ensaios de pesquisa competitivos nos quais anticorpos neutralizantes capazes de se ligar a um polipéptido desta divulgação, p.ex. um polipéptido do presente invento, competem especificamente

com um composto de teste pela ligação ao polipéptido. Deste modo, os anticorpos podem ser utilizados para detectar a presença de qualquer péptido que partilhe um ou mais determinantes antigénicos com NgR. Estudos de ligação competitiva radiomarcada são descritos em Lin et al., *Antimicrob. Agents Chemother.* 41: 2127-2131, 1997.

Noutros casos, os polipéptidos desta divulgação, incluindo os do presente invento, são empregues como ferramenta de investigação para identificação, caracterização e purificação de proteínas reguladoras que interagem. Marcadores apropriados são incorporados nos polipéptidos desta divulgação através de vários métodos conhecidos na arte e os polipéptidos são utilizados para capturar moléculas que interagem. Por exemplo, as moléculas são incubadas com os polipéptidos marcados, lavadas para remover polipéptidos não ligados e o complexo polipeptídico é quantificado. Os dados obtidos utilizando diferentes concentrações de polipéptido são utilizados para calcular valores para o número, a afinidade e a associação do polipéptido com o complexo proteico.

Os polipéptidos marcados são também úteis como reagentes para a purificação de moléculas com as quais os polipéptidos interajam incluindo, mas não se limitando a, inibidores. Num caso de purificação por afinidade, um polipéptido é covalentemente ligado a uma coluna cromatografia. As células e as suas membranas são extraídas e vários subcomponentes celulares são passados pela coluna. As moléculas ligam-se à coluna em virtude da sua afinidade com o polipéptido. O complexo polipeptídico é recuperado da coluna, dissociado e a molécula recuperada é sujeita a sequenciação proteica. Esta sequência de aminoácidos é então utilizada para identificar a molécula capturada ou para desenhar oligonucleótidos degenerados para clonagem do gene correspondente a partir de uma biblioteca de ADNc apropriada.

Alternativamente, podem ser identificados compostos que exibam propriedades semelhantes ao ligando para o NgR desta divulgação, mas que sejam menores e exibam uma semi-vida maior que o ligando endógeno no corpo de um humano ou animal. Quando é desenhado um composto orgânico, é utilizada uma molécula de acordo com o presente invento como um composto "guia". O

desenho de miméticos para compostos farmacêuticamente activos conhecidos é uma abordagem bem conhecida no desenvolvimento de produtos farmacêuticos baseados em tais compostos "guia". O desenho de miméticos, a síntese e os testes são geralmente utilizados para evitar a pesquisa de um grande número de moléculas para uma propriedade alvo. Para além disso, dados estruturais que derivam da análise das sequências de aminoácidos deduzidas codificadas pelos ADN da presente divulgação são úteis para desenhar novos fármacos, mais específicos e portanto com uma potência farmacológica mais elevada.

A comparação da sequência proteica da presente divulgação com as sequências presentes em todas as bases de dados disponíveis mostrou uma homologia significativa com a porção transmembranar de receptores acoplados à proteína G. Concordantemente, pode ser utilizada modelação por computador para desenvolver uma putativa estrutura terciária das proteínas do presente invento com base na informação disponível do domínio transmembranar de outras proteínas. Assim, podem ser desenhados novos ligandos baseados na estrutura prevista de NgR.

Esta divulgação refere-se ainda a novos agentes identificados através dos ensaios de pesquisa acima descritos e a utilizações destes para tratamentos tal como descrito acima.

Composições e Composições Farmacêuticas

Num caso particular, as novas moléculas identificadas através de métodos de pesquisa de acordo com esta divulgação são moléculas orgânicas de baixo peso molecular, caso em que uma composição ou composição farmacêutica pode ser preparada destas para administração oral ou parentérica. As composições, ou composições farmacêuticas, compreendendo as moléculas de ácido nucleico, vectores, polipéptidos, anticorpos e compostos identificados pelos métodos de pesquisa aqui descritos, compreendem tipicamente a molécula de ácido nucleico, proteína ou anticorpo e um transportador farmacêuticamente aceitável. Tal como aqui se utiliza, pretende-se que "transportador farmacêuticamente aceitável" inclua qualquer um e todos os solventes, meios de dispersão, revestimentos, agentes

antibacterianos e antifúngicos, agentes isotónicos e de retardamento da absorção e semelhantes, compatíveis com a administração farmacêutica. A natureza do transportador ou outros ingredientes dependerá da via específica de administração e da molécula particular a administrar. Exemplos de técnicas e protocolos que são úteis neste contexto encontram-se, *inter alia*, em "Remington's Pharmaceutical Sciences", 16^a ed., (1980) Osol, A (Ed.). Exemplos preferidos de tais transportadores ou diluentes incluem, mas não se limitam a, água, solução salina, solução de Ringer, solução de dextrose e albumina de soro humano a 5%. Lipossomas e veículos não aquosos tais como óleos fixados podem também ser utilizados. A utilização de tais meios e agentes para substâncias farmacêuticamente activas é bem conhecida na arte. Excepto na medida em que qualquer meio convencional ou agente é incompatível com o composto activo, a utilização deste nas composições é contemplada. Compostos activos suplementares podem também ser incorporados nas composições.

Uma composição farmacêutica desta divulgação é formulada para ser compatível com a sua via de administração pretendida. Exemplos de vias de administração incluem a oral e parentérica (p.ex., administração intravenosa, intradérmica, subcutânea, por inalação, transdérmica (tópica), transmucosa e rectal). As soluções ou suspensões utilizadas para aplicação parentérica, intradérmica ou subcutânea podem incluir os seguintes componentes: um diluente estéril tal como água para injeção, solução salina, óleos fixados, polietilenoglicóis, glicerina, propilenoglicol ou outros solventes sintéticos; agentes antibacterianos tais como álcool benzílico ou metilparabenos; antioxidantes tais como ácido ascórbico ou bissulfito de sódio; agentes quelantes tais como ácido etilenodiaminotetraacético; tampões tais como acetatos, citratos ou fosfatos, e agentes para o ajustamento da tonicidade tais como cloreto de sódio ou dextrose. O pH pode ser ajustado com ácidos ou bases, tais como ácido clorídrico ou hidróxido de sódio. A preparação parentérica pode ser fechada em ampolas, seringas descartáveis ou frasco multidose feitos de vidro ou plástico.

Composições farmacêuticas adequadas para utilização injectável incluem soluções aquosas estéreis (quando solúveis

em água) ou dispersões e pós estéreis para a preparação extemporânea de soluções ou dispersões estéreis injectáveis. Para administração intravenosa, os transportadores adequados incluem solução salina fisiológica, água bacteriostática, Cremophor ELTM (BASF, Parsippany, NJ) ou solução salina tamponada com fosfato (PBS). Em todos os casos, a composição tem de ser estéril e deve ser fluida até ao ponto em que exista fácil capacidade de encher uma seringa. Tem de ser estável sob as condições de fabrico e armazenamento e tem de ser conservada contra a acção contaminante de microrganismos tais como bactérias e fungos. O transportador pode ser um solvente ou meio de dispersão contendo, por exemplo, água, etanol, poliol (por exemplo, glicerol, propilenoglicol e polietilenoglicol líquido, e semelhantes) e misturas adequadas destes. A fluidez correcta pode ser mantida, por exemplo, através da utilização de um revestimento tal como lecitina, através da manutenção do tamanho da partícula necessário no caso de dispersão e através da utilização de tensioactivos. A prevenção da acção de microrganismos pode ser alcançada através de vários agentes antibacterianos e antifúngicos, por exemplo, parabenos, clorobutanol, fenol, ácido ascórbico, timerosal e semelhantes. Em muitos casos, será preferível incluir agentes isotónicos, por exemplo, açúcares, poliálcoois tais como manitol, sorbitol, cloreto de sódio na composição. A absorção prolongada das composições injectáveis pode ser obtida através de inclusão na composição de um agente que atrase a absorção, por exemplo, monostearato de alumínio e gelatina.

Soluções injectáveis estéreis podem ser preparadas através da incorporação do composto activo (p.ex., uma proteína NgR ou anticorpo anti-NgR) na quantidade necessária num solvente apropriado com um ou uma combinação de ingredientes enumerados acima, conforme necessário, seguido de esterilização por filtração. Geralmente, as dispersões são preparadas através de incorporação do composto activo num veículo estéril que contém um meio de dispersão básico e os outros ingredientes necessários dos enumerados acima. No caso de pós estéreis para a preparação de soluções injectáveis estéreis, os métodos de preparação são secagem em vácuo e secagem por congelação que produz um pó do ingrediente activo

mais qualquer ingrediente adicional desejado a partir de uma solução destes previamente esterilizada por filtração.

As composições orais incluem geralmente um diluente inerte ou um transportador edível. Podem ser encerradas em cápsulas de gelatina ou comprimidas em comprimidos. Para fins de administração terapêutica oral, o composto activo pode ser incorporado com excipientes e utilizado na forma de comprimidos, trociscos ou cápsulas. As composições orais podem também ser preparadas utilizando um transportador fluido para utilizar como colutório, em que o composto no transportador fluido é aplicado oralmente e bochechado e expectorado ou engolido. Agentes de ligação farmacêuticamente compatíveis e/ou materiais adjuvantes podem ser incluídos como parte da composição. Os comprimidos, pílulas, cápsulas, trociscos e semelhantes podem conter qualquer um dos seguintes ingredientes ou compostos de natureza semelhante: um aglomerante tal como celulose microcristalina, goma adragante ou gelatina; um excipiente tal como amido ou lactose, um agente desintegrante tal como ácido algínico, Primogel ou amido de milho; um lubrificante tal como estearato de magnésio ou Sterotes; um deslizante tal como dióxido de silicone coloidal; um agente adoçante tal como sacarose ou sacarina; ou um agente aromatizante tal como hortelã-pimenta, metilsalicilato ou aroma de laranja.

Para administração por inalação, os compostos são distribuídos na forma de um *spray* de aerossol a partir de um recipiente ou dispensador pressurizado que contenha um propulsor adequado, p.ex., um gás tal como dióxido de carbono ou um nebulizador.

A administração sistémica pode também ser através de meios transmucosos ou transdérmicos. Para administração transmucosa ou transdérmica, são utilizados na formulação penetrantes apropriados para a barreira ser permeada. Tais penetrantes são geralmente conhecidos na arte, e incluem, por exemplo, para administração transmucosa, detergentes, sais biliares e derivados de ácido fusídico. A administração transmucosa pode ser alcançada através da utilização de *sprays* nasais ou supositórios. Para administração transdérmica, os

compostos activos são formulados em unguentos, pomadas, géis ou cremes tal como são geralmente conhecido na arte.

Os compostos podem também ser preparados na forma de supositórios (p.ex., com bases de supositório convencionais tais como manteiga de cacau e outros glicéridos) ou enemas de retenção para distribuição rectal.

Num caso, os compostos activos são preparados com transportadores que protegerão o composto contra uma rápida eliminação do corpo, tal como uma formulação de libertação controlada, incluindo implantes e sistemas de distribuição microencapsulados. Podem ser utilizados polímeros biodegradáveis, biocompatíveis, tais como vinilacetato de etileno, polianidridos, ácido poliglicólico, colagénio, poliortoésteres e ácido poliláctico. Os métodos para preparação de tais formulações serão aparentes dos peritos na arte. Os materiais podem também ser obtidos comercialmente em Alza Corporation e Nova Pharmaceuticals, Inc. Podem também ser utilizadas suspensões lipossómicas (incluindo lipossomas para células infectadas com anticorpos monoclonais para antigénios virais) como transportadores farmacêuticamente aceitáveis. Estas podem ser preparadas de acordo com métodos conhecidos na arte, por exemplo, tal como descrito na Patente U.S. N.º 4522811. É especialmente vantajoso formular composições orais ou parentéricas na forma de unidade de dosagem para uma fácil administração e uniformidade da dosagem. A forma de unidade de dosagem tal como aqui se utiliza refere-se a unidades fisicamente discretas adequadas como dosagens unitárias para o sujeito a tratar; cada unidade contendo uma quantidade predeterminada de composto activo calculada para produzir o efeito terapêutico desejado em associação com o transportador farmacêutico requerido. As especificações para as formas de unidade de dosagem desta divulgação são ditadas pelas e directamente dependentes das características únicas do composto activo e do efeito terapêutico particular a alcançar.

As moléculas de ácido nucleico desta divulgação, incluindo as do presente invento, podem ser inseridas em vectores e utilizadas como vectores de terapia génica. Os vectores de terapia génica podem ser distribuídos a um sujeito através de qualquer uma de várias vias, p.ex., tal como

descrito na Patente U.S. N.º 5703055. A distribuição pode assim incluir também, p.ex., injeção intravenosa, administração local (ver Patente U.S. N.º 5328470) ou injeção estereotáctica (ver p.ex., Chen et al., *Proc. Natl. Acad. Sci. USA* 91: 3054-3057, 1994). A preparação farmacêutica do vector de terapia génica pode incluir o vector de terapia génica num diluente aceitável, ou pode compreender uma matriz de libertação lenta na qual o veículo de distribuição génica está embebido. Alternativamente, quando o vector de distribuição génica completo pode ser produzido intacto a partir de células recombinantes, p.ex., vectores retrovirais, a preparação farmacêutica pode incluir uma ou mais células que produzam o sistema de distribuição génica.

As composições farmacêuticas podem ser incluídas num recipiente, embalagem ou dispensador juntamente com instruções para administração.

A dosagem destes compostos de baixo peso molecular dependerá do estado de doença ou da condição a tratar e de outros factores clínicos tais como peso e condição do humano ou animal e da via de administração do composto. Para tratamento de humanos ou animais, podem ser administrados entre aproximadamente 0,5 mg/kg do peso corporal a 500 mg/kg do peso corporal do composto. A terapia é tipicamente administrada a dosagens inferiores e é continuada até ser observado o resultado terapêutico desejado.

Outro aspecto da presente divulgação é a utilização da sequência nucleotídica de NgR aqui divulgada para identificação de homólogos do Nogo-R noutros animais, incluindo mas não se limitando a humanos e outros mamíferos e invertebrados. Qualquer uma das sequências nucleotídicas aqui divulgadas ou qualquer porção destas, pode ser utilizada, por exemplo, como sonda para pesquisar bases de dados ou bibliotecas de ácidos nucleicos, tais como, por exemplo, bibliotecas de ADN genómico ou ADNc, para identificar homólogos utilizando procedimentos de pesquisa bem conhecidos dos peritos na arte. Concordantemente, podem ser identificados homólogos possuindo pelo menos 50%, de preferência pelo menos 60%, de maior preferência pelo menos 70%, de maior preferência pelo menos 80%, de maior preferência pelo menos 90%, de maior

preferência pelo menos 95%, e de maior preferência pelo menos 100% de homologia com sequências de NgR.

Os presentes compostos e métodos, incluindo moléculas de ácido nucleico, polipéptidos, anticorpos, compostos identificados pelos métodos de pesquisa aqui descritos, têm uma variedade de aplicações farmacêuticas e podem ser utilizados, por exemplo, para tratar ou prevenir crescimento celular desregulado, tal como crescimento de células de cancro e tumores. Num caso particular, as presentes moléculas são utilizadas em terapia génica. Para uma revisão dos procedimentos de terapia génica, ver p.ex., Anderson *Science* 256: 808-813, 1992.

A presente divulgação também ensina um método de agonização (estimulação) ou antagonização da actividade associada a um parceiro de ligação natural de NgR num mamífero compreendendo a administração ao referido mamífero de um agonista ou antagonista de um dos polipéptidos acima divulgados numa quantidade suficiente para efectuar o referido agonismo ou antagonismo. Um caso da presente divulgação, é então um método de tratamento de doenças num mamífero com um agonista ou antagonista da proteína da presente divulgação compreendendo a administração do agonista ou antagonista a um mamífero numa quantidade suficiente para agonizar ou antagonizar funções associadas a NgR.

Métodos de determinação das dosagens de compostos a administrar a um paciente e modos de administração de compostos a um organismo são divulgados no Pedido U.S. com o N.º de Série 08/702282, apresentado a 23 de Agosto de 1996, e na publicação de patente internacional número WO 96/22976, publicada a 1 de Agosto de 1996, ambos aqui incorporados por referência na sua totalidade, incluindo quaisquer desenhos, figuras ou tabelas. Os peritos na arte apreciarão que tais descrições são aplicáveis à presente divulgação e podem ser facilmente adaptadas a ela.

A dosagem correcta depende de vários factores tais como o tipo de doença a tratar, a composição particular a utilizar e o tamanho e a condição fisiológica do paciente. As doses terapêuticamente eficazes para os compostos aqui descritos

podem ser estimadas inicialmente a partir da cultura celular e modelos animais. Por exemplo, uma dose pode ser formulada em modelos animais para alcançar um intervalo de concentração em circulação que tenha inicialmente em conta a CI_{50} tal como determinada em ensaios de cultura celular. Os dados do modelo animal podem ser utilizados para determinar mais precisamente doses úteis em humanos.

A semi-vida plasmática e a biodistribuição do fármaco e dos metabolitos no plasma, em tumores e órgãos principais podem também ser determinadas para facilitar a selecção dos fármacos mais apropriados para inibir um distúrbio. Tais medições podem ser realizadas. Por exemplo, pode ser efectuada análise de HPLC no plasma de animais tratados com o fármaco e a localização de compostos radiomarcados pode ser determinada utilizando métodos de detecção tais como raios X, varrimento CAT e MRI. Os compostos que mostram potente actividade inibidora nos ensaios de pesquisa, mas têm fracas características farmacocinéticas, podem ser otimizados através da alteração da estrutura química e de novos testes. A este respeito, compostos que apresentem boas características farmacocinéticas podem ser utilizados como modelo.

Estudos de toxicidade podem também ser realizados através da medição da composição celular do sangue. Por exemplo, podem ser realizados estudos de toxicidade num modelo animal adequado como se segue: (1) o composto é administrado a ratinhos (um ratinho de controlo não tratado deverá também ser utilizado); (2) amostras de sangue são periodicamente obtidas através da veia caudal de um ratinho em cada grupo de tratamento; e (3) as amostras são analisadas quanto às contagens de glóbulos vermelhos e brancos, à composição celular do sangue e à percentagem de linfócitos *versus* células polimorfonucleares. Uma comparação dos resultados para cada regime de dosagem com os controlos indica se está presente toxicidade.

Na terminação de cada estudo de toxicidade, mais estudos podem ser realizados através do sacrifício dos animais (de preferência, de acordo com as linhas orientadoras da *American Veterinary Medical Association, Report of the American Veterinary Medical Assoc. Panel on Euthanasia, J. Am. Vet. Med*

Assoc. 202: 229-249, 1993). Os animais representativos de cada grupo de tratamento podem então ser examinados através de necrópsia grosseira quanto a provas imediatas de metástase, doença invulgar ou toxicidade. São observadas anomalias grosseiras no tecido e os tecidos são examinados histologicamente. Compostos que causem uma redução no peso corporal ou nos componentes sanguíneos são menos preferidos, tal como o são compostos que possuam um efeito adverso nos órgãos principais. Em geral, quanto maior for o efeito adverso menos preferido é o composto.

Para o tratamento de cancros a dose diária esperada de um agente farmacêutico hidrófobo é de entre 1 a 500 mg/dia, de preferência 1 a 250 mg/dia, e de maior preferência 1 a 50 mg/dia. Os fármacos podem ser distribuídos menos frequentemente desde que os níveis plasmáticos da porção activa sejam suficientes para manter a eficácia terapêutica. Os níveis plasmáticos devem reflectir a potência do fármaco. Geralmente, quanto mais potente o composto menores os níveis plasmáticos necessários para alcançar eficácia.

Transcritos de ARNm de NgR foram verificados no cérebro e coração. SEQ ID NOs: 1 e/ou 3 permitirão, tal como detalhado acima, a pesquisa dos neurotransmissores/hormonas/ligandos endógenos que activem, agonizem ou antagonizem NgR e de compostos com potencial utilidade no tratamento de distúrbios incluindo distúrbios do SNC (p.ex., AVC) e distúrbios degenerativos tais como os associados a desmielinização.

Por exemplo, a activação do receptor NgR pode mediar a prevenção de crescimento de neurites. A inibição será benéfica tanto em lesão cerebral crónica como aguda. Ver, p.ex., Donovan *et al.*, *J. Neurosci.* 17: 5316-5326, 1997; Turgeon *et al.*, *J. Neurosci.* 18: 6882-6891, 1998; Smith-Swintosky *et al.*, *J. Neurochem.* 69: 1890-1896, 1997; Gill *et al.*, *Brain Res.* 797: 321-327, 1998; Suidan *et al.*, *Semin. Thromb. Hemost.* 22: 125-133, 1996.

Farmacogenómica

Os agentes ou moduladores que têm um efeito estimulador ou inibidor na actividade de NgR (p.ex., expressão do gene de NgR), tal como identificado através de um ensaio de pesquisa

aqui descrito podem ser administrados a indivíduos para tratar (profilacticamente ou therapeuticamente) distúrbios (p.ex., uma condição de doença tal como um distúrbio de desmielinização) associado a actividade de NgR aberrante. Em conjunto com tal tratamento, a farmacogenómica (i.e., o estudo da relação entre o genótipo de um indivíduo e a resposta desse indivíduo a um composto ou fármaco estranho) do indivíduo pode ser considerada. As diferenças no metabolismo de produtos terapêuticos podem conduzir a toxicidade grave ou insuficiência terapêutica através de alteração da relação entre a dose e a concentração sanguínea do fármaco farmacologicamente activo. Assim, a farmacogenómica do indivíduo permite a selecção de agentes eficazes (p.ex., fármacos) para tratamentos profiláticos ou terapêuticos baseados numa consideração do genótipo do indivíduo. Tal farmacogenómica pode ainda ser utilizada para determinar as dosagens apropriadas e os regimes terapêuticos. Concordantemente, a actividade da proteína NgR, expressão do ácido nucleico de NgR ou o teor da mutação dos genes de NgR num indivíduo podem ser determinados para deste modo seleccionar o agente ou agentes apropriados para tratamento terapêutico ou profiláctico do indivíduo.

A farmacogenómica lida com variações hereditárias clinicamente significativas na resposta a fármacos devido à disposição alterada do fármaco e acção anormal em pessoas afectadas. Ver, p.ex., Eichelbaum, *Clin. Exp. Pharmacol. Physiol.* 23: 983-985, 1996 e Linder, *Clin. Chem.* 43: 254-266, 1997. Em geral, dois tipos de condições farmacogenéticas podem ser diferenciadas. Condições genéticas transmitidas como um único factor alterando o modo como os fármacos actuam no corpo (acção alterada do fármaco) ou condições genéticas transmitidas como factores únicos alterando o modo como o corpo actua nos fármacos (metabolismo alterado do fármaco). Estas condições farmacogenéticas podem ocorrer como defeitos raros ou como polimorfismos. Por exemplo, deficiência em glicose-6-fosfato-desidrogenase (G6PD) é uma enzimopatia herdada vulgar na qual a principal complicação clínica é hemólise após ingestão de fármacos oxidantes (anti-maláricos, sulfonamidas, analgésicos, nitrofuranos) e consumo de favas.

Como caso ilustrativo, a actividade de enzimas metabolizadoras de fármacos é um determinante principal tanto da intensidade como da duração da acção do fármaco. A descoberta de polimorfismos genéticos de enzimas metabolizadoras de fármacos (p.ex., N-acetiltransferase 2 (NAT 2) e enzimas CYP2D6 e CYP2C19 do citocromo P450) proporcionou uma explicação quanto ao porquê de alguns pacientes não obterem os efeitos esperados dos fármacos ou mostrarem uma resposta exagerada aos fármacos e toxicidade grave após tomarem a dose padrão e segura de um fármaco. Estes polimorfismos são expressos em dois fenótipos na população, o metabolizador extensivo (ME) e o metabolizador fraco (MF). A prevalência do MF é diferente entre diferentes populações. Por exemplo, o gene codificando para CYP2D6 é altamente polimórfico e foram identificadas várias mutações no MF, as quais conduziam todas à ausência de CYP2D6 funcional. Os metabolizadores fracos de CYP2D6 e CYP2C19 experimentam muito frequentemente uma resposta exagerada a fármacos e efeitos secundários quando recebem doses padrão. Se um metabolito for a porção terapêutica activa, o MF não mostra uma resposta terapêutica, tal como demonstrado para o efeito analgésico da codeína mediado pelo seu metabolito morfina formado através de CYP2D6. No outro extremo estão os designados metabolizadores ultra-rápidos que não respondem a doses padrão. Recentemente, a base molecular do metabolismo ultra-rápido foi identificada como sendo devida a amplificação do gene de CYP2D6.

Assim, pode ser determinada a actividade da proteína NgR, a expressão do ácido nucleico de NgR ou o teor das mutações dos genes de NgR num indivíduo para deste modo seleccionar agente ou agentes apropriados para tratamento terapêutico ou profilático do indivíduo. Adicionalmente, podem ser utilizados estudos farmacogenéticos para aplicar genotipagem de alelos polimórficos codificando enzimas metabolizadoras de fármacos para a identificação de um fenótipo de capacidade de resposta ao fármaco do indivíduo. Este conhecimento, quando aplicado à dosagem ou selecção de fármacos, pode evitar reacções adversas ou insuficiência terapêutica e aumentar assim a eficiência terapêutica ou profilática quando se trata um sujeito com um modulador de NgR, tal como um modulador identificado através de um dos ensaios de pesquisa exemplares aqui descritos.

Monitorização da Eficácia Clínica

A monitorização da influência de agentes (p.ex., fármacos, compostos) na expressão ou actividade de NgR (p.ex., a capacidade para modular proliferação e/ou diferenciação celular aberrante) pode ser aplicada não só em pesquisa de fármacos básicos, mas também em ensaios clínicos. Por exemplo, a eficácia de um agente determinada através de um ensaio de pesquisa tal como aqui descrito para aumentar a expressão do gene de NgR, os níveis proteicos ou sobrerregular a actividade de NgR, pode ser monitorizada em ensaios clínicos de sujeitos exibindo menor expressão do gene de NgR, menores níveis proteicos ou actividade subregulada de NgR. Alternativamente, a eficácia de um agente determinada através de um ensaio de pesquisa para diminuir a expressão do gene de NgR, dos níveis proteicos ou subregular a actividade de NgR, pode ser monitorizada em ensaios clínicos de sujeitos exibindo maior expressão de gene de NgR, maiores níveis proteicos ou actividade de NgR sobrerregulada. Em tais ensaios clínicos, a expressão ou actividade de NgR e, de preferência, de outros genes que tenham sido implicados, por exemplo, numa doença ou distúrbio, pode ser utilizada como uma "leitura" ou um marcador da capacidade de resposta imunitária de uma determinada célula.

Por exemplo, podem ser identificados genes, incluindo de NgR, que sejam modulados em células através de tratamento com um agente (p.ex., composto, fármaco ou molécula pequena) que module a actividade de NgR (p.ex., identificada num ensaio de pesquisa tal como aqui descrito). Assim, para estudar o efeito de agentes em distúrbios de desmielinização, por exemplo, num ensaio clínico, podem ser isoladas células e o ARN preparado e analisado quanto aos níveis de expressão de NgR e outros genes implicados no distúrbio. Os níveis de expressão génica (i.e., um padrão de expressão génica) podem ser quantificados através de análise *Northern blot* ou RT-PCR, tal como aqui descrito, ou alternativamente através de medição da quantidade de proteína produzida por um dos métodos tal como aqui descrito ou através de medição dos níveis de actividade de NgR ou de outros genes. Deste modo, o padrão de expressão génica pode servir como marcador, indicativo da resposta fisiológica das células ao agente. Concordantemente, este estado de resposta pode ser

determinado antes, e a vários pontos durante o tratamento do indivíduo com o agente.

Num caso, a divulgação proporciona um método para monitorização da eficácia do tratamento de um sujeito com um agente (p.ex., um agonista, antagonista, proteína, péptido, peptidomimético, ácido nucleico, molécula pequena ou outro fármaco candidato identificado através dos ensaios de pesquisa aqui descritos) compreendendo os passos de (i) obtenção de uma amostra pré-administração de um sujeito antes da administração do agente; (ii) detecção do nível de expressão de uma proteína NgR, ARNm ou ADN genómico na amostra pré-administração; (iii) obtenção de uma ou mais amostras pós-administração do sujeito; (iv) detecção do nível de expressão ou actividade da proteína NgR, ARNm ou ADN genómico nas amostras pós-administração; (v) comparação do nível de expressão ou actividade da proteína NgR, ARNm ou ADN genómico na amostra pré-administração com a proteína NgR, ARNm ou ADN genómico na amostra ou amostras pós-administração; e (vi) alteração da administração do agente ao sujeito concordantemente. Por exemplo, uma maior administração do agente pode ser desejável para aumentar a expressão ou actividade de NgR até níveis superiores aos detectados, i.e., para aumentar a eficácia do agente. Alternativamente, uma menor administração do agente pode ser desejável para diminuir a expressão ou actividade de NgR até níveis inferiores aos detectados, i.e., para diminuir a eficácia do agente.

Métodos de Tratamento

A presente divulgação ensina métodos tanto profiláticos como terapêuticos de tratamento de um sujeito em risco de (ou susceptível a) um distúrbio ou possuindo um distúrbio associado a expressão ou actividade aberrante de NgR.

Doenças e distúrbios que são caracterizados por maiores níveis ou actividade biológica (relativamente a um sujeito não sofrendo da doença ou distúrbio) podem ser tratados com Produtos terapêuticos que antagonizem (i.e., reduzam ou inibam) a actividade. Os Produtos terapêuticos que antagonizam a actividade podem ser administrados de um modo terapêutico ou profilático. Produtos terapêuticos que podem ser utilizados incluem, mas não se limitam a: (i) um polipéptido NgR, ou

análogos, derivados, fragmentos ou homólogos deste; (ii) anticorpos para um péptido NgR; (iii) ácidos nucleicos codificando um péptido NgR; (iv) administração de ácido nucleico anti-sentido e ácidos nucleicos que sejam "disfuncionais" (i.e., devido a uma inserção heteróloga dentro das sequências de codificação para um péptido NgR) são utilizados para "knockout" da função endógena de um péptido NgR através de recombinação homóloga (ver, p.ex., Capecchi, *Science* 244: 1288-1292, 1989); ou (v) moduladores (i.e., inibidores, agonistas e antagonistas, incluindo imitadores de péptidos adicionais da divulgação ou anticorpos específicos para um péptido desta divulgação) que alterem a interacção entre um péptido NgR e o seu parceiro de ligação.

As doenças e distúrbios que são caracterizados por menores níveis ou actividade biológica (relativamente a um sujeito não sofrendo da doença ou distúrbio) podem ser tratados com Produtos terapêuticos que aumentem (i.e., são agonistas) a actividade. Os Produtos terapêuticos que sobrerregulam a actividade podem ser administrados de um modo terapêutico ou profilático. Produtos terapêuticos que podem ser utilizados incluem, mas não se limitam a, um péptido NgR ou análogos, derivados, fragmentos ou homólogos deste; ou um agonista que aumente a biodisponibilidade.

Os maiores ou menores níveis podem ser facilmente detectados através de quantificação do péptido e/ou ARN, através da obtenção de uma amostra de tecido do paciente (p.ex., a partir de tecido de biopsia) e seu ensaio *in vitro* quanto aos níveis de ARN ou péptido, estrutura e/ou actividade dos péptidos expressos (ou ARNms de um péptido NgR). Métodos que são bem conhecidos na arte incluem, mas não se limitam a, imunoensaios (p.ex., através de análise *Western blot*, imunoprecipitação seguida de electroforese em gel de dodecilsulfato de sódio (SDS)-poliacrilamida, imunocitoquímica, etc.) e/ou ensaios de hibridação para detectar a expressão de ARNms (p.ex., ensaios *Northern*, *dot blots*, hibridação *in situ*, etc.).

A divulgação ensina um método para prevenção, num sujeito, de uma doença ou condição associada a uma expressão ou actividade aberrante de NgR, através da administração ao

sujeito de um agente que module a expressão de NgR ou pelo menos uma actividade de NgR. Os sujeitos em risco de uma doença que seja causada ou tenha a contribuição de uma expressão ou actividade aberrante de NgR podem ser identificados através, por exemplo, de qualquer um ou de uma combinação de ensaios de diagnóstico ou prognóstico tal como aqui descrito. A administração de um agente profilático pode ocorrer antes da manifestação dos sintomas característicos da aberração de NgR, para que seja prevenida uma doença ou distúrbio ou, alternativamente, seja atrasada na sua progressão. Dependendo do tipo de aberração de NgR, por exemplo, pode ser utilizado um agonista de NgR ou antagonista de NgR para tratamento do sujeito. O agente apropriado pode ser determinado com base em ensaios de pesquisa aqui descritos.

A divulgação ensina também métodos de modulação da expressão ou actividade de NgR para fins terapêuticos. O método modulador desta divulgação envolve pôr em contacto uma célula com um agente que module uma ou mais das actividades da actividade da proteína NgR associadas à célula. Um agente que modula a actividade da proteína NgR pode ser um agente tal como aqui descrito, tal como um ácido nucleico ou uma proteína, um ligando cognato de ocorrência natural de uma proteína NgR, um péptido, um peptidomimético de NgR ou outra molécula pequena. Num caso, o agente estimula uma ou mais actividades da proteína NgR. Exemplos de tais agentes estimuladores incluem a proteína NgR activa e uma molécula de ácido nucleico codificando NgR que foi introduzida na célula. Noutro caso, o agente inibe uma ou mais das actividades da proteína NgR. Exemplos de tais agentes inibidores incluem moléculas de ácido nucleico de NgR anti-sentido e anticorpos anti-NgR. Estes métodos moduladores podem ser efectuados *in vitro* (p.ex., através de cultura da célula com o agente) ou, alternativamente, *in vivo* (p.ex., através da administração do agente a um sujeito). Como tal, a presente divulgação proporciona métodos de tratamento de um indivíduo afligido com uma doença ou distúrbio caracterizado por expressão ou actividade aberrante de uma proteína NgR ou molécula de ácido nucleico. Num caso, o método envolve a administração de um agente (p.ex., um agente identificado através de um ensaio de pesquisa aqui descrito), ou combinação de agentes que modula

(p.ex., sobrerregula ou subregula) a expressão ou actividade de NgR. Noutro caso, o método envolve a administração de uma proteína ou molécula de ácido nucleico de NgR como terapia para compensar a expressão ou actividade reduzida ou aberrante de NgR.

Terapia Génica

Mutações no gene de NgR que resultem em perda da função normal do produto do gene de NgR estão subjacentes a estados de doença humana de NgR. A divulgação ensina terapia génica para restaurar a actividade de NgR para tratar esses estados de doença. A distribuição de um gene de NgR funcional a células apropriadas é efectuada *ex vivo*, *in situ*, ou *in vivo* através da utilização de vectores, e mais particularmente vectores virais (p.ex., adenovírus, vírus adeno-associados ou um retrovírus), ou *ex vivo* através da utilização de métodos de transferência física de ADN (p.ex., lipossomas ou tratamentos químicos). Ver, por exemplo, Anderson, *Nature*, suplemento a 392(6679): 25-20, 1998. Para revisões adicionais da tecnologia de terapia génica ver Friedmann, *Science* 244: 1275-1281, 1989; Verma, *Sci. Am.* 68-84, 1990; e Miller, *Nature* 357: 455-460, 1992. Alternativamente, está contemplado que noutros estados de doença humana, a prevenção da expressão, ou inibição da actividade de NgR será útil no tratamento de estados de doença. Está contemplado que a terapia anti-sentido ou terapia génica pode ser aplicada para regular negativamente a expressão de NgR.

A presente divulgação ensina métodos tanto profiláticos como terapêuticos de tratamento de um sujeito em risco de (ou susceptível a) um distúrbio ou possuindo um distúrbio associado a expressão ou actividade aberrante de NgR.

As doenças e distúrbios que são caracterizados por maiores níveis ou actividade biológica (relativamente a um sujeito não sofrendo da doença ou distúrbio) podem ser tratados com Produtos terapêuticos que antagonizem (i.e., reduzam ou inibam) a actividade. Os Produtos terapêuticos que antagonizam a actividade podem ser administrados de um modo terapêutico ou profilático. Produtos terapêuticos que podem ser utilizados incluem, mas não se limitam a, (i) um polipéptido NgR, ou análogos, derivados, fragmentos ou

homólogos deste; (ii) anticorpos para um péptido NgR; (iii) ácidos nucleicos codificando um péptido NgR; (iv) administração de ácido nucleico anti-sentido e ácidos nucleicos que sejam "disfuncionais" (i.e., devido a uma inserção heteróloga dentro das sequências de codificação de sequências de codificação para um péptido NgR) são utilizados para "*knockout*" da função endógena de um péptido NgR através de recombinação homóloga (ver, p.ex., Capecchi (1989), acima); ou (v) moduladores (i.e., inibidores, agonistas e antagonistas, incluindo imitadores de péptidos adicionais desta divulgação ou anticorpos específicos para um péptido desta divulgação) que alterem a interacção entre um péptido NgR e o seu parceiro de ligação.

As doenças e distúrbios que são caracterizados por menores níveis ou actividade biológica (relativamente a um sujeito não sofrendo da doença ou distúrbio) podem ser tratados com Produtos terapêuticos que aumentem (i.e., são agonistas) a actividade. Os Produtos terapêuticos que sobrerregulam a actividade podem ser administrados de um modo terapêutico ou profilático. Produtos terapêuticos que podem ser utilizados incluem, mas não se limitam a, um péptido NgR, ou análogos, derivados, fragmentos ou homólogos deste; ou um agonista que aumente a biodisponibilidade.

Maiores ou menores níveis podem ser facilmente detectados através da quantificação do péptido e/ou ARN, através da obtenção de uma amostra de tecido do paciente (p.ex., a partir de tecido de biopsia) e seu ensaio *in vitro* quanto aos níveis de ARN ou péptido, estrutura e/ou actividade dos péptidos expressos (ou ARNs de um péptido NgR). Métodos que são bem conhecidos na arte incluem, mas não se limitam a, imunoensaios (p.ex., através de análise *Western blot*, imunoprecipitação seguida de electroforese em gel de dodecilsulfato de sódio (SDS)-poliacrilamida, imunocitoquímica, etc.) e/ou ensaios de hibridação para detectar a expressão de ARNm (p.ex., ensaios *Northern*, *dot blots*, hibridação *in situ*, etc.).

A divulgação ensina um método para prevenção, num sujeito, de uma doença ou condição associada a uma expressão ou actividade aberrante de NgR, através da administração a um sujeito de um agente que module a expressão de NgR ou pelo

menos uma actividade de NgR. Os sujeitos em risco de uma doença que seja causada ou tenha a contribuição de uma expressão ou actividade aberrante de NgR podem ser identificados através, por exemplo, de qualquer um ou de uma combinação de ensaios de diagnóstico ou prognóstico tal como aqui descrito. A administração de um agente profilático pode ocorrer antes da manifestação dos sintomas característicos da aberração de NgR, para que seja prevenida uma doença ou distúrbio ou, alternativamente, seja atrasada na sua progressão. Dependendo do tipo de aberração de NgR, por exemplo, pode ser utilizado um agonista de NgR ou antagonista de NgR para tratamento do sujeito. O agente apropriado pode ser determinado com base em ensaios de pesquisa aqui descritos.

A divulgação ensina também métodos de modulação da expressão ou actividade de NgR para fins terapêuticos. O método modulador da divulgação envolve pôr em contacto uma célula com um agente que module uma ou mais das actividades da actividade da proteína NgR associada à célula. Um agente que modula a actividade da proteína NgR pode ser um agente tal como aqui descrito, tal como um ácido nucleico ou uma proteína, um ligando cognato de ocorrência natural de uma proteína NgR, um péptido, um peptidomimético de NgR ou outra molécula pequena. Num caso, o agente estimula uma ou mais actividades da proteína NgR. Exemplos de tais agentes estimuladores incluem a proteína NgR activa e uma molécula de ácido nucleico codificando NgR que foi introduzida na célula. Noutro caso, o agente inibe uma ou mais das actividades da proteína NgR. Exemplos de tais agentes inibidores incluem moléculas de ácido nucleico de NgR anti-sentido e anticorpos anti-NgR. Estes métodos moduladores podem ser efectuados *in vitro* (p.ex., através de cultura da célula com o agente) ou, alternativamente, *in vivo* (p.ex., através da administração do agente a um sujeito). Como tal, a presente divulgação proporciona métodos de tratamento de um indivíduo afligido com uma doença ou distúrbio caracterizado por expressão ou actividade aberrante de uma proteína NgR ou molécula de ácido nucleico. Num caso, o método envolve a administração de um agente (p.ex., um agente identificado através de um ensaio de pesquisa aqui descrito), ou combinação de agentes que modula (p.ex., sobrerregula ou subregula) a expressão ou actividade

de NgR. Noutro caso, o método envolve a administração de uma proteína NgR ou molécula de ácido nucleico como terapia para compensar a expressão ou actividade reduzida ou aberrante de NgR.

O presente invento não deve ser limitado no âmbito pelas concretizações específicas aqui descritas. De facto, várias modificações do presente invento para além das descritas aqui tornar-se-ão aparentes dos peritos na arte a partir da descrição anterior e da figura anexa.

A seguinte Tabela 5 contém as sequências de polinucleótidos e polipéptidos exemplares da divulgação.

Tabela 5

A seguinte sequência de ADN NgR2 <SEQ ID NO:1> foi identificada em humanos:

```

ATGCTGCCCCGGGCTCAGGCGCCTGCTGCAAGCTCCCGCCTCGGCCTGCCTCCTGCTGATG
CTCCTGGCCCTGCCCCTGGCGGCCCCAGCTGCCCCATGCTCTGCACCTGCTACTCATCC
CCGCCCACCGTGAGCTGCCAGGCCAACAACCTTCTCCTCTGTGCCGCTGTCCCTGCCACCC
AGCACTCAGCGACTCTTCCTGCAGAACAACCTCATCCGCACGCTGCGGCCAGGCACCTTT
GGGTCCAACCTGCTCACCCTGTGGCTCTTCTCCAACAACCTCTCCACCATCTACCCGGGC
ACTTTCCGCCACTTGCAAGCCCTGGAGGAGCTGGACCTCGGTGACAACCGGCACCTGCGC
TCGCTGGAGCCCGACACCTTCCAGGGCCTGGAGCGGCTGCAGTCGCTGCATTTGTACCGC
TGCCAGCTCAGCAGCCTGCCCCGGCAACATCTTCCGAGGCCTGGTCAGCCTGCAGTACCTC
TACCTCCAGGAGAACAGCCTGCTCCACCTACAGGATGACTTGTTGCGGGACCTGGCCAAC
CTGAGCCACCTCTTCCTCCACGGGAACCGCCTGCGGCTGCTCACAGAGCACGTGTTTCGC
GGCCTGGGCAGCCTGGACCGGCTGCTGCTGCACGGGAACCGGCTGCAGGGCGTGACCGC
GCGGCCTTCCGCGGCCTCAGCCGCCTCACCATCCTCTACCTGTTCAACAACAGCCTGGCC
TCGCTGCCCCGGCGAGGCGCTCGCCGACCTGCCCTCGCTCGAGTTCCTGCGGCTCAACGCT
AACCCCTGGGCGTGCGACTGCCGCGCGCGGCCGCTCTGGGCCTGGTTCCAGCGCGCGCGC
GTGTCCAGCTCCGACGTGACCTGCGCCACCCCCCGGAGCGCCAGGGCCGAGACCTGCGC
GCGCTCCGCGAGGCCGACTTCCAGGCGTGTCCGCCCCGCGCACCCACGCGGCCGGGCAGC
CGCGCCCGCGGCAACAGCTCCTCCAACCACCTGTACGGGGTGGCCGAGGCCGGGGCGCCC
CCAGCCGATCCCTCCACCCTCTACCGAGATCTGCCTGCCGAAGACTCGCGGGGGCGCCAG
GGCGGGGACGCGCCTACTGAGGACGACTACTGGGGGGGCTACGGGGGTGAGGACCAGCGA
GGGGAGCAGATGTGCCCCGGCGCTGCCTGCCAGGCGCCCCCGACTCCCGAGGCCCTGCG
CTCTCGGCCGGGCTCCCCAGCCCTCTGCTTTGCCTCCTGCTCCTGGTGCCCCACCACTC

```

A seguinte sequência de aminoácidos <SEQ ID NO:2> é a sequência de aminoácidos prevista derivada da sequência de ADN de SEQ ID NO:1:

M L P G L R R L L Q A P A S A C L L L M L L A L P L A A P S C
P M L C T C Y S S P P T V S C Q A N N F S S V P L S L P P S T
Q R L F L Q N N L I R T L R P G T F G S N L L T L W L F S N N
L S T I Y P G T F R H L Q A L E E L D L G D N R H L R S L E P
D T F Q G L E R L Q S L H L Y R C Q L S S L P G N I F R G L V
S L Q Y L Y L Q E N S L L H L Q D D L F A D L A N L S H L F L
H G N R L R L L T E H V F R G L G S L D R L L L H G N R L Q G
V H R A A F R G L S R L T I L Y L F N N S L A S L P G E A L A
D L P S L E F L R L N A N P W A C D C R A R P L W A W F Q R A
R V S S S D V T C A T P P E R Q G R D L R A L R E A D F Q A C
P P A A P T R P G S R A R G N S S S N H L Y G V A E A G A P P
A D P S T L Y R D L P A E D S R G R Q G G D A P T E D D Y W G
G Y G G E D Q R G E Q M C P G A A C Q A P P D S R G P A L S A
G L P S P L L C L L L L V P H H L

A seguinte sequência de ADN NgR3 <SEQ ID NO:3> foi identificada em ratinho:

ATGTCCTGGCAGTCTGGAACACAGTGACACAATCTCCCGTGCAGGCTGCTCAGGTCTCA
GGGTGCTGTGTGGAATTGCTGCTGTTGCTGCTCGCTGGAGAGCTACCTCTGGGTGGTGGT
TGTCCTCGAGACTGTGTGTGCTACCCTGCGCCCATGACTGTCAGCTGCCAGGCACACAAC
TTTGCTGCCATCCCGAGGGCATCCAGAGGACAGTGAGCGCATCTTCTGTCAGAACAAT
CGCATCACCTTCTCCAGCAGGGCCACTTCAGCCCCGCCATGGTCACCCTCTGGATCTAC
TCCAACAACATCACTTTTCATTGCTCCCAACACCTTCGAGGGCTTTGTGCATCTGGAGGAG
CTAGACCTTGGAGACAACCGACAGCTGCGAACGCTGGCACCCGAGACCTTCCAAGGCCTG
GTGAAGCTTCACGCCCTCTACCTCTATAAGTGTGGACTGAGCGCCCTGCCCGCAGGCATC
TTTGGTGGCCTGCACAGCCTGCAGTATCTCTACTTGCAGGACAACCATATCGAGTACCTC
CAAGATGACATCTTTGTGGACCTGGTCAATCTCAGTCACTTGTTTCTCCATGGTAACAAG
CTATGGAGCCTGGGCCAAGGCATCTTCCGGGGCCTGGTGAACCTGGACCGGTTGCTGCTG
CATGAGAACCAGCTACAGTGGGTTCAACACAAGGCTTTCATGACCTCCACAGGCTAACC
ACCCTCTTTCTCTTCAACAACAGCCTCACTGAGCTGCAGGGTGACTGTCTGGCCCCCTG
GTGGCCTTGGAGTTCCTTCGCCTCAATGGGAATGCTTGGGACTGTGGCTGCCGGGCACGT
TCCCTGTGGGAATGGCTGCGAAGGTTCCGTGGCTCTAGCTCTGCTGTCCCCTGCGCGACC
CCCGAGCTGCGGCAAGGCCAGGATCTGAAGCTGCTGAGGGTGGAGGACTTCCGGAACCTGC
ACAGGACCAAGTGTCTCCTCACCAGATCAAGTCTCACACGCTTACCACCTCTGACAGGGCT
GCCCCGAAGGAGCACCATCCGTCCCATGGGGCCTCCAGGGACAAAGGCCACCCACATGGC
CATCCGCCTGGCTCCAGGTCAAGTTACAAGAAGGCAGGCAAGAACTGCACCAGCCACAGG
AACCAGGAACCAGATCTCTAAGGTGAGCTCTGGGAAAGAGCTTACCGAACTGCAGGACTAT
GCCCCGACTATCAGCACAAGTTCAGCTTTGACATCATGCCACCGCACGACCCAAGAGG
AAGGGCAAGTGTGCTCGCAGGACCCCCATCCGTGCCCCCAGTGGGGTGCAGCAGGCATCC
TCAGGCACGGCCCTTGGGGCCCCACTCCTGGCCTGGATACTGGGGCTGGCAGTCACTCTC
CGC

A seguinte sequência proteica <SEQ ID NO:4> é a proteína deduzida de SEQ ID NO:3:

MSWQSGTTVTQSPVQAAQVSGCCVELLLLLL
AGELPLGGGCPRDCVCYPAPMTVSCQAHNFA
AIEGIPEDSERIFLQNNRITFLQQGHFSPA
MVTLWIYSNNITFIAPNTFEGFVHLEELDLG
DNRQLRTLAPETFQGLVKLHALYLYKCGLSA
LPAGIFGGLHSLQYLYLQDNHIEYLQDDIFV
DLVNLSHLFLHGNKLWSLGQGIFRGLVNLDL
LLLHENQLQWVHHKAFHDLHRLTTFLFNNS
LTELQGDCLAPLVALEFLRLNGNAWDCGCRA
RSLWEWLRRFRGSSSAVPCATPELRQGQDLK
LLRVEDFRNCTGPVSPHQIKSHTLTTSDDRA
RKEHHPSHGASRDKGHPHGHPPGSRSGYKKA
GKNCTSHRNRNQISKVSSGKELTELQDYAPD
YQHKFSFDIMPTARPKRKKGKCARRTPIRAPS
GVQQASSGTALGAPLLAWILGLAVTLR

A seguinte sequência proteica <SEQ ID NO:5> é NgR1 de humanos:

MKRASAGGSRL LAWVLWLQAWQVAAPCPGA
C
CYNEPKVTTSCPQQGLQAVPVGIPAASQRI
FLHG NRISHVPAASFRA CRNL TLWLHS NVL
ARIDAA AFTGLALLEQLDLS DNAQLRSV DPA
TFHGLGRLHTLHLDR CGLQELGPGLFRGLAA
LQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLHQNRVAHV
HPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRLNDNPWVCD CRARPLWAWLQKFR
GSSSEVPCSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAA
DKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPLTAVRPEG
EPPGFPTSGPRRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSEGS GALPSLTCSLTPLGLALV
LWTVLGPC

A seguinte sequência de aminoácidos <SEQ ID NO:6> é uma Sequência de Consenso de NgR baseada em homologia com NgR1

**C P X X C X C Y X X P X X T X S C X X X X X X X X P X
 X X P X X X X R X F L X X N X I X X X X X X F X X X X X X X L W X
 X S N X X X X I X X X X F X X X X X L E X L D L X D N X X L R
 X X X P X T F X G L X X L X L X L X X C X L X X L X X X X F X
 G L X X L Q Y L Y L Q X N X X X X L X D D X F X D L X N L X H
 L F L H G N X X X X X X X X X F R G L X X L D R L L L H X N X
 X X X V H X X A F X X L X R L X X L X L F X N X L X X L X X X
 X L A X L X X L X X L R L N X N X W X C X C R A R X L W X W X
 X X X R X S S S X V X C X X P X X X X G X D L X X L X X X D X
 X X C X X X X X P X X P X X X X X X X X X X X X X X X X X X
 X X X X X X X X X X X X X X X X X X X X X X X X X X X X X X
 P P X X X S X X X X X X X X X X X X X X X X X X X X X X X X
 X X X X X X X X X X X X X X X X X X X X X X X X X X R X X X
 X X X X X X X X X X X X X X X X X X X X X X X L X X X X X
 X X X X X L**

A seguinte sequência proteica <SEQ ID NO:7> é o domínio activo de 66 aminoácidos de Nogo:

**R I Y K G V I Q A I Q K S D E G H P F R A Y L E S E V A I S E
 E L V Q K Y S N S A L G H V N C T I K E L R R L F L V D D L V
 D S L K**

A seguinte sequência proteica <SEQ ID NO:8> é a sequência de aminoácidos da NgR2 madura:

**C P M L C T C Y S S P P T V S C Q A N N F S S V P L S L P P S
 T Q R L F L Q N N L I R T L R P G T F G S N L L T L W L F S N
 N L S T I Y P G T F R H L Q A L E E L D L G D N R H L R S L E
 P D T F Q G L E R L Q S L H L Y R C Q L S S L P G N I F R G L
 V S L Q Y L Y L Q E N S L L H L Q D D L F A D L A N L S H L F
 L H G N R L R L L T E H V F R G L G S L D R L L L H G N R L Q
 G V H R A A F R G L S R L T I L Y L F N N S L A S L P G E A L
 A D L P S L E F L R L N A N P W A C D C R A R P L W A W F Q R
 A R V S S S D V T C A T P P E R Q G R D L R A L R E A D F Q A
 C P P A A P T R P G S R A R G N S S S N H L Y G V A E A G A P
 P A D P S T L Y R D L P A E D S R G R Q G G D A P T E D D Y W
 G G Y G G E D Q R G E Q M C P G A A C Q A P P D S R G P A L S
 A G L P S P L L C L L L L V P H H L**

A seguinte sequência proteica <SEQ ID NO:9> é a sequência de aminoácidos da NgR3 madura:

**CPRDCVCYPAPMTVSCQAHNFAAIPEGIPED
SERIFLQNNRITFLQQGHFSPAMVTLWIYSN
NITFIAPNTFEGFVHLEELDLGDNRQLRTLA
PETFQGLVKLHALYLYKCGLSALPAGIFGGL
HSLQYLYLQDNHIEYLQDDIFVDLVNLSHLF
LHGNKLWSLGQGIFRGLVNLDRLLLHENQLQ
WVHHKAFHDLHRLTTLFLFNNSLTELGDCCL
APLVALEFLRLNGNAWDCGCRARSLWEWLRR
FRGSSSAVPCATPELRQGQDLKLLRVEDFRN
CTGPVSPHQIKSHTLTTSDDRAARKEHHP SHG
ASRDKGHPHGHPPGSRSGYKKAGKNCTSHRN
RNQISKVSSGKELTELQDYAPDYQHKFSFDI
MPTARPKRKKGKCARRTPIRAPSGVQQASSGT
ALGAPLLAWILGLAVTLR**

A seguinte sequência de aminoácidos <SEQ ID NO:10> é um motivo de cisteína conservado (Domínio de cisteína 1) da NgR e homólogos baseada na Sequência de Consenso:

CPXXCXCYXXPXXTXSC

A seguinte sequência de aminoácidos <SEQ ID NO:11> é um motivo de cisteína conservado (Domínio de cisteína 2) da NgR e homólogos baseada na Sequência de Consenso:

**NXWXCXCRARXLWXWXXXXRXSSSXVXCXXP
XXXXGDXDLXXLXXDXXXC**

A seguinte sequência de aminoácidos <SEQ ID NO:12> é um domínio rico em Leucina conservado da NgR e homólogos baseados na Sequência de Consenso:

**RXFLXXNXIXXXXXXXXXFXXXXXXXXXLWXXSN
XXXXIXXXXFXXXXXLXLXDNXXLXXXX
PXTFXGLXXLXLXLCXLXLXXXXFXGLX
XLQYLYLQXNXXXXLXDDXFXDLXNLXHLFL
HGNXXXXXXXXXXFRGLXXLDRLLLHXNXXXX
VHXXAFXXLXRLXXLXLFXNXLXXLXXXXLA
XLXXLXXLRL**

A menos que indicado em contrário, X é qualquer aminoácido. Por exemplo, onde indicado X pode não ser um aminoácido. As características adicionais do presente invento serão aparentes a partir dos seguintes Exemplos. Os Exemplos 1-5 são reais, enquanto os restantes Exemplos são proféticos.

Tal como mostrado pelos seguintes Exemplos, um gene codificando novos NgR foi identificado através de análise computacional dos dados da sequência de ADN. As proteínas codificadas por NgR2 e NgR3 têm uma putativa sequência sinal, oito domínios de repetição ricos em leucina numa região rica em leucina conservada (SEQ ID NO:12), uma região rica em cisteína não conservada (SEQ ID NO:10) N-terminal à região rica em leucina, um segundo domínio rico em cisteína (SEQ ID NO:11) C-terminal à região rica em leucina, e um putativo local de ligação a glicofosfatidilinositol (ligação a GPI). NgR2 e NgR3 diferem da sequência de NgR anteriormente identificada. Os homólogos de NgR, quando comparados com NgR conhecidos, mostram uma sequência de consenso (SEQ ID NOs:6). Os putativos NgR2 e NgR3 maduros são mostrados na Tabela 5 como SEQ ID NOs: 8 e 9, respectivamente.

Exemplo 1: Pesquisa por Tblastn da base de dados HTG

A sequência proteica para a NgR humana (NgR1) (SEQ ID NO:5) foi utilizada para pesquisar a base de dados genómicos de elevada tiragem (HTG) cuja utilização é familiar dos peritos na arte. A base de dados HTG é uma parte de GenBank, uma base de dados de sequências genéticas compreensiva de NIH, que inclui uma colecção anotada de todas as sequências de ADN publicamente disponíveis (*Nucleic Acids Res.* 28: 15-8, 2000). A base de dados HRG inclui sequências obtidas a partir de ADN genómico. Dentro do ADN genómico, os genes são tipicamente codificados através de múltiplos segmentos de ADN designados exões. Assim quando se alinha uma sequência de ADNc (ou uma sequência proteica codificada por uma sequência de ADNc) com uma sequência genómica, a sequência será quebrada em segmentos dependendo do número de exões no gene.

O algoritmo BLAST, que significa *Basic Local Alignment Search Tool* (Ferramenta Básica de Pesquisa Local de Alinhamentos) é adequado para determinação da semelhança de sequências (Altschul et al., *J. Mol. Biol.* 215: 403-410, 1990). O suporte lógico para realização de análises BLAST está publicamente disponível através do *National Center for Biotechnology Information* (<http://www.ncbi.nlm.nih.gov/>). O algoritmo básico BLAST envolve primeiro a identificação de pares de sequências de elevado registo (HSPs) através da identificação de palavras curtas de comprimento W na sequência

de busca que coincidam ou satisfaçam algum registo limiar com valor positivo T quando alinhadas com uma palavra com o mesmo comprimento numa sequência da base de dados. T é referido como o limiar do registo da palavra de vizinhança (Altschul *et al.*, *supra*). Estes acertos iniciais de palavras de vizinhança actuam como sementes para iniciação de pesquisas para encontrar HSPs contendo-as. Os acertos de palavras são prolongados em ambas as direcções ao longo de cada sequência até o registo do alinhamento cumulativo poder ser aumentado. A extensão dos acertos de palavras em cada direcção é parada quando: 1) o registo do alinhamento cumulativo cair fora numa quantidade de X do seu valor máximo alcançado; 2) o registo cumulativo cair para zero ou abaixo, devido a acumulação de um ou mais alinhamentos de resíduos com registo negativo; ou 3) ser alcançado o final de uma das sequências. Os parâmetros do algoritmo Blast W , T e X determinam a sensibilidade e a velocidade do alinhamento. O programa Blast utiliza como defeito um comprimento de palavra (W) de 11, alinhamentos da matriz de registos BLOSUM62 (ver Henikoff *et al.*, *Proc. Natl. Acad. Sci. USA* 89: 10915-10919, 1992) (B) de 50, expectativa (E) de 10, $M=5$, $N=4$, e uma comparação de ambas as cadeias.

O algoritmo BLAST (Karlin *et al.*, *Proc. Natl. Acad. Sci. USA* 90: 5873-5787, 1993) e Gapped BLAST efectuam uma análise estatística da semelhança entre duas sequências. Uma medida da semelhança proporcionada pelo algoritmo BLAST é a probabilidade da menor soma ($P(N)$), que proporciona uma indicação da probabilidade de uma coincidência entre duas sequências nucleotídicas ou de aminoácidos ocorrer por acaso. Por exemplo, um ácido nucleico é considerado semelhante a um gene de NgR ou ADNc se a probabilidade da menor soma na comparação do ácido nucleico de teste com um ácido nucleico de NgR for de menos de cerca de 1, de preferência menos de cerca de 0,1, de maior preferência menos de cerca de 0,01, e de ainda preferência menos de cerca de 0,001.

Para pesquisar a base de dados HTG com a sequência da proteína NgR, utilizámos uma variação do algoritmo BLAST conhecido como o programa tblastn, que compara uma sequência proteica de busca contra uma base de dados de sequências nucleotídicas traduzidas dinamicamente em todos os enquadramentos de leitura (*J. Mol. Biol.* 215: 403-410, 1990;

Nucleic Acids Res. 25: 3389-3402, 1997). Os resultados da pesquisa por tblastn indicaram a presença de genes na base de dados com uma identidade significativa com NgR. Para além de descobrir acertos com clones genómicos que contêm os genes de NgR humano e de ratinho, verificámos acertos com clones onde a identidade não era elevada, mas era mesmo assim muito significativa. Foram encontrados três clones humanos (Números de Entrada: AC068514, AC016869, AC013606) com um valor de e de $4e^{-43}$ e foi encontrado um clone de ratinho (N.º de Entrada AC021768) com um valor de e de $1e^{-78}$. Os três clones humanos pareciam todos codificar o mesmo gene, assim as análises posteriores foram confinadas a AC013606.

Exemplo 2: Previsão da sequência proteica de NgR2 humana (AC013606)

A sequência proteica de NgR humana alinhou com duas regiões da sequência traduzida da sequência nucleotídica AC013606, indicando que o novo gene era codificado por pelo menos dois exões. Para definir o gene completo, utilizámos o programa de computador GENSCAN™ (*J. Mol. Biol.* 268: 78-94, 1997) que pode identificar estruturas exão/intrão completas de genes no ADN genómico. A previsão de genes através de GENSCAN™ continha sete exões. Através da comparação destes exões previstos com a NgR, foi concluído que o novo gene humano contém dois destes exões e uma parte de outro (contendo a metionina iniciadora). O ADNc previsto (ARNm) codificado por estes três exões foi montado a partir de AC013606 (HTG11; depositada em Março de 2000; comprimento = 143899; GenBank edição 118.0; SEQ ID NO:15) através da combinação de nucleótidos dos três exões cujas coordenadas são: 123292-123322 (exão 1); 130035-130516 (exão 2); e 138589-139335 (exão 3). A sequência para esta sequência de ADNc é SEQ ID NO:1 (sequência nucleotídica da NgR2 humana; AC013606). A tradução deste ADNc proporciona a sequência proteica da NgR2 humana (SEQ ID NO:2).

Utilizámos a sequência proteica da NgR2 humana como sequência de busca contra a base de dados de EST humanas. Foram encontrados vários acertos de elevada significância indicando que o ARNm de NgR2 é expresso em vários tecidos incluindo cérebro fetal. Para além disso, duas destas EST proporcionavam apoio à estrutura do exão que deduzimos. Uma

EST (N° de Entrada: GB_EST19:AI346757) contém 565 nucleótidos correspondendo aos aminoácidos 84-271 da NgR2 humana (SEQ ID No:4). Esta abrange o segundo intrão situado entre os aminoácidos 171 e 172, e proporciona uma prova positiva para o processamento dos exões 2 e 3 ao nível do ARNm. Outra EST (GB_EST26:AI929019) contém 545 nucleótidos, parte dos quais corresponde aos aminoácidos 1-75 da NgR2 humana (SEQ ID NO:2). Esta abrange o primeiro intrão localizado entre os aminoácidos 10 e 11 e proporciona uma prova positiva para o processamento dos exões 1 e 2 ao nível do ARNm.

Exemplo 3: Previsão da sequência proteica de NgR3 de ratinho (AC021768)

A sequência proteica de NgR humana alinhou com apenas uma região da sequência traduzida da sequência nucleotídica AC021768, indicando que a maior parte do novo gene de ratinho era codificada por um grande exão. No entanto, após inspecção, faltava à proteína codificada por este exão uma metionina iniciadora. Para definir o gene completo, utilizámos o programa de computador GENSCAN tal como descrito acima. A previsão do gene através de GENSCAN continha dois exões; o grande encontrado através de inspecção visual e um curto na extremidade 5' que proporcionou uma metionina iniciadora. O ADNc previsto (ARNm) codificado por estes dois exões foi montado a partir de AC021768 (HTG14; depositado em Março de 2000; comprimento = 215980; GenBank edição 118.0; SEQ ID NO:16) através da combinação de nucleótidos dos dois exões cujas coordenadas são: o complemento de 164265-164325 (exão 1); e o complemento de 155671-156992 (exão 2). A sequência para esta sequência de ADNc é SEQ ID NO:3 (sequência nucleotídica de NgR3 de ratinho; AC021768). A tradução deste ADNc proporciona a sequência proteica de NgR3 de ratinho (SEQ ID NO:4).

Utilizámos a sequência proteica de NgR3 de ratinho como sequência de busca contra a base de dados de ESTs de ratinho. Foi encontrado um acerto de elevada significância indicando que o ARNm de NgR2 é expresso no coração. Este EST (GB_EST20:AI428334) contém 463 nucleótidos, parte dos quais corresponde aos aminoácidos 45-193 de NgR3 de ratinho (SEQ ID NO:4).

Exemplo 4: Semelhança entre os NgR

Um alinhamento entre NgR1 e os dois novos receptores é mostrado na Fig. 1A-1B. As semelhanças entre estas proteínas incluem:

- (1) O programa SignalP, que localiza a posição de clivagem da sequência sinal, prevê uma clivagem antes da primeira cisteína conservada em todas as proteínas. Assim, a proteína madura em todos os casos terá uma cisteína no terminal N.
- (2) Todas as proteínas contêm oito Repetições Ricas em Leucina (LRR). As LRRs são motivos de sequência curtos presentes em várias proteínas com diversas funções e localizações celulares. Estas repetições estão habitualmente envolvidas em interações proteína-proteína. Cada LRR é composta por uma unidade beta-alfa.
- (3) Todas as três proteínas contêm um domínio N-terminal de repetições ricas em leucina (LRRNT), em que quatro cisteínas são conservadas. As LRRs são frequentemente flanqueadas por domínios ricos em cisteína em ambos os seus terminais N e C.
- (4) Todas as três proteínas contêm um domínio LRR C-terminal (LRRCT). Os LRRCTs das três proteínas NgR podem ser distinguidos dos de outras proteínas contendo LRRs, através do padrão de triptofanos e cisteínas que são completamente conservados neste domínio.
- (5) Todas as três proteínas contêm uma cisteína conservada no quarto domínio LRR.
- (6) Todas as três proteínas contêm um potencial local de glicosilação conservado no oitavo domínio LRR.
- (7) NgR2 e NgR3 têm um terminal C hidrófobo, tal como NgR1, uma indicação de que provavelmente também sofrem uma modificação semelhante a NgR1, onde uma porção GPI está covalentemente ligada a um aminoácido C-terminal. Isto permite que a proteína permaneça presa à célula.

Exemplo 5: Preparação de Proteínas Nogo

Foi desenvolvido um ensaio de ligação de Nogo que utiliza um método amplamente utilizado no exame da função de orientação axonal de semaforina e efrina (Flanagan & Vanderhaeghen, *Annu. Rev. Neurosci.* 21: 309-345, 1998; Takahashi *et al.*, *Cell* 99: 59-69, 1999). Envolve a fusão de uma porção fosfatase alcalina (AP) placentária segregada com o ligando em questão para proporcionar um agente de ligação ao receptor biologicamente activo que pode ser detectado com um ensaio colorimétrico extremamente sensível. Para Nogo, é criado um vector de expressão codificando um péptido sinal, uma etiqueta His6 para purificação, AP e o domínio activo de 66 aminoácidos de Nogo. A proteína de fusão pode ser purificada a partir do meio condicionado de células transfectadas em quantidades de miligramas. Esta proteína é biologicamente activa como um agente de colapso do cone de crescimento com uma CE_{50} de 1 nM.

Alternativamente, pode ser preparada uma proteína de fusão glutathione-S-transferase-Nogo (GST-Nogo). Para GST-Nogo, é criado um vector de expressão (p.ex., um vector pGEX) codificando um péptido sinal, GST, e o domínio activo de 66 aminoácidos de Nogo. A GST-Nogo pode ser purificada a partir do meio de cultura e utilizada como uma proteína de fusão de GST, ou a GST pode ser clivada a partir da porção Nogo da proteína de fusão com uma enzima que reconheça o local de clivagem específico desenhado entre a porção GST e a porção Nogo da proteína de fusão. Tais locais são parte dos vectores de GST comercialmente disponíveis. Os locais de clivagem específicos e as enzimas podem ser utilizados de acordo com as especificações do fabricante.

Verificou-se que AP-Nogo é de facto ligeiramente mais potente que GST-Nogo, talvez porque a proteína é sintetizada numa célula eucariótica em vez de uma célula procariótica.

A ligação de Nogo a homólogos de NgR imobilizados pode ser efectuada num ensaio tipo ELISA no qual AP-Nogo é deixada reagir com um homólogo do receptor imobilizado. A especificidade da ligação pode ser demonstrada num ensaio de ligação competitiva utilizando quantidades crescentes de GST-

Nogo no tipo de ensaio para mostrar uma quantidade decrescente de ligação de AP-Nogo (conforme julgado no ensaio colorimétrico).

Exemplo 6: Ensaio de ligação em Células COS Transfectadas

Os homólogos da presente divulgação podem ser utilizados em estudos de transfecção em células COS para demonstrar a ligação de Nogo. Especificamente, sequências nucleotídicas codificando NgR2 e NgR3 podem ser transfectadas em células COS utilizando um vector adequado. As células COS-7 não transfectadas não se ligam a AP-Nogo. No entanto, a transfecção de células COS com sequências de ácido nucleico codificando NgR torna-las-á capazes de se ligarem a Nogo. O ALP sozinho não se ligará com qualquer afinidade estável a estas células transfectadas, indicando que qualquer afinidade de Nogo para NgR2 ou NgR3 seria devida aos 66 aminoácidos derivados de Nogo. Para além disso, a afinidade específica de Nogo para as proteínas NgR2 ou NgR3 pode ser testada em ensaios de deslocamento de AP-Nogo utilizando GST-Nogo. NgR2 e/ou NgR3 podem também ligar-se a homólogos de Nogo, que podem também ser testados utilizando este ensaio.

Exemplo 7: A Expressão de NgR em Linhas Celulares Humanas utilizando Northern Blot e uma Sonda Iniciada Aleatoriamente

Um *Northern blot* é adquirido a partir de uma fonte comercial ou são corridas amostras de ARN de células de interesse num gel de agarose e transferidas para uma membrana utilizando qualquer uma das técnicas bem conhecidas para *Northern blotting*. A membrana é sondada com um fragmento de NgR2 (SEQ ID NO:1) ou NgR3 (SEQ ID NO:3). A sonda é preparada a partir de 50 ng de ADNc marcado através do método de iniciação aleatória (Feinberg e Vogelstein, *Anal. Biochem.* 132: 6-13, 1983). A hibridação é realizada a 68°C durante 1 hora em solução ExpressHyb™ (Clontech, N.º de Cat. 8015-1) seguido de lavagem com SSC 2×/SDS a 0,05% à temperatura ambiente e duas lavagens com SSC 0,1×/SDS a 0,1% a 50°C. A expressão de NgR2 e/ou NgR3 pode ser avaliada através da presença de uma banda com o tamanho apropriado na membrana.

Exemplo 8: Clonagem de ADNc correspondendo a NgR

Para obter o clone inteiro correspondendo a NgR2 a partir de uma biblioteca de ADNc, pode ser utilizado o seguinte método. É gerada uma biblioteca de ADNc utilizando métodos padrão a partir de um tecido que se saiba conter NgR2. Um tal tecido foi identificado no Exemplo 2. 1×10^6 unidades formadoras de placas fágicas da biblioteca de ADNc podem ser pesquisadas em duplicado em filtros OPTITRANTM. Os filtros são hibridados com oligonucleótidos marcados com ^{32}P que são gerados a partir das ESTs correspondentes a porções de NgR2. A reacção de hibridação pode consistir em 400 µl de tampão de pesquisa de placas (Tris 50 mM pH 7,5, NaCl 1 M, pirofosfato de sódio a 0,1%, Polivinilpirrolidona a 0,2% e Ficoll a 0,2%) contendo sulfato de Dextrano a 10% e 100 µg/ml de ARNt e 80 pmol de cada oligonucleótido marcado com ^{32}P a 65°C de um dia para o outro. Os filtros são lavados duas vezes com SSC 2x/SDS a 1% e duas vezes com SSC 1x/SDS a 1% e expostos a película. Os positivos em duplicado são purificados. O ADN de cada um destes clones é analisado através de digestão com enzimas de restrição seguida de electroforese em gel de agarose e *Southern blotting*. Os filtros são hibridados com os oligonucleótidos marcados com ^{32}P utilizados para a hibridação original para confirmar que as inserções hibridam com a sonda. A inserção é então sequenciada para confirmar que representa o ADNc para NgR2. Métodos semelhantes podem ser utilizados para gerar um clone inteiro correspondendo a NgR3.

Alternativamente, um clone inteiro de NgR2 ou NgR3 pode ser obtido por um vulgar perito na arte empregando técnicas de PCR convencionais.

Exemplo 9: Análise de Hibridação para demonstrar expressão de NgR no cérebro

A expressão de NgR em mamíferos, tais como o rato, pode ser investigada através de histoquímica de hibridação *in situ*. Para investigar a expressão no cérebro, por exemplo, são preparadas criossecções coronais e sagitais de cérebro do rato (20 µm de espessura) utilizando um crióstato Reichert-Jung. Secções individuais são montadas por descongelação sobre lâminas silanizadas, livres de nucleases (CEL Associates, Inc., Houston, TX) e armazenadas a -80°C. As secções são

processadas a começar com pós-fixação em paraformaldeído a 4% frio, lavadas em solução salina tamponada com fosfato (PBS), acetiladas utilizando anidrido acético em tampão trietanolamina, e desidratadas através de uma série de lavagens em álcool a 70%, 95% e 100% à temperatura ambiente. Subsequentemente, as secções são deslipidadas em clorofórmio, seguido de re-hidratação através de exposição sucessiva a álcool a 100% e 95% à temperatura ambiente. Lâminas de microscópio contendo criossecções processadas são deixadas secar ao ar antes da hibridação. Outros tecidos podem ser ensaiados de um modo semelhante.

Uma sonda específica de NgR pode ser gerada utilizando PCR. Após amplificação por PCR, o fragmento é digerido com enzimas de restrição e clonado em pBluescript II clivado com as mesmas enzimas. Para produção de uma sonda específica para a cadeia com sentido de NgR, um fragmento clonado de NgR clonado em pBluescript II pode ser linearizado com uma enzima de restrição adequada, que proporciona um substrato para transcritos escoados marcados (i.e., ribossondas de ARNc) utilizando o promotor T7 com origem no vector e ARN-polimerase T7 comercialmente disponível. Uma sonda específica para a cadeia anti-sentido de NgR pode também ser facilmente preparada utilizando o clone de NgR em pBluescript II através da clivagem do plasmídeo recombinante com uma enzima de restrição adequada para gerar um substrato linearizado para a produção de transcritos de ARNc escoados marcados utilizando o promotor T3 e a polimerase cognata. As ribossondas podem ser marcadas com [³⁵S]-UTP para produzir uma actividade específica de cerca de $0,40 \times 10^6$ cpm/pmol para ribossondas anti-sentido e de cerca de $0,65 \times 10^6$ cpm/pmol para ribossondas da cadeia com sentido. Cada ribossonda pode ser subsequentemente desnaturada e adicionada (2 pmol/ml) a tampão de hibridação que contém formamida a 50%, dextrano a 10%, NaCl 0,3 M, Tris 10 mM (pH 8,0), EDTA 1 mM, solução de Denhardt 1x, e ditiotreitol 10 mM. Lâminas de microscópio contendo criossecções sequenciais de cérebro podem ser independentemente expostas a 45 µl de solução de hibridação por lâmina e podem ser colocadas lamelas silanizadas sobre as secções a expor à solução de hibridação. As secções são incubadas de um dia para o outro (15-18 horas) a 52°C para permitir que ocorra hibridação. Uma série equivalente de criossecções é então

exposta a ribossondas de ARNc específicas de NgR com sentido ou anti-sentido.

Após o período de hibridação, as lamelas são lavadas das lâminas em SSC 1×, seguido de tratamento com ARNase A envolvendo a exposição das lâminas a 20 µg/ml de ARNase A num tampão contendo Tris-HCl 10 mM (pH 7,4), EDTA 0,5 M e NaCl 0,5 M durante 45 minutos a 37°C. As criossecções são então sujeitas a três lavagens de elevado rigor em SSC 0,1× a 52°C durante 20 minutos cada. Após as séries de lavagens, as criossecções são desidratadas através de exposição consecutiva a acetato de amónio a 70%, 95% e 100% em álcool, seguido de secagem ao ar e exposição a película Kodak BioMaxTM MR-1. Após 13 dias de exposição, a película é revelada e qualquer sinal significativo de hibridação é detectado. Com base nestes resultados, as lâminas contendo o tecido que hibridou, tal como mostrado através de autorradiogramas da película, são revestidas com emulsão de seguimento nuclear Kodak NTB-2 e as lâminas são armazenadas no escuro durante 32 dias. As lâminas são então reveladas e contrastadas com hematoxilina. As secções revestidas com emulsão são analisadas microscopicamente para determinar a especificidade da marcação. O sinal é determinado como sendo específico se os grãos auto-radiográficos (gerados através de hibridação da sonda anti-sentido) estiverem claramente associados a corpos celulares corados com violeta de cresil. Os grãos auto-radiográficos verificados entre corpos celulares indicam ligação não específica da sonda.

Nalguns casos, tais como utilizando uma sonda para detectar um homólogo de NgR numa espécie heteróloga, para alcançar uma hibridação óptima, pode ser necessário diminuir as condições de rigor. Tais condições são bem conhecidas dos vulgares peritos na arte e são proporcionados exemplos acima.

A expressão de NgR no cérebro proporciona uma indicação de que os moduladores da actividade de NgR têm utilidade para tratamento de distúrbios neurológicos. Algumas outras doenças para as quais os moduladores de NgR podem ter utilidade incluem depressão, ansiedade, doença bipolar, epilepsia, neurite, neurastenia, neuropatia, neuroses e semelhantes. Pretende-se que a utilização de moduladores de NgR, incluindo

ligandos de NgR e anticorpos anti-NgR, para tratar indivíduos possuindo tais estados de doença seja um aspecto da divulgação.

Exemplo 10: Análise Northern Blot de ARN de NgR com uma Sonda gerada por PCR

As hibridações de *Northern blot* podem ser efectuadas para examinar a expressão do ARNm de NgR. Um clone contendo pelo menos uma porção da sequência de SEQ ID NO:1 pode ser utilizado como sonda. Iniciadores específicos do vector são utilizados em PCR para gerar um fragmento da sonda de hibridação para marcação com ^{32}P . A PCR é efectuada como se segue:

Mistura:	1 μl	plasmídeo contendo NgR
	2 μl	iniciador directo (10-50 pM)
	2 μl	iniciador inverso (10-50 pM)
	10 μl	tampão de PCR 10 $\times$ (tal como o proporcionado com a enzima, Amersham Pharmacia Biotech)
	1 μl	dNTP 10 mM (tal como #1969064 de Boehringer Mannheim)
	0,5 μl	polimerase Taq (tal como #27-0799-62, Amersham Pharmacia Biotech)
	83,5 μl	água

A PCR é efectuada num Termociclador utilizando o seguinte programa:

94°C	5 min	30 ciclos
94°C	1 min	
55°C	1 min	
72°C	1 min	
72°C	10 min	

O produto de PCR pode ser purificado utilizando o Estojó de Purificação de PCR QIAquick (#28104) de Qiagen e a marcação radioactiva com ^{32}P -dCTP (#AA0005/250, Amersham Pharmacia

Biotech) pode ser feita através de iniciação aleatória utilizando *"Ready-to-go DNA Labeling Beads"* (#27-9240-01) de Amersham Pharmacia Biotech. A hibridação é realizada em *Human Multiple Tissue Northern Blot* de Clontech tal como descrito no protocolo do fabricante, ou num *Northern Blot* preparado correndo amostras de ARN de células de interesse num gel de agarose e *blotting* numa membrana utilizando qualquer um dos protocolos de *Northern blotting* conhecidos. Após exposição de um dia para o outro num ecrã *Molecular Dynamics Fosfor Imager* (#MD146-814) as bandas de um tamanho apropriado são visualizadas.

Exemplo 11: Expressão Recombinante de NgR em Células Hospedeiras Eucarióticas

A. Expressão de NgR em Células de Mamífero

Para produzir proteína NgR, um polinucleótido codificando NgR é expresso numa célula hospedeira adequada utilizando um vector de expressão adequado e técnicas padrão de engenharia genética. Por exemplo, uma sequência codificando NgR descrita na Tabela 4 é subclonada no vector de expressão comercial pzeoSV2 (Invitrogen, San Diego, CA) e transfectada para células de Ovário de Hamster Chinês (CHO) utilizando o reagente de transfecção FuGENE6TM (Boehringer-Mannheim) e o protocolo de transfecção proporcionado no folheto do produto. Outras linhas celulares eucarióticas, incluindo células de rim embrionário humano (HEK 293) e COS, são igualmente adequadas. As células expressando estavelmente NgR são seleccionadas através de crescimento na presença de 100 µg/ml de zeocina (Stratagene, LaJolla, CA). Como alternativa a FuGENE6TM, o vector de expressão pode possuir o gene para a di-hidrofolato-redutase (*dhfr*) e a selecção de clones com pressão do fármaco metotrexato (MTX) permite uma transformação estável de células CHO. Opcionalmente, NgR pode ser purificado a partir das células utilizando técnicas cromatográficas padrão. Para facilitar a purificação, é criado anti-soro contra uma ou mais sequências peptídicas sintéticas que correspondem a porções da sequência de aminoácidos de NgR e o anti-soro é utilizado para purificar por afinidade Nogo-R. O NgR pode também ser expresso enquadrado com uma sequência etiqueta (p.ex., poli-histidina, hemaglutinina, FLAG) para facilitar a purificação. Para além disso, será apreciado que muitas das utilizações para

polipéptidos NgR, tais como os ensaios descritos abaixo, não requerem purificação de NgR a partir da célula hospedeira.

B. Expressão de NgR em células CHO

Para expressão de NgR em células de ovário de hamster chinês (CHO), é preparado um plasmídeo possuindo a sequência de codificação de NgR relevante, utilizando um vector que também possui o marcador seleccionável di-hidrofolato-redutase (DHFR). O plasmídeo é transfectado para células CHO. A selecção sob pressão do fármaco MTX permite a preparação de transformantes estáveis de um NgR (NgR2 ou NgR3) num plasmídeo de expressão possuindo um marcador seleccionável tal como DHFR.

C. Expressão de NgR em células 293

Para expressão de NgR em células de mamífero 293 (células de rim embrionário primárias, humanas transformadas), é preparado um plasmídeo possuindo a sequência de codificação de NgR relevante, utilizando o vector pSecTag2A (Invitrogen). O vector pSecTag2A contém a sequência líder da cadeia IgK de murídeo para secreção, o epitopo c-myc para detecção da proteína recombinante com o anticorpo anti-myc, uma polihistidina C-terminal para purificação com cromatografia de quelato de níquel, e um gene resistente à zeocina para selecção de transfectantes estáveis. O iniciador directo para amplificação deste ADNc de NgR é determinado através de procedimentos de rotina e contém de preferência uma extensão a 5' de nucleótidos para introduzir o local de clonagem *HindIII* e nucleótidos coincidindo com a sequência de NgR. O iniciador inverso é também determinado através de procedimentos de rotina e contém de preferência uma extensão a 5' de nucleótidos para introduzir um local de restrição *XhoI* para clonagem e nucleótidos correspondendo ao complemento inverso da sequência de NgR. As condições de PCR são 55°C como temperatura de ligação. O produto de PCR é purificado em gel e clonado nos locais *HindIII*-*XhoI* do vector.

O ADN é purificado utilizando colunas de cromatografia Qiagen e transfectado em células 293 utilizando meios de transfecção DOTAPTM (Boehringer Mannheim, Indianapolis, IN). As células transfectadas transientemente são testadas para

expressão após 24 horas de transfecção, utilizando *western blots* sondados com anticorpos anti-His e anti-péptido NgR. As células transfectadas permanentemente são seleccionadas com Zeocina e propagadas. A produção da proteína recombinante é detectada a partir tanto das células como do meio através de *Western blots* sondados com anticorpos anti-His, anti-Myc ou anti-péptido NgR.

D. Expressão Transiente de Nogo-R em células COS

Para expressão do NgR em células COS7, uma molécula polinucleotídica possuindo uma sequência nucleotídica de SEQ ID NO:1 pode, por exemplo, ser clonada no vector p3-CI. Este vector é um plasmídeo derivado de pUC18 que contém o promotor HCMV (citomegalovírus humano)-intrão localizado a montante da sequência de poliadenilação da bGH (hormona de crescimento bovina) e um local de clonagem múltipla.

O iniciador directo é determinado através de procedimentos de rotina e contém de preferência uma extensão a 5' que introduz um local de restrição *XbaI* para clonagem, seguido de nucleótidos que correspondem a uma sequência nucleotídica de SEQ ID NO:1. O iniciador inverso é também determinado através de procedimentos de rotina e contém de preferência uma extensão a 5' de nucleótidos que introduz um local de clonagem *SalI* seguido de nucleótidos que correspondem ao complemento inverso de uma sequência nucleotídica de SEQ ID NO:1.

A PCR consiste num passo inicial de desnaturação de 5 min a 95°C, 30 ciclos de 30 seg de desnaturação a 95°C, 30 seg de ligação a 58°C e 30 seg de extensão a 72°C, seguido de 5 min de extensão a 72°C. O produto de PCR é purificado em gel e ligado nos locais *XbaI* e *SalI* do vector p3-CI. Esta construção é transformada em células de *E. coli* para amplificação e purificação do ADN. O ADN é purificado com colunas de cromatografia Qiagen e transfectado para células COS7 utilizando reagente LipofectamineTM de BRL, seguindo os protocolos do fabricante. Quarenta e oito e 72 horas após a transfecção, o meio e as células são testados quanto a expressão recombinante.

O NgR expresso a partir da cultura de células COS pode ser purificado através da concentração do meio de crescimento celular até cerca de 10 mg de proteína/ml e purificação da proteína, por exemplo, através de cromatografia. O NgR purificado é concentrado até 0,5 mg/ml num concentrador Amicon com uma membrana YM-10 ajustada e armazenado a -80°C . NgR3 pode também ser expresso utilizando este método e a sequência nucleotídica de SEQ ID NO:3 ou SEQ ID NO:13.

E. Expressão de NgR em Células de Insecto

Para expressão de NgR num sistema de baculovírus, uma molécula polinucleotídica possuindo uma sequência nucleotídica de SEQ ID NO:1, 3 ou 13 pode ser amplificada por PCR. O iniciador directo é determinado através de procedimentos de rotina e contém de preferência uma extensão a 5' que adiciona o local de clonagem *NdeI*, seguido de nucleótidos que correspondem a uma sequência nucleotídica de SEQ ID NO:1 (ou SEQ ID NO:3 ou SEQ ID NO:13, respectivamente). O iniciador inverso é também determinado através de procedimentos de rotina e contém de preferência uma extensão a 5' que introduz o local de clonagem *KpnI*, seguido de nucleótidos que correspondem ao complemento inverso de uma sequência nucleotídica de SEQ ID NO:1 (ou SEQ ID NO:3 ou SEQ ID NO:13, respectivamente).

O produto de PCR é purificado em gel, digerido com *NdeI* e *KpnI*, e clonado nos locais correspondentes do vector pACHTL-A (Pham-dngen, San Diego, CA). O vector de expressão pACHTL contém o promotor forte de poli-hedrina do vírus da poli-hedrose nuclear de *Autographa californica* (AcMNPV), e uma etiqueta 6×His a montante do local de clonagem múltipla. Um local de proteína-quinase para fosforilação e um local de trombina para excisão da proteína recombinante que precede o local de clonagem múltipla estão também presentes. Claro que, muitos outros vectores de baculovírus podiam ser utilizados em vez de pACHTL-A, tais como pAc373, pVL941 e pAcIM1. Outros vectores adequados para a expressão de polipéptidos NgR podem ser utilizados, desde que a construção do vector inclua sinais localizados apropriadamente para transcrição, tradução, e tráfico, tais como um AUG enquadrado e um péptido sinal, conforme necessário. Tais vectores são descritos em Luckow et al., *Virology* 170: 31-39, entre outros.

O vírus é criado e isolado utilizando métodos padrão de expressão de baculovírus, tais como os descritos em Summers *et al.*, (1987) "A Manual of Methods for Baculovirus Vectors and Insect Cell Culture Procedures", Texas Agricultural Experimental Station Buletin N.º 1555.

Num caso preferido, pAcHLT-A contendo o gene de NgR é introduzido no baculovírus utilizando o estojo de transfecção "BaculoGold™" (Pharmlingen, San Diego, CA) utilizando métodos estabelecidos pelo fabricante. Isolados individuais do vírus são analisados quanto à produção de proteína através de radiomarcagem de células infectadas com ³⁵S-metionina 24 horas após a infecção. As células infectadas são colhidas 48 horas após a infecção, e as proteínas marcadas são visualizadas através de SDS-PAGE. Os vírus exibindo elevados níveis de expressão podem ser isolados e utilizados para expressão em maior escala.

Para expressão de um polipéptido NgR em células Sf9, uma molécula polinucleotídica possuindo a sequência nucleotídica de SEQ ID NO:1 (ou SEQ ID NO:3 ou SEQ ID NO:13) pode ser amplificada por PCR utilizando os iniciadores e métodos descritos acima para expressão de baculovírus. O ADNc de NgR é clonado no vector pAcHLT-A (Pharmlingen) para expressão em Sf9 de insecto. A inserção é clonada nos locais *NdeI* e *KpnI*, após eliminação de um local *NdeI* interno (utilizando os mesmo iniciadores descritos acima para expressão em baculovírus). O ADN é purificado com colunas de cromatografia Qiagen e expresso em células Sf9. Experiências preliminares de *Western blot* de placas fágicas não purificadas são testadas quanto à presença de proteína recombinante do tamanho esperado que reagia com o anticorpo específico de NgR. Estes resultados são confirmados após mais optimização da purificação e expressão em células HiG5.

F. Expressão de formas solúveis de NgR2 e NgR3 como proteínas de fusão NgR-Ig

Para gerar uma proteína de fusão NgR2-Ig, podem ser utilizados métodos padrão tal como descritos na literatura (p.ex., Sanicola *et al.*, *Proc. Natl. Acad. Sci. USA.* 94: 6238-6243, 1997). Por exemplo, um fragmento de ADN codificando NgR2

sem a sequência codificando o terminal C hidrófobo (sinal âncora de GPI) pode ser ligado a um fragmento de ADN codificando o domínio Fc de IgG1 (que pode ser IgG1 humana), e o fragmento quimérico pode ser clonado num vector de expressão para gerar um plasmídeo. O plasmídeo pode então ser transfectado para células de ovário de hamster chinês para gerar uma linha celular estável produtora da proteína de fusão. A proteína de fusão é então purificada a partir de meio condicionado utilizando métodos padrão. Por exemplo, o meio condicionado clarificado da linha celular pode ser carregado por gravidade directamente sobre Sepharose com Proteína A. A coluna pode então ser lavada com cinco volumes de coluna de cada um de PBS, PBS contendo NaCl 0,5 M e fosfato de sódio 25 mM, NaCl 100 mM (pH 5,0). A proteína ligada pode então ser eluída com NaH_2PO_4 25 mM, NaCl 100 mM (pH 2,8) e imediatamente neutralizada com uma fracção de 1/10 do volume de Na_2BPO_4 0,5 M (pH 8,6).

Métodos semelhantes podem ser utilizados para gerar uma proteína de fusão NgR3-Ig.

Exemplo 12: Sistema de Aprisionamento por Interação/Dois Híbridos

Para ensaiar quanto a proteínas que interagem com NgR, pode ser utilizado o método de pesquisa de bibliotecas de aprisionamento por interacção/dois híbridos. Este ensaio foi descrito pela primeira vez em Fields *et al.*, *Nature* 340: 245, 1989, que é aqui incorporado por referência na sua totalidade. Um protocolo está publicado em "Current Protocols in Molecular Biology", 1999, John Wiley & Sons, NY e Ausubel, F. M. *et al.*, 1992, "Short Protocols in Molecular Biology", quarta edição, Greene and Wiley-interscience, NY. Estão disponíveis estojos em Clontech, Palo Alto, CA (Matchmaker Two-Hybrid System 3).

Uma fusão das sequências nucleotídicas codificando todo ou um NgR parcial e o domínio de ligação ao ADN de GAL4 do factor de transcrição de levedura (DNA-BD) é construída num plasmídeo apropriado (i.e., pGBKT7) utilizando técnicas padrão de subclonagem. Semelhantemente, uma biblioteca de fusões do domínio activo de GAL4 (AD) é construída num segundo plasmídeo (i.e., pGADT7) a partir de ADNc de potenciais proteínas de

ligação a NgR (para protocolos sobre formação de bibliotecas de ADNc, ver Sambrook *et al.*, 1989, "Molecular Cloning: A Laboratory Manual", segunda edição, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). A construção da fusão DNA-BD/NgR é verificada através de sequenciação, e testada quanto à activação autónoma do gene repórter e toxicidade celular, ambas as quais impediriam uma análise de dois híbridos com sucesso. Controlos semelhantes são efectuados com a construção AD/fusões da biblioteca para assegurar a expressão em células hospedeiras e a falta de actividade de transcrição. As células de levedura são transformadas (ca. 10⁵ transformantes/mg de ADN) com ambos os plasmídeos de NgR e da fusão de biblioteca de acordo com o procedimento padrão (Ausubel, *et al.*, 1992, "Short Protocols in Molecular Biology", quarta edição, Greene and Wiley-interscience, NY. A ligação *in vivo* de DNA-BD/NgR com AD/proteínas da biblioteca resulta em transcrição de genes repórter do plasmídeo de levedura específicos (i.e., *lacZ*, *HIS3*, *ADE2*, *LEU2*). As células de levedura são plaqueadas em meio sem nutrientes para pesquisar quanto a expressão dos genes repórter. As colónias são duplamente ensaiadas quanto à actividade de β -galactosidase após crescimento em meio suplementado com Xgal (5-bromo-4-cloro-3-indolil- β -D-galactósido) (o ensaio de filtro da actividade de β -galactosidase está descrito em Breeden *et al.*, *Cold Spring Harb. Symp. Quant. Biol.* 50: 643, 1985). Os plasmídeos AD-biblioteca positivos são recuperados a partir dos transformantes e reintroduzidos na estirpe de levedura original bem como outras estirpes contendo proteínas de fusão DNA-BD não relacionadas para confirmar interacções proteicas NgR/biblioteca específicas. O ADN inserido é sequenciado para verificar a presença de um enquadramento de leitura aberto fundido com AD de GAL4 e para determinar a identidade da proteína de ligação a NgR.

Exemplo 13: Anticorpos para Nogo-R

São empregues técnicas padrão para gerar anticorpos policlonais ou monoclonais para o receptor NgR e gerar fragmentos de ligação ao antigénio destes ou variantes destes úteis, incluindo variantes "humanizadas". Tais protocolos podem ser verificados, por exemplo, em Sambrook *et al.*, (1989), acima, e Harlow *et al.*, (Eds.), *Antibodies A*

Laboratory Manual; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1988). Num caso, polipéptidos NgR recombinantes (ou células ou membranas celulares contendo tais polipéptidos) são utilizados como antigénio para gerar os anticorpos. Noutro caso, um ou mais péptidos possuindo sequências de aminoácidos correspondendo a uma porção imunogénica de NgR (p.ex., 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 ou mais aminoácidos) são utilizados como antigénio. Péptidos correspondendo a porções extracelulares de Nogo-R, especialmente porções extracelulares hidrófilas, são preferidos. O antigénio pode ser misturado com um adjuvante ou ligado a um hapteno para aumentar a produção de anticorpos.

A. Anticorpos Policlonais ou Monoclonais

Como um protocolo exemplar, NgR recombinante ou um fragmento sintético deste é utilizado para imunizar um ratinho para criação de anticorpos monoclonais (ou um mamífero maior, tal como um coelho, para anticorpos policlonais). Para aumentar a antigenicidade, os péptidos são conjugados com Hemocianina da Lapa (Pierce), de acordo com as recomendações do fabricante. Para uma injeção inicial, o antigénio é emulsionado com Adjuvante Completo de Freund e injectado subcutaneamente. A intervalos de duas as três semanas, alíquotas adicionais de antigénio de NgR são emulsionadas com Adjuvante Incompleto de Freund e injectadas subcutaneamente. Antes da injeção final de reforço, uma amostra de soro é tomada dos ratinhos imunizados e ensaiada através de *western blot* para confirmar a presença de anticorpos que imuno-reagem com NgR. O soro dos animais imunizados pode ser utilizado como anti-soro policlonal ou utilizado para isolar anticorpos policlonais que reconheçam NgR. Alternativamente, os ratinhos são sacrificados e o seu baço removido para geração de anticorpos monoclonais.

Para gerar anticorpos monoclonais, os baços são colocados em 10 ml de RPMI 1640 livre de soro e são formadas suspensões de células individuais através de moagem dos baços em RPMI 1640 livre de soro suplementado com L-glutamina 2 mM, piruvato de sódio 1 mM, 100 unidades/ml de penicilina e 100 µg/ml de estreptomicina (RPMI) (Gibco, Canadá). As suspensões celulares são filtradas e lavadas através de centrifugação e ressuspensas em RPMI livre de soro. Timócitos tomados de três

ratinhos Balb/c ingênuos são preparados de um modo semelhante e utilizados como Camada Alimentadora. Células de mieloma NS-1, mantidas em fase logarítmica em RPMI com soro fetal bovino (FBS) a 10% (Hyclone Laboratories, Inc., Logan, UT) durante três dias antes da fusão, são também centrifugadas e lavadas.

Para produzir fusões de hibridoma, as células do baço dos ratinhos imunizados são combinadas com células NS-1 e centrifugadas, e o sobrenadante é aspirado. O sedimento celular é deslocado batendo no tubo, e 2 ml de PEG 1500 a 37°C (50% em HEPES 75 mM, pH 8,0) (Boehringer-Mannheim) são misturados no sedimento, seguido da adição de RPMI livre de soro. A partir daí, as células são centrifugadas, ressuspensas em RPMI contendo FBS a 15%, hipoxantina de sódio 100 µM, aminopterina 0,4 µM, timidina 16 µM (HAT) (Gibco), 25 unidades/ml de IL-6 (Boehringer-Mannheim) e $1,5 \times 10^6$ timócitos/ml e plaqueadas em 10 placas de cultura de tecidos de 96 poços de fundo plano Corning (Corning, Corning, NY).

Nos dias 2, 4 e 6 após a fusão, 100 µl de meio são removidos dos poços das placas de fusão e substituídos por meio fresco. No dia 8, as fusões são pesquisadas através de ELISA, testando a presença de IgG de ratinho que se liga a NgR. Os poços de fusão selecionados são ainda clonados através de diluição até serem obtidas culturas monoclonais produtoras de anticorpos anti-NgR.

B. Humanização de anticorpos monoclonais anti-NgR

O padrão de expressão de NgR tal como aqui relatado e o potencial dos NgR como alvos para intervenção terapêutica sugerem indicações terapêuticas para inibidores de NgR (antagonistas). Os anticorpos neutralizantes de NgR compreendem uma classe de produtos terapêuticos úteis como antagonistas de NgR. A seguir estão protocolos para melhorar a utilidade de anticorpos monoclonais anti-NgR como produtos terapêuticos em humanos através da "humanização" dos anticorpos monoclonais para melhorar a sua semi-vida sérica e torná-los menos imunogênicos em hospedeiros humanos (i.e., para prevenir uma resposta de anticorpos humanos a anticorpos anti-NgR não humanos).

Os princípios de humanização foram descritos na literatura e são facilitados pelo arranjo modular de proteínas de anticorpo. Para minimizar a possibilidade de ligação do complemento, é preferido um anticorpo humanizado do isotipo IgG4.

Por exemplo, é alcançado um nível de humanização através da geração de anticorpos quiméricos compreendendo os domínios variáveis de proteínas de anticorpo não humanas de interesse com os domínios constantes de moléculas de anticorpo humanas. (Ver, p.ex., Morrison *et al.*, *Adv. Immunol.*, 44: 65-92, 1989). Os domínios variáveis de anticorpos anti-NgR neutralizantes de NgR são clonados a partir do ADN genómico de um hibridoma de células B ou de ADNc gerado a partir de ARNm isolado a partir do hibridoma de interesse. Os fragmentos do gene da região V são ligados a exões codificando domínios constantes do anticorpo humano e a construção resultante é expressa em células hospedeiras de mamífero adequadas (p.ex., mieloma ou células CHO).

Para alcançar um nível de humanização ainda maior, apenas as porções dos fragmentos génicos da região variável que codifiquem regiões determinantes de complementaridade que se ligam ao antigénio ("CDR") dos genes de anticorpo monoclonal não humano, são clonados em sequências de anticorpo humano. (Ver, Jones *et al.*, *Nature* 321: 522-525, 1986; Riechmann *et al.*, *Nature* 332: 323-327, 1988; Verhoeven *et al.*, *Science* 239: 1534-1536, 1988; e Tempest *et al.*, *Bio/Technology* 9: 266-271, 1991). Se necessário a estrutura de folha β do anticorpo humano rodeando as regiões CDR3 também é modificada para espelhar mais a estrutura tridimensional do domínio de ligação ao antigénio do anticorpo monoclonal original. (Ver Kettleborough *et al.*, *Protein Engin.* 4: 773-783, 1991; e Foote *et al.*, *J. Mol. Biol.* 224: 487-499, 1992).

Numa abordagem alternativa, a superfície de um anticorpo monoclonal não humano de interesse é humanizada através da alteração de resíduos da superfície seleccionados do anticorpo não humano, p.ex., através de mutagénese dirigida ao local, enquanto se mantêm todos os resíduos interiores e de contacto

do anticorpo não humano. Ver Padlan, *Mol. Immunol.* 28: 489-498, 1991.

As abordagens anteriores são empregues utilizando anticorpos monoclonais anti-NgR neutralizantes de NgR e os hibridomas que os produzem para gerar anticorpos neutralizantes de NgR humanizados úteis como produtos terapêuticos para tratar ou paliar condições em que a expressão de NgR ou a sinalização de NgR mediada pelo ligando seja prejudicial.

C. Anticorpos Humanos Neutralizantes de NgR de Apresentação Fágica

Anticorpos neutralizantes de NgR humanos são gerados através de técnicas de apresentação fágica tais como as descritas em Aujame *et al.*, *Human Antibodies* 8: 155-168, 1997; Hoogenboom, *TIBTECH* 15: 62-70, 1997; e Rader *et al.*, *Curr. Opin. Biotechnol.* 8: 503-508, 1997. Por exemplo, regiões variáveis do anticorpo na forma de fragmentos Fab ou fragmentos Fv de cadeia simples ligados são fundidas com o terminal amino da proteína menor de revestimento pIII de fago filamentoso. A expressão da proteína de fusão e a incorporação desta no revestimento maduro do fago resulta em partículas fágicas que apresentam um anticorpo à sua superfície e contêm o material genético codificando o anticorpo. Uma biblioteca fágica compreendendo tais construções é expressa em bactérias e a biblioteca é pesquisada quanto a anticorpos do fago específicos de NgR utilizando NgR marcado ou imobilizado como sonda do antígeno.

D. Anticorpos humanos neutralizantes de NgR de ratinhos transgênicos

Anticorpos neutralizantes de NgR humanos são gerados em ratinhos transgênicos essencialmente tal como descrito em Bruggemann *et al.*, *Immunol. Today* 17: 391-397, 1996 e Bruggemann *et al.*, *Curr. Opin. Biotechnol.* 8: 455-458, 1997. Ratinhos transgênicos possuindo segmentos do gene V humano na configuração da linha germinativa e que expressam estes transgenes no seu tecido linfóide são imunizados com uma composição de NgR utilizando protocolos de imunização convencionais. São gerados hibridomas utilizando células B dos

ratinhos imunizados utilizando protocolos convencionais que são pesquisados para identificar hibridomas segregando anticorpos humanos anti-NgR (p.ex., tal como descrito acima).

Exemplo 14: Ensaios para Identificar Moduladores da Actividade de NgR

Expostos abaixo estão vários ensaios não limitantes para identificação de moduladores (agonistas e antagonistas) da actividade de NgR. Entre os moduladores que podem ser identificados por estes ensaios estão compostos ligandos naturais do receptor; análogos sintéticos e derivados de ligandos naturais; anticorpos, fragmentos de anticorpo e/ou compostos semelhantes a anticorpos derivados de anticorpos naturais ou de bibliotecas combinatórias semelhantes a anticorpos; e/ou compostos sintéticos identificados através de pesquisa de elevada tiragem de bibliotecas; e semelhantes. Todos os moduladores que se ligam a NgR são úteis para a identificação de NgR em amostras de tecido (p.ex., para fins de diagnóstico, fins patológicos e semelhantes). Moduladores agonistas e antagonistas são úteis para sobre-regulação e sub-regulação da actividade de NgR, respectivamente, para tratar estados de doença caracterizados por níveis anormais da actividade de NgR. Os ensaios podem ser efectuados utilizando putativos moduladores simples e/ou podem ser efectuados utilizando um agonista conhecido em combinação com antagonistas candidatos (ou *vice versa*).

A. Ensaios de AMPc

Num tipo de ensaio, os níveis de adenosina-monofosfato cíclico (AMPc) são medidos em células transfectadas com NgR que foram expostas a compostos moduladores candidatos. Os protocolos para os ensaios de AMPc foram descritos na literatura. (Ver, p.ex., Sutherland *et al.*, *Circulation* 37: 279, 1968; Frandsen *et al.*, *Life Sciences* 18: 529-541, 1976; Dooley *et al.*, *J. Pharmacol. Exp. Therap.* 283: 735-41, 1997; e George *et al.*, *J. Biomol. Screening* 2: 235-40, 1997). Um protocolo exemplar para um tal ensaio, utilizando um Ensaio de Activação de Adenilil-Ciclase FlashPlate® de NEW™ Life Science Products, é exposto abaixo.

Resumidamente, a sequência codificando NgR (p.ex., um ADNc ou ADN genómico sem intrões) é subclonada num vector de expressão comercial, tal como pzeoSV2 (Invitrogen), e transientemente transfectada para células de Ovário de Hamster Chinês (CHO) utilizando métodos conhecidos, tais como o protocolo de transfecção proporcionado por Boehringer-Mannheim quando se fornece o reagente de transfecção FuGENE 6. As células CHO transfectadas são semeadas em microplacas de 96 poços do estojo de ensaio FlashPlate®, que são revestidas com cintilante sólido ao qual se ligou anti-soro para AMPc. Para controlo, alguns poços são semeados com células CHO de tipo selvagem (não transfectadas). Outros poços na placa recebem várias quantidades de uma solução padrão de AMPc para utilizar na criação de uma curva padrão.

Um ou mais compostos de teste (i.e., moduladores candidatos) são adicionados às células em cada poço, com água e/ou meio/diluyente livre de composto servindo como controlo ou controlos. Após tratamento, o AMPc é deixado a acumular-se nas células durante exactamente 15 minutos à temperatura ambiente. O ensaio é terminado através da adição de tampão de lise contendo AMPc marcado com [¹²⁵I], e a placa é contada utilizando um contador de cintilações de microplacas de 96 poços Packard Topcount™. AMPc não marcado das células lisadas (ou dos padrões) e quantidades fixas de [¹²⁵I]-AMPc competem pelo anticorpo ligado à placa. Uma curva padrão é construída, e os valores de AMPc para os desconhecidos são obtidos através de interpolação. As alterações nos níveis intracelulares de AMPc das células em resposta à exposição a um composto de teste são indicadoras da actividade moduladora de NgR. Moduladores que actuem como agonistas de receptores que se acoplam ao subtipo G_s das proteínas G estimularão a produção de AMPc, conduzindo a um aumento mensurável de 3-10 vezes nos níveis de AMPc. Agonistas de receptores que se acoplem ao subtipo G_{i/o} das proteínas G inibirão a produção de AMPc estimulada por forskolina, conduzindo a uma diminuição mensurável nos níveis de AMPc de 50-100%. Moduladores que actuem como agonistas inversos reverterão estes efeitos nos receptores que são constitutivamente activos ou activados através de agonistas conhecidos.

B. Ensaios de Aequorina

Noutro ensaio, as células (p.ex., células CHO) são transientemente co-transfectadas tanto com uma construção de expressão de NgR como com uma construção que codifica a fotoproteína apoaequorina. Na presença do cofactor celenterazina, a apoaequorina emitirá uma luminescência mensurável que é proporcional à quantidade de cálcio intracelular livre (citoplasmático). (Ver geralmente, Cobbold, *et al.*, "Aequorin measurements of cytoplasmic free calcium," Em: McCormack J.G. e Cobbold P.H., eds., "Cellular Calcium: A Practical Approach". Oxford, IRL Press (1991); Stables *et al.*, *Anal. Biochem.* 252: 115-26, 1997; e Haugland, "Handbook of Fluorescent Probes and Research Chemicals". Sexta edição. Molecular Probes, Eugene, OR (1996)).

Num ensaio exemplar, NgR é subclonado no vector de expressão comercial pzeoSV2 (Invitrogen) e transientemente co-transfectado juntamente com uma construção que codifica a fotoproteína apoaequorina (Molecular Probes, Eugene, OR) para células CHO utilizando o reagente de transfecção FuGENE 6 (Boehringer-Mannheim) e o protocolo de transfecção proporcionado no folheto do produto.

As células são cultivadas durante 24 horas a 37°C em MEM (Gibco/BRL, Gaithersburg, MD) suplementado com soro fetal bovino a 10%, glutamina 2 mM, 10 U/ml de penicilina e 10 µg/ml de estreptomicina, tempo ao fim do qual o meio é mudado para MEM livre de soro contendo celenterazina 5 µM (Molecular Probes, Eugene, OR). A cultura é então continuada durante mais duas horas a 37°C. Subsequentemente, as células são desprendidas da placa utilizando VERSEN (Gibco/BRL), lavadas e ressuspensas a 200 000 células/ml em MEM livre de soro.

São preparadas diluições dos compostos moduladores de NgR candidatos em MEM livre de soro e dispensadas para poços de uma placa de ensaio de 96 poços opaca a 50 µl/poço. As placas são então carregadas num luminómetro de placas de microtítulo MLX (Dynex Technologies, Inc., Chantilly, VA). O instrumento é programado para dispensar 50 µl de suspensões celulares a cada poço, um poço de cada vez, e ler imediatamente a luminescência durante 15 segundos. Curvas de resposta à dose para os moduladores candidatos são construídas utilizando a área sob a

curva para cada pico de sinal luminoso. Os dados são analisados com SlideWrite, utilizando a equação para um ligando de um local, e são obtidos os valores de CE_{50} . Alterações na luminescência causadas pelos compostos são consideradas indicativas da actividade moduladora. Moduladores que actuem como agonistas em receptores que se acoplam ao subtipo Gq das proteínas G dão um aumento na luminescência de até 100 vezes. Moduladores que actuem como agonistas inversos reverterão este efeito em receptores que são constitutivamente activos ou activados através de agonistas conhecidos.

C. Ensaio do Gene Repórter de Luciferase

A fotoproteína luciferase proporciona outra ferramenta útil para ensaio de moduladores da actividade de NgR. Células (p.ex., células CHO ou células COS 7) são transientemente co-transfectadas tanto com uma construção de expressão de NgR (p.ex., NgR em pzeoSV2) como com uma construção repórter que inclua um gene para a proteína luciferase a jusante de um local de ligação de um factor de transcrição, tal como o elemento de resposta a AMPc (CRE), AP-1, ou NF-kappa B. Os níveis de expressão da luciferase reflectem o estado de activação dos eventos de sinalização. (Ver geralmente, George *et al.*, *J. Biomol. Screening* 2: 235-240, 1997; e Stratowa *et al.*, *Curr. Opin. Biotechnol.* 6: 574-581, 1995). A actividade da luciferase pode ser quantitativamente medida utilizando, p.ex., reagentes do ensaio de luciferase que estão comercialmente disponíveis em Promega (Madison, WI).

Num ensaio exemplar, células CHO são plaqueadas em caixas de cultura de 24 poços a uma densidade de 100000 células/poço um dia antes da transfecção e cultivadas a 37°C em MEM (Gibco/BRL) suplementado com soro fetal bovino a 10%, glutamina 2 mM, 10 U/ml de penicilina e 10 µg/ml de estreptomicina. As células são co-transfectadas transientemente com uma construção de expressão de NgR e uma construção repórter contendo o gene da luciferase. Os plasmídeos repórter CRE-luciferase, AP-1-luciferase e NF-kappaB-luciferase podem ser adquiridos em Stratagene (Legally, CA). As transfecções são efectuadas utilizando o reagente de transfecção FuGENE 6 (Boehringer-Mannheim) de acordo com as instruções do fornecedor. As células transfectadas com a construção repórter sozinha são utilizadas como controlo.

Vinte e quatro horas após a transfecção, as células são lavadas uma vez com PBS pré-aquecido até 37°C. MEM livre de soro é então adicionado às células sozinho (controlo) ou com um ou mais moduladores candidatos e as células são incubadas a 37°C durante cinco horas. Posteriormente, as células são lavadas uma vez com PBS gelado e lisadas através da adição de 100 µl de tampão de lise por poço do estojo de ensaio de luciferase fornecido por Promega. Após incubação durante 15 minutos à temperatura ambiente, são misturados 15 µl do lisado com 50 µl de solução de substrato (Promega) numa placa de 96 poços branca opaca e a luminescência é lida imediatamente num contador de cintilações e luminescência Wallace modelo 1450 MicroBeta (Wallace Instruments, Gaithersburg, MD).

Diferenças na luminescência na presença *versus* ausência de um composto modulador candidato são indicativas de actividade moduladora. Os receptores que são constitutivamente activos ou activados através de agonistas dão tipicamente uma estimulação de 3-20 vezes a luminescência em comparação com células transfectadas com o gene repórter sozinho. Moduladores que actuem como agonistas inversos reverterão este efeito.

D. Medição do cálcio intracelular utilizando FLIPR

Alterações nos níveis de cálcio intracelular são outro indicador reconhecido da actividade do receptor e tais ensaios podem ser empregues para pesquisar moduladores quanto a actividade de NgR. Por exemplo, células CHO estavelmente transfectadas com um vector de expressão de NgR são plaqueadas a uma densidade de 4×10^4 células/poço em placas de 96 poços de parede negra Packard especialmente desenhadas para discriminar sinais de fluorescência emanando dos vários poços da placa. As células são incubadas durante 60 minutos a 37°C em PBS modificado por Dulbecco (D-PBS) contendo 36 mg/l de piruvato e 1 g/l de glicose com a adição de soro fetal bovino a 1% e um de quatro corantes indicadores de cálcio (Fluo-3TM AM, Fluo-4TM AM, Calcium GreenTM-1 AM ou Oregon GreenTM 488 BAPTA-1 AM), cada um a uma concentração de 4 µM. As placas foram lavadas uma vez com D-PBS modificado sem soro fetal bovino a 1% e incubadas durante 10 minutos a 37°C para remover corante residual da membrana celular. Adicionalmente, é efectuada uma

série de lavagens com D-PBS modificado sem soro fetal bovino a 1% imediatamente antes da activação da resposta de cálcio.

Uma resposta de cálcio é iniciada através da adição de um ou mais compostos agonistas do receptor candidatos, ionóforo de cálcio A23187 (10 μ M; controlo positivo) ou ATP (4 μ M; controlo positivo). A fluorescência é medida através de FLIPR de Molecular Device com um laser de árgon (excitação a 488 nm). (Ver, p.ex., Kuntzweiler et al., *Drug Dev. Res.* 44: 14-20, 1998). O obturador da câmara detectora é regulado para 2,5 e a duração da exposição é de 0,4 milissegundos. A fluorescência basal das células é medida durante 20 segundos antes da adição do agonista candidato, ATP ou A23187, e o nível de fluorescência basal é subtraído do sinal da resposta. O sinal de cálcio é medido durante aproximadamente 200 segundos, tomado leituras a cada dois segundos. O ionóforo de cálcio A23187 e ATP aumentam o sinal de cálcio 200% acima dos níveis basais. Em geral, NgR activados aumentam o sinal de cálcio pelo menos cerca de 10-15% acima do sinal basal.

E. Ensaio de Ligação a [³⁵S]-GTP γ S

É também possível avaliar se NgR sinaliza através de uma via medida por proteína G. Como os receptores acoplados à proteína G sinalizam através de proteínas G intracelulares cuja actividade envolve a ligação de GTP e hidrólise para produzir GDP ligado, a medição da ligação do análogo de GTP não hidrolisável [³⁵S]-GTP γ S na presença e ausência de moduladores candidatos proporciona outro ensaio para a actividade moduladora. (Ver, p.ex., Kowal et al., *Neuropharmacology* 37: 179-187, 1998).

Num ensaio exemplar, células estavelmente transfectadas com um vector de expressão de NgR são criadas em caixas de cultura de tecidos de 10 cm até subconfluência, lavadas uma vez com 5 ml de solução salina tamponada com fosfato livre de Ca²⁺/Mg²⁺ gelada e raspadas para 5 ml do mesmo tampão. As células são sedimentadas através de centrifugação (500 $\times$ g, 5 minutos), ressuspensas em tampão TEE (Tris 25 mM, pH 7,5, EDTA 5 mM, EGTA 5 mM) e congeladas em azoto líquido. Após descongelação, as células são homogeneizadas utilizando um homogeneizador Dounce (1 ml de TEE por placa de células) e

centrifugadas a 1000 $\times g$ durante 5 minutos para remover os núcleos e as células não destruídas.

O sobrenadante do homogenato é centrifugado a 20 000 $\times g$ durante 20 minutos para isolar a fracção membranar e o sedimento de membranas é lavado uma vez com TEE e ressuspensão em tampão de ligação (HEPES 20 mM, pH 7,5, NaCl 150 mM, MgCl₂ 10 mM, EDTA 1 mM). As membranas ressuspensas podem ser congeladas em azoto líquido e armazenadas a -70°C até à utilização.

Alíquotas de membranas celulares preparadas tal como descrito acima e armazenadas a -70°C são descongeladas, homogeneizadas e diluídas em tampão contendo HEPES 20 mM, MgCl₂ 10 mM, EDTA 1 mM, NaCl 120 mM, GDP 10 μ M, e ascorbato 0,2 mM, a uma concentração de 10-50 μ g/ml. Num volume final de 90 μ l, os homogenatos são incubados com concentrações variáveis de compostos moduladores candidatos ou GTP 100 μ M durante 30 minutos a 30°C e depois colocados em gelo. A cada amostra, foram adicionados 10 μ l de guanosina 5'-O-(3[³⁵S]tio)trifosfato (NEN, 1200 Ci/mmol; [³⁵S]-GTP γ S) até uma concentração final de 100-200 pM. As amostras são incubadas a 30°C durante mais 30 minutos, é adicionado 1 ml de HEPES 10 mM, pH 7,4, MgCl₂ 10 mM a 4°C e a reacção é parada através de filtração.

As amostras são filtradas sobre filtros Whatman GF/B e os filtros são lavados com 20 ml de HEPES 10 mM gelado, pH 7,4, MgCl₂ 10 mM. Os filtros são contados através de espectroscopia de cintilações líquidas. A ligação não específica de [³⁵S]-GTP γ S é medida na presença de GTP 100 μ M GTP e subtraída do total. São seleccionados compostos que modulem a quantidade de [³⁵S]-GTP γ S que se liga nas células, em comparação com as células de controlo não transfectadas. A activação de receptores através de agonistas dá um aumento de até cinco vezes na ligação de [³⁵S]-GTP γ S. Esta resposta é bloqueada por antagonistas.

F. Libertação de [³H]-Ácido Araquidónico

A activação de NgR pode também potenciar a libertação de ácido araquidónico nas células, proporcionando ainda outro ensaio útil para moduladores da actividade de NgR. (Ver,

p.ex., Kanterman *et al.*, *Mol. Pharmacol.* 39: 364-369, 1991). Por exemplo, células CHO que estejam estavelmente transfectadas com um vector de expressão de NgR são plaqueadas em placas de 24 poços a uma densidade de 15000 células/poço e criadas em meio MEM suplementado com soro fetal bovino a 10%, glutamina 2 mM, 10 U/ml de penicilina e 10 µg/ml de estreptomicina durante 48 horas a 37°C antes de utilizar. As células de cada poço são marcadas através de incubação com [³H]-ácido araquidónico (Amersham Corp., 210 Ci/mmol) a 0,5 µCi/ml em 1 ml de MEM suplementado com HEPES 10 mM, pH 7,5 e albumina de soro bovino livre de ácidos gordos a 0,5% durante 2 horas a 37°C. As células são então lavadas duas vezes com 1 ml do mesmo tampão.

Os compostos moduladores candidatos são adicionados em 1 ml do mesmo tampão, sozinhos ou com ATP 10 µM e as células são incubadas a 37°C durante 30 minutos. O tampão sozinho e as células falsamente transfectadas são utilizados como controlos. Amostras (0,5 ml) de cada poço são contadas através de espectroscopia de cintilações líquidas. Agonistas que activem o receptor conduzirão a potenciação da libertação estimulada por ATP de [³H]-ácido araquidónico. Esta potenciação é bloqueada por antagonistas.

G. Velocidade de Acidificação Extracelular

Ainda noutro ensaio, os efeitos dos moduladores candidatos da actividade de NgR são ensaiados através da monitorização de alterações extracelulares no pH induzidas pelos compostos de teste (ver, p.ex., Dunlop *et al.*, *J. Pharmacol. Toxicol. Meth.* 40: 47-55, 1998). Num caso, células CHO transfectadas com um vector de expressão de NgR são semeadas em copos de cápsulas de 12 mm (Molecular Devices Corp.) a 4×10^5 células/copo em MEM suplementado com soro fetal bovino a 10%, L-glutamina 2 mM, 10 U/ml de penicilina, e 10 µg/ml de estreptomicina. As células são incubadas neste meio a 37°C em CO₂ a 5% durante 24 horas.

As velocidades de acidificação extracelular são medidas utilizando um microfisiómetro Cytosensor (Molecular Devices Corp.). Os copos de cápsulas são carregados nas câmaras do sensor do microfisiómetro e as câmaras são perfundidas com tampão de corrida (MEM livre de bicarbonato suplementado com

L-glutamina 4 mM, 10 unidades/ml de penicilina, 10 µg/ml de estreptomicina, NaCl 26 mM) a um caudal de 100 µl/minuto. Os agonistas candidatos ou outros agentes são diluídos no tampão de corrida e perfundidos através de uma segunda via de fluido. Durante cada ciclo de bombagem de 60 segundos, a bomba funciona durante 38 segundos e está desligada durante os restantes 22 segundos. O pH do tampão de corrida na câmara do sensor é registado durante o ciclo de 43-58 segundos e a bomba é reiniciada aos 60 segundos para iniciar o ciclo seguinte. A velocidade de acidificação do tampão de corrida durante o tempo de registo é calculada através do programa Cytosoft. Alterações na velocidade de acidificação são calculadas através da subtracção do valor limiar (a média de 4 medições da velocidade imediatamente antes de um modulador candidato) da medição da velocidade mais elevada obtida após a adição de um modulador candidato. O instrumento seleccionado detecta 61 mV/unidade de pH. Moduladores que actuem como agonistas do receptor resultam num aumento na velocidade de acidificação extracelular em comparação com a velocidade na ausência do agonista. Esta resposta é bloqueada através de moduladores que actuem como antagonistas do receptor.

Exemplo 15: mNgR3 não se liga a hNogo-A(1055-1120)

Para testar funcionalmente o NgR3 de ratinho (daqui em diante, mNgR3) quanto à sua capacidade para se ligar a hNogo-A(1055-1120), foi criado um vector de expressão de ADNc para uma proteína mNgR3 marcada com um epitopo myc. O ADNc de NgR3 de ratinho foi amplificado por PCR a partir de ADNc de cérebro adulto de ratinho, da sequência sinal até ao codão de terminação e ligado no vector pSecTag2 para que o vector codifique uma sequência sinal seguida de uma etiqueta myc seguida da sequência de mNgR3 madura. Este plasmídeo foi transfectado para células COS07, e a expressão de uma proteína etiquetada com myc do tamanho previsto foi verificada através de análise *immunoblot*. Estudos de ligação de fosfatase alcalina-hNogo-A(1055-1120) e imuno-histologia de myc foram conduzidos tal como descrito (Fournier *et al.*, *supra*).

As células expressando mNgR3 expressam a proteína etiquetada com myc mas a ligação a AP-hNogo-A(1055-1120) não foi observada sob as condições empregues (Fig. 8).

Exemplo 16: Identificação de ADNc de NgR3 humano parcial e sequências proteicas

O programa tblastn foi utilizado para pesquisar o homólogo humano de NgR3 de ratinho. A sequência proteica de NgR3 de ratinho (SEQ ID NO:4) foi utilizada para pesquisar uma base de dados de etiquetas de sequências expressas (EST) humanas privada de Incyte produzindo um acerto altamente significativo: Molde Incyte ID 190989.1. Esta sequência (937 nucleótidos) contém um enquadramento de leitura aberto de 312 aminoácidos no segundo enquadramento inverso que exibe 88% de identidade com os resíduos 66 a 381 de NgR3 de ratinho (SEQ ID NO:4), indicando fortemente que é parte do homólogo de NgR3 humano.

Uma busca de SEQ ID NO:4 contra a base de dados de EST humanas pública em Genbank também produziu um acerto com uma EST de 465 pb (número de Entrada: R35699; número da Versão: R35699.1; GI: 792600). Existem várias deleções e inserções de um único nucleótido dentro desta sequência que causam erros de mudança de enquadramento. Toda a sequência de confiança contida nesta EST pública está presente na EST de Incyte (Molde ID 190989.1).

Para obter mais sequência nucleotídica que prolongaria a sequência de aminoácidos na extremidade do terminal carboxilo, o clone N.º 38319 do Consórcio I.M.A.G.E., que corresponde ao N.º de entrada Genbank R35699, foi adquirido em Incyte Genomics Inc. e sujeito a mais análise da sequência de ADN. Este clone consiste num fragmento *NotI/HindIII* contendo a sequência de interesse, clonada nos locais *NotI/HindIII* do vector Lafmid BA (<http://image.llnl.gov/image/html/libs/lafmidBA.shtml>). O clone foi recebido como uma punção de agar, que foi riscada em placas de LB-agar contendo 50 µg/ml de ampicilina para isolar colónias individuais. Foram criadas seis colónias em meio LB com antibiótico e foi preparado ADN plasmídico utilizando o Sistema de Purificação de ADN Promega *Wizard Plus Miniprep* (Promega #A7500). Estes ADN foram subsequentemente digeridos com as enzimas de restrição *NotI* e *HindIII* para confirmar que os clones continham uma inserção. A inserção de um isolado foi sequenciada utilizando uma combinação de iniciadores específicos do vector e específicos

do gene produzindo uma sequência nucleotídica parcial de NgR3 humano de 1176 nucleótidos (SEQ ID NO:13). Uma tradução desta sequência proporciona uma sequência parcial para NgR3 humano de 392 aminoácidos (SEQ ID NO:14).

A sequência nucleotídica de SEQ ID NO:13 difere da sequência de EST de Incyte em três posições. As posições nucleotídicas 12-13 em SEQ ID NO:13 são CG, enquanto os nucleótidos correspondentes no Molde ID 190989.1 de Incyte são GT (i.e., as posições 12-13 do complemento do Molde ID 190989.1 de Incyte). Adicionalmente, a posição 641 em SEQ ID NO:13 é um C, enquanto o nucleótido correspondente na sequência do Molde ID 190989.1 de Incyte é um A (i.e., a posição 641 do complemento do Molde ID 190989.1 de Incyte). Isto resulta em duas mudanças nos aminoácidos quando se compara SEQ ID NO:14 com o ORF codificado pelo Molde ID 190989.1 de Incyte; SEQ ID NO:14 contém uma valina na posição 5, enquanto o ORF codificado pelo Molde ID 190989.1 de Incyte contém uma leucina; SEQ ID NO:14 contém uma alanina na posição 214, enquanto o ORF codificado pelo Molde ID 190989.1 de Incyte contém um ácido glutâmico.

A sequência nucleotídica de SEQ ID NO:13 difere da sequência da EST pública (número de Entrada: R35699; Versão número: R35699.1; GI: 792600) em duas posições (dentro dos primeiros 200 nucleótidos da sequência de confiança). As posições nucleotídicas 12-13 em SEQ ID NO:13 são CG, enquanto os nucleótidos correspondentes na EST pública são GT (i.e., posições 12-13 da EST pública; número de Entrada: R35699; Versão nº: R35699.1; GI: 792600). Isto conduz à alteração de um único aminoácido quando se compara SEQ ID NO:14 com o ORF codificado pela EST pública: SEQ ID NO:14 contém uma valina na posição 5, enquanto o ORF codificado pela EST pública contém uma leucina.

Uma análise Bestfit da sequência de aminoácidos humana parcial com a sequência de aminoácidos de ratinho inteira indica que a sequência de aminoácidos de NgR3 humana está completa na extremidade do terminal carboxilo e que partilham 89,54% de identidade. Um alinhamento de todas as proteínas NgR é mostrado na Figura 9. Embora na sequência de aminoácidos de NgR3 humana estejam em falta os primeiros 25 aminoácidos, pode

ser determinado que a proteína NgR3 humana contém as seguintes características em comum com as outras sequências de NgR: (1) oito domínios de Repetições Ricas em Leucina (LRR); (2) um domínio LRR carboxi-terminal (LRR-CT); (3) uma cisteína conservada no quarto domínio LRR; (4) um potencial local de glicosilação conservado no oitavo domínio LRR; e (5) um terminal carboxilo hidrófobo.

Tal como apreciarão os peritos na arte, poderão ser feitas numerosas alterações e modificações às concretizações preferidas do presente invento sem se sair do espírito do presente invento. Pretende-se que todas essas variações recaiam dentro do âmbito do presente invento.

Chave para a Listagem das Sequências:

SEQ ID NO:1	sequência de ADNc de NgR2 humana derivada da sequência genómica AC013606
SEQ ID NO:2	sequência de aminoácidos de NgR2 humana
SEQ ID NO:3	sequência de ADNc de NgR3 de ratinho derivada de AC021768
SEQ ID NO:4	uma sequência de aminoácidos de NgR3 de ratinho
SEQ ID NO:5	uma sequência de aminoácidos de NgR1 humana
SEQ ID NO:6	uma sequência de aminoácidos de consenso para NgR
SEQ ID NO:7	resíduos de aminoácidos #1055-1120 de hNogoA (Nogo-66)
SEQ ID NO:8	uma sequência de aminoácidos de NgR2 humana madura
SEQ ID NO:9	uma sequência de aminoácidos de NgR3 de ratinho madura
SEQ ID NO:10	uma sequência de aminoácidos do LLRNT de NgR de consenso
SEQ ID NO:11	uma sequência de aminoácidos do domínio LRRCT de NgR de consenso
SEQ ID NO:12	uma sequência de aminoácidos do domínio LRR de NgR de consenso
SEQ ID NO:13	uma sequência nucleotídica de NgR3 humana parcial
SEQ ID NO:14	uma sequência de aminoácidos de NgR3 humana parcial
SEQ ID NO:15	uma sequência genómica codificando uma sequência de NgR2 humana
SEQ ID NO:16	uma sequência genómica (cadeia complementar) codificando um NgR3 de ratinho
SEQ ID NO:17	uma sequência de aminoácidos de NgR1 de ratinho
SEQ ID NO:18	uma sequência de consenso para o domínio NTLRRT de NgR
SEQ ID NO:19	uma sequência de aminoácidos do domínio LRRCT de NgR de consenso

LISTAGEM DE SEQUÊNCIAS

<110> BIOGEN, INC.

YALE UNIVERSITY

STEPHEN M.STRITTMATTER,

RICHARD L.CATE,

DINAH W.Y.SAH,

<120> HOMÓLOGOS DO RECEPTOR DE NOGO

<130> A116PCT

<140>

<141>

<150> 60/238,361

<151> 2000-10-06

<160> 16

<170> PatentIn Ver. 2.1

<210> 17

<211> 1260

<212> ADN

<213> *Homo sapiens*

<400> 1

```

atgctgcccc ggctcaggcg cctgctgcaa gctccccgct eggcctgcct cctgctgatg 60
ctcctggccc tgccccctggc ggccccccagc tgccccatgc tctgcacctg ctactcatcc 120
ccgcccaccc tgagctgcca ggccaacaac ttctcctctg tgccgctgtc cctgccaccc 180
agcactcagc gactcttctc gcagaacaac ctcatccgca cgctgcggcc aggcaccttt 240
gggtccaacc tgctcaccct gtggctcttc tccaacaacc tctccaccat ctaccgggc 300
actttccgcc acttgcaagc cctggaggag ctggacctcg gtgacaaccg gcacctgcgc 360
tcgctggagc ccgacacctt ccagggcctg gagcggtgc agtcgctgca tttgtaccgc 420
tgccagctca gcagcctgcc cggcaacatc ttccgaggcc tggtcagcct gcagtacctc 480
tacctccagg agaacagcct gctccacctc caggatgact tgttcgcgga cctggccaac 540
ctgagccacc tcttcctcca cgggaaccgc ctgcggtgc tcacagagca cgtgtttcgc 600
ggcctgggca gcctggaccg gctgctgctg cacgggaacc ggctgcaggc cgtgcaccgc 660
gcggccttcc gcggcctcag ccgcctcacc atcctctacc tgttcaacaa cagcctggcc 720
tcgctgcccg gcgaggcgct ccgcgacctg ccctcgtctg agttcctgcg gctcaacgct 780
aaccctggg cgtgcgactg ccgcgcgcgc ccgctctggg cctgggtcca gcgcgcgcgc 840
gtgtccagct ccgacgtgac ctgcgccacc ccccggagc gccagggcgc agacctgcgc 900
gcgctccgcg aggcgcgactt ccaggcgtgt ccgcccgcgc caccacgcgc gccgggcagc 960
cgcgcccgcg gcaacagctc ctccaaccac ctgtacggg tggccgaggc cggggcgccc 1020
ccagccgata cctccacctt ctaccgagat ctgectgccg aagactcgcg ggggcgcag 1080
ggcggggacg cgctactga ggacgactac tgggggggct acgggggtga ggaccagcga 1140
ggggagcaga tgtgccccgc cgctgcctgc caggcgcccc cggactccgc aggcctgcgc 1200
ctctcgccgc ggctccccag ccctctgctt tgctcctgc tcctggtgcc ccaccaacct 1260

```

<210> 2

<211> 420

<212> PRT

<213> *Homo sapiens*

<400> 2

Met Leu Pro Gly Leu Arg Arg Leu Leu Gln Ala Pro Ala Ser Ala Cys
 1 5 10 15
 Leu Leu Leu Met Leu Leu Ala Leu Pro Leu Ala Ala Pro Ser Cys Pro
 20 25 30
 Met Leu Cys Thr Cys Tyr Ser Ser Pro Pro Thr Val Ser Cys Gln Ala
 35 40 45
 Asn Asn Phe Ser Ser Val Pro Leu Ser Leu Pro Pro Ser Thr Gln Arg
 50 55 60
 Leu Phe Leu Gln Asn Asn Leu Ile Arg Thr Leu Arg Pro Gly Thr Phe
 65 70 75 80
 Gly Ser Asn Leu Leu Thr Leu Trp Leu Phe Ser Asn Asn Leu Ser Thr
 85 90 95
 Ile Tyr Pro Gly Thr Phe Arg His Leu Gln Ala Leu Glu Glu Leu Asp
 100 105 110
 Leu Gly Asp Asn Arg His Leu Arg Ser Leu Glu Pro Asp Thr Phe Gln
 115 120 125
 Gly Leu Glu Arg Leu Gln Ser Leu His Leu Tyr Arg Cys Gln Leu Ser
 130 135 140
 Ser Leu Pro Gly Asn Ile Phe Arg Gly Leu Val Ser Leu Gln Tyr Leu
 145 150 155 160
 Tyr Leu Gln Glu Asn Ser Leu Leu His Leu Gln Asp Asp Leu Phe Ala
 165 170 175
 Asp Leu Ala Asn Leu Ser His Leu Phe Leu His Gly Asn Arg Leu Arg
 180 185 190
 Leu Leu Thr Glu His Val Phe Arg Gly Leu Gly Ser Leu Asp Arg Leu
 195 200 205
 Leu Leu His Gly Asn Arg Leu Gln Gly Val His Arg Ala Ala Phe Arg
 210 215 220
 Gly Leu Ser Arg Leu Thr Ile Leu Tyr Leu Phe Asn Asn Ser Leu Ala
 225 230 235 240
 Ser Leu Pro Gly Glu Ala Leu Ala Asp Leu Pro Ser Leu Glu Phe Leu
 245 250 255
 Arg Leu Asn Ala Asn Pro Trp Ala Cys Asp Cys Arg Ala Arg Pro Leu
 260 265 270
 Trp Ala Trp Phe Gln Arg Ala Arg Val Ser Ser Ser Asp Val Thr Cys
 275 280 285
 Ala Thr Pro Pro Glu Arg Gln Gly Arg Asp Leu Arg Ala Leu Arg Glu
 290 295 300
 Ala Asp Phe Gln Ala Cys Pro Pro Ala Ala Pro Thr Arg Pro Gly Ser
 305 310 315 320
 Arg Ala Arg Gly Asn Ser Ser Ser Asn His Leu Tyr Gly Val Ala Glu
 325 330 335
 Ala Gly Ala Pro Pro Ala Asp Pro Ser Thr Leu Tyr Arg Asp Leu Pro
 340 345 350
 Ala Glu Asp Ser Arg Gly Arg Gln Gly Gly Asp Ala Pro Thr Glu Asp
 355 360 365
 Asp Tyr Trp Gly Gly Tyr Gly Gly Glu Asp Gln Arg Gly Glu Gln Met

370		375		380
Cys Pro Gly Ala Ala	Cys Gln Ala Pro Pro Asp	Ser Arg Gly Pro Ala		
385	390	395		400
Leu Ser Ala Gly Leu	Pro Ser Pro Leu Leu Cys	Leu Leu Leu Leu Val		
	405	410		415
Pro His His Leu				
	420			

<210> 3

<211> 1383

<212> ADN

<213> *Mus sp.*

<400> 3

```

atgtcttggc agtctggaac cacagtgaca caatctcccg tgcaggctgc tcagggtctca 60
gggtgctgtg tggaaattgct gctgttgctg ctcgctggag agctacctct gggtgggtgt 120
tgtctctgag actgtgtgtg ctacctgctg cccatgactg tcagctgcca ggacacaaac 180
tttgctgcca tcccggaggg catcccagag gacagtgagc gcatcttcct gcagaacaat 240
cgcacacact tcctccagca gggccacttc agccccgcca tggtcacctt ctggatctac 300
tccaacaaca tcactttcat tgctccaac accttcgagg gctttgtgca tctggaggag 360
ctagaccttg gagacaaccg acagctgcga acgctggcac ccgagacctt ccaaggcctg 420
gtgaagcttc acgccctcta cctctataag tgtggactga gcgccctgcc cgcaggcatc 480
tttggtgccc tgcacagcct gcagtatctc tacttgagg acaaccatat cgagtacctc 540
caagatgaca tctttgtgga cctgttcaat ctcagtcact tgtttctcca tggtaacaag 600
ctatggagcc tgggccaagg catcttccgg ggctgtgtga acctggaccg gttgctgtct 660
catgagaacc agctacagtg gggtcaccac aaggctttcc atgacctcca caggctaacc 720
acctcttttc tcttcaacaa cagcctcact gagctgcagg gtgactgtct ggccccctg 780
gtggccttgg agttccttcg cctcaatggg aatgcttggg actgtggctg ccgggcacgt 840
tccctgtggg aatggctgcg aaggttccgt ggccttagct ctgctgtccc ctgcgcgacc 900
cccagctgc ggcaaggcca ggatctgaag ctgctgaggg tggaggactt ccggaactgc 960
acaggaccag tgtctcctca ccagatcaag tctcacacgc ttaccacctc tgacagggct 1020
gcccgaagg agcaccatcc gtcccatggg gcctccaggg acaaaggcca cccacatggc 1080
catccgcctg gctccaggtc aggttacaag aaggcaggca agaactgcac cagccacagg 1140
aaccggaacc agatctctaa ggtgagctct gggaagagc ttaccgaact gcaggactat 1200
gccccgact atcagcaciaa gttcagcttt gacatcatgc ccaccgcacg acccaagagg 1260
aagggcaagt gtgctcgag gaccccatc cgtgccccca gtggggtgca gcaggcatcc 1320
tcaggcacgg cccttggggc cccactcctg gcctggatac tggggctggc agtcactctc 1380
cgc

```

<210> 4

<211> 461

<212> PRT

<213> *Mus sp.*

<400> 4

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

Arg	Ile	Thr	Phe	Leu	Gln	Gln	Gly	His	Phe	Ser	Pro	Ala	Met	Val	Thr	
				85					90					95		
Leu	Trp	Ile	Tyr	Ser	Asn	Asn	Ile	Thr	Phe	Ile	Ala	Pro	Asn	Thr	Phe	
				100					105					110		
Glu	Gly	Phe	Val	His	Leu	Glu	Glu	Leu	Asp	Leu	Gly	Asp	Asn	Arg	Gln	
				115					120					125		
Leu	Arg	Thr	Leu	Ala	Pro	Glu	Thr	Phe	Gln	Gly	Leu	Val	Lys	Leu	His	
				130					135					140		
Ala	Leu	Tyr	Leu	Tyr	Lys	Cys	Gly	Leu	Ser	Ala	Leu	Pro	Ala	Gly	Ile	
				145					150					155		
Phe	Gly	Gly	Leu	His	Ser	Leu	Gln	Tyr	Leu	Tyr	Leu	Gln	Asp	Asn	His	
				165					170					175		
Ile	Glu	Tyr	Leu	Gln	Asp	Asp	Ile	Phe	Val	Asp	Leu	Val	Asn	Leu	Ser	
				180					185					190		
His	Leu	Phe	Leu	His	Gly	Asn	Lys	Leu	Trp	Ser	Leu	Gly	Gln	Gly	Ile	
				195					200					205		
Phe	Arg	Gly	Leu	Val	Asn	Leu	Asp	Arg	Leu	Leu	Leu	His	Glu	Asn	Gln	
				210					215					220		
Leu	Gln	Trp	Val	His	His	Lys	Ala	Phe	His	Asp	Leu	His	Arg	Leu	Thr	
				225					230					235		
Thr	Leu	Phe	Leu	Phe	Asn	Asn	Ser	Leu	Thr	Glu	Leu	Gln	Gly	Asp	Cys	
				245					250					255		
Leu	Ala	Pro	Leu	Val	Ala	Leu	Glu	Phe	Leu	Arg	Leu	Asn	Gly	Asn	Ala	
				260					265					270		
Trp	Asp	Cys	Gly	Cys	Arg	Ala	Arg	Ser	Leu	Trp	Glu	Trp	Leu	Arg	Arg	
				275					280					285		
Phe	Arg	Gly	Ser	Ser	Ser	Ala	Val	Pro	Cys	Ala	Thr	Pro	Glu	Leu	Arg	
				290					295					300		
Gln	Gly	Gln	Asp	Leu	Lys	Leu	Leu	Arg	Val	Glu	Asp	Phe	Arg	Asn	Cys	
				305					310					315		
Thr	Gly	Pro	Val	Ser	Pro	His	Gln	Ile	Lys	Ser	His	Thr	Leu	Thr	Thr	
				325					330					335		
Ser	Asp	Arg	Ala	Ala	Arg	Lys	Glu	His	His	Pro	Ser	His	Gly	Ala	Ser	
				340					345					350		
Arg	Asp	Lys	Gly	His	Pro	His	Gly	His	Pro	Pro	Gly	Ser	Arg	Ser	Gly	
				355					360					365		
Tyr	Lys	Lys	Ala	Gly	Lys	Asn	Cys	Thr	Ser	His	Arg	Asn	Arg	Asn	Gln	
				370					375					380		
Ile	Ser	Lys	Val	Ser	Ser	Gly	Lys	Glu	Leu	Thr	Glu	Leu	Gln	Asp	Tyr	
				385					390					395		
Ala	Pro	Asp	Tyr	Gln	His	Lys	Phe	Ser	Phe	Asp	Ile	Met	Pro	Thr	Ala	
				405					410					415		
Arg	Pro	Lys	Arg	Lys	Gly	Lys	Cys	Ala	Arg	Arg	Thr	Pro	Ile	Arg	Ala	

	420		425		430
Pro Ser Gly Val Gln Gln Ala Ser Ser Gly Thr Ala Leu Gly Ala Pro					
435		440		445	
Leu Leu Ala Trp Ile Leu Gly Leu Ala Val Thr Leu Arg					
450		455		460	

<210> 5

<211> 473

<212> PRT

<213> *Homo sapiens*

<400> 5

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

260										265					270						
Ala	Trp	Leu	Gln	Lys	Phe	Arg	Gly	Ser	Ser	Ser	Glu	Val	Pro	Cys	Ser						
		275					280					285									
Leu	Pro	Gln	Arg	Leu	Ala	Gly	Arg	Asp	Leu	Lys	Arg	Leu	Ala	Ala	Asn						
		290				295					300										
Asp	Leu	Gln	Gly	Cys	Ala	Val	Ala	Thr	Gly	Pro	Tyr	His	Pro	Ile	Trp						
305					310					315					320						
Thr	Gly	Arg	Ala	Thr	Asp	Glu	Glu	Pro	Leu	Gly	Leu	Pro	Lys	Cys	Cys						
				325					330					335							
Gln	Pro	Asp	Ala	Ala	Asp	Lys	Ala	Ser	Val	Leu	Glu	Pro	Gly	Arg	Pro						
			340					345						350							
Ala	Ser	Ala	Gly	Asn	Ala	Leu	Lys	Gly	Arg	Val	Pro	Pro	Gly	Asp	Ser						
		355					360					365									
Pro	Pro	Gly	Asn	Gly	Ser	Gly	Pro	Arg	His	Ile	Asn	Asp	Ser	Pro	Phe						
		370				375					380										
Gly	Thr	Leu	Pro	Gly	Ser	Ala	Glu	Pro	Pro	Leu	Thr	Ala	Val	Arg	Pro						
385					390					395					400						
Glu	Gly	Ser	Glu	Pro	Pro	Gly	Phe	Pro	Thr	Ser	Gly	Pro	Arg	Arg	Arg						
				405					410					415							
Pro	Gly	Cys	Ser	Arg	Lys	Asn	Arg	Thr	Arg	Ser	His	Cys	Arg	Leu	Gly						
			420					425					430								
Gln	Ala	Gly	Ser	Gly	Gly	Gly	Gly	Thr	Gly	Asp	Ser	Glu	Gly	Ser	Gly						
		435					440					445									
Ala	Leu	Pro	Ser	Leu	Thr	Cys	Ser	Leu	Thr	Pro	Leu	Gly	Leu	Ala	Leu						
		450				455					460										
Val	Leu	Trp	Thr	Val	Leu	Gly	Pro	Cys													
465					470																

 $\langle 210 \rangle$ 6

$\langle 211 \rangle$ 440

<212> PRT

<213> Sequência Artificial

 $\langle 220 \rangle$

<223> Descrição de Sequência Artificial: Sequência de consenso

 $\langle 220 \rangle$

```
<221> MOD RES
```

$\langle 222 \rangle$ (3) . . (4)

<223> Aminoácido variável

 $\langle 220 \rangle$

<221> MOD RES

 $\langle 222 \rangle$ (6)

<223> Aminoácido variável

 $\langle 220 \rangle$

<221> MOD_RES

$\langle 222 \rangle$ (9) . . (10)

<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (12)..(13)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (15)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (18)..(25)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (27)..(29)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (31)..(34)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (36)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (39)..(40)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (42)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (44)..(50)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (52)..(59)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (62)..(63)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (66).. (69)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (71)..(74)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (76)..(80)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (83)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (87)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (90)..(91)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (94)..(96)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (98)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (101)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (104)..(105)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (107)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (109)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (111)..(112)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (114)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (116)..(117)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (119)..(122)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (124)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (127)..(128)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (136)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (138)..(141)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (193)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (146)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (148)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (151)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (154)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (162)..(170)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (175)..(176)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (184)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (186)..(189)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (192)..(193)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (196)..(197)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (199)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (202)..(203)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (205)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (208)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (210)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (212)..(213)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (215)..(218)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (221)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (223)..(224)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (226)..(227)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (232)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (234)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (236)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (238)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (243)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (246)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (248)..(251)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (253)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (257)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (259)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (261).. (262)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (264).. (267)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (269)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (272)..(273)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (275)..(277)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (279)..(281)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (283)..(287)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (289)..(290)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (292)..(328)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (330)..(341)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (344)..(346)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (348)..(399)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (901)..(428)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (431)..(439)
<223> Aminoácido variável

<400> 6

Cys Pro Xaa Xaa Cys Xaa Cys Tyr Xaa Xaa Pro Xaa Xaa Thr Xaa Ser
1 5 10 15

Cys Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Pro Xaa Xaa Xaa Pro Xaa Xaa
 20 25 30
 Xaa Xaa Arg Xaa Phe Leu Xaa Xaa Asn Xaa Ile Xaa Xaa Xaa Xaa Xaa
 35 40 45
 Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Trp Xaa Xaa Ser
 50 55 60
 Asn Xaa Xaa Xaa Xaa Ile Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa
 65 70 75 80
 Leu Glu Xaa Leu Asp Leu Xaa Asp Asn Xaa Xaa Leu Arg Xaa Xaa Xaa
 85 90 95
 Pro Xaa Thr Phe Xaa Gly Leu Xaa Xaa Leu Xaa Leu Xaa Leu Xaa Xaa
 100 105 110
 Cys Xaa Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Phe Xaa Gly Leu Xaa Xaa
 115 120 125
 Leu Gln Tyr Leu Tyr Leu Gln Xaa Asn Xaa Xaa Xaa Xaa Leu Xaa Asp
 130 135 140
 Asp Xaa Phe Xaa Asp Leu Xaa Asn Leu Xaa His Leu Phe Leu His Gly
 145 150 155 160
 Asn Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Arg Gly Leu Xaa Xaa
 165 170 175
 Leu Asp Arg Leu Leu Leu His Xaa Asn Xaa Xaa Xaa Xaa Val His Xaa
 180 185 190
 Xaa Ala Phe Xaa Xaa Leu Xaa Arg Leu Xaa Xaa Leu Xaa Leu Phe Xaa
 195 200 205
 Asn Xaa Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Leu Ala Xaa Leu Xaa Xaa
 210 215 220
 Leu Xaa Xaa Leu Arg Leu Asn Xaa Asn Xaa Trp Xaa Cys Xaa Cys Arg
 225 230 235 240
 Ala Arg Xaa Leu Trp Xaa Trp Xaa Xaa Xaa Xaa Arg Xaa Ser Ser Ser
 245 250 255
 Xaa Val Xaa Cys Xaa Xaa Pro Xaa Xaa Xaa Xaa Gly Xaa Asp Leu Xaa
 260 265 270
 Xaa Leu Xaa Xaa Xaa Asp Xaa Xaa Xaa Cys Xaa Xaa Xaa Xaa Xaa Pro
 275 280 285
 Xaa Xaa Pro Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 290 295 300
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 305 310 315 320
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Gly Xaa Xaa Xaa Xaa Xaa Xaa
 325 330 335
 Xaa Xaa Xaa Xaa Xaa Pro Pro Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa Xaa
 340 345 350
 Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa

355	360	365
Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Xaa
370	375	380
Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Arg
385	390	395 400
Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Xaa
405	410	415
Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Xaa
420	425	430
Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Xaa	Xaa Xaa Xaa Xaa Xaa Xaa
435	440	

<210> 7
 <211> 66
 <212> PRT
 <213> *Homo sapiens*

<400> 7

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

<210> 8
 <211> 390
 <212> PRT
 <213> *Homo sapiens*

<400> 8

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

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

<210> 9

<211> 421

<212> PRT

<213> *Mus sp.*

<400> 9

Cys Pro Arg Asp Cys Val Cys Tyr Pro Ala Pro Met Thr Val Ser Cys
 1 5 10 15
 Gln Ala His Asn Phe Ala Ala Ile Pro Glu Gly Ile Pro Glu Asp Ser
 20 25 30
 Glu Arg Ile Phe Leu Gln Asn Asn Arg Ile Thr Phe Leu Gln Gln Gly
 35 40 45
 His Phe Ser Pro Ala Met Val Thr Leu Trp Ile Tyr Ser Asn Asn Ile
 50 55 60
 Thr Phe Ile Ala Pro Asn Thr Phe Glu Gly Phe Val His Leu Glu Glu
 65 70 75 80
 Leu Asp Leu Gly Asp Asn Arg Gln Leu Arg Thr Leu Ala Pro Glu Thr
 85 90 95
 Phe Gln Gly Leu Val Lys Leu His Ala Leu Tyr Leu Tyr Lys Cys Gly
 100 105 110
 Leu Ser Ala Leu Pro Ala Gly Ile Phe Gly Gly Leu His Ser Leu Gln
 115 120 125
 Tyr Leu Tyr Leu Gln Asp Asn His Ile Glu Tyr Leu Gln Asp Asp Ile
 130 135 140
 Phe Val Asp Leu Val Asn Leu Ser His Leu Phe Leu His Gly Asn Lys
 145 150 155 160
 Leu Trp Ser Leu Gly Gln Gly Ile Phe Arg Gly Leu Val Asn Leu Asp
 165 170 175
 Arg Leu Leu Leu His Glu Asn Gln Leu Gln Trp Val His His Lys Ala
 180 185 190
 Phe His Asp Leu His Arg Leu Thr Thr Leu Phe Leu Phe Asn Asn Ser
 195 200 205
 Leu Thr Glu Leu Gln Gly Asp Cys Leu Ala Pro Leu Val Ala Leu Glu
 210 215 220
 Phe Leu Arg Leu Asn Gly Asn Ala Trp Asp Cys Gly Cys Arg Ala Arg
 225 230 235 240
 Ser Leu Trp Glu Trp Leu Arg Arg Phe Arg Gly Ser Ser Ser Ala Val
 245 250 255
 Pro Cys Ala Thr Pro Glu Leu Arg Gln Gly Gln Asp Leu Lys Leu Leu
 260 265 270
 Arg Val Glu Asp Phe Arg Asn Cys Thr Gly Pro Val Ser Pro His Gln
 275 280 285
 Ile Lys Ser His Thr Leu Thr Thr Ser Asp Arg Ala Ala Arg Lys Glu
 290 295 300
 His His Pro Ser His Gly Ala Ser Arg Asp Lys Gly His Pro His Gly
 305 310 315 320
 His Pro Pro Gly Ser Arg Ser Gly Tyr Lys Lys Ala Gly Lys Asn Cys
 325 330 335
 Thr Ser His Arg Asn Arg Asn Gln Ile Ser Lys Val Ser Ser Gly Lys

340	345	350
Glu Leu Thr 355	Glu Leu Gln Asp Tyr 360	Ala Pro Asp Tyr Gln His Lys Phe 365
Ser Phe Asp Ile Met 370	Pro Thr Ala Arg 375	Pro Lys Arg Lys Gly Lys Cys 380
Ala Arg Arg Thr Pro 385	Ile Arg Ala Pro Ser Gly Val 390 395	Gln Gln Ala Ser 400
Ser Gly Thr Ala Leu 405	Gly Ala Pro Leu Leu Ala Trp 410	Ile Leu Gly Leu 415
Ala Val Thr Leu Arg 420		

<210> 10

<211> 17

<212> PRT

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: Sequência de consenso

<220>

<221> MOD_RES

<222> (3)..(4)

<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (6)

<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (9).. (10)

<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (12).. (13)

<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (15)

<223> Aminoácido variável

<400> 10

Cys	Pro	Xaa	Xaa	Cys	Xaa	Cys	Tyr	Xaa	Xaa	Pro	Xaa	Xaa	Thr	Xaa	Ser
1				5						10					15

Cys

<210> 11
<211> 50
<212> PRT
<213> Sequência Artificial

<220>
<223> Descrição de Sequência Artificial: Sequência de consenso

<220>
<221> MOD_RES
<222> (2)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (4)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (6)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (11)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (14)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (16)..(19)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (21)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (25)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (27)
<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (29)..(30)

<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (32)..(35)

<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (37)

<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (40)..(41)

<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (43)..(45)

<223> Aminoácido variável

<220>

<221> MOD_RES

<222> (47)..(49)

<223> Aminoácido variável

<400> 11

Asn	Xaa	Trp	Xaa	Cys	Xaa	Cys	Arg	Ala	Arg	Xaa	Leu	Trp	Xaa	Trp	Xaa
1				5					10					15	
.															
Xaa	Xaa	Xaa	Arg	Xaa	Ser	Ser	Ser	Xaa	Val	Xaa	Cys	Xaa	Xaa	Pro	Xaa
			20					25				30			
.															
Xaa	Xaa	Xaa	Gly	Xaa	Asp	Leu	Xaa	Xaa	Leu	Xaa	Xaa	Xaa	Asp	Xaa	Xaa
			35			40						45			
.															
Xaa	Cys														
	50														

<210> 12

<211> 196

<212> PRT

<213> Sequência Artificial

<220>

<223> Descrição de Sequência Artificial: Sequência de consenso

<220>

<221> MOD_RES

<222> (2)

<223> Aminoácido variável

<220>

<221> MOD_RES
<222> (5)..(6)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (8)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (10)..(16)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (18)..(25)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (28)..(29)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (32)..(35)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (37)..(40)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (42)..(46)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (49)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (53)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (56)..(57)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (60)..(62)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (64)

<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (67)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (70)..(71)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (73)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (75)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (77)..(78)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (80)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (82)..(B3)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (85)..(88)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (90)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (93)..(94)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (102)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (104)..(107)
<223> Aminoácido variável

<220>
<221> MOD_RES

<222> (109)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (112)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (114)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (117)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (120)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (128)..(136)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (141)..(142)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (150)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (152)..(155)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (158)..(159)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (162)..(163)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (165)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (168)..(169)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (171)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (174)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (176)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (178)..(179)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (181)..(184)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (187)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (189)..(190)
<223> Aminoácido variável

<220>
<221> MOD_RES
<222> (192)..(193)
<223> Aminoácido variável

<400> 12

```

Arg Xaa Phe Leu Xaa Xaa Asn Xaa Ile Xaa Xaa Xaa Xaa Xaa Xaa Xaa
 1          5          10          15

Phe Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Leu Trp Xaa Xaa Ser Asn Xaa
 20          25          30

Xaa Xaa Xaa Ile Xaa Xaa Xaa Xaa Phe Xaa Xaa Xaa Xaa Xaa Leu Glu
 35          40          45

Xaa Leu Asp Leu Xaa Asp Asn Xaa Xaa Leu Arg Xaa Xaa Xaa Pro Xaa
 50          55          60

Thr Phe Xaa Gly Leu Xaa Xaa Leu Xaa Leu Xaa Leu Xaa Xaa Cys Xaa
 65          70          75          80

Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Phe Xaa Gly Leu Xaa Xaa Leu Gln
 85          90          95

Tyr Leu Tyr Leu Gln Xaa Asn Xaa Xaa Xaa Xaa Leu Xaa Asp Asp Xaa
100          105          110

Phe Xaa Asp Leu Xaa Asn Leu Xaa His Leu Phe Leu His Gly Asn Xaa
115          120          125

Xaa Xaa Xaa Xaa Xaa Xaa Xaa Xaa Phe Arg Gly Leu Xaa Xaa Leu Asp
130          135          140

Arg Leu Leu Leu His Xaa Asn Xaa Xaa Xaa Xaa Val His Xaa Xaa Ala
145          150          155          160

Phe Xaa Xaa Leu Xaa Arg Leu Xaa Xaa Leu Xaa Leu Phe Xaa Asn Xaa
165          170          175

Leu Xaa Xaa Leu Xaa Xaa Xaa Xaa Leu Ala Xaa Leu Xaa Xaa Leu Xaa
180          185          190

Xaa Leu Arg Leu
195

```

<210> 13

<211> 1176

<212> ADN

<213> *Homo sapiens*

<400> 13

```

gagggcatcc ccgtaggacag cgagcgcgtc ttctgcaga acaaccgcat cggcctcctc 60
cagcccggcc acttcagccc cgccatggtc accctgtgga tctactcgaa caacatcacc 120

tacatccacc ccagcacctt cgagggcttc gtgcacctgg aggagctgga cctcggcgac 180
aaccggcagc tgcggacgct ggcacccgag accttcagg gcctgggtgaa gcttcacgcc 240
ctctacctct acaagtgtgg gctcagcgcc ttgccggccg gcgtctttgg cggcctgcac 300
agcctgcagt acctctacct gcaggacaac cacatcgagt acctccagga cgacatcttc 360
gtggacctgg tcaacctcag ccacctgttt ctccacggca acaagctgtg gagtctggggc 420
ccgggcacct tccggggcct ggtgaacctg gacctcttt tgctgcacga gaaccagctg 480
cagtgggtcc accacaaggc attccacgac ctccgcaggc tgaccacctt cttcctcttc 540
aacaacagcc tctcggagct gcagggtgag tgcctggccc cgctgggggc cctggagttc 600
ctccgcctca acggcaaccc ctgggactgt ggttgtcgcg cgcgctccct gtgggaatgg 660
ctgcagaggt tccggggctc cagctccgct gtccctgtg tgtccctgg gctgcggcac 720
ggccaggacc tgaagctgct gagggccgag gacttcgga actgcacgg accagcgcc 780
ccgcaccaga tcaagtcaca cagctcacc accaccgaca gggccgcccg caaggaacac 840
cactcaccac acggcccccac caggagcaag ggccaccgc acggcccccg gcccgccac 900
aggaagccgg ggaagaactg caccaacccc aggaaccgca atcagatctc taaggcgggc 960
gccgggaac aggcccccga gctgccagac tatgcccag actaccagca caagttcagt 1020
tttgacatca tgcctacggc ccggcccaag aggaagggca agtggtgccg caggacccc 1080
atccgtgccc ccagcggggt gcagcaggcc tctcggcca gttccctggg ggcctccctc 1140
ctggcctgga cactggggct ggcggtcact ctccgc 1176

```


<210> 14

<211> 392

<212> PRT

<213> *Homo sapiens*

<400> 14

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

Asp Cys Gly Cys Arg Ala Arg Ser Leu Trp Glu Trp Leu Gln Arg Phe
 210 215 220
 Arg Gly Ser Ser Ser Ala Val Pro Cys Val Ser Pro Gly Leu Arg His
 225 230 235 240
 Gly Gln Asp Leu Lys Leu Leu Arg Ala Glu Asp Phe Arg Asn Cys Thr
 245 250 255
 Gly Pro Ala Ser Pro His Gln Ile Lys Ser His Thr Leu Thr Thr Thr
 260 265 270
 Asp Arg Ala Ala Arg Lys Glu His His Ser Pro His Gly Pro Thr Arg
 275 280 285
 Ser Lys Gly His Pro His Gly Pro Arg Pro Gly His Arg Lys Pro Gly
 290 295 300
 Lys Asn Cys Thr Asn Pro Arg Asn Arg Asn Gln Ile Ser Lys Ala Gly
 305 310 315 320
 Ala Gly Lys Gln Ala Pro Glu Leu Pro Asp Tyr Ala Pro Asp Tyr Gln
 325 330 335
 His Lys Phe Ser Phe Asp Ile Met Pro Thr Ala Arg Pro Lys Arg Lys
 340 345 350
 Gly Lys Cys Ala Arg Arg Thr Pro Ile Arg Ala Pro Ser Gly Val Gln
 355 360 365
 Gln Ala Ser Ser Ala Ser Ser Leu Gly Ala Ser Leu Leu Ala Trp Thr
 370 375 380
 Leu Gly Leu Ala Val Thr Leu Arg
 385 390

<210> 15

<211> 143899

<212> ADN

<213> *Homo sapiens*

<220>

<221> modified_base

<222> (2044)..(2144)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (6609)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (6625)..(6729)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (14153)..(14252)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (19512)..(19611)

<223> a, t, c, g, outra ou desconhecida

<220>
 <221> modified_base
 <222> (22595)..(22694)
 <223> a, t, c, g, outra ou desconhecida

 <220>
 <221> modified_base
 <222> (27825)..(27924)
 <223> a, t, c, g, outra ou desconhecida

 <220>
 <221> modified_base
 <222> (34953)..(35052)
 <223> a, t, c, g, outra ou desconhecida

 <220>
 <221> modified_base
 <222> (40783)..(40882)
 <223> a, t, c, g, outra ou desconhecida

 <220>
 <221> modified_base
 <222> (49000)..(49099)
 <223> a, t, c, g, outra ou desconhecida

 <220>
 <221> modified_base
 <222> (62884)..(62983)
 <223> a, t, c, g, outra ou desconhecida

 <220>
 <221> modified_base
 <222> (75528)..(75627)
 <223> a, t, c, g, outra ou desconhecida

 <220>
 <221> modified_base
 <222> (87944)..(88043)
 <223> a, t; c, g, outra ou desconhecida

 <220>
 <221> modified_base
 <222> (111030)..(111129)
 <223> a, t, c, g, outra ou desconhecida

<400> 15

aagcacatac	aggtgacatt	acagaactga	cagttatgcc	aggcactgta	cttagccct	60
ataccatcct	caaacagctg	tatgatgtag	attgggtatt	aacccatta	ataacaaaag	120
tacagggaac	aaagtgactt	tccaaaggtc	atgccattca	aaggagggtg	aatcttaggt	180
tggacgcagg	ctgtctgact	ctggagtctg	aggtgttaat	gctgcctcct	ccatgggaac	240
agcccaagt	aaaaacagct	gatccactct	tcatttactt	ggcatctgtg	ctaagctggt	300
ccctgagcca	agctctgagc	aacagaaaca	gaagctctgc	attaggagct	tgtgagcatg	360
tcaatgccgg	gtaaaaggagt	gctggaaaac	gctgggatgg	ccgccgagca	ctaggccggt	420
gaaggtgggc	tctgtgtgac	tggttcctct	acactctggc	ctggctgcct	gcaggaagaa	480
gatcaagctg	agtgggctgg	ccctggacca	caaggtgaca	ggtgacctct	tctacacca	540
tgtgaccacc	atgggccaga	ggctcagcca	gaaggccccc	agcctggagg	acggttcgga	600
tgccttcatt	tcacccaggt	atgttcgggg	cacctcagaa	aaccttcctg	agagttagtg	660
tctggtcaag	gtgccgcctt	tgggggatag	tgatggtggg	tcctcatatt	cagttagcac	720
tcattggtga	gtattttatt	gcacccctct	tcagtccctt	caacacccca	tgatgtaggt	780
ggggcatgct	cctcatttac	agatgggcac	atcaaagctc	agctaacgct	gggaagttca	840
gattcagggt	taccctgctg	gattcctggg	attggggagg	gaggagcttc	caaaatgggg	900

acaaggtctc	tgggcctgtc	gggtagctgg	tttccctcagg	gccccttgca	acctctgagc	960
ttattgcac	agggtgcagcc	aggcccgtga	gcctcctggc	aggggtcctc	cacacctggc	1020
tgtcttttgc	cccctgctgg	tcacaggagg	agctgcagca	cctgctggg	ctgcttctca	1080
ggagggtaca	tgaagatccc	aggaccgcca	gctccatgat	aagtggaaag	agctccttgg	1140
agtcaggagc	gggagttgag	gagtttgagt	cctgctctcc	agttataggg	tatgtgactt	1200
gtgtagatca	cctaacccttg	ctcttgattt	ccttacctct	taaactagca	ctaaaagcac	1260
cccacaaact	gtaaagtttag	ttgtgatgat	tgaatgacac	catgggtgtg	gaagctcttt	1320
gtaaagtgc	aaacgggtgtg	cagtttgagg	gtggttaccc	ccagtgcga	ttctcaggag	1380
gcaacatggc	taagggcacg	agctggagtt	aggctgacct	gctgcttcca	gcctgtgag	1440
cttgagcaag	tcaatttaact	tcctgagctg	cagtttcctc	atcagtataa	tgtgataagg	1500
atagggttgt	tgtgaagattt	tattaaatgg	ggtaataaat	gtcaagtatg	tagcccatag	1560
tgagtgtctc	agagtttttt	tcctttgttt	ctttccccc	cgccccgaga	tggagcctta	1620
ctctgttgcc	caggctggag	tgcatgtggc	tgatcttggc	tcactgcaac	ctccgcctcc	1680
cgggttcaag	caatttcctc	gcctcagcct	cccaaatagc	tgggactaca	ggcgtgcacc	1740
accattctcg	gctaattttt	gtatctttag	tagagacggg	gtttcaccat	gttggccagg	1800
ctgggtctga	actcctgacc	tcatgatgct	cctgcctcag	cccccgaaag	ttttgggatt	1860
acaagtgtga	gcccccgctg	cctgccaggt	tttttttttt	tttttttttt	tgtaaaacac	1920
ccacagggtg	ttgctgttgc	ctgggctgga	gtgcggtagt	gcaatcatag	ttcactgcag	1980
ccttgacctc	ctgggctcaa	gtgatcctcc	tgccctcagc	tcctgagtag	ctgggaatac	2040
aggnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	2100
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnttttgta	ttttaagtag	2160
agacagggtt	ttcccaatgt	tgcccaaggc	tggtctaaaa	ctcccaacct	caggtagacc	2220
accacacctc	gcctcccaaa	gtactgggat	tacaggcgtg	agacaccgtg	cccagccagg	2280
aggcttattt	tcttgataaa	ttaccagtc	tcaggatatt	ctctacagcg	atgcaagaac	2340
agcctaatac	atccaggctc	agcatcagtg	gaccaggtg	ggagagctta	agatgtcaag	2400
gtctgaatgc	cgcttccaca	cacctttggg	acctagggac	tcctctctct	ttctcttttt	2460
cagtagaaga	tgttatcttc	tcctttctct	gaccagtagt	tggtgatggt	ttcagagata	2520
gtttttcagt	caagatatat	ttcagtggct	tcactgagcc	caagttccct	cgccctctca	2580
ggactttatt	tcctttgttc	tagaagaggg	ataacacata	ttttctaagg	tggttgtgag	2640
attaaggagg	ctggtaaccg	gtgggtgcata	aggacaggat	agagcaatgg	tgagaccact	2700
caaaaagcga	aaagttgacc	tgcgagggtg	acacttatca	aatcagcaca	cagtgggagc	2760
ggaagggaat	tcctcatca	gttacaatat	ttggagagtg	caagttatag	aaaaccagct	2820
cctggccggg	cgcggtgggt	catgcctata	atcccagcac	tttggggagg	tgaggcagg	2880
ggatcacag	gtcaggagtt	caagaccagc	ctgaccaacg	tggtgaaacc	ccacctctac	2940
taaaaataca	aaattagctg	ggcgtgggtg	tgtgtgcctg	taactcgagc	tactcaggag	3000
gctgaggcac	gagaatcact	tgaaaccggg	aggtggagtt	tgcaagtgag	cgagatcgca	3060
ccactgcact	ccagcctggg	caacagagcg	agactccatc	tcaaacgaaa	aaaaaaaag	3120
aaagaaaacc	cagctctaac	tggtctaaac	agtaagaaga	tctattatat	tatccatctc	3180
aggcagcagc	aagccagag	gtaggggact	ccaaggttgg	ttgatccagg	gcctaacgat	3240
gtcatcaaa	accaggttc	tttctgtctc	ggcacctctg	tctgcagggc	cagcttcac	3300
ctaagccaga	ttgttcttgt	cttgattaca	agtttgctgc	tgggccagca	gacgtgcct	3360
gcctccctgt	tcctcttcag	aagtagaaag	tgccctctcc	ccagtcagtg	aatgaaagag	3420
tttccctttt	gtctgggatt	gcttaggtcc	accacactga	agccaatgac	tgtaaccagg	3480
aaggtaatat	acactgattg	tcttaagtca	gggttctctg	gccagctctg	ggcaaggagt	3540
gtgatactgt	catgattgtc	ttgggctcat	cagggcagct	ctgcagatga	gatcaaaact	3600
caagctacat	tattctgaac	agtgggaagt	aggaaaagaga	cattttggga	gatacaaaac	3660
acaatgtcta	tcccatatcc	ctaggtcag	gtcacagtgt	cttggttggg	catcaaatgt	3720
agaaaaagaa	agactgtcca	tccatttatc	tacctattca	tctggttttt	gatttttttt	3780
aaattttatt	ttaagacatt	ctcactctgt	caccagact	ggagtgcagt	ggtttgatca	3840
tggtcatgg	cagcctcaac	ctcccaggct	caagtgaacc	tcccatgctc	aagtgatcct	3900
cctacctcag	cctcccaagt	agctagaact	aaaggtgcat	gccaccacgc	tcagttaatt	3960
tttgcatttt	ttgtagagat	ggggtttcgt	catgatgccc	atgctagtct	ggaattcctg	4020
aactcaagca	atatgcctgc	ctttgcctcc	caaaatgctg	ggattgtagg	catgagccac	4080
tgctcctggc	tcactctgtt	aataatttat	gaaacaacta	ctgggtgctg	agcacggggc	4140
caggggctgg	agatctagca	gggaccagge	agatctctgc	caagtgcgtg	gtttcttaaa	4200
ggttttgtct	ataattcccc	ttttcttttc	tccttctgtt	ttttcttttt	ctttctttct	4260
ttctttcttt	tttttttttt	gagacagagt	ctcactctgt	tacccaggct	ggagtgcagt	4320
ggtgcgatct	cagctcactg	caacctctgc	ctcctgggtt	caagcgattc	tcctgcctca	4380
gcctcccag	tagctgggac	tcagggcgcc	tgccaccatg	cccggcta	ttttgtgttt	4440
ttagttagaga	ctgggtttca	cacatattgg	caggctggtc	ttgaactcct	gacctgtgta	4500
tcgcgccgct	tcggcctccc	acagtgcctg	gattacaggg	gtgagccacg	gcgccagcc	4560
agtttccctt	ttcaatgagg	cctccctgac	ctccatactc	tactcctcca	cctggccac	4620
tcagctctac	tttttcttcc	ccatagcact	caagacctcc	taacatacta	cgtaagttat	4680
ttatttacta	ggcttactgt	gtattgtctg	tcttctctca	ctagaatgta	aactccatga	4740

gaatagaaat	ttttgccttt	ttatttagtg	tgggtgtctgc	agcccttggc	ttagtccctg	4800
gcatacaaca	gtcactccac	ccacagttgc	tgaataagtg	actaaaggtc	cctgccctca	4860
tattgttatg	agggagtggt	catgttggtta	gagaaaaatc	tgaggcaca	taaaatttta	4920
tagagttaa	gttttctttt	ttaagcaatc	cacgaattgg	ggtagtttca	gaggtagttt	4980
ttcagtcacg	acgtatttca	atggcttcac	tgagcccaag	ttctttcacc	tctctaggac	5040
tttatttcct	tatttctaga	acggggataa	cacatagttc	ataaggcagt	tatgagagta	5100
agggagctgg	tatggggtga	tgcataagga	caggatagag	cagtggtag	accgctcaga	5160
tgacaaagcg	tcagagacca	gtatttaoga	cggaaatgtg	gaagcatgat	aaagaaatta	5220
tttgggctgg	gcacaatgac	tcacaactaa	taaaactttg	ggaggccaag	gtgggaggat	5280
cacttgactt	gcagaaggtc	aaggctgcag	tgagctgtga	ttttgccact	gcactccagc	5340
ctgggtcaaca	gagtgagacc	ctgggtcgaa	acgttatttg	attggttaca	gttatacagt	5400
tgcccttattt	ggctctattcc	atttgaaagt	tcctagttct	ataattttaa	gtttgttggc	5460
tgtttctgat	tgggttaagct	taagttttgt	tttcttttaa	tacagttaag	tgcccataaa	5520
tgacattttg	gtcaaggaca	gaccacatat	acagtgggtg	tcccataaga	ttataatgga	5580
gctgaacatg	ctctattgtc	tatggcgtag	tggctcctgat	gttgtagcgc	aatgcattag	5640
ttatatgttt	gtggcaatgc	tggtgtaaac	acacctactg	cactgccagt	gatataaaag	5700
aatagcacat	acagtatat	atagtacata	atatctgata	atgataatac	ataactatat	5760
tactggttta	tatatttact	atattatttta	tctttatttt	atttttgaga	cagagtctca	5820
ttctgtcacc	caggctggag	tgcaagtggc	cgactcttgg	tcaccgcaac	ctccgcttcc	5880
tgggttcaag	tgattctcct	gcctcagttc	cctgagtagc	tgggattaca	gggtgtgcacc	5940
atgacaccct	gctaataatgt	tttgtatttt	tagtagagat	ggggtttcac	catgttggcc	6000
aggtctggtc	tgaactactg	acctcaagtg	atcaccccg	ctcggcttcc	caaagtgtcg	6060
ggattacagg	cgtgagccac	cacgcatggc	ctatttataa	ttattttaga	gtgtacgct	6120
tatacttata	aaaaaaagct	aactgtcaaa	cagcctcggt	caggtccttc	aacagatatt	6180
ccagaagaca	ttgttatcat	aggagatgac	agctccgtgc	atattattgt	ccctgaaaac	6240
cttctagtg	ggaagtggaa	gacagtgata	ttgatgatag	gaccagtg	aggcctaggc	6300
taatgtgtgt	gtttgtgtct	ttgcttttaa	caagaaagtt	taaaaagtta	aaataaaata	6360
caaaaaattt	taaatagaaa	aaagctgccc	aggaacaatg	gctcacacct	gtaatccacc	6420
cattcgggga	ggccaaggtg	ggtggattgc	ttgagctcag	gagttcaaga	ccagcctggg	6480
caacatgggt	aaaccccatc	tctacaaaaa	atacaaaaa	tagccgggtg	tgggtggcatg	6540
oggtctatagt	tccagcta	cgaggggctg	aggtgggagg	atcactgggg	gggaggtggt	6600
tgaggtgna	gtgagctgtg	attgnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	6660
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	6720
nnnnnatatc	ttaaaaaaat	ttttttttat	ttttgagaca	gaatttctct	cttgttgccc	6780
aggctggagt	gcaatggcgc	tatctcagct	cagggcaacc	tccacctcct	gggttcaagc	6840
gattctcctg	cttagcctc	ccaggtacag	gcgcccgcca	ccatgctcgg	ctaatttttg	6900
tatttttagt	agagatgggg	tttcaccatg	ttgtccaggc	tggctctgaa	atcctgcctc	6960
aggtgatcca	ccccctcgg	cctcccaaa	tgctggaatt	tacaggcgtg	agccactgtg	7020
cctggcctcc	tttacatttt	tttaaattta	attttaattt	tttaattttt	aatttctcat	7080
atataatata	ttttaagact	agccaagtga	agcagtgga	gtggaaaagg	aactggtttt	7140
gatcaatagg	tgtaaacacc	actgcaactg	gaccagccta	ttttacattc	ctgttagcag	7200
tgatgaggg	tcacttttct	tgtagcctca	acaatatgtg	tcgttgccca	tctttttttt	7260
tttttttttt	tttttttttg	agatggagtc	tcactctgtt	gcctaggctg	gaatgcaatg	7320
gcatgatctc	agctcaactg	aacctccgcc	tcccagggtc	aagtgtattc	tgtgtctcag	7380
cctcctgagt	agatgggatt	acaggcgctc	accaccacgc	ccggctaatt	ttttgtattt	7440
tcagttagaga	tgggggtttc	ccatgttggc	caggttgggt	tcgaactcct	gacctcaagt	7500
gtacggccca	cctcgccctc	ccaaagtgtc	gggattacag	gcatgagcca	ccgcgccggg	7560
cctgcccatc	ttttttttgt	tatagccatc	ctagtggatg	taaaagtttt	ttgtgatttt	7620
gatttgtgtt	tccttactga	tcaatgatgt	tgagcatctt	ttcctgtgct	tattggcttt	7680
tggtatatct	ttggagaaa	gtctattcag	gtcctttgcc	cacttttaaa	ttagggttatc	7740
tttctattac	tgagatgtaa	gagttcttta	tgttctagat	ataagtctcc	tacatatgat	7800
ttgtaaaaat	tttcttcca	ttattgggtt	gtctttcact	ttcttttgg	gtccttttagt	7860
gcacaacagt	ttttaatatt	gaagtccaat	tttctatttt	tctcttttgc	cacttgatc	7920
ttggtgtcat	gtttaaggaa	ctattgccta	atctcaggtc	acaaagattt	acacctgtgt	7980
ttccttcttt	ccttccttcc	ttccttccct	ccttcttccc	ctccctccct	ctctccctcc	8040
ctccctctct	ccctccctcc	ctccttccct	tctcctccct	ctccctccct	ccttccctcc	8100
ttccttccct	ccttccctcc	ttccttccct	ccttccctcc	tcttctctgc	ggaatcttgc	8160
tctgtcacc	aggctggagt	gtagtggcac	gatcttggct	cactgcaacc	tctgcctcct	8220
gggttcaagc	aattctcctg	cctcagcctc	ctgagtagct	gggactacag	gcacacacca	8280
ccatgccacg	ctaatttttg	tatttttagt	agagacgggg	tttcaccaca	ttggccagga	8340
tggtttcgat	ctcctgacct	cgtgatccac	ccgcttgggc	ctcccaaaag	gctgggattg	8400
caggtgtgag	ccaccatgcc	cggcctgtgt	tttcttagag	ttttgtagtt	ttagctctta	8460
tagtttagatc	cttgatccat	tttgagttga	ttttgtatat	agtgtgagat	atccacctgg	8520
tgttgtaaat	tgcccagaag	tgggtatgct	tctaaatctg	gctgttaggg	attactagag	8580

gtgaccaaag	tgaatttttt	ctttgtttct	tttttttttt	ggagacagag	tctccgtcac	8640
ccaggctgga	gtgcaatggc	ttcatcttgg	ctcagtgcaa	cctctgcctt	ctggtttcaa	8700
gcagtttccc	tgccctcagac	tccctgagtag	ctgggtattac	aggcgtgtac	caccatgctt	8760
ggctaatttt	tgtattttta	gtaaagatgc	agtttcacct	gttggccagg	cttttctgga	8820
actcccggcc	tcaagtgtac	catctgcctc	tacctcccaa	agtgtctggg	ttacaggtgg	8880
gagccaccgt	gcccagtcct	tttctcagaa	tttatttgtt	tttttttgtt	ttgtttcatt	8940
tttgagatag	ggctctcactc	tgctcagctag	gcaggagttc	agtgggtgtga	tcattgtctgc	9000
agccttgaac	ttctggactc	acgtgatctt	cccacctcag	cctcctgagt	agctaggatt	9060
acagggcatgt	gcttccacac	ctggctaatt	ttttaatttt	ctaggactta	tttgtccatt	9120
cttgcaaaagc	aggggtacaac	atgcctatct	ctacctacct	ctcttccctt	caagggactc	9180
cacgcaaaat	ccttgaggct	ctcgggctga	ctgtgggtgc	tggtgcctga	tctgcctcag	9240
tcattgctgca	tgatcaaaaag	tgctccgtttt	ctgcttcttg	gaactttatt	cactttgggt	9300
gtcagttctt	ctctgcagtg	tcccaagaac	acagaattag	accaggaatc	tgtgttgcca	9360
tagtgtgtgg	aaagaggcag	acttccaaact	ccgctatgtg	ctgttggtgg	attgaagctt	9420
aattttcttt	ctatctttctt	ttcttttctt	ttcttttttt	ttttttggag	atggaatctc	9480
gctctgttgc	ccaggctgga	gtgcagtggt	gcgactctac	ctcactgcaa	cctccgctc	9540
ccagggtcaa	gcgattctcc	tgccctcagcc	tcctgagtag	ctgggattac	aggtgcatgc	9600
caccatgccc	ggctaatttg	tgtaatttta	gtagaaacag	tgtttcacca	tattgtctag	9660
gctgggtctcg	acctcctcac	ctcaggtgat	ccacccgctt	tgccctccca	aagtgtcggg	9720
attacaggcg	tgagccaccg	tgccctggcac	tttaattttt	taatacctca	attaccccat	9780
atggtaaaat	gggactagta	atccatacct	tatagcgctg	ttgtgaaaat	gaaatgaggg	9840
taagcagata	aaatttcaga	ctacggatgg	gattgttact	acattctgaa	cctggccttg	9900
ctgttatctg	ctatgtgacc	ttatcttctc	tggtatctcca	ttctttccaa	gtctataaaa	9960
caaagtggac	aattgtcaac	ctttcttcca	aagagcaatg	atttaaggat	caaagtatgt	10020
catttaacaa	aaataatgaag	agctcaacaa	atgaggaact	cattattatt	attacaatta	10080
ttattatttt	agaaataggg	tcttggtctc	ttgcctaggc	tggaagtccag	tggtataaac	10140
acagctcaat	gcatcttcag	cctcctggat	acaagtgtac	ctcatgtctc	atccccctaa	10200
gtagctggga	ccacaggcat	gtaccaccac	gcacggctaa	ttttttat	tttattttta	10260
ttttttgaga	cagctctgtc	ttgtcgccca	gactggagtg	cagcagcgca	atcaccgctc	10320
actgcaacct	cgcctcctcg	ggttcaagtg	attctgctgc	ctcaacctcc	caagttagctg	10380
ggattacagg	cctgtgccac	catgcccggc	taattttttt	gtatttttgg	taaagacggg	10440
gtttcaccat	gttgcccagg	ctgatctaga	acccctggcc	tcaagtgtac	cccctttctt	10500
ggcctcctaa	agtgcctagga	ttacaggcgt	gagcctctgc	acctggcctc	ggctaatttt	10560
ttattttttg	tagagacagg	ttctcactat	gttgccaggg	ctggctctga	actcctgggc	10620
tcaagtgtac	ttcccacctc	agcctcccaa	agtgtctgaga	ttacagatgt	gagccactgt	10680
gcctggcctg	gaactcat	ttgaagcat	cactagtatc	aactttgggg	ttacctggcc	10740
acatcctctg	acctacctat	aagggatca	cagctaaccg	agcctctgtt	tctcagaatt	10800
taggcagaag	cagttcaatt	tatcacaac	tactctatat	ccagcataag	tgcccaata	10860
aaacaattgc	taaagttctt	taggcattta	ctgtttgtta	gttagatatt	tagtcctcac	10920
tacaaatctg	tgatacaggt	attattttta	tttaacccat	tttatagaag	agaaacctga	10980
agctcagaga	tgctaagtaa	cttgtgcaag	gtcacacagc	tagtaaataa	agggcagagt	11040
aaagattttg	tttccacattg	gactccagaa	cctttctact	gggactcatg	ggaatagtgt	11100
ggatgtccct	gaccttcagt	ggcccagggc	tctcctgggg	gaatccagcc	atagacaaga	11160
caccagcgag	agcccaatcc	taagattttt	tttgtttgtt	tttgagacaa	ggtctcactc	11220
tgtcaccaga	ctggagtgca	gtggcatgat	caatgctcac	tgcaaccttg	atctcccagg	11280
ctcaagcaat	cctcccacct	cagcctcctg	agtagcttgg	actacaggtg	cacaccacca	11340
cacctgacta	attttaaaat	tttattttaat	taattactta	ctattatttt	ttgagacagg	11400
gtatcacttt	gtcacccaag	ctggactgca	atgggtgtgt	ctcagctcat	tgctgctctc	11460
acctcccagg	ttcaagtgtat	cctcccacct	cagcctctgg	agttgcaggg	actgcagggtg	11520
tgccccaacta	tgctcagcta	atgtttttat	tttttgtata	gatggggtct	cactatgttg	11580
ccagggtctag	tctcaaaactc	ttggactcaa	gcgctcctcc	tgtcttggcc	tcccaaagtg	11640
ccgggattac	aggcataaac	caccacaccc	aacccttaag	gtgtttttgc	tgaatgtgac	11700
catgtcagag	gcaggaaaagg	gaagcatcat	gggttagga	aaggaacact	gagcagggag	11760
acaaagaaaa	tggtatcatt	ttgtgagtgt	tcgtgtgtgt	tgatgtgtgt	acaattctca	11820
gagccagcct	ctcaggttgt	tgagaccaca	gtcccattt	ccagatgag	ataatggagc	11880
ctcagagagt	ttctgcagca	cagctagtgg	aattagaatt	tgaacccggc	tcttccagac	11940
tccaggtctg	tcacaacccat	cccaaaccta	gtcatttgca	gtttaccttc	atgattttac	12000
catttccctt	tgccatagct	agtgttattt	acttaataat	tctttttgaa	tcagtctgct	12060
taaaaaaaa	tagcttcatt	ctaaagtgt	atattcttgg	aatatcgggt	ttgctgttac	12120
ccacccccac	acgttatata	tatacatgta	tgtttctaat	acatatatat	gtacgtatat	12180
acgtgtatcg	ttttttgtta	ttttttttgt	tggtgttagt	tttttttaga	tggaagtctc	12240
ctctgtagcc	caggctggag	tgcaagtgtg	tgatttcggc	tactgggaac	ctctgcctcc	12300
tggtgtcaag	cgattctcct	gcctcagcct	ctggagtagc	tggtgattaca	ggcaccaccc	12360
actacacccg	gctaattgttt	gtatttttag	tagagacagg	gtttcaccat	gttggccagg	12420

tgggtcttga	actcctgac	tcaagtgatc	cacctgcttt	ggcttcccaa	agtgtctgga	12480
ttataggtgc	gagctactgc	ggctggccaa	tgtatgtttt	taatacacat	tcaataacg	12540
aataactatg	aaacctgaaa	aactgctcca	tgttacttcc	tgaacccatc	ttgagtgtctc	12600
acatgtctgt	cataccacat	attgggaaac	actgctttcc	ctggcttcca	agcccgagctt	12660
aatcaactgtc	ccatccctatg	cttcgcttta	tttgtctata	aatgttgggg	ttgggggttg	12720
atgccaaaga	ccttttctgt	tgtcattaac	atggacacag	ctctaagagg	tcttggcatc	12780
ttgggctggc	tctcctttta	gttcagaatt	tggattttta	tccaactact	cagagtgtac	12840
aagccttcct	tatgaatgaa	ctcgttggtc	aaactcataa	aaggctgac	gataaaacag	12900
gaatgaatgt	atgaattgac	actaagtcat	tagcatttca	cgggaatgga	ttctccgtta	12960
gtggaagagc	acatgtcctt	tctggcactg	atgtgtgctt	gggaaactta	ctgagctaac	13020
tggcccatgt	aacacagagg	ccctttgggt	cagtggaaaa	ctgttgactt	tggagattat	13080
cttgagtttg	aatctgagcc	tgccctgtaag	aagctggcta	actgaattgc	tttgcttctt	13140
ggacccttac	catttataaa	atggggacca	ttgtactcac	cctttagggt	tattgcatgg	13200
attaaatggg	attctctata	gaaaaatattg	gcacaaagta	gggtgaaatt	tgacagctag	13260
tgggattgtt	tgtgagggaa	attgtcattt	gattatcaaa	gacttaggag	caggaaacag	13320
gtctaatttc	gggactgcaa	atggaaatgc	cagctgaggc	caggcatttg	ctaataattg	13380
ggtaaaagcag	ggcaggtgta	gaatagcaat	gtctgggaat	taaaagagag	gtgaggacgt	13440
gtatgacctt	gagaaggcaa	gccctggcaa	aaggggatgg	cctccactca	gctacagtca	13500
tgccctagatc	ttctaacctt	ttatttttat	ttttattttt	tgagacggag	ccttgctctg	13560
tcacccagggc	tggagtgcag	tggcgcgac	tgggctcact	gcaagctccg	cctcccggtt	13620
tcacgcccatt	ctcctgectc	agcctcccaa	gtagctggga	ctacggggcg	ccaccacctt	13680
gcccggctaa	tttttttttt	gtatttttag	tagagatggg	gtttcacctg	gttagccagg	13740
atggtctctga	tctcctgact	ttgtgattta	ccctccttgg	cctcccaaag	tgctgggatt	13800
acaggcttga	gccaccgcac	ctggccgac	ttctaacttt	ttaaagagaa	gcaagacatc	13860
tggatttttt	tgtgataact	cctgatttta	aactggcacc	caattataat	ttacaacact	13920
ataaagggtca	acattggccag	cagagcaaaa	catgggtggg	ggcaactgct	ggtcacccgt	13980
gtgcagcctc	tggctctaaa	tcactcttct	atctctctct	gctttacgca	ttgtcccagc	14040
acagtgtctgt	tgtatagtaa	atatccagta	agtgggtgta	gaatgaataa	accaatgcag	14100
ataaacctgt	agagaggccg	ggcacagttg	ctcatgtctg	taatctcagc	acnnnnnnnn	14160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	14280
gggtagatca	cctgaggtca	ggagttcaag	accagcctgg	ccaatatagt	gaaaccccg	14340
ctctacaaaa	ataaaaaaat	tatctgggca	tgattgcagg	tgccctctaa	cccagctact	14400
cgggaggctg	aggccggaga	attgcttgaa	cctgggaggc	ggaggttgta	gtgagccgag	14460
atcatgccat	tgcactccag	cctaggtgac	ggagcaagat	tctgtctcaa	aaaaaaaaaa	14520
aaaaaaaag	aaaaaaagaa	agaaaaagaa	acaatgaatg	agtgtgaggc	tcattggtagt	14580
attggttctt	gagagtagcc	aaccttattg	gtcatcccag	ccacgaagtg	aaatggtacc	14640
cctggcttgg	gccaatgaat	gaggaagaat	aatggcaaat	gggggtctat	gcctccaccc	14700
tccaccactc	gggaggtctc	aagcttgaaa	tccagtgaac	aggttttttag	gtcctggacc	14760
tggccagtc	tccctacagtc	aagtagataa	gtggagggtt	tggctccgtt	ggctacggag	14820
atagtgatca	aggccgttac	tctgcaatca	gactcagaaa	tggcctctca	gttacttctc	14880
catittgtgg	tcttttgtaa	gagcagagaa	gaggaaggaa	tttaggtctt	ctcacccctt	14940
gggctgectg	tccctgctcc	ctgagccatg	gagggctggg	gtggaatatg	gggaataaat	15000
ctgtactttt	tttttttttt	ttttttgaga	cagagtctcg	ctccgtcgcc	caggctggag	15060
tgcctgtggc	tgatctctgc	tcacagcagc	atctgcctcc	cgggttcaag	ttattcttcc	15120
acctcagctc	cctgagtagc	tgggattaca	gggtgccacc	accacgccc	gctaattttt	15180
gtatttttag	tagagacagg	gtttcactgt	gttgggcagg	ctggtctcaa	atacctgacc	15240
tcagggtgatc	caccgcgaca	tgccctccaa	agtgtgggaa	ttacaggcat	gagccaccgt	15300
gcccggctct	accaatctgc	acattttaat	tgacaagggt	caccctccac	tcattgtgcca	15360
ggcatagtct	tgagaagcat	cccacaagga	tgccctctgag	ttcacccctga	caagtccact	15420
agctcttggc	agagacatct	ggcaaatcca	aggcttgaga	catgctggcc	tctctttaaa	15480
gtgcagcaaa	ttttgtctag	agcttgggtca	gttaaaattt	tgtatgtttg	ttttgcatca	15540
atttcaattt	ttaagaaatg	ttgcattaaa	atgttatatta	tcttgaatag	taaatttctt	15600
agtgtccctt	taattttctta	gtgtgtctga	gttgagagcc	tcccctgcct	gattctagtc	15660
cagaccctgg	ggtgacagaa	gactgggtgg	agatgggagg	tgaggagggg	agtgttgggt	15720
ggagaggatg	atctacagag	tgcctggagag	actctgtatg	gagcttttca	tgctgcctgt	15780
ttgccagccc	tgaagctatg	ccttgagggt	gggcaagggt	gcataatccta	gatcagagat	15840
cctcaactgg	ggccattttt	ctccccagag	gacatttgga	aacatgtgga	gacatttttg	15900
atcatctcg	gggttgggga	gaggggctac	tgacatctgg	tgagttagaga	ccagagggac	15960
cattaaactt	tctacaacgc	ccagggacagc	ccctccacaa	taaagagtta	tttgacctca	16020
catattaata	gcacaaagtt	gaggaacctt	gatctagatc	cacagcacag	aagaaaggat	16080
gtagattttt	cacacattaa	agatgagaaa	gcttgtgcct	gtaatccctg	tgactcagga	16140
ggctgtggca	ggaggattac	ttgagccagc	gaattcaggg	ttacagtga	ctatcatcgc	16200
agcactgcac	tccagcctgg	gtgacagagc	aagattttgt	ctcttaaaaa	aaaaaaagat	16260

gaggacaggg	acagtggttc	atgcctgtaa	tcccagcatt	ttgggaggcc	gaagtgggtg	16320
gatcacgagg	tcaggagttc	aagaccagcc	tggccagcat	agtgaacc	catctctact	16380
aaaaatacaa	aaaattagcc	agctacttgg	gaggtgagg	caggagaagc	gcttgaaccc	16440
gggaggtgga	gcttgacgtg	agccaaaatc	ttgccattgc	actccagcct	ggcgagacaga	16500
gcaagactcc	gtctcaaaaa	gaaaaaaa	aaagatgaga	aagaggaagg	gagagaaaaa	16560
agagagagag	gaaagaaa	gagaagggtt	tggagtcaaa	aagacttaga	aattccagtt	16620
cttccacttc	ccatggaacc	ttggcaagtt	gccttctctc	tttctctgaa	tctcacattt	16680
tgctctctgt	aagtaggggt	ggtacctggt	ggagatgatg	cggagatgag	ggtgaggggt	16740
gtgttgacac	ctatgcccc	aggatgggtg	agagcttggg	agcactgaac	ctccctttcc	16800
cctcttgttt	cttcccccca	ttgtctccca	ccagctccct	gggatctcca	cttcaactctc	16860
tgggattcca	ccagcaggag	gctactcctg	gagttaaggc	gtgttggtta	gactggggca	16920
ttttaggggg	cataaataat	aattatgcct	ggacaatgga	cataacatct	agggccttct	16980
gaagcaaaac	aggtgtggg	gtacccaaac	aaggcagtag	gccccaggag	gcagggtccct	17040
gcagctccag	cagagagcag	ggcacagggt	tgagaagact	gagcaaatct	cattatcagc	17100
tcctttgtcc	ccactctgt	cctggagcaa	tcattctggc	ctctccccc	ttccccaaaa	17160
accaggtata	aaggctgctt	ctggccccctg	aagccagagg	cactgagagt	ggaggtctca	17220
gactcttggg	aggtgagttc	ttttctgggt	gcccaggcag	gaccagtgtg	ggccctggga	17280
agaagcagca	cctcataggg	caaacacgta	ggaggcctgt	ccttaggaac	atcatagcta	17340
agcagactcg	tccccgcagg	ggcaggagtc	tgggctaagg	gtgatactgg	agagcagcaa	17400
cggagactgg	aagacaaaatg	aaatttggta	cctgagttat	ccctccacc	attccttttc	17460
tagactctcc	agctcagggt	ctgttcatgg	caagaggaga	aagcaatctt	gtttgctctt	17520
taatcaaaaca	attaacaaa	tattccctct	atactatgtg	ccaggggcta	tactagacac	17580
acaagacag	ccccagaag	gacggtggag	tagtgcctc	gctaaaagac	agtagatatg	17640
caatgcctct	tgctcctgcc	ctttctctg	ctgggaacag	tttctgctct	tcatctgggt	17700
aagtctctcc	cttccctcct	catgcgtctt	tccctttttt	ccttttctct	acactccctc	17760
ccccccgctt	ttatttgcac	tcatgaggcc	aggaccacag	ccttccctct	ttagctgata	17820
cagctcatct	ccgttaagat	atcacttgga	ctcagaactg	taacctggaa	ctttctcttt	17880
tttgtttgat	ttttttttgt	tggtgtgtgt	tttgtttttt	ttttgtttg	ttttttgttt	17940
tgttttgaga	cggagtctcg	ctctgttgcc	caggctggag	tgagtggtcg	cgatctcggc	18000
tcaccacaaa	ctccgcctcc	cgggttcaag	caattctctc	gcctcagcct	cctgagtagc	18060
tgggactaca	ggcacatgcc	accacgectg	gctaactctt	gtatttttag	tagagatggg	18120
gtttcaccaat	atttgccagg	ctggtctcaa	actcctaacc	ttgtgattcg	cccgccccgg	18180
cctcccaaag	tgctgggatt	acaggcgtga	gccaccgcac	ccggcaaaact	gtaacctgaa	18240
cttctcagaag	gaaaaaccac	ccacctgtta	agatgaaggg	ctggtgactg	ccccaggctt	18300
ctcacacgtg	ctttctccca	ccttcaaaac	acacactcgt	ggtgtcggcc	agaagtcagg	18360
ttcttgtcca	tttgtgggtg	tgacccgaga	gatctctcct	tacctaacac	caaggaaatc	18420
ctccagtcct	gtcttcaggt	ggaattccta	ggaagactcg	agcgacgttg	ctggagctgt	18480
ccacggtgct	ggaactagga	agctcttgac	ctgatggcag	gttacctctt	cttcccagag	18540
aatgatgccc	cccatctgga	gagcctagag	acacaggcag	acctaggcca	ggatctggat	18600
agttcaaaag	agcaggagag	agacttggct	ctgacggagg	aggtgattca	ggcagaggga	18660
gaggagggtca	aggcttctgc	ctgtcaagac	aactttgagg	atgagggaagc	catggagtcg	18720
gacccagctg	ccttagacaa	ggacttccag	tgccccaggg	aagaagacat	tggtgaagtg	18780
cagggaagtc	caaggtgcaa	gatctgccgc	tacctattgg	tgcggaactcc	taaaactttt	18840
gcagaagctc	aggtaaagtag	tagggaggct	actgcggagg	acctggggga	aaagagagta	18900
cattcagctc	tctgttccct	attcatattag	gctagtgggt	ctcaaaagcct	cgcatgcata	18960
agaatcacct	ggagtgtgtg	ttaaaacaca	gctttctggg	cctcaactgc	acgacttctg	19020
atttaggagg	gctgaggtga	agcotgagaa	tttgcattta	caacaaatcc	ccaggtagtg	19080
atgatattgt	tggtctgggg	agaaccaccg	atttaacaaa	aaggcttttg	tgtagaaac	19140
gcctgtgtta	aattctggtt	ctgcctttta	ttagctgtgt	tacctgggca	agttgctttg	19200
cctttcaaaag	ctttagcacc	ttcatttgta	aaacgaagat	atatagcacc	aacttcttag	19260
agttgtgtgtg	agcattaaat	gagataatac	atgaaaagtg	tttggaaatg	tactgggct	19320
gtaataaaact	ctcaataaagc	ggtggttata	attattatga	gtattatcat	ttcctgtagg	19380
attgtcctga	cagctaatta	agaagcaaaa	gataggatta	agggaggcaa	gtaggtttat	19440
ttttaacctg	aaaagggatg	ccgggctcct	gcctggagac	tcagaaactt	gaaataaatg	19500
agagggaatt	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	19560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	19620
gttagcacat	agccagaaca	tctagaaggg	gtggtaggag	tggggattag	aggttccagc	19680
tggaggcaat	ggcacttgca	aaggctttgt	tgaagtggcg	taagtgtgga	ggtggagcat	19740
tcaggaaaag	agagcttcag	cttcagtgtg	gctggagtgc	tgggtgtgaa	gagaggtgaa	19800
gatgaggctt	ggaggctggg	cagattttgc	tccaaaagag	cttgggtgaa	tgtgataagg	19860
agtttggtt	ttctcctact	aaggacaaca	gcaaaactatt	gaagagttaa	aatcggttcag	19920
tgacaatgac	acgtttgcgt	tttgggtggt	cactcgagct	gccagccagg	tagacagtgg	19980
cagaagatgg	aagataaagc	actaaagggt	gatgaggcag	gaagccagtg	aggagagaaa	20040
ggggacgatg	tgagtgcacg	taaatcattt	gttgggttgc	tattgtgtgc	taagctctgt	20100

gctaaattct	tcacgtgtat	tatttcagct	aatccatcta	acaactctgt	aaggcaggta	20160
caatcggtcc	cagctgaaga	agctgaggct	ctcaaaagct	agtaacttgc	ctaagttcat	20220
gcagcatgca	agttgtccag	ccaggattct	aacttagaca	ccagaggcca	cttttaacca	20280
ctgctctagg	actgggggaa	atgggtcccta	gtgagatag	tgtcgagttt	catatttcat	20340
tcaacaatat	tgttgccctg	ctacatgtga	agagctgtgg	aaagcgccca	aagtgaagta	20400
gatccctatg	agcaagtggg	atgggggtgg	agtggacagt	aggagggtcg	gaacacacat	20460
aaaagggtat	aagaaataac	aattaggccg	gccagggtg	gtggctcacg	cttttaatcc	20520
cagcactttg	ggaggcccgag	gagggtggat	cacttgaggc	caggagtttg	agaccagccc	20580
ggccaacatg	gtaaaacccc	atctctacta	aaaatgcaaa	aattagctgg	gctgggtggg	20640
cacgcctgta	atcccagcta	cttgggaggc	tgaggcacga	gaatcacttg	aaccaggag	20700
gcagagggtta	cagtgaactg	agattgcacc	actctactcc	agcctgggag	acagagtgtg	20760
accctgtctc	aaaaaaagaa	aacaaaacaa	gtaggtaact	tctgccatag	ggaggattca	20820
taaagtgtca	gtcctcaggt	gcatttttgc	ttatcagttt	taaaaatcag	agaatgtctc	20880
aaagaattag	gagtgtcagct	tcttttgaaa	atttgggcca	gaagcgggtg	ctcacgcctg	20940
taatcccagc	actttggggag	gctgagggtg	gtagatcacc	cgaggctcagg	agttggagac	21000
cagcctgacc	aacatggcga	aaccccgtat	ctactaaaaa	tacaaaaatt	agctgggctg	21060
gtggtgcatg	cctttagttc	cagctactca	ggatgctgag	gcatgagaat	cacttgaacc	21120
cgggaggcag	gggttacagt	gaaatgagat	tgcaccactg	cactctagcc	tgggagacag	21180
agcaagacc	tgcctcgaaa	aaaagaaaaa	gaaaattttg	aagatctgac	aacagttgac	21240
ctgcattcct	gctcggcaac	agcctgattg	tggatgggga	gaggctcagt	tgtctgccaa	21300
acctcccatc	ctgatgtct	tccctcgctg	tcacatctg	cttgacatgt	aggcatttgg	21360
tgtgtgcctt	ctgctctggg	tgcccagatg	aattggatgc	tatatgagaa	aacattctgt	21420
aaatgtcttg	tggtaggcaa	cctcaaagat	cactggggcc	tccaatgate	cctccttctt	21480
ggatattcat	cctgtgtata	atcctctccc	ttgagtgtgt	actacacctg	gatacttgtt	21540
tctaataaac	agaacacagc	aagggtaatg	ggatgctact	tctaagggtta	aattacaaga	21600
gtgtaaaagt	tgtcttgttt	gtttccctct	cttgatcttc	ctctcattct	ctctctctcc	21660
ctctctctca	ctttcttact	gtcttgtcct	tccctttgtt	tactctgatg	aagcaagcta	21720
gcaagcatcc	atgttgtgag	ctgacctatg	aagaggccca	tgtgggtggt	aggaaactgag	21780
tgccagcctc	accagcaag	gaactgagtc	actcatcata	tgggtgagct	tggagacaaa	21840
tccctcccca	cttgagcttt	cagatgacgg	cagccctggc	tgatgctttg	caggcttgtg	21900
agagaccctg	agacagaaca	ctcagctaag	ctatacccta	tctctcgaga	tagagtataa	21960
tacatgtagt	tttaagctac	tatgttttgg	gataatttgt	tactcagcaa	tagataacca	22020
atacatatac	catgtacata	actgtttcag	ttgtctgaga	ctatatattag	tcattttaca	22080
cctacatcaa	gaatgtgtca	ggcaccattc	caggtacttg	gaatacatca	attaacagaa	22140
taggtaaaga	ggccaggcat	agggtcacca	tctataatcc	cagcactttg	ggaggccag	22200
gtggggagac	tgtctgagcc	caggagttga	gaccagcctg	ggtaaaaatg	tgagacactg	22260
tctcaactaa	aaaaaaaaaa	aattagttgg	gcacagtggc	acatgcctgt	ggtgccagct	22320
gctcaggagg	ctgaggtggg	aagatcgctt	gagcccagga	gtttgaagct	ccagtgaacc	22380
acggtcacaa	aactgcactc	tagcctgagc	aacagaaaaa	gaccctgtct	caattaaaaa	22440
aaaaaataaa	aaagggaag	aaagaaaaaa	ataggtaaag	atccttgatt	cttgccctct	22500
tggaaactct	attctagagg	gggatgggtt	ttcacagtag	aagtctgtgt	tgacagcgct	22560
gtttaaagct	ccttcagcat	ctggggaaaa	ggttnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22620
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	22680
nnnnnnnnnn	nnnnattttt	tagagatagg	gtcttgctat	gttgcccacc	aggctggtct	22740
tgaactcctg	ggctcaagca	atcctcctgc	ctcagcctcc	tgagttagctg	ggaatacagg	22800
tgtgcaccac	catgcctggc	ttatttcata	tatatatatt	tttatatata	tgtatatatta	22860
tatatataaa	tatatatata	atttctgtat	ataaataaat	aaatatatat	atatatatat	22920
ttttagagat	agggtcctgc	tatgttgacc	accaggtctt	gaactcctgg	gctcaagtga	22980
tcctcctacc	tctgcctttc	aaagtgttgg	gattacaggc	gtgagccatg	gcacctaaact	23040
gagttatttt	taccacacga	agcataggac	atacatccaa	aaatgttctg	agctgagcaa	23100
gagcctggag	gcaagtgaat	ctgaactttc	ccgtctttga	agaaaccagt	ctctctccaa	23160
agtcacatag	ttagtgtcac	tccccccaag	aactgcatga	gctgggacaa	tcagagggca	23220
gtggaaggte	tggggctcag	gggcgcccc	tgtgtctctc	ccagggtctg	tccccctacg	23280
caagagcctc	tgtctcccca	ctttcctgtg	gagcctcctc	accatgggca	tgaccagcct	23340
goggatcatc	ttctacatgg	ctgctgtgaa	caagatgctg	gagtaccttg	tgactggtgg	23400
ccaggagcat	ggtgaggcac	cgctgaggcc	cctgggggtt	gggggcacag	gcgggtcacc	23460
ctggctgagc	tccccctacc	atacgtttcc	ctaccacag	agacaaatga	acagcaacaa	23520
aagggtggcag	agacaggtag	ggctatgaaa	gcaggggcct	ggctcacgcc	caccccactg	23580
caaccgcctt	ctcagggggc	gggactcctc	taggcctggg	cccacccagg	taaccctttt	23640
gtgggatgta	agagtctggg	ttcagaggaa	ggctattttg	gtgctctctg	gcctccgctg	23700
gaaggggtga	tagtgtccac	ttcctgaccc	cactgccttt	cactgccttt	ccatcctgac	23760
ccagttgggt	tctactcctc	cgctctcggg	gccatgcagc	tgttgtgcct	tctcacctgc	23820
cccctcattg	gctacatcat	ggactggcgg	atcaaggact	gctgtgacgc	cccaactcag	23880
ggcactgtcc	tcggagatgc	cagggtgacct	gcctgtacag	ggatgggtgac	agcaagtggg	23940

caggcagtg	ttttcattt	ctctgtgcgt	ttacatccag	cagcttgttg	ctttctccca	24000
agaaccctg	gagatcaggg	gtaccctccc	attttacaga	tgaggaaact	gaggctagga	24060
agggacctgg	cttgcttaat	aataagaata	gctaatgcag	agtgcctgact	gtgcacttgg	24120
caccttgcc	tgtttagtcc	tacaacacct	ctttgaggta	gatgcgttaa	tatcttcatt	24180
ttgcagttga	ggaaaccgag	gtacaggggt	gcacagttag	gtcattccacc	caagatcaca	24240
cagctttcag	tggcagcctc	cagaacctgt	gttataagg	tacacgctaa	agtcttgtaa	24300
gggctagaat	aggtagagtt	ggtatattag	atatttattg	ctgtataaca	aatcacccca	24360
aggcttggca	ttttaaaaca	acaaacactt	ctcatctcat	acagtttctg	acagtcagaa	24420
atcaggggaga	gactcagccg	gctgattctg	agtcacagtc	tctcatgaag	acatagtcag	24480
gctgtcagcc	agggctgcag	tcctctgaag	ggctgactgg	ggttggagaa	tctatgtcag	24540
ttcaattacc	cccatggcct	ctccatagg	ctgctcagga	cacagcacct	gctttccctt	24600
gagcaagagg	gctaagcgac	agagacccc	tatcttctct	cacataatct	cagacgtagc	24660
ataccatcac	ttctgttacg	ttctattata	ggcacagagc	aacctgata	tactgtggaa	24720
ggagactgga	caaagcaggg	gaataccagg	agggcaggatc	cttgagggct	gtcttgttgg	24780
ctggagacca	ccattgaggg	tttttttttt	tttttttatt	gagacagtct	tgctctgtcg	24840
cccaggtctg	agtgcagtg	cacgatctca	gctcactgca	acctctgcct	cccaggttca	24900
agcgattctc	ctgcctcagc	ctcccagta	gctgggattc	accatggagt	cttgaaccca	24960
gattctgtga	ctgcttttgc	tctttttgtg	ttcatccaaa	cagtcctctg	ttatccaaag	25020
aggatgggag	aaagagactg	ggagagaagg	aaatccagtg	gcctccctcc	ctgctagcag	25080
agcctggccc	tggeactgag	ccttccctct	ctacctctg	ctcctaattg	tgagggtccc	25140
ctagcagggc	ccctctgtcc	aggacacatg	ggcgcctgt	cctcacccca	gcctactgac	25200
ctctctcctg	ggctggcctc	agtgcccttg	attgtgccc	agagaggaag	cgctggacag	25260
tcaggccaag	ctgctgtccc	caggagggca	tctgcttatg	tctagggcag	ggacaccttc	25320
ctgtagagct	acggtgtgag	acgttccac	gtcttccac	ttccacacct	ccttcccttc	25380
cttggttctc	aaaccttcaa	gtgtgcatga	gaatcactta	gtgggggata	tttgtccaaa	25440
tgcatgattg	cagatatccc	cgctgagatt	ctgagggccg	agatgaggcc	tgtgaatctg	25500
catgtttaaga	aagcacccgc	tttgatgcgt	gtgtcattgg	gtaggggagc	aacactttga	25560
gaaacatgga	gctagagaac	gtgggtttct	atgggtttcc	catagaaaca	tggattttctg	25620
tggtttctgc	tgccctgaca	tcgaaggcac	atctgaagg	ggaggggcca	ggccaagaac	25680
caggagtgct	tggaacgta	gaggcagcag	ccagtgaact	ccgtactcc	tcagggacgg	25740
gggtgttacc	aaatccatca	gaccaagcta	ctgcaagatc	caaaagctca	ccaatgccat	25800
cagtgccctc	acctgacca	acctgctgct	tgtgggtttt	ggcatcacct	gtctcatcaa	25860
caacttacac	ctccaggtac	ccaccttcat	ccttcccttc	tccttgcttc	ccgaggctcc	25920
tcacaaggga	tggtccatcc	agcacctgcc	ttccaggaag	cgagttctg	gtcttctgat	25980
ctggatctat	tttccgggtt	ctccaggaag	tggttctagt	agattgggtt	ggcgaggggg	26040
tggaattga	ggcccagttg	gcctcttccg	cctacccctc	cttctccag	cctccacaca	26100
ctctcctaac	ctcttcactc	tctctttttg	gttttagttt	gtgacctttg	tcttgccacac	26160
cattgttctga	ggtttcttcc	actcagcctg	tgggagtctc	tatgctgcag	tgtgagtctg	26220
ttgggctgaa	atgccttctc	gagctttgca	accgtgatca	gagaacccca	gggaagggtt	26280
gggagggccc	caggcatccc	ctaattgcacc	tctctctgag	acctctgat	ggcagggagc	26340
tcaattcctt	aaaggcagcc	tatcctgctg	taattgactc	ccctgtttgg	agtccttccct	26400
tagagggaagc	tgaataacct	ggcttgatga	cactttggtt	ctatgtctgc	tggttgaaac	26460
ggcccccaga	atggcctccc	ctccatgccc	acctgaaga	aatttcccaa	gggcagccat	26520
ttgccttata	attttcctct	tcatgttgga	cagtccccac	ttgcatctct	ctcctgggtt	26580
ccctgtctgg	gcgtgctga	gggactctcc	cctgtgtatg	tgatggagta	acaggacatt	26640
acaataatga	tgacaaaatg	acaaccatta	tcaagtgtct	cgttgggtga	ggcagcaggg	26700
aggatccttg	accatcactc	cctgagttca	gcctcactgc	agcgggtctg	gcagagggca	26760
gctctcttcc	cttcatctgc	tcaaggccaga	accttgagtg	ttccttgatg	tttctctccc	26820
tcacactcca	tgttcactcc	atcctcagta	cagccagcag	cagcttctac	acacccccaa	26880
tctgaccctt	ctgtgcacct	ccactgtctg	ctctccagtc	ctagccacca	acatctctag	26940
cctggattat	tgtggcagcc	tttagtctcc	cacatctgcc	ctggccccgc	tgtctcagtc	27000
tatttttaac	acaggggctg	cagtcaacct	tcaggacata	agtctcttca	catcaactctg	27060
tggtgtcctg	tctcatctgt	ctcagagtaa	aagccaaagg	ctttactatg	gcctaaaaag	27120
ccttgcaagc	tctggcccca	gcacttcaact	ccccctctagc	tccccctcct	ccattgttca	27180
ctctgccaca	gccacagtgc	ttcctagtgc	tccggaagtc	tcaagtgtgt	ttcctgcttg	27240
gcactcttgc	atgtactagt	ccctgtttct	agaacattct	tctccagata	tctgcaaggt	27300
gcccactctt	accttctctc	cttcttcagg	tcttccctg	actgtcctct	tctcagtgag	27360
gcctcccttg	gctgtcccat	gtacaattgc	aacctcccta	ctgcccgtct	ctctgcttgg	27420
tttttctcag	cgtttatcac	taacactctg	cctatctctt	gcttattgtc	tgaccgccac	27480
ctgctccatg	ggaatgccac	ctcctcgatg	gcaggaatct	gttgacttgc	ttgatcgtgg	27540
tatctccagc	acctagagca	gtgcttgcca	catagtaggt	tctcagctaa	atgtttgttg	27600
acagaataca	gtggacagtc	ctgcgaggtc	aatgccatcc	ctgttattag	tggagggaagt	27660
ggggctcagg	gagtttgagc	cacttgccaa	tatcacacat	acaggaggtg	tgagaaccca	27720
gctcagtggc	cctgaagttg	gagcatttgc	cctcaaggct	ggggacccaa	gagcccatgc	27780

aaagagcccg	aacgcttaag	caccaccctg	cctggccagc	ggggnnnnnn	nnnnnnnnnn	27840
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27960
accacaatta	ctttttgcacc	cacataaata	gttaccatgg	gctgagcatg	gtggctcagg	28020
cctgcaatcc	cagcactttg	ggaggtgag	ccaggcggat	cacttgaggc	caggagtcca	28080
agaccagcct	ggccaaacatg	gtgaaacccc	gtctccacta	aaaaatacaa	aaattagctg	28140
ggtgtggttg	cgctgcctg	taatcccgag	tattcaggag	gcagagggtg	cagttcactg	28200
aaatcatgcc	actgcactcc	agcctgggag	acagaaatgag	actctgtctc	aaaaataaat	28260
aaataaataa	ataaatattt	accatgtttt	gaccacctgt	tatgtgccaa	ctgtattact	28320
taaaaaacacc	catgggaggc	tgggcacagt	ggctcacgcc	tgtaatcgga	cactttggaa	28380
gggcaagcgg	ggaggatccc	ttaaggccag	gagttcaaaa	ccagcctagg	taacacagta	28440
agccctgtct	ctacaaaaaa	taaaaaaatt	aactgggcat	ggtggtgtgt	gcctgtaacc	28500
ccagctcctc	gggagggcaga	gggagagggt	cgcttgagcc	cagcagtttt	aggttgagct	28560
gagccaggac	caagacacta	cactccagcc	tgagtgcag	agcaagacac	tgccctctaaa	28620
caaaacaaaca	aacaaaagcg	acctgtgggt	aggtaggaac	aggctcatag	tacagatgag	28680
aaagcagagc	ttggagggct	caagcgattt	gccaaagcaga	ggtccaagcc	gaggtctctc	28740
tgaatccaaa	tctaatccg	tctatcatat	caccacagcc	ctctctgccc	caggagagat	28800
ctctgccccac	tccagccact	cacgtgtaat	tgacttcctc	aggggcagga	aaggcttcga	28860
tgggccagtt	gaggtgagc	ttcagaaaga	taaggcaggc	caggccagac	caggtgaaca	28920
tgatgaccac	gaaggccaca	ccggcatcgt	agatcagctg	tgagaggagg	gggcaggccc	28980
gtgggggaga	ctgcctggcc	ccagacccca	ccaaggtaga	tcccaggcct	cagaggcctt	29040
aaagaagttc	tctctcccc	ttgtccttgt	gcccaatttg	cagatgagga	aaccaagacc	29100
agaagttag	atccagactc	agaagaccga	tcattccttt	ttctttttca	cttgaggccc	29160
cctagagagc	tatgaaatag	tctccacaaa	gctgaagtt	gctggccact	ggctcaaaaat	29220
atctctgaaa	tttccattat	cttaaaaaaa	tacatacatt	tttgccctatg	actccacaaa	29280
cattcatggt	catgttcgca	caaaaatgtc	catttcagag	tacgtacaaa	ggaaacttag	29340
tgctctaggt	ttaccgggccc	taatcggtgt	tatcctgccc	cttctctggca	cattcccccag	29400
gggaaaaggc	aaaccagac	tgctcatgct	cagccttttc	tcaccttttc	caggctcctc	29460
cacgtgcaac	aactgggggg	gttggggaga	gggaggtgca	agtgtctctc	ccaagggctc	29520
tcaaccccag	ggcaggtaag	ttctcaattg	aatgagattc	tgtgcaaatg	tgtcagccct	29580
tcttatggaa	gaagctgatg	caccatctgt	cctcttgtcc	tccccatacc	atctgaccag	29640
gataattaat	gtctgtctct	ccctcaggct	cctgtcctca	cctttttctc	tgagctcttg	29700
gaccttggtg	ccttttctct	cctaggggca	ggaagagagc	tcaaaggggc	acaccccaca	29760
atgtgtggag	gtaagatctg	gctcttcaaa	cactacttca	gttgaaaaga	agggagaaact	29820
gccacccctc	catgcctgcc	caccagaaaca	actgatggcc	ccccaccaca	tgccgtctct	29880
caaaactcct	tgagacact	gagcaaaagt	accttcttta	gtactctttg	taaaagtcaa	29940
aacgggtatgc	agtgttggtac	tgcccaccgt	ggaggttgag	gagcatggca	tggtctcaag	30000
ggtcctttga	tatttgacag	aggaaattga	ggccccatc	ttgcaactgag	ctaaaacttt	30060
ggtccctctg	cttcgaggtg	caccaggttg	acctgtccag	gatccagcct	ggcataaaact	30120
cacttttgta	ccttggaaca	aaccacccat	cctctctgga	aggtgtggaa	aaatgtggcc	30180
ccaaaggctg	aataaagcca	gagagtcagg	gaccttgaac	gcatgtgaag	gggctggact	30240
tgattctgta	ggtgaagcta	aaccactgaa	ggtttttcag	cagtgtgtga	gccagttccc	30300
catctgagat	cttcttgga	gtcacgtgag	tgacagagta	cagagaaaaa	gaatcagagg	30360
cagggagacc	agctgagaaa	gcttgctgtg	gccagggaga	gaggggggag	gcctgcattg	30420
ggatgatgac	agagaaagga	gagcggagaa	gtcagaccgg	tggttcagca	ctagctgtctg	30480
ctcactcgcc	cccacccggt	tcttgtgtca	agacaaaaag	aaaaccagg	tgccctcata	30540
ccttgattcc	tggaacgta	atggcagaag	aggcgtaaga	gccaatcatg	agggccatta	30600
acgtggagcg	caggttccca	aacatgttgg	gcagctgagg	agggaaagca	gcacccatga	30660
ggtggggaca	ccgtgaccct	tgcccagcat	tcccagccct	gctccatata	atagctccag	30720
gagacgcaga	agaaaagccc	caaggtaaaa	caaacagaaa	aatcaatgtg	ggaaactgta	30780
ctctgcccc	tgccatacaca	gtcacagtgc	ccttttagctt	caaaaaggct	cccagacacc	30840
cctcagagag	acattttgtt	aattttgttt	aattccaggt	ttcccaagtt	tgttacgtaa	30900
cacctctgaa	aaacacatgg	aataggtgct	taagaaacac	tgatcttggt	tggtgcagct	30960
ggctcatgcc	tgtaatccca	gcactttggg	aagccgaagc	tggtgggag	cttgagggtca	31020
ggagtccaag	accagcctgg	acaacatggt	gaaaccccat	ctccaccaa	aatacaaaaa	31080
ttagctaggc	atggtggcat	gcgcctgtaa	tcccacctac	tccagaggct	gaggcaggag	31140
aatcgcttga	acctgggagg	tgagggtg	agtgagccga	gatcgaccca	ctgcacttta	31200
gctgggtga	cagagcgaga	ctatgtcccc	accccccaaa	aaaaaagaaa	agaaaagaaa	31260
gaaacagtga	tcttgtccaa	ccattttgag	atgagacaa	tgagacccag	ggaggaaaaag	31320
tgtactcaag	ttcacagagc	acattaatgg	ctttctcccc	attgtcgttg	ttccagccct	31380
aaaccaaggg	tgtgacctg	gctgtgtccc	ggtaataggc	agtgcctctt	aaccctctcg	31440
gttgacgtcc	cagccaggtt	tctgccta	caggacaaat	cacatcctgg	gaggtgaggg	31500
tggaataaag	ggagggaact	gagccagggc	agacagtctc	cagaggaggt	ggctctgacg	31560
cagagcaggg	tcgaaccca	caccaggaga	gaatttaatt	gatcatgtgt	tccactcacc	31620

tgcctcagcc	aagccctcag	ggcaggggaa	ggcaaagtca	ggatgccctt	cgcacacacc	31680
ctctctgggc	cccaccatcc	tccccaagtc	actagatccc	acagctgaga	aggaccttag	31740
gatccgtaca	aagcctaaac	acactccaca	gagggggaaa	ctgagactct	gaagggaggc	31800
ctcaacagct	ctggtaaaaa	aggcggttag	gccggggcca	gtgggtcaca	cctgtaatcc	31860
cagcactttg	ggagggccgag	gcgggtggat	tgcctgagct	caggagttcg	cgaccagcct	31920
gagcaacacg	gtgaaacccc	gtctccacta	aaatacgaag	aaattagccg	ggcgtggagg	31980
cgtgcacctg	tagtcccagc	tactcgggag	gctgaggcag	gagaattgct	tgaacctagg	32040
aggcagaggt	tgcagtgagc	cgagatcgcg	ccactgcact	ccagcctggg	cgcacactcg	32100
agctctcggt	tcaaccacaa	atgggtgtta	aggaataaac	aacacataca	actaaattat	32160
ccttccccct	tccccgaag	tggctggctc	aggaataaac	tctaccact	caggcagagg	32220
ttttcctgca	ccctgcatcc	gtgaggcacc	actgccagg	acgccaggga	aggctgccag	32280
gcctggagag	gggcaggggc	ccctccccct	caagggggcca	caaacgctgt	ctgcgcccag	32340
taccgtgggt	aaggcgaagg	cggccggcta	acccggggct	ggcggccttg	cagcgtgcgt	32400
ggcaacagca	gctggggccc	caagactcag	cacgggacgt	cctcgtccaa	gtctggggcca	32460
agagcagcgg	cccagggggc	ggggccggcc	agagggagcg	gggagaggct	gagggggcgt	32520
gccagcgccg	gaccccgcca	ttggctggag	attacaggag	gcggggacat	agcagggagg	32580
agcctctgga	caagcccccac	cggcccgcca	gggagggctc	gaggtcaaga	gccggagaga	32640
agggatttag	ggccctgggc	caagttgcac	agcagggaga	aggggctgcg	cagagggggc	32700
gggagaaagg	gatccgcttc	cttcctttag	agctgtgaaa	tgtccccggt	tggaaataaa	32760
ggcggtgct	ggggagaggt	gaaattcagc	caaaaccacc	cagtcaggca	gcccttctca	32820
gagataaaca	gtccgagcca	gcccggccag	gaaccttccc	ctccaacctc	cctaagcctt	32880
taacactcct	aagcctttaa	cgcgtttaca	cactcacata	aataaacaca	ctttgagcaa	32940
cacacataca	ccactcacca	catgtaatat	gtcaagccat	gtgcacgacg	aggtgtcgac	33000
aatttcatat	ggttcaacct	agtacactca	caaacacacc	taccaactca	tggctttcac	33060
agggacgggg	tcacacaccc	actctccacc	gacatggcaa	gcgtgcacac	gctatctcaa	33120
gctgtccctt	ccccctcaag	atcatgttac	ccagttttat	tttcttccca	gcacctatga	33180
cgactgacat	aattttattag	tttacttgtt	tattgggtta	tctgtgcccc	tcaccccaaa	33240
aatgtaacct	ccagcaggga	ggatgactcg	gtcagtcctg	attgtgctgt	agtccaggac	33300
ctagaaacaga	gctccatgga	cattcatggg	ctctgtacac	acaaacacac	acattaacat	33360
acaccccgac	acacagcctc	atccacacac	acacagcctc	acacctgctc	tttgacgcca	33420
cctgcacagt	ttctcacaca	ctcacttgat	ctagtgatct	gcgtccacag	gccccctccc	33480
cagcccactc	atactgccct	caccccactc	actctgccct	caccccactc	gggggaaactc	33540
tgtgtccagg	ccaggcctgt	gacactcacc	gtgagtgaag	tgaacgttag	gcagatgcca	33600
ccaaagccat	tcagggacag	gcgcaggaa	atcaacggag	acagagctgg	aaagggggaaa	33660
gcagcagatg	agggcatttg	gggagctgtg	ggaagccaag	ggcgggagct	ggggtaaaaca	33720
tccgccttca	tcccacctat	tcttttcttg	tggggccaca	agaggacaga	caactcacct	33780
tccacgtccc	gggagggccag	ggccatgagg	gtgcaggacg	cagtgaagca	ggcactgtgg	33840
agacacaggg	aagggcgagg	ggttggcctg	tgagcacccc	ccctcccttc	cccctgcagc	33900
acggctccctg	tcctcccggt	ccccatagcc	cagccacctc	acctgccaac	cagccgcacg	33960
ggctgggggc	caaagcggtc	catgaggatc	cccagtgcca	gggtgggtggc	gctgagcacg	34020
aagggaaccaa	tgggtgaagcc	cagggtgagc	atctcgctct	gctggtcaca	gcctggccac	34080
ctgcgcgtgct	catcctgggt	ggtgttggtg	ctgctctcag	ctgaggaggg	ggaaggggagg	34140
gctcagcaca	tgacaccagg	aacagctggg	cacaggagac	agcagcccac	agtcaggcgg	34200
cctgctttca	aatcccattgc	caagtgcctt	tgggggtacc	ctagagtcaac	atctcctctg	34260
atggggctgc	tccagaaatg	gcagccatta	gtacctgacc	ctgggagagt	cttgtgcaca	34320
cacagcctga	ggcttcaact	agctcaaatg	aaatactgga	cataaaagta	tttactaaat	34380
tgtaatatgc	actcagtggtc	caagcttagg	gggttggtga	cccccaacaa	gaagtggccc	34440
catatctaga	ggcaaaaggca	aaggcagtga	gtggtactct	aatggctata	acaagaattc	34500
attaaatagg	cccggcggtg	tagttcatgc	ctgtaatccc	accactgtgg	gaggctgaga	34560
caggcagatc	gctttagcct	acaagtttga	gaccagcctg	ggcaacatgg	taaaacccca	34620
tctctaaata	aaaaaaagaa	atttagaaa	aacactaaaa	cttagaggaa	gctttcccga	34680
taaatgatag	tctgataaaa	taatagctaa	tacttattga	gcacttaact	atgctccagg	34740
cactgttata	agtcaagttaa	taaagtatcc	cggtccctag	gtgatgaagc	tgaggcacag	34800
aatgagaacc	aggcaactgcc	ctccagtccc	ctctagaagt	ccacttgag	gacttgctct	34860
taacggtaaa	ctgccaaactt	ggagttgtga	caagttaagg	agaaaagcta	gtgataggag	34920
acaaagggct	gcttcgcttt	actcaatgct	cannnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34980
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35040
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	35100
ttccttctct	ccctccatct	ttccacaccc	tggatcatc	atacagaagc	agagaggact	35160
gcacttggtg	aaagtttcaa	ttctcctgtg	tggagaggtg	agcactgagg	aaggggtggg	35220
ggctgtcaaa	ggagactttac	ccaatctttc	cagcccaacca	atcccttgcc	cagtgtttct	35280
ataaaataag	ggccttttgc	atctgattta	agtaggaagc	tgattcctga	gcccctcaga	35340
tctgtgtaat	cagatagcta	agggggccct	ggaatctgca	ttttagcaag	cggaggtggg	35400
ttatgaagca	ctccaaagtt	tgagacacgc	ctcaaaggtg	gagtgggtct	gtggggggca	35460

gaaaggaaaa	tgcaaaagggg	gaaggggtca	cacttgggga	aggtttcaga	caataccgag	35520
tggaaagggg	gatgccaggt	gtggggagta	acagatagag	gaggcaaggt	gagtggagac	35580
caagccagac	cggggaggag	ggggccacag	ccaaggtgag	acaggtcagc	agccagaaac	35640
cgaagcagac	acttgcaggg	tgcaccccgc	cctctcttcg	tgccaatctg	agaccgagga	35700
cgtggagacc	ctggagagcc	cccaaccttg	tttctggggg	gtgggtcaga	gaggaagcct	35760
ctcatccccg	ggcaccagcg	gccttcccgg	gaggtccaac	acgcagatac	ctggataggg	35820
gtccatcaact	ccccgggcca	gagccacca	acgtctctcg	aggtccgacc	ttgtccctcc	35880
ttctacccca	cagtccccag	tcctagctca	tctgcataaa	gtccaatta	acatgttttt	35940
cctttgctat	ttgcgatccc	agaactcgtt	ccccaccccg	agcccgtttc	ccgcgcgttc	36000
ctcgcctatg	ggagggcggc	cccattaacc	ctcgcgaccc	gggcccgtcc	tggcggtcct	36060
gaccccgcca	ccccgtcccg	cggcgggggg	ctgggggtga	ggggcgcgcc	ctggggcaga	36120
ggattgcgcy	gcagggctctg	ccacagggca	gaggccaggg	ctctccggga	aaaaggcgag	36180
cgcataatag	cccccttttc	tgggaaaaga	cggggagggg	ggctttctct	gggagactcc	36240
aggtctcgaa	attcctcgtt	ccctatctct	cggcccccg	acccctctct	ctccccgcca	36300
cgcacctctct	ccccctccca	gccatctgtt	ccactccgca	gcgcgcgac	aaacacgggt	36360
ccagctcgct	tccgcctctg	cccagccccc	tccccaaagg	ccggggagtg	ggggagttag	36420
cagacgcctct	tctcctagga	ggcgggaatt	tctgcctcca	tctccacccg	gggtccggct	36480
ggccagaggg	aagcttcgag	acccccaccc	aaccaccacc	accgttcgca	gggcccgtga	36540
ggctgcagat	aacgcttgca	aggacgggag	tcggggaggg	tgtaggcgca	gtttaaagga	36600
cgggcagagc	aagccccggg	agggcgaggg	ggttttccct	cccgggtcgc	cgcccccgcc	36660
accctcgagg	ccagcccgag	ccacgcagcg	ccgcctgcgg	ggcacacca	ggacctggcg	36720
cgcacgtggc	gcttaccccc	acccccgggt	ccgctcctgg	ctcgcgtcca	gcctccccag	36780
actattcgca	aattgaggat	cccggacaca	gagtgacag	accccgcaa	gcctactgaa	36840
agccagccga	acccgctggg	gggtgctagc	caattctgat	tttgtacttt	acaaaaacaa	36900
aaaaagtgat	tgttggaaag	cgggagctct	ggctcagagc	agcagggatc	tgcgatgtga	36960
ctttgccaa	tctccagacc	cctgaggaca	ggttttctta	tctgaaaaac	gaggggacag	37020
tctctcttat	taacttctca	agagaaaaca	agacaaaggg	agggaaaatg	gcttagctgg	37080
aatgctgtct	tacagagcca	accttttgag	gtgggggaga	tggccaaggg	ctctgaggtc	37140
actcttgccc	ccaggagcag	ctgagaaccg	gaaagaagct	tgggacctcc	tttctgcaga	37200
gctatccttt	ccacagactg	ccgaggttcc	aaattgagct	ccaccacct	acactgtgtg	37260
cccttgggtg	tgtgccttaa	cctctctggg	cttgtttcct	acagcgacaa	gaaagaatga	37320
caacaccaac	ctcttaggct	atagtttgga	taaaatgaga	tagctgtgta	gaacagacag	37380
atcctaaacc	aatgttagtt	ttccctcat	ttggggactt	gctctaacct	ccagggtcta	37440
tgtcccagag	gcacaagcag	gtgcagggct	ggataaataa	ggtatgtctt	tctgcaggat	37500
ctcttgtcct	cactgatggt	gtcttctctt	gatatagata	attttaaagc	ttcacgttat	37560
ttattttatt	actttaaagc	ctcactttaa	tgttaaagg	aaatgtaaat	atagtataac	37620
aaggaagctc	aaaatttgca	taaagtttta	agataaaaata	ggagactcca	aaaaagtgtt	37680
actttcgcca	ggccctaggg	atgctatggt	gggaagtttg	agtcatacct	tagcattctt	37740
tctaaagcat	tctgtcctaa	tctctgttat	ggagaaaagc	cagcttctcg	gatgtacccc	37800
aaatcctggg	aagtaggggg	caggagctgg	actccctcca	agcactaagg	gcagggcatg	37860
gttgggaaca	gggaggtgag	ccagacagcc	agaggcgaa	gggctggcat	gccaaagcgt	37920
ctagttaatg	ccagctgag	cctgggtgaa	gaaggatggg	ggtgtgggga	agacaccccc	37980
caccaaccgc	caaagacagg	cgcacaccag	ccagtctctc	acttcccttt	ttatttctct	38040
taagacttgc	aagcagcagc	accagagagg	gaacctgcct	tcttggccct	ggaaagggcc	38100
gacccccaac	ccctaaccac	ggacacagct	ggcacctcag	gcccccttcc	ttctgaaag	38160
agggctgtgt	ctctctcaca	ttcacacata	cacagacaca	tgcagtgtgt	cacactcatg	38220
gcacatggga	cctcaggggt	agcctgtttg	ccgatcccc	caagaggtag	caggaggcag	38280
accgctagaa	ggagataaga	ggcacctcgg	tctctctcaa	cccaggagg	aagaaagctc	38340
aaacctctta	ggatagggac	tgtcttcagt	caatggagcg	ttgacttagg	ggcggttttt	38400
gaaggttttt	tttctctctt	tttgagctct	ttacaaaaat	agaacttctc	ttggtattta	38460
taaatctacg	gccatggctc	tatgtgcatg	ttacaggtag	aaaagccata	tggggcactc	38520
cttttgggtg	ctcaggcctt	gattgcctgt	catccaggtc	ccttgggtctg	agaagtctat	38580
gcggtcacct	cagagccgct	aagcaccttc	agtggggcca	tcccattggc	ggcgtactcc	38640
tgtctggagcc	gggcacggta	atagaagagg	taggaaggca	acaggaatcc	caggagttag	38700
aatagcagga	ggcccagatt	cacctttagg	gcaaggagag	agaaacagag	tcaagttagt	38760
agtcatctgc	ccttagcctc	ccacagggag	gagaaggcgg	ccatttttct	ccagggtcctg	38820
agccagaata	aatacagcta	gtacttatta	tgtgtagtca	ttgttccacc	agtatctcac	38880
ttaatgttca	gcaattctgc	aaagtggctg	agatgagact	tctcaggtag	aacaagtggc	38940
agggcctggg	gggtgccac	accatatggc	actcactagg	taggtatgag	gaaggcacag	39000
cactgtagga	gtctgggctg	gtcaggctgc	tcccgaatg	gggccttctg	ggctcacccc	39060
tctgaccttt	ggagatgtta	accaatggga	tcccgttcag	ggtggcgaga	ggaggctctc	39120
agacacagtt	caaggaaactg	ggatgcacag	cctggtggac	agaaggcttg	gaaggccag	39180
gacacgcggg	ctctgactcg	gttcacatcc	cactctgcac	tactcactgt	gtgactttgg	39240
gcaaatatg	gcaattctta	ctgagtgcct	ccttctcagg	gctgttgtgg	cgaagatgta	39300

agttaaaaaa	aagtatgcat	catgcttagc	acatagtgag	tgcttggtaa	atagaagcag	39360
ttatttcac	acaattcttt	gggaggagg	tttacgtgtg	ggtggcccca	cagggcagat	39420
gaaagatcag	cgtcagggag	gcagatgagt	tcaatgtaag	gaaaagactt	actaacagca	39480
gcagggtgc	ctcgtgcagg	agtgggtgcc	ctaccactga	gggtatctaa	gctaagagg	39540
aagggtcccc	tttcaggggt	gctggagaca	ggatcccaca	ctaggtagaa	ctggattgga	39600
ccaatggtgc	ctgaacacag	gcccagagt	caggactggc	cacttcacaa	agcacctgga	39660
gtttactaaa	aacagactcc	taggagggtca	ggcactgtgg	ctcacgcctg	taacccagc	39720
actctgggag	gccaaaggtga	gaagatcatt	tgaggccagg	agtttaagac	tagcctgtgc	39780
aacatggcaa	gacctgtgtt	atctgtacaa	aatttttttt	taaaaaat	gccaggtatg	39840
gtagccatca	cctgtggttg	cagctactca	gaaggctggg	gccggaggat	cgcttgagcc	39900
caggaaatcag	aggctgcagt	gagctgtgat	tttaccaccg	cactccagac	tgggcaacag	39960
aacaagacac	cttctctaca	aaaaaaaaaa	aacaataggg	ccgggcgcgg	tggctaaggc	40020
atgtaatccc	agcacttttg	gaggctgagg	agggcagatc	acgaggtcgg	gagatcgagg	40080
ccatcctggc	tagcacggtg	aaaccccgtc	tctactaaaa	atccaaaaaa	aaaaaaaaaa	40140
ttagctgggc	gtggtggttg	gcgcctgtgg	tccagctac	ttgagaggct	gaggcaggag	40200
aatggcatga	accggggagg	cgagactgtc	agttagccga	gatcgacca	ctgactcca	40260
gcctgggcaa	cagaatgaga	ctccgtctca	aaaaataaaa	ataaaaataa	ataaataaat	40320
aaaataacaa	taaattaaaa	acaaaaacag	actcctacgg	tcaggctgag	atatcctgat	40380
tcaggggact	ggggaatctg	tatttttaac	actcctgtag	gggttctaaa	aggcagacaa	40440
cttggaacc	tgagatttag	agacctctga	ggtgcctctg	gctgagatga	gtgagggatg	40500
gcaccacata	caaggcccta	cccctgcccc	caggagagtg	gctcctgtct	ccccacacc	40560
aacctcgct	ctcacccaga	aggctctctc	tttcaggggg	cccaccatcc	ccatgaaaag	40620
tggtctgtga	agcaaggcga	acacagcact	ggtgagggac	tgaggccctg	tcagcgtccc	40680
aaaagggggt	ggatgggaac	ctgtccccc	aacgggagat	caaagggtgg	tgggggcctt	40740
tcagcccgag	caagaacttt	ttcttttct	tcccaacatg	ggnnnnnnnn	nnnnnnnnnn	40800
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	40860
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	40920
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	40980
aagaaaaaaa	aatcttaag	aataaggata	taaaagaaaga	aaatatattt	gtgtagctgt	41040
tcaatgtttg	tatttcaagc	caagtgttat	tacaaaaacag	tcaaaagtgt	ttaaaaat	41100
aaaagtttat	aaagtaaaaa	agctaagtaa	gctagggtta	atttttttat	cgaacaaaga	41160
aaaatatctt	tgtataaact	tagttagtgc	taagtgtaca	ttgtttttat	tttatttatt	41220
ttttattttt	ttgaaatgga	gtttcactct	tggtgcccga	gctggagtgc	aatggcatga	41280
tcttggtctca	cggaagctgc	tgtctcctgg	gttcaagcga	ttctcctgcc	tcagcctccc	41340
aagttagctg	gattataggt	accgccacc	atgcatggct	agtttctttg	catttttttt	41400
ttgaaatgga	attttgctct	ttgacccagg	ctggagtga	atgggtgcaat	ctgggctaaa	41460
tgcaacctcc	acctcccagg	ttcaagagat	tctcctgcct	cagcctcctg	agtagctggg	41520
attacaggga	tgacacacca	cactcggcta	atttttgtat	ttttagttag	gacagggttc	41580
tcaactaaag	agaaccatgt	tgccagggt	ggtctagaat	tcctgacctc	aggtgatcca	41640
cccacctcgg	cctcccaaag	tgctgggatt	gcaggcatga	gccaccatgc	ccagccagta	41700
tacagtgttt	ataaagcctc	cagttagtga	cagcaatgtc	ctagaccttc	acattcactt	41760
actactcact	cactcactca	cccagagcaa	ctgcagctcc	tgcaagctgc	atgcatgata	41820
agtgccctat	ataggtgaac	cattttttta	tattttatac	tatatatttt	ctgcaccttt	41880
tctatgatta	gctacacaaa	tgcttaccat	tggtttacaa	ctgcctacag	taatcagtac	41940
agtactatgt	atgggtttgt	agcctaggct	ataccatgtt	gcctacgtgt	gtagtcgtct	42000
atactgtcta	gtttgtacac	tetatcatgt	ttgcataaag	ataaaatcac	ctaatagcac	42060
attttctctga	gtgtattcct	gttggttaag	aacacatgta	taaacattta	caagaaatag	42120
ctcaaatctt	tttttctttt	gatacagggt	cttgctttgt	cacccaggct	ggagtgcagt	42180
ggcgcaatct	cggcgcactg	cgacatctac	ctcccgggt	caatcgattc	tccggcctta	42240
gcctcctgag	tagttaggac	tacaggcacg	caccaccacg	cctggccta	ttttttgtat	42300
ttttatttaag	agatgggggt	ttgccatgtt	ggctaggctg	gtctcgaa	cctgacctca	42360
ggtgatctgc	ccgccttggc	ctcccaacat	gctgggatta	caggcatgag	ccacctgcc	42420
cagccattac	gttttttttg	ttgtttaatt	tttttttttt	taagagacag	attctcactc	42480
tgatcatcaag	gctggagtgc	aatggcacia	ccatagctca	ctgcagctc	caactcctgg	42540
gctcaaggga	ccctcctgcc	tcagccttcc	cagtaactga	gactacaggt	gtgagccacc	42600
atgctcagct	aattattttt	tatcttttat	tttttgtaga	gggggggtct	ttctatgttg	42660
ctcagggttg	tctcaaacctc	ctgggctcaa	tcaattctcc	tgctttggcc	tcccaaagg	42720
ctgggattac	aggtgtgagc	ctgaaaaact	tctagtgtgg	aagtggaaaga	taggccagg	42780
ccacttatgt	tttcaagtta	agcaagggtt	aggtcactta	tgaagcctga	ctagttttgt	42840
ttgcttaagg	gatctgcagg	cctgacctcg	gttttcattt	gttttaacag	tgtctatgtg	42900
tatgtgtgtg	tttatgtacg	tgcatgatgg	ggggaaagct	cagaaatcaa	gtaagccaaa	42960
cacaaacatg	taattataag	caggataaaa	ttctatgatg	agaagatag	ggccacggga	43020
gagtacttgt	gccagctcgg	tgatcaggaa	caatgtcctt	tggaagtgta	catttgagcc	43080
atgccctgaa	gtacggtagg	agttgggttag	gggtgaggca	gtaagaccca	gagctggggc	43140
ttcctgcaca	agctcagctg	ggcactgagg	accagtgga	ctctgctaca	gggcagtgag	

gagcagaaag	gctgaggaag	gctgggtgtg	gtggctcaca	cttghtaatcc	cagggtcttg	43200
agaggctgat	gggggaaaaat	cggtagagct	caggagtttg	agaccagcct	gagcaacata	43260
gcaagactcc	atccctgtaa	aaagctttta	aaaatttagct	gggtgtgggtg	gtatgcatct	43320
gcagtctcag	ctactcaaga	ggctggggta	aggatttgctt	gagcctagga	ggtggagcgt	43380
gcagtgcgcc	acgattgtgc	cactgtactc	caacctagga	gacaaagcga	gatcctgtct	43440
caaaactgaa	tgaataggct	gtgtgcgggtg	gctcactcct	gtaatcccag	cactttttgga	43500
ggctgaggtg	ggtggatcac	ctgtgatttg	gagtttgaga	ccagcctggc	caatatggtg	43560
aaacccgata	caaaaattaa	ctgggcatgg	tggctcacat	ctgtaattcc	agctactcgg	43620
gagggctgag	catgagaatg	tcttgaaacc	ggggggcaga	gggtgcagtg	agctgagatc	43680
gcaccactgc	actccagcct	gggagacagc	gagactccat	ctcaaaaaaa	aaataataat	43740
aataacaatt	aaaaaaaaat	taaaaaggcca	gggagcactg	gcagcctgtc	caagggttca	43800
ggtcaacttta	gtaaaaggag	aacaatggct	cctcccagga	cctctgggat	ctcagcattg	43860
atacgacagt	catggaaatg	ctagggccca	ggcagaccat	ctcagggaaa	acaagtggct	43920
ctgccctggc	ttggccactt	cctggccctc	tgcattgccc	agggtctcag	caccaagctg	43980
ttctcagtga	ttagctctca	tttagtgcca	gggtctctcg	gcttacatcc	tacgatgacg	44040
atgggaatgca	ggagaaatgg	ggctgtgata	gccagagctg	aggggtttga	atctcatgag	44100
atgttcatgg	agccctggga	gggagctcag	tgcaagtcca	tttctctttt	ttggttgaga	44160
tggggctcag	aggaggaagg	acttgttcaa	agacacacag	ggagtgttcc	agtgtgggac	44220
ggagggtttat	ggagaaaagg	tgaccatcca	aggcttgga	aaagatcatg	acttcgacca	44280
gcaagcctca	actctgtaga	cttggtgggg	gccaggccct	cccaaacaca	cctgacaggt	44340
gtctgtggtc	ttggggacat	tgtcgtctcc	cttctgtctg	atgctctgct	gtccctctcc	44400
catgaagcgt	atctcttcgc	cgtcccccac	ccttgctgag	agaggatggg	ttctctctg	44460
accaatactg	aagatcttta	gtaaagtctt	cttttttttc	attttctgaa	agtccctctc	44520
ttgagaaatc	aggacaagtg	agtcagggcc	aggacaaaaa	acagtgtggg	acgagtgtgg	44580
tggtcaccgc	ctgtaatccc	agcactttgg	gaggccaagg	tggcggatca	cttgaggtca	44640
tgagtttgag	actagcctgg	ccaacatggg	gaaacctcgt	ctctacaaaa	tacaaaaaatt	44700
agccagggcgt	ggtgtgtgat	gcctgtaatc	ccagctattc	gggaggctga	ggcaggagaa	44760
tcacatgaac	ccaggaggcg	gaggttgcat	cgagctgaaa	ttggggccact	gcactctggc	44820
ctcttgggca	cagagccaga	ctacctctca	aaacaaaaac	aaaaacaaac	gacaaacagt	44880
gtagactttg	tgtttttctc	aaaagcactg	tcaagccagt	gcccgcagca	gtgggcctag	44940
acacctccag	tcttgccctc	gggtcagttt	ccagcctccc	tggaacattc	ccccaggat	45000
gtgtactttt	tgattgtcct	aaatccagag	tctgtggcct	gacctggttt	gtcacagctc	45060
tcagtccctc	cccatcccca	atcccaggga	gccgcagggtg	tgtgcagaag	aggcacacca	45120
cactcaatac	atcttgcctc	ctcgtgggac	ccaatccatt	ggcttgggtg	tgtacagact	45180
gagcctcatt	atagccgttc	gttccctgtt	acctttccag	atcaatctgc	cagcttggct	45240
tctccgagtt	tcgcttgtca	gcatttctcc	aatcccatca	tgtactttgg	acctctttgt	45300
tggtgtggctt	gctttatctg	aaattttcag	atttgacttc	aggctctctc	tttgtccctt	45360
aatatggctt	aatgggtgac	ctgtccagg	gtagagaaaa	tattgaggag	ccctgacttt	45420
gaggtgcaca	agttagaggg	ttagacaagt	ccagccacaa	ccagcccaag	ctgcagtgtg	45480
gggaggccgt	tccagctgct	ccacgggttg	gggtggagca	tacagggaag	cttccctctt	45540
gctgcagccc	aggtgttctg	gctgccctag	ctgcctgggt	ttggtagaag	aaagaaaagg	45600
ctctgtctctg	actgttcaac	taatggcact	atgagattgc	acataattaa	cctgggtctg	45660
ctcttccaaa	agccttgggc	ctctgactgc	aacatggagt	ctgggtatca	ctccccatcc	45720
ctgcgccact	cacctgctct	ggcgctaggc	gtgtgcttaa	tcacttaatt	tctctgtgct	45780
gcctcttagt	tatcacttcc	cctgatccca	aatacttacc	agggtgtggga	tgcacactga	45840
ctagttaactc	cttgagggtg	tctgcttctc	accggggact	ccgaaaccaa	acgaaaagca	45900
aggccaagcc	cagcctaaag	gacgcttcc	acatgacttc	aggcttgcgg	gggctggagc	45960
gtgggggttg	caatggagtt	ggggggggct	cagggagggg	atgtggaaat	gctttgtctt	46020
gcaaaactcta	gagaaccgtg	taaataggag	tgattattct	gtcccttccc	tttctttcca	46080
acaggaatca	gcattcccaca	gcccatgttc	agctatgaag	aatggaaact	gaggctccgg	46140
gaggggtata	gggaggagcc	agcagggtct	tgagttcata	ttagtgcct	ttcctccata	46200
ggcacactctg	tgttttcttt	tattttatct	tgaatttaatt	tttttttttt	tttggcagag	46260
tcttgcctctg	tcgcccaggc	tgaggtgcag	tggcgcggtc	tcagttcact	gcaatctccg	46320
cctcctgggt	tcaagtgatt	ctcctgcctc	agcctcccga	gtagctggga	ttacaggtgt	46380
acaccaccac	accagctga	tttttgcaat	tttagtagag	acagggttcc	acagtgttgg	46440
ccaggcttgt	cttgaaatcc	tgacctcaag	tgatctgcta	gcctcggcct	cccaaagtgc	46500
tggtattata	ggtgtgagcc	actgcgctcg	gccacatctg	tgttttaaat	gagaggaaa	46560
gggataatgt	gcattttgtg	gaagcttggg	ccgtttgtgt	ctaggactct	tatgatcttc	46620
ataagttttc	cccaggggag	gacactgttc	cacttaggga	gtcaggaccc	ccagtcctta	46680
caagattcag	cctctcaaaa	tgagagacag	agttccaggc	ctgggctggg	ttctgttcac	46740
actagggagag	ggcaagtggg	tggtgtttgg	gatgtgggga	agtattatga	aaacagagat	46800
gctccaattc	ctagtgtatg	gaaaccatta	agctacttgg	catcttaaaa	ccaagagcgg	46860
ttcaagttct	gagattgtta	acacacctta	caacacggcc	gccgttatta	ggaagaagct	46920
ctgtttgatg	acgtcccaca	ctgtgggtac	ctttatgaac	aggaatttgc	tttttcaaat	46980

ccccagagaag	taagattaaaa	gttggctgtt	ctccatcctt	gaaaaatttg	gttttaggg	47040
gaattcaaga	atgactgacc	atacagaatg	gggagcaaac	ttgggaagaa	agaaggcaca	47100
gttcagagct	ctcccaatag	tcacccctga	actgcacccg	gaccatcagt	tatctctgtg	47160
ggtagagctc	aggaatctaa	aatccatttt	aaaattaaag	tatatcgggg	ctgggcgcgg	47220
ctggctcatgc	ctgtaatccc	agcacttttg	gaggccgagg	tgggaggatc	acgaggtcag	47280
gagtttgaga	ccagcctggc	cacatgggtg	aaccccgctc	ctactaacia	tacaaaaatt	47340
agccaggcat	gggtggcagac	acctgttagtc	ccagctattc	ggaaggctga	gtcagaagaa	47400
ttgcttgaac	ctgggaggca	gaggttgag	taagccaaga	ttgtgccact	gcactccagc	47460
ctgggcaaca	gaggagact	ctgtctcaaa	aaaaaaaaaa	aaaaaattaa	agtatgtcat	47520
acatactggt	acaggcacag	acctaagtg	tacagcccaa	tgaattttta	cacatctata	47580
cagctatata	actaccacct	atatcaagac	acattccagg	aactcagact	ccatcatacc	47640
cctcctcagc	agaggttaaca	gacccacacc	tctcctgctc	cggtggtaat	taaccactat	47700
tctaactttt	ctatcaatta	gttttgccca	ttcttgagct	tcacacagat	atacattgtc	47760
aggcatgatg	actcatgcct	gtaattctcag	cactttggga	ggccgagacg	ggagtatcac	47820
ttgagcccgag	gagttggaga	ctactctgga	caacatagtg	agaccccgga	ctctacaaaa	47880
aaaataaatt	agctgggtcat	gggtgggtcgt	gcctgtagtc	ttagctatct	gagacgctga	47940
gagaggagaa	tctcttgagc	ctgggaggtt	gaggctgaag	tgagccgtga	ttgcaccact	48000
gcactgcagc	ctaggtgaca	gagtgagatt	ctgcctcaaa	aaagaaaaaa	tatggccggg	48060
cgcggtggct	caagcctgta	atcccgagac	tttgggaggc	caaggcgggc	ggatcacgag	48120
gtcaggagat	ggagaccatc	ctggctaaca	cggtgaaacc	ctgtctctac	taaaaatata	48180
aaaaaagaaa	gaaaaaaaaa	ttagccaggc	atggtggcgg	gctctttag	tcccagttac	48240
ttgggaggct	gaggcaagag	aatggtgtga	acccgggagg	cagagcttgc	agtgaagccga	48300
gatcgaccca	ttgcactcca	gcctgggcga	cagagtaaga	ctctgtctca	aaaaaaaaaa	48360
ggaaaaagaa	aaaatatata	tacatttgtt	actttttggc	atctggttta	ttttgctcaa	48420
tatcacatct	cgcaaatata	tctacactgt	gtgtatgaaa	ggttggttct	ttttgtgtgt	48480
atgcagtatt	ccgtcgtgtg	actacgggac	aatttgctta	tccgtattcc	tatcggtggg	48540
catttgggct	gttaccaggt	tctggctgtt	atgaataaag	ttgctatgga	tattcttgtta	48600
cactacttct	ggtgagcgt	tgcactcatt	tgccttatgt	aaatatcttg	ggtggaatta	48660
cttgatcata	aggtagggtg	gttggccttg	taatgtgctg	acttggttat	gctgaattcc	48720
cttttttgtg	tatttctggt	tagagcggaa	catgagggtg	tctcttcagg	gaatctggag	48780
ggtggaaggg	aagcaggagt	cggtttctgg	ctcacacatg	ttgtgactga	actgctggtg	48840
cacctggttg	gcatggagct	ggcttctcct	ttggcgttgc	ctactgttgg	ggcaggtgtg	48900
tatgtggtta	gctccatgca	atgaacccgg	gcttctgcaa	aatacattaa	caacgcacaga	48960
gacacaaaaa	gctgatgtgg	atttaaaggc	ttcagttcan	nnnnnnnnnn	nnnnnnnnnn	49020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	49080
nnnnnnnnnn	nnnnnnnnnn	cagcagtggt	tctcaactga	gcatagtttt	gcctcagagg	49140
ggacattttg	taatgtctgc	agacattttt	tgattgtcac	agcccagccg	agaaggtact	49200
actagtatct	ttttggtaga	ggctagagag	gctgctaaac	atctaaacat	gcacaggaca	49260
ggcctctgta	acaaaaaagt	atccagtcac	aaatgtccac	agtgttgaga	ggtttaggta	49320
agtaggcgct	aaaacataag	gagactgtgc	ctgagagcaa	gaaggagtaa	ttggaaagtg	49380
ctggtgtgag	tagctctggg	ttttagaaag	ctcatttttg	ctgctttag	acagtgcac	49440
agaggtggag	gaggttggtg	agactggagg	cagggaaggt	aatgtgggag	ccactgaaat	49500
gatccaggtg	aaaaacggtc	agcaggtgac	taggaaagtg	gcagaggcaa	tggggatggg	49560
tggctggatg	agatgggtgaa	gaaagcacta	taactaacta	atgtgtggat	gatgggcagg	49620
aggggtgaag	gatgaccaga	gtcctgcctt	gcaggtctag	ttggaagggt	atggtttctc	49680
ctgagaaagt	gaccacaaaa	agtgaagcag	gtttgtgcgt	gtgtgtgtgt	gtgtgtgtgt	49740
gtgtgtgttg	agttcagctc	gagatgtgtt	ggactcacaa	tgtccatggg	acatccaagt	49800
ggagaagcat	cttgggtgac	catatgtgtg	agtctgcagc	tcagaaacag	gcctggggct	49860
ggagatgaag	acttgggaat	gatctgcgta	tatattttgt	agcttgagcc	acaagagttag	49920
atgacataaac	ccgtggtggg	tgtgcagaat	taggagagac	gtgcaccaag	aagccaggtg	49980
atccccaata	tttaaccatc	tggaagaata	agaggagcct	gccaaacagaa	attgggaggg	50040
aatggccaca	aaggctactg	agaagggaag	cagttcttaa	gaagggggaa	gtgaagaggt	50100
atcactactg	cagaggtcaa	gtaggataag	aactgaagaa	tgtctgttgg	gtttggcaat	50160
ggggtagtca	gtgggcacct	gggcaaaagc	agttttgggt	gagcaatagg	gataacagaa	50220
acaagactgc	tatggtaaga	ggaggaagag	gggtgttagg	aagtggccag	cgagtctaca	50280
ccacttgctg	gaggagcttg	gctttgggtg	aaagcagaga	agccagctca	ctcattgact	50340
taacctccaa	gaaacacaaa	atcatccata	tctgggctca	aattccagca	ctaccaggag	50400
atgggtggcc	cctagaaatg	ccatcccact	tctcctctgc	ttatcctatc	ctatctgtca	50460
gtctgttgag	cccaggctaa	gcgctacctc	ctcaagcaag	ccttctctgc	ctgccgtcac	50520
actttaagt	atcctgacaa	cactgaaaat	gtgtgtctct	tccattcatg	ttagtctac	50580
acttctgagt	atctcctcaa	tatatgtcct	tggtttacta	atatgctcgt	tctgtttgcc	50640
ttatttatca	gctaccttaa	acctccctgc	aactagagat	tctctttaag	tatttgttga	50700
ataaatgaat	gaatcaatcg	atgatccaga	gcctggtaga	ggcttgtgtc	catgggtgag	50760
gaggtcaga	aaatacctgt	agaatcgaaa	taaatgcatg	tgtgctctga	tctaaactca	50820

gctaaacttt	ctccaggggg	ttaaagttcaa	gttgattagt	caattgatta	attaattcat	50880
tatgtaattg	aaaaactcct	tctatgacct	gggcagagtt	ataggcagtg	aacaagacag	50940
acaaggctcct	tgttgtcatg	aagtttgctt	tctgaaggag	agagataata	aacaagaaac	51000
cagtaagaaa	gcaaagattat	atcatttttg	taaatgttct	tgtggaataa	aatgtgatga	51060
tgtgtaacaa	aagtaccaa	taggagagtg	gggtgggtgg	gcttctttta	gaaagagttc	51120
tcggagaaag	cttatctgag	gaggtggcct	tttaaccagt	acaaatgctt	tagcttggcc	51180
agtggagctg	ggaccaggat	gacaagggtc	acttgctcatg	ccagtgagtt	tgagcttgta	51240
gacaagagcc	tgatcatgaa	agactttgca	gatgggtgta	atgggtttgg	gttaattgct	51300
actatgtggg	aagactttga	atgggaagca	tggggacaat	ggcctgtgat	acatgttatc	51360
aaatatggtc	gcaggggcta	gtgaggtggc	agcagagata	gggagaagta	gacggactgg	51420
ggaaggtaga	agatggggca	ggggagggca	ttactgcaa	gacatattcc	ttctaagctc	51480
actgagtgtt	catggtctct	gggagcagag	gttcctggag	gggaaagagg	ataatgtcac	51540
ttcctgagga	agcgggaaga	acccatctga	gacgtgggga	ctgtgctggt	tcgtttctaa	51600
ggggccttcc	agatctcaca	tgccaatcgt	cttgggtctat	gtcaattggt	ggggcatcca	51660
aatggggaac	tgttgtccag	gccgatttca	cagaacaacc	gccagtgcca	tatctccga	51720
gccattccac	cttgagtggt	ctgtagctct	ttcaccagct	tttatctgcc	ccgtggggat	51780
gttggccaa	ccagtgtaac	aagcagttga	tcagccccag	agatcaggtc	cctggagtct	51840
gtcacttttc	tgaggggtgg	gagagaatcc	tgagagcaga	catgtaacta	gaaggggcac	51900
ctggccttcc	atggtctgag	ggagagaaatg	gtgggatctc	tgccctgaat	caaacctccc	51960
tttctcagtg	tccatcttac	ctctctgctg	taccttcgtt	attttccagc	agctcctcag	52020
cccgttccct	tgggaccctt	ctctgccaat	ccctacaccc	actgtaaat	tcaccgtggg	52080
agggagatgg	gccttgagg	ctgtattagt	cttctattct	gcataacaaa	ttgcctcaaa	52140
tttagcagct	tcaaaacact	catgtttatt	agctcatcgt	gagttcatca	gcagtggtgg	52200
cccagcatgg	ctaggttttc	tgctcagggt	ctcacaaggc	taaaatcaag	atgttgtctg	52260
ggctgtgtgc	tcatctggag	tttagggttc	tcttccaggc	tcacgtggtt	gtggcagaat	52320
tctgttccct	ggagttgcag	ggctgaggtc	ctgttttctt	gctgactgtc	agatgagggc	52380
tgctctcagg	tcctcgaggc	tgcccacatt	gcttgccacg	tgctgtgtct	tttccatcct	52440
tgaagccagt	gatggagaat	ttcccttgga	ttgaatcac	cacatgggtg	gactctctga	52500
cttcaggaa	agagccctgt	ctcttttatg	ggatcacctg	attagatcat	acccatagag	52560
ggcagttcct	tttccctaaa	gtcaactgtg	gcattgtaaca	tcacacaacc	acaggagtaa	52620
aatccatcat	atttacagtc	ccagggatta	tgccagctgc	accaggggac	aactgaattc	52680
tgccgtgcac	aaggggccaag	caggacttta	ttggtgaaga	acagtggaat	gtcattcttg	52740
gttcttccag	aaaaaaatca	ctcagtaaa	ttagagggtc	tcttgccctt	tgggaaagtca	52800
tcaaaagaat	tcatggaggg	tttggacctt	cacccataga	acatcacacc	atgttttcta	52860
taattgcagg	gttctatggc	ccttgaagcc	tattcatagt	ttccaggttg	aaaagctctg	52920
ctgcagggtg	tggggaggga	tgccaggtgga	ggtgagggtc	gaatagtgtg	agctgcatat	52980
ctggagctgt	ggtggttttt	ttagtcttta	agctgtcatg	tgttgggggt	tgggcatggg	53040
agggggcatc	caagagctcc	ttggtattga	caccatctcc	aaggtgatct	ctgctctgcc	53100
tggtgcacac	atgtttttct	cctgttgcaa	cagccacttc	ttgtagaaga	gcagaccctc	53160
cagtaccagg	tctgaccttg	gacagcttgt	accaggagct	acagcacact	cccccaaac	53220
cctaaagtgt	ggatgagccc	cccagaaatt	agatcagaaa	agattaaatg	cagaggtgat	53280
ctgtcagggt	ccctttggaa	gtgctgggat	ggagaggatt	gactgagttc	gtttagggaac	53340
ctccaagctc	tgtagtaact	ttagggtcag	aaaggaggat	gcctaagatt	caggatcctg	53400
cagtgatgag	tcaacatttc	ttggggaaag	aggcagggtc	gaggattaaa	cggagatgat	53460
gggtatcgtt	ctcttgctca	aaggcactgg	accccaaggc	ctccagctct	tcgctcccat	53520
ttgaaattca	agtcctgagc	acaccacagt	tgtgatgcag	ggaaagaatg	tgcttatcag	53580
agagcctggg	caagtggggc	ccttgtgagt	accgttcaac	ctcatttatg	tcattggcac	53640
caaaagtaga	catcagcttc	ttgaaagttt	gattaatgct	ggtcacactc	aaagaccctg	53700
ggtagcattc	atttactaag	caattactaa	ataccagttt	ctgtgctaaa	tgctgcatca	53760
gtcaggggtc	ttaattggcag	gcagcagaaa	ctctccttgg	ctgatctaa	tagaaaaatc	53820
caggactgaa	aggaaacgga	gtagctcatg	aaattgcagg	aaggggccgga	aaaccagaca	53880
tgagagccaa	gtcagggtgc	agaacagggtc	tagggaggat	cccaactgctg	ctgagacctc	53940
gacctgtgtg	ctggcaccga	ggatgttgta	gggctcagac	cctggatcaa	tgatctctgc	54000
agtgcctctg	tggttactgc	aactccaggga	actcaatctt	gtcaacgcca	ccgccaagaa	54060
gaggccttct	tgccctccat	ctttttggtc	actagctcca	gattcaaaat	cttgaataga	54120
tgcttcttct	ctttgataga	gcccagtcac	atgcgttagc	tgcaaggaa	gctgaaaatc	54180
tattaggaac	ttttgtcttc	aaaaatgaga	ggcctgtcct	ccaccaagat	ccataggaaa	54240
tggaatccaa	gaaaccacag	gaagggtgga	ggtgactggg	cagctcacag	catgcatgct	54300
acatgtgaat	tatctcattc	atttctcaca	ctaccagtg	aggtaggtat	tgctatccct	54360
acttcataaa	tgatgatag	aggtacagaa	agtttaagga	acttgccag	gacacgcac	54420
gcagctatta	agtgttagac	ccagtcattt	tgagtctgac	ttggactgtc	tgactccaga	54480
agccaccttc	tcagacactg	ctgtataact	ccagtgaatg	ttgatgaaat	tttcagggtt	54540
gctaaagctg	ggatttcaga	tcctggattg	tatgacctaa	aagagagact	tcctaggag	54600
tgagggtccc	tgaacagtc	actggtttcc	aagaatgggc	tcctctcat	caccttatga	54660

cagtaatcct	ctgtccaaca	gccaaagagg	tcctgtgggg	agggcttgca	gatgggagtg	54720
cgcagagccc	agctcaaagc	tcctgactag	gctcttgttg	agtattcctt	tgattcctgc	54780
ttctgtcttt	ttaaatcaat	ggagacaggg	gaggggttatc	tccatcctcg	gctcaagatg	54840
aaatgcatcg	ttcctcgttt	ttctcattcc	ttcccaatgt	gtgtactgtt	aacttttagtt	54900
atgaaggaaa	ttacagtgtc	ctgtgcatat	accaagggctg	tccaacctcc	acacctttgc	54960
tcaagctggt	ccttctactt	gaaatgcctg	tttctctccc	ttctaattgc	atctttccat	55020
ccaggtagga	atcagctcct	tggttcatgg	agccttttct	gctctgtttt	actatgcatg	55080
gacttccttc	tgaattagca	gaggatgttt	cctagcttgg	tcttaacctt	tctccttttg	55140
tttgacctca	atttactcat	cttacaattt	aggttgtaag	ctaattgaat	acaggatcta	55200
tgttcactc	tgattttatc	tccacctgga	tagcatcatt	tttgacacac	aagcaggcat	55260
atgggagggg	agagaagttt	ggtgccagaa	agaactggat	ttgaattcta	accctgttgt	55320
ttacgtgagt	acgttactta	accattaatt	acttcaatgt	atatattatta	agtacctact	55380
atgtgccggg	cactgtacta	agcaccaagg	ataccaatgg	gagtaaagag	atgcagcctt	55440
caccatcacg	aaggaagaca	gatgttaatc	cattaaccaa	gtaatctcac	aagaaaagta	55500
aaatgactaa	ctgataagga	caagcccctg	gagctacaag	aggggtgata	cagggcatcg	55560
atccaataag	ggcagtggtg	cggggagatc	aggagccaca	cagagcctgg	gttgtctcac	55620
ttggaaaatg	gggtatcaac	cacctacctc	actaggtttt	taaaatcagg	ttaaagtagg	55680
taatacttgc	catgaacagt	attttgttga	ttgatgattg	attgaaacgg	agtcctcactc	55740
tctcgcccaa	gctggagtgc	agtgggtgca	tctcagctca	ctgcaacctc	tacttcctgg	55800
gttcaagtga	ttctcctgcc	tcagactccc	aagtagctgg	gattacagge	agccacctct	55860
atgcctgact	aatttttcta	tttttagtag	agacaaaggct	ttgcaatggt	gaccaggctg	55920
gtctcaacct	cctgacctca	aaagatccac	ccacctcagc	ctcccaaagt	gctgggatca	55980
caggcatgag	ccactgcac	cagccacttg	ccatgcatgg	catttaaaaa	tgttcagtaa	56040
atgttaccat	aatgaaggct	ggtagggttg	ccaactgagt	ggtctgattc	agaaggaaag	56100
aagttagaca	tacgtgaaca	tttctgttac	ttgaagatcc	tcaggacagt	gactcctaga	56160
ccatccttcc	atcacagtca	gctgggaagc	ttttaaaaaa	atgcagacat	ctgaccttca	56220
cgctagacct	attagccaag	cagaagtttc	tgggcagggc	atctgcata	ttttaaaaat	56280
ctttaataag	gcagcctcaa	aattacagat	tcagcacgca	tttaccataa	ccactgaaga	56340
aatgcaaaag	tataaaaaga	agataaacaa	caatctgtct	cctgctttct	tccctctcct	56400
ccctgtcttc	tggaggcaac	aaggtcaact	atttgggtgtg	attcctttta	gcattccctc	56460
catcaatggt	cacataagga	tgctcacaga	taagcaccta	tgcggggggtt	ttttttttcc	56520
ttgtaaaact	atccacatac	taaatacttt	cctcagtatc	ttgccttttt	tacttctatg	56580
tcacagaaac	atctcttcag	gtttatagat	acaggtccag	ctcttctttt	catagccata	56640
taacattctg	tagaatagag	aggacacatt	ttactcagtg	tccgattgat	ggatatcaat	56700
attgttttca	tttctacaaa	tagtcaagga	ataacataac	tctgtaaaag	ttttattact	56760
tttaggcgca	tttatgccta	aaggatagtc	tcaaaagagt	gaaactgatc	aaatgtgcat	56820
ttttttattt	taataggtat	ggacagattt	gttctcaaaa	tgtttgtggc	agttcaaaac	56880
accagtaaaa	caggggagat	atgtattttg	gaaaagcacc	caaggcgatt	ctgaagtgtg	56940
gtccaggata	agaaccattg	ccagagctg	ttccagatgg	cccttgggtt	cctgaagtgg	57000
gtatcgggag	agaaatcttc	actgaatgaa	tgagtgggct	ccccagggaa	gtgatgaaat	57060
ggctccttat	agccttgcta	tctccctctg	acagaggcaa	actctctctc	cctgggggaa	57120
gttccctcaa	ggcctctata	taagaagtct	ttgtgagagg	aagcaaaaga	ggacctgggc	57180
tttgggaaga	tctaaagacc	caggaaagtc	tctgggtggg	tgagtggctt	ctctgctgtg	57240
gtggagctgg	tgacagttta	ttctcccagg	aggtccctgg	ctgtggctga	cagtttcttg	57300
agggctggca	ggcgtctacc	tgtggcttcc	aggttatgag	gatgtcagca	ggggcagcct	57360
tcactctctg	ccttgcacat	tccttctcgg	ggatgtgaaa	gtgctccttg	gctggggaaa	57420
ggagatggtg	gagacatgga	ggagggtgtg	ggtggcttct	tgaactctga	ggaggggaca	57480
taccttctaa	gtcctatgtg	ttcctaggaa	agccaataat	cattgcttct	cccgctttt	57540
ttatgtcata	gactctgagg	gacccattaa	gtacaaacaa	ataagcgtaa	tagtcccttc	57600
tttactctcg	ggcctgaagg	aaagccagcc	tcagccaccc	ctcaggggtt	gctgcgttct	57660
gtttagaaga	aggtccttgc	gtcctggatc	ctggagcatc	aggagctggg	cttggcatga	57720
gcttttctgg	cccatcctga	tttctattca	ggccttcttt	ttctccacct	cactcccacg	57780
gtccctaat	ggtgtgattg	tgatgtgtgt	gcatgtgtgt	ctgtgtgtgt	caatgacaaa	57840
ctgtgttctc	cggtgcagga	taaagccaag	atgaaaactcc	ccttacttct	ggctcttcta	57900
tttggggcag	tttctgctct	tcactaaagt	aagtgttttt	tgccctcagg	cttctcttct	57960
ctgttttttc	cctttctatg	gtagatgggg	tcagagttac	acacccaccc	ccttctttga	58020
tcgtcttcta	tttctgaatt	tctgtgtgct	taaaagggatg	gggactctat	ggccaggagt	58080
tgaaaggatt	tctcaaggcg	tctgttatgt	ctgtggctct	gggtctactg	tgacattccc	58140
aattttgtcc	tttctccatt	atgcttactt	tgagcttact	gagtgccttc	tctcctttaa	58200
ctctcttagc	atcgccatga	agtaggtggg	attgtatacc	catttcacag	aaatacagct	58260
ggtggatgat	ggaaccagta	cccaagccca	tgactgcccg	actctaagtc	catgctctta	58320
accaccttga	ccttgtcagg	cagcttgggt	tccctcata	gagactgggt	tccaggttcc	58380
ccttcccagg	cagagttgag	cactctgatg	cccagggcaa	ggtgtgagct	gtctgtgggt	58440
ctggggagga	acaaggggag	atgtgaagga	aggacactta	gctatcctcc	ctgccagggt	58500

ctgagacttc	cacctttgag	accccttttg	gtgctaagac	gctgcctgag	gatgaggaga	58560
caccagagca	ggagatggag	gagacccctt	gcagggagct	ggaggaagag	gaggagtggg	58620
gctctggaag	tgaagatgcc	tccaagaaag	atggggctgt	tgagtctatc	tcagtgccag	58680
atatggtgga	caaaaacctt	acgtgtcctg	aggaagagga	cacagtaaaa	gtgggtgggca	58740
tcctcgggtg	ccagacctgc	cgctacctcc	tggtgagaag	tcttcagacg	tttagtcaag	58800
cttggttgag	tggcctatgg	ctgagggctga	ggtgggagca	tggaaacgggt	gtgggatatg	58860
ccccagcat	tgctatcact	ggctcttttt	cccattgagg	gccctggggg	tgctcagtaga	58920
acctgagcct	cagagagggtg	ttggggtaag	aggggagggc	cacctacaaa	cagaagtgtg	58980
attttggtct	ccaaccttca	aatggttgtg	gcaggggagg	gagggaatga	attgtgggga	59040
ctcaagaccc	atgtgaattc	atgtaggaag	gatgctccat	tctttgtcct	ttatcctgcc	59100
ctgtagttta	cttgccggag	gtgctacagg	ggcaacctgg	tttccatcca	caacttcaat	59160
attaattatc	gaatccagtg	ttctgtcagc	gcgctcaacc	agggctcaagt	ctggaattgga	59220
ggcaggatca	caggctcggg	aagagaagtg	tgaacactaa	atggggtgca	cctgctgac	59280
tcagccagca	ctcagcttgc	atcagatttg	tctgtttttc	tctgtataaa	tctccagaag	59340
aaccagggat	agatggacac	ccacagacaa	cactgagggg	gctgcctggg	cattcagggg	59400
agagctaagg	atttagaatc	aggaggtttg	ggtccaagtt	cctttccatc	tctcactatc	59460
tatgtaactt	aagtttagctg	ggcatggtgg	tgcatgtctg	taatcctagc	tacttgggag	59520
gctgaggcag	gagagtcact	ggaacctggg	agacagaggt	tgcggtgagc	cgagatggag	59580
ccattgcact	ccagcctggg	caacaagagc	gaaactccgc	ctcaaaaata	aataaataaa	59640
taataaaaaat	aaaaaaaaaa	ttaaaacaag	accatgagtt	tgtttctctc	tctctaggat	59700
gagttggcaa	cccttggtct	accttttggt	agggctggaa	ggacaagcct	gtcactggga	59760
tgcatagaat	ctgatgtgta	taattgcccgt	ggatcagcat	ttcagatgac	taggacagtt	59820
cccatcatgg	tccagcaggg	aagggcccat	tgccccgtgg	gcagcagaaa	gagctggcag	59880
atacggggcc	aggtctgctt	ctctgccttc	cctctgcccc	atcccttctt	ccccctcttg	59940
tttctccagg	gtcgtgcag	acgctttcag	tggttgtagc	gcagccgctg	gaactttgca	60000
tactgggtcg	ctcaccagcc	ctggctcccg	ggtggtcact	gcgtggccct	gtgtaccgga	60060
ggtagggtgg	ggctggggat	gaacgatgga	aaggctctgg	agatgggaag	tgcccccaag	60120
aggagatgct	acaaagagcc	tgaccttttg	tgggagaggc	ttcctgggtc	ttttatatat	60180
tctgactcca	cagcagtgtg	tgggtgggaa	aagaggccct	cctgtgggtt	gagttgggat	60240
ggacaagagg	ctgaaagtcc	ctttctgttc	tgccctcaca	ggaggccact	ggcgtcgagc	60300
ccactgcctc	agaagacttc	ctttcatctg	ttcctactga	gctggtccca	gccagcagtt	60360
cagagctgcc	ctctcctggg	cagctgcctc	ccctcctctg	cttgccatcc	ctccctccac	60420
ctccctgcaa	taaaatgggt	tttactgaaa	tggatttatt	ttctcctctg	atcgcggtac	60480
cactctgctt	agccctcatt	gaaacttctt	ccttatcatc	tctcccccac	ccacaacttt	60540
catagaagtg	tcagaagcta	ctactccttg	aggaggagga	tggagggtgg	agttgggtct	60600
atggagcctt	ttggagatgg	aggaatgggc	tcagctagtt	ctcttcatag	aacacctgat	60660
tactgggcac	ctgcatagtg	ctgccaggac	ctttcaaggt	tgtaggtaga	ctcccaatgy	60720
cccagtttgc	atctctgtaa	ccaaaggcct	ttctctcttc	tctctccaac	cccagaactg	60780
tggttggttt	tatatgtaag	gaagttaaca	tgtccctggg	aacagtccac	aacattcagg	60840
aatgaatgta	taagtaccgc	aatccccggc	ccctcaagtg	gaataaatct	aacatgtatt	60900
gggcaccatt	tcccagtggc	ctgctgtggt	agttggcctt	attccatgca	tttttatggg	60960
ctgccttccc	ttcctcaact	gcattctctg	ctccttctca	ctctctgcaa	ctcccaataa	61020
aacacttgta	cgcaactccc	tctctcagga	tctccttctg	gggaaacctg	atataagaca	61080
gcttgccatg	cgctcagactc	tgaatgaggc	ctgggaatac	aagacatagt	cctctggcac	61140
ttgggatata	tggttatttg	taacataggc	acaaaaacat	ctactagtgt	ttatcgctta	61200
ttgagcacc	acaacatacc	ccctgctgtg	gcaggcacct	tgccatagatg	acctcatgtg	61260
atcaataatt	atgagcccta	ttttacagaa	ccaggctcag	agaagttagg	atctgtcaaa	61320
agacttgccc	aagactgaac	ctctaaatgc	aactcatatt	gaaattcaac	tctgctccaa	61380
agcatgttac	tttaaccctt	gtgcttttac	agctggctac	tctcccttta	tggtcacacg	61440
gggatgaagc	acgggggggag	gaaagccaga	ctgtctcact	cctgggttca	tcttgggaca	61500
caggacacca	gccagctgg	aggtgaggga	gctttaatca	gaggggaggg	aggaaggcat	61560
tctcaacccc	ttctgtacta	gggaggtcag	cagaagaaaa	taattcaatg	ttctaaagcc	61620
atttttttct	ccagcattcc	tccaattcat	agatcttcat	atgggattag	gggctcagag	61680
aggggtgaaa	caagaactct	attttttttg	agtgtggtat	agagaaggga	tgctacttct	61740
ctaaggtcac	atagtaagtt	gagaaagaga	gagaaatcaa	actcaggttc	atttcaacta	61800
ttgttcacac	agaatctgtt	gatttcaaag	atgggtgact	atgggttcat	ccctgtggtg	61860
agtgtgtgta	ggatgcagct	gaggtggaac	tttctactct	tgccctcttg	gactttatat	61920
tctgggtgtg	aaaggcattg	cttcccttat	ttcaatatata	acaacaaagg	gtaataatat	61980
ttcccattta	ttaagcattt	actaggtgtc	aggtactgtg	ctaaatgtta	ggtgaacttt	62040
gtcttgttcc	tcataaatct	ctgccgctgt	gggtgtgtac	tttgacagaa	gtttgacttc	62100
cagttccacg	agatcttctt	tgggggagta	atatcaagaa	ggggcacgaa	ggaagctgca	62160
gggtccttag	teccatcctg	tatctcgacc	taggcatgtt	tacattgttg	cattcactgt	62220
gaagtttccc	tgagcagtc	actctatagt	gtgctttata	ggagcacatt	gtacatccat	62280
tgaaaaattt	ttcttggccg	ggcacggtgg	ctcatgtctg	taatcccagc	actttgggag	62340

gccgagacag	gccgatcacc	tgaggtcggg	agtttgagac	ctgcctgacc	aacatggaga	62400
aaccccgctc	ctactaaaaa	tacaaaaaaa	ttagccgggt	gtggtggcac	atgcctgtaa	62460
tcccagctac	taggagggtt	gaggetggag	aatcgcttga	accctggagg	cgaaggttgc	62520
agtgaagcga	gatcgtgcc	ttgcactcca	gcctgggcaa	caagagcgaa	actccgtctc	62580
aaaagaaaga	aagagatttt	ttctttttct	taaaaagtaa	aaatcatgaa	ataaggggac	62640
tgggctaata	ttccaaaata	tgggtttgtg	tgtgaatttt	cctctccagt	aagatactaa	62700
ctaagctctg	tgaactgttt	tatctatggt	tctttatcat	tgaatccctg	gagttcctta	62760
caactgtgcg	agcacagagt	aggggctcaa	tcaacagtgc	actcattgct	ttttcataga	62820
caaggggccac	cctcaactcaa	ctcatgtgcc	aggcatagtt	ctgagagctt	tgtttaagct	62880
gatnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnggtcact	aatttggtat	63000
agagtattgt	tcatgtattt	cattttattt	tgtttctgat	ttttaaagat	tgttttactt	63060
gttttctctc	tattattatt	ttattttatt	tgtaaaacat	ttacatatca	gacatttaca	63120
ttttcccaaa	ggtaaaactg	tgaacaaga	tatattcaaa	gaagtttact	ttccctctct	63180
gtttcttgta	ccccttttcc	tcttctttag	gttaaccattt	ttattttttt	aaatataaac	63240
attgtgtagg	tgtatataca	tgtattagtc	tgttttcatg	ctgctgataa	agacctatct	63300
gagactggga	agaaaaagag	gtttaattgg	acttacagtt	ccacatggct	ggcaaggcct	63360
cagaatcatg	gcaggagggtg	aaaggcactt	cttacacggt	ggtggcaaga	gaaaaatgag	63420
gaagaatcaa	aagtgggaaac	ccctgataaa	cccatcagat	ctcgtgagat	ttattcacta	63480
tcacaagaat	agcgtgggaa	agactggccc	ccatgattca	gttaccctcc	cccactgggt	63540
cccacccaca	atacgtggga	attctgggag	atataattca	acgtgagatt	tgggtgggga	63600
cacagccaaa	ccatatcaat	acatttccct	ctcttttaga	taaaaggtag	tatactgtat	63660
acactattct	gcagagtttt	tttttttttt	gatgtaactc	tatcctgagg	gtgctctgta	63720
gcagggacct	ctcatgcctt	ttaaccactg	cctgggtctc	cattacatgg	ctgcagcata	63780
gttgccacag	cattcctgta	ctgatgacta	tttggattgt	ttccagtctt	ttgctattac	63840
cagtagtggt	acaaagagga	tctggctaca	tgttcagggt	ggggaggggc	agatgtgtag	63900
cctgtcagga	gggtatttga	gtaatccatg	actgagttaa	tggtagttta	aagctaggat	63960
gagtcagtgg	gggtggagag	aagtgggcac	atttgaatga	tatgtaggag	gtgaatgatc	64020
agcatttattg	atgagtttga	ggtggggcat	gtggggaaag	gattcgagga	tgactccag	64080
gtttctgttg	ggacagtggg	tggatagtgg	ctcctccctc	ttttccaatc	ttccttggcc	64140
cttcgctgac	ttctgttggg	ttggcctaca	gagagcttct	ttttcctctc	tgttcgcccc	64200
ggttcctcca	ctttggcggt	ggccctctgc	tcgacggtgc	cttcgctggc	cctgacatcc	64260
ctgctgtgcc	tgggcttcgc	cctctgtgcc	tcagtcccca	tcctccctct	ccagtaacctc	64320
acctctcatc	tgcgaagtgt	cagccgctcc	ttcctctatg	ggagcaacgc	ggccttccctc	64380
acccttgctg	aagtggcctt	ggggcgggct	ctgtggagac	ggacacactg	gggcaaagag	64440
aagctggagg	taaaagaaat	gggaggcaag	gcggggcctg	gaggcagtc	ggtgcgggag	64500
actgggtttct	ggggcaggtg	tggagggggt	gagaccagag	gtgggtggga	ggatagaaca	64560
ttctctcact	ttagccttta	catctgcggt	gccctctccc	tctgttttct	acctggtgaa	64620
ctcgtattca	tcctctgagg	cccacttctc	tttcagttct	ccagggaaga	aatggaaaag	64680
tgtcttccct	tctttgtgcc	cttagtactc	tagtcttact	tcctttgcta	gtgcgtgcat	64740
tgtctggcat	gccatccatt	tacatgcctg	tcttttcttt	cctgggtgcag	cctgcatgag	64800
ggtcctgtct	gtttttccag	ggccccgat	gtgccttctt	ctgggttctg	tgggtcaa	64860
gtctgagcag	agctgaagag	ggaaaggcca	gacaggtgtg	gttgaggggc	aggcctagga	64920
caggggagct	ggggacaagc	ggccgacagc	ccccagaggc	caggcttctg	cttgagggga	64980
gggtccctga	agctcaactg	aacccctctg	gtttctctcc	ccagtttccc	ttcagagcac	65040
tttggaagc	tctttgggct	ggtgatggcc	ttgtcggctg	tgggtgtctct	gctccagttc	65100
cccatcttca	ccctcatcaa	aggctccctt	cagaatgacc	cattttacgt	gagtactggg	65160
aggatgggga	tccttgccag	gaggcctggg	ccttaggcct	tggctgcccc	aaatctggct	65220
gtgatggcct	gggtatgtag	catggtgcag	cttcccaaag	ggtctgtgtt	attcaagtat	65280
ttggggcaaa	agtatttgtg	tgtgtgggga	aacagacatt	ctggagtagg	gtggggaatt	65340
ctcacgaaac	ttcaagcaaa	atcctgagac	ctcaaagggtg	tttccctgctt	gtggtgagtg	65400
cagccccacc	ctggcctctc	ccctaggccc	acacagggtt	tccacagttg	gccccaggga	65460
caggacctct	gtgctttcac	ctctgtgtcc	ttacacctgg	agggatgctc	tgaggctcctg	65520
ctctaggagg	tggctcgtgag	tctcctgctc	tttgagaaa	ctgaggtcca	aagaggttac	65580
ttacgtgttc	agaggcacca	gctaaggagc	aaaagtcaac	tttgaattct	gtgttttgac	65640
tactgcacag	ctctatttgc	ctcatttttt	atttttaaag	cagcaaatct	tagaatagga	65700
gtttaaatcc	atcaatttga	gaaaagaaag	actaaatggt	ttttgttttt	gttttgagga	65760
cacgatcttg	ctttgtcacc	caggctggag	tgcagtggca	caatctcggc	tcactgcagc	65820
ctcgatctcc	tggactcaag	cgatcctctc	atctcagcct	cctgagtagc	tgacactaca	65880
ggcatgtgcc	accatgccaa	gcttatttta	ttttattttt	ttgatagaca	ctggggttct	65940
gctatgttgc	ctgggctggg	tttgaattcc	tggcctcaag	cgatccaccc	gtctctgctc	66000
tccaaaatgc	tgtgattaca	ggcgtgaacc	actgtgcatg	gccaaaagag	taaaacttgaa	66060
atctgaggcg	aatgacttga	ttgtgacatc	aggtgacctc	gtaatcagct	gtgtatttcta	66120
gctggtgcct	ctaccagctt	cccattgtgac	cttgaacatg	tcattgaatg	ctcgctaggc	66180

ctctgtttct	ttatctgtga	aatgggcttg	atattcctcc	tctaccccaa	ccgatagtgc	66240
agaatgaaaa	gtaactgaaa	gtccttccct	cagggcacca	tagtgtctgg	gtgaaaaagta	66300
gaatataaac	tcggtagact	tctgggtccct	tcatttgtca	tggaaatggac	cagtgtctgc	66360
ttcatttgagc	aacagttctg	ttgttcagaa	ttcctggatt	tcacctcact	tctgtctctcc	66420
ctgcaggtga	atgtgatgtt	catgcttgcc	attctcttga	cattcttcca	cccccttctg	66480
gtatatcggg	aatgccgtac	ttggaaagaa	agtccctctg	caattgcata	gttcagaagc	66540
cctcactttt	cagccccgag	gatggttttg	ttcatcttcc	accacctttg	aggacctcgt	66600
gtcccaaaaag	actttgccta	tcccagcaaa	acacacacac	acacacacac	acacacaaaa	66660
taaagacaca	caaggacgtc	tgcgacgcaa	gaaaagaatc	tcagttgcc	agcagattga	66720
tatcacacag	actcaaaagca	aaggcatgtg	gaacttcttt	atttcaaaac	agaagtgtct	66780
ccttgcaactt	agccttgcca	gacccttgac	tccaggggag	atgacctggg	ggagggaagt	66840
tgtcaactat	ttcttttaggc	ctgtttggct	cogaagccta	tatgtgcctg	gatcctctgc	66900
cacgggttaa	attttcaggt	gaagagtga	gttgtcatgg	cctcagctat	gcttccctggc	66960
tctccctcaa	gagtgcagcc	ttggctagag	aactcacagc	tctgggaaaa	agaggagcag	67020
acagggttcc	ctgggcccag	tctcagccca	gccactgatg	ctggatgacc	ttggcctgac	67080
cctggctctgg	tctcagaatc	acttttccca	tctgtaaaaat	tgagatgaat	tttggtgttg	67140
aaagtctctc	ctggagcaga	tgtcctagaa	ggttttagga	atagtgcacg	agtcaggcca	67200
ccccaggggc	catgggagcc	agctgacctg	cttgaccgaa	ggatttctga	cagactatct	67260
ttggggatgt	tttcaagaag	ggatataaag	tatttacttt	gggcattta	aagaaaaatt	67320
cttcgggaa	taattttata	gaaaaataaa	gcttctgtgt	ctaaggcaac	tactgtttcc	67380
atctctctag	gctttgggcc	ggggctgtgt	gtgtgtgtgt	gtgtgtgttt	gtgtgtatgt	67440
gtatgtttct	gaggaggccc	taccctggca	tgagagggta	gggaatctgg	ctacacatct	67500
agtgtggcag	ctggaccag	aggtggggca	ggaaccctga	ctatgattca	ccccgttgtt	67560
cctgggatgt	gggcccagag	acttctctcc	ccaggaaacc	ctctgcttcc	tcttctctct	67620
cacatcctta	actaaactta	gcagaacctt	actcctcact	acacaccccc	agctagaagc	67680
gctgggtgaa	actagaaatt	octagtgtga	gtttcaattc	tgccctcag	cagctgggca	67740
agccccctaa	ccactctgag	tactagtctc	cccacctgca	aagtgcagtt	aatcatttct	67800
atctctgatg	gcgattgtga	gaatgtaaa	tcattgcac	tgccctagc	atggtaggag	67860
cacatgaggg	ttgtctctg	tgcttactca	tgaccttgg	ggaggacggg	ggcaagagg	67920
gagaagtga	gggtgcagga	ggagagatgg	caggtgggtg	ggatgggaga	atctggggca	67980
cacctgtctgt	ctcattccca	ccttgctagg	agagggacta	ggaaagaaca	gtgggaggca	68040
gggggatggg	ggtggaagcc	aggggggtgg	aggcaggttc	atccatccat	tcattcaaca	68100
aatgtttatt	gagcacctgc	cacgtgtcag	gccctgtcct	gggtgctggg	gctataaaga	68160
tgcaaaaggg	tctgaaaccc	agctcttctc	tcttctctgt	gatgtcgggg	tgtaatttcc	68220
aggggcccag	agcctgggtg	tgagggcgga	caccaaagtt	ctagtgggtg	ctatttagcag	68280
ggtttaaatc	taattggatg	atttggctct	gttaccctgc	tcaaaaagctt	tcagcagctc	68340
ccactgtctc	acaggacaaa	aatccagatg	ctagcctggc	attcaaggct	gtcactagt	68400
tgatctcaac	ctctccctct	ccctctttac	ctcctaccaa	cagcggggca	gagcccacc	68460
ctgtggacca	agattcccag	tctctgggtc	tgtgtgtgca	ccagttctct	tgctgtgggtg	68520
gctcaccctg	cctcagcttg	tgaaatccat	ctggtctgct	gggatcctgc	tcaaaatgtc	68580
atcttctcca	aaaatcatta	ctcaggcttt	ccagcatgtc	tgagtccctg	gcacttggct	68640
acacccttcc	tggtgactgg	catttgcttc	cacatcatga	ccctcccacc	ccttgctctg	68700
gcagcatact	ccaggaggca	aggtctgttc	tcgctgggct	ctaattaatc	tgtgtctacc	68760
atccacatgg	taccagctaa	ttcttgttga	atgaatgatc	ggtgaatgag	tggtattctg	68820
ttttggcctc	agaaccaatt	agaaggagcc	agaaaaacac	atgggggtgg	gggaggtgca	68880
gtgtgggtga	gtggaaaaaa	acccttctgg	aaatctcagc	tctgtcactt	actttgtcag	68940
ctctgtgact	ttggatggac	cacttctttg	tcagtatggt	gggagaaata	gacatgcctc	69000
tctgggctgt	tgtaaggatt	acaaattagg	tcgagtgtct	ggcatgtggt	gggttgaaac	69060
gatcacagct	agcattacag	atgatataat	aaagccaaaa	aaagatgcct	aatgtccacc	69120
agttgtgtaa	cggacaaagg	aatgtacca	tatttgggat	attatttggc	aatcaaaaaa	69180
agtactgaca	cctgtctaaa	cacggatgaa	tcttggaaaac	attagactaa	gtgaaagag	69240
ccagacacaa	gaaactgcta	atgattccat	ttaaataatga	aatatcgggc	cagggtgcag	69300
tggtctatgc	ctgtaatccc	agcactttgg	gatgccagg	tggtcagatc	acttgaggcc	69360
aggagtctgt	gaccagcctg	gccaacatgg	cgaacccccg	tctctactaa	aaattagccg	69420
agtgtagtgg	catgcacctg	taatccagc	tacttgggtg	gctgaggcac	aagaattggg	69480
tgagcctggc	aggtggaggt	tgcatgagc	caagatcgtg	ccactgcact	ccagcctgga	69540
tgacacagtg	aggttccgtc	tcaaaaaaaa	aaaaaaaaaa	ggaaaaagaa	aaaaagaaat	69600
ttccagaata	ggccaatctg	tagaggcaga	aagtagattc	atgattgggt	aggcctgggt	69660
gtggaggcca	tggttagtga	tggtcaatgg	ggagggggtt	tcttttgggg	tgatgaaat	69720
gggtggactt	atggatatgt	aattatacct	caataaaact	gttattttaa	ggaagaaaag	69780
atgcctggat	tccccaggaa	gtgtacagta	gacttctgtg	agaatcagaa	atgatttctg	69840
gggaagatgg	gcgagaggag	agtaagtggg	agaagtgaac	acgtgcgcaa	ctctcatcgt	69900
tctgccctga	gagccttctc	cctgcaactt	tatttattta	tttattttga	aacaggttct	69960
cactctgtta	ccctggctgg	agtgcagtgg	tgtgatctca	gctcactgca	gcctcgacct	70020

gccaggetca	agcaatcctc	ctgtttgagc	tcctgagtag	ctgggactac	aggcgcatgc	70080
caccacatct	ggctaatctt	ttattttatt	atttatttat	ttatagagat	tggggagtct	70140
cactctgttg	ctcaggctgg	tgtaaaacgc	ctggactcaa	gtgatccctc	caccttggcc	70200
tcccaaagtg	ttgggattat	gggtgtgagc	cactgtacct	ggcacctcct	gcaacttctt	70260
cctcaagtgg	aaccaatgag	gaagcaagca	actcagagct	ttcacaagtt	ttgatttcaa	70320
tcagcaacgg	gcttccaatg	caacccttct	ctcctgtaac	cagcctcagt	agagaggaaac	70380
tggaggtgaa	ttggccccc	tcacaccccc	acagtgcaca	gctgggccc	tccatcaggg	70440
ggagaaacaca	tgccgtgtaa	gggacagcca	acagcataaa	atagggaattg	tgtgatgatc	70500
ccttttaagc	ctattcagcc	cagggaaagt	catatgatca	gccccatttc	atagatgaag	70560
aaagtcaggt	tcaccattta	gcacattgtg	gggctggtat	ttaaaccagg	tctgtctggc	70620
tcccaaggtc	acattcattt	agacattacc	tttactttac	atttcttctt	cttttcttct	70680
tcttcttctt	cttcttcttc	ttcttcttct	tcttcttctt	cttcttcttc	ttcttcttct	70740
tcttcttctt	cttcttcttc	ttcttcttct	tcttcttctt	cttcttcttc	cttcttcttc	70800
tcttcttctt	cttcttcttc	ttcttcttct	cttcttcttc	ttcttcttct	tcttcttctt	70860
cttcttcttc	ttcttctctt	tcttcttctt	tcttcttctt	cttttttttt	tgagggtggg	70920
tcttcttctt	tgtaatgcag	atcatcatac	ctaaatgcag	ccttgaactc	cttgaactc	70980
ctggccttaa	gcaatccccc	tgccctcgcc	tccaaaagtg	ccaagatttc	aggcatgagc	71040
caccatggcc	agcctgcatt	tattctcttg	taagaaagat	atcattttaa	acagacgaga	71100
aaataaagag	ggacatgaaa	aagacgcact	accattaat	ggaccactca	gagataatca	71160
tggttaacat	gttgggtatg	tccctcccg	catttgactg	gatgtatgtg	ataattttaa	71220
tgatctcata	agcctttctt	tatgtaatca	aatagtagcc	aaaaacatga	ttttaaatgg	71280
ctgctcaca	ccccatctcg	tggttctgcc	acgccttgtt	tatccccatc	cacccccatc	71340
tccctttccc	cttccctgcc	tgtgtggggg	tcctagatga	cggtagccca	gagggcagcc	71400
ttggtcagca	gattggagag	tgcaataaat	aaaaacactc	agaaggcgag	ctgttgtcaa	71460
tggggcttat	cacaaaagag	caccttggga	tattccagag	aatgacctca	taccgcttaa	71520
gtactatcca	taactgtgtg	ctaactgtac	tttagctgaa	ggtgtctggc	ggtcctggcc	71580
aggtgctgct	aagaacactt	ctattctgtg	agaatcagag	atgatttcta	gggaaaatgg	71640
gcagagagga	gtaagcagga	gaaacaaccc	acaggcacag	ctctcatctt	tctgcccgtg	71700
gagccttctt	cctgccacgt	ggtttgtgtt	gtttgtgtgt	ttgtttgttt	cagataggg	71760
ctcactctgt	cacccaggct	ggagtgtagt	ggcaagatca	tggctcactg	aagcctcgac	71820
ctcccaggct	caagcagtc	tccccaat	caaagcttgg	agtgatggtc	ccagtgggta	71880
tgtctaggag	ccctttttcc	tgccagcccc	tcaggggatt	gatgactctc	aaatgcttca	71940
ggtgtgacat	gggcacagca	gtgagtcatt	cctctgacat	tctttgggaa	gaacattttc	72000
catccaggct	tccaggcata	agatccagtc	ctctggtgat	aaggagttca	cagacaggac	72060
aatgtctgag	tgtatcttaa	acccaggacc	atggcttgtg	ttcacaccag	accctccagg	72120
gatttttagg	tgttttgggt	gtttgtttgt	ttgtttgttt	gttttttgag	acagagtctc	72180
tctctgtcgc	caggctggag	tgcatgtggc	cgatctcagc	tactgcaac	cttcgcctcc	72240
cggttcaagc	gatttctctg	tctcagcctc	ctgagtagct	gggactacag	gtgtgcacca	72300
ccacaccggg	ctaatttttg	tatttttaat	agagactgtg	tttcaccatg	ttggacagga	72360
tggctctgat	ctcttgacct	cgtgatcctc	ccgcctcggc	ctcccaaat	actgggatta	72420
cagggcatgag	ccaccgtggc	ccgcccaatt	ttgagttttt	atgttctaat	cccaaacatc	72480
tgtctcacagg	ccctcagca	tattctttcc	tgggtccagt	gtcacctccc	aggcctgcag	72540
gctggctaga	gcagtagggt	gtgtgggaaa	gctctgggct	ttgcaggcac	tgtacagctg	72600
tgtgacctta	accaccctga	acctcagttt	cctcacctgt	aatggaaata	ggtaccacgg	72660
cagtttgttg	caaggactag	agagtaacct	tgggaataaa	aggtagcagc	agcttgggct	72720
ctggagatgg	actgtccaag	accaaactcc	agttcctccc	cacacaagct	ctggcactta	72780
gattccctggt	acctccgctg	cttcatctgt	aaaaatggagt	aacaatagga	atactttata	72840
gagttgttaag	gattgagtgg	ctggatgaac	gtcaagcact	tcaaggggga	cctggcatgt	72900
agtgaagtgt	caatataaac	cacctgggctt	gtagcagggt	tgctgtgtgt	ggctgcaggt	72960
gttattagta	acatctgtgt	gcccttcaga	gcgtgcacca	cacttcacac	cttgtggagt	73020
ctggaaatgcc	actatttatag	ttcaggatag	aaaacctccc	tgcaagcact	cgctttagct	73080
tgtctccacc	gaacaaaaca	acacaagttc	tttattactt	ggaatgggaa	aacttcaaa	73140
gcaaaaaaaa	aaaaagactt	tgcagttacc	ccaaatctta	agccaaagtc	aatgaaaaat	73200
atcaatcttc	atattcaatt	tttgcgatac	ttttgtctcc	ccagcagtc	atggagagaa	73260
tccaagcaca	cagaaatgtc	aattaccagg	ggcagggcta	tgaattcctt	tcagagccct	73320
gggctgggga	agagtgcagg	cagacagatc	tgggtcctgt	tatcacgttc	ttagattggg	73380
tgtccttgta	ggagtcatga	agcatcttag	tgcctttgtt	tgctacctat	aatgcctacc	73440
tcagagagta	ataaggataa	gtaaggctct	acgtgaaaag	tgctcgcccc	tggcacatag	73500
taggtccttc	attaatggca	gctaactaatt	tttattacat	acgcacaaatc	acattacagg	73560
tcaagtacgc	tacatgacag	tgaacagttt	ttttgttttg	tttgttttga	gacagagtct	73620
cgctctgtca	cccaggctgg	agtgcagtg	cacgatcttg	gctcacccga	acttctgcct	73680
tcaagcaatt	ctcctgtctc	agcctcccga	gtagctggga	ttacaggcat	gtgccaccac	73740
gccagcta	tttttttggt	atttttagta	gagacggggt	ttcacccat	tggccagact	73800
ggtctcaaac	tcctgacctt	gtgatctgcc	caactcagac	tcccaaagt	ctgggattac	73860

tggcatgagc	caccgcacct	ggctgtgaaa	cagttttatt	gtgtttctgt	ggaatgtgtc	73920
ctacccaacc	tatagctaac	tcctatagtt	ccctcagttc	tcagctcaga	tatccctccc	73980
tttctgtact	gttacctagt	actggttttc	atagcaccag	gtacctctct	ggcatagagc	74040
ttgtcacagt	tgcaagttta	tgtaccatca	taggatttta	aaaatattca	gttgtgtctt	74100
ccattaggct	ttcatttggg	aactccacgc	aggcagcagc	tgatatattt	gtattgccta	74160
ctgtatccct	agaactttgt	accctactta	gcacagaatg	gaggctcagt	aaatactgga	74220
catgagagag	agagagagag	agagaggaga	gggagagaga	gagagagaga	ttcaacctac	74280
aatcccgact	ctgagcttct	agttccctga	tggtgaggac	tgtgatgtgt	ctcacacggt	74340
aatgagcaact	tatgcagaag	aggctcagaa	aatttctcct	catggccaac	ggaagactta	74400
gagttctttt	ccaagctcca	cgttttgctg	gcatgcaaaa	tttggactat	cacttaagtt	74460
ttccaagcct	tgctttttct	atccctaaca	taggacaata	ttcagcattg	ttgtttgttt	74520
gttgggggca	ccatgtttca	ggcacttagt	agattattgt	accaccacat	ttcaattggt	74580
cctcctcaag	ccctgcaaca	tctgtgaggt	ggtcatcctt	aacaactcac	agatgagcaa	74640
caggagactg	gggggatgag	ggaactgccca	aggaggtcca	gcttatgggc	agcagagcca	74700
agaatggaac	cagggtcttt	tatttttttt	tttttttatt	tttatttttt	aaccagggtc	74760
ttttaacatc	cgaggaccac	attctttgtg	ctttccaaat	catcacctgc	cccatgcaac	74820
ttacagggta	agttacatta	aacaacgtat	gtaaatggct	ttgtgctagt	tattcaccac	74880
cacaggggaa	gtgagtcacg	gacaagagtg	cagccgctcc	attcggatcc	tggctctgac	74940
acttacctgg	aaaatgactt	aaccattccc	aggatcagct	gtttgtctgt	aatttaggta	75000
gtttaatggc	acttgtgtcc	tagagttggt	tagaaggttg	aataatatgg	agcacttaac	75060
atacttagca	cctagaaaca	cttcctaata	attagttgct	gctgtgtgta	tcgttattaa	75120
aaattctgcc	taagatctca	tttcagggag	cccaactcaa	tctttgacaa	gcttaaacaa	75180
aaattgcttt	tcttcattta	ttcacttaca	cagcaaacat	gaattgagcc	tgtactgtgt	75240
ttccagaact	gtgcaggacc	agagaggcac	aggtgaagga	agcaaggctc	tggctctact	75300
ggggaacac	caagaagatt	gctacaatga	ggtgggaaga	gggctggact	agagagaagc	75360
cctgattagt	gtccttgcta	cctttctctg	ggagagccaa	ggcaggcttc	ctggaagagg	75420
tgatccttgg	ctgaaacttc	gatgaagaaa	aggaaagagc	gcagtgggta	gggaggaaag	75480
ggcattctgg	gcagatgaaa	tgacatgtga	caaaaatagg	gtgatcannn	nnnnnnnnnn	75540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	75600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	75660
tgctcctaaa	aaatataaaa	aaaaaagtc	agaagaacat	ttgggtctca	ctctgtggcc	75720
caggctggag	tatagtggca	caatcatagc	tcactgcaca	ttcaaaactc	tggcctcaag	75780
tgatcctcct	gccttagcct	tgaataaagc	ttggattaca	gatgagccac	cacaccagc	75840
cagaccattt	ttcataatag	ccaaaatgtg	aaaacaaccc	aaatctccat	caactgacaa	75900
atggataaat	agaatggtgg	ttgatccata	caatggagta	tttactcagc	ataaaaaga	75960
agtcctgata	catgctacaa	ggatgaacct	cgaaaacatt	atgctaagtg	aaagcagcca	76020
atcacaaaag	gctacatatt	acaagattcc	atttaaatga	aatgttcaga	ataggtaaat	76080
ctaactttta	tcacaggcaa	agctatgaca	ggaaatagat	gagtgggtgc	ctagtgtgtg	76140
ggggcagagc	tggggggtgag	gcgagtgagt	actgctaagt	gtacagagtt	acttttgggg	76200
ataaaagaa	tggtctgaaa	tggactctgg	tgatggttgc	actactctga	acatactaaa	76260
actgttaaat	tataatactg	aaatgggtga	cttgtgaggg	atggaaatta	tatcttaata	76320
aagctgtttt	acataattta	aatgcagggt	aatgcagggt	gagggatgag	ccctctaaa	76380
agaagcagga	gtttgaggag	gttctaaata	ttgtgtgtgt	ggtagctgag	catataaatt	76440
tgtcagacct	catcaaaatg	tatgatgtaa	tcttcaagaa	agttgatttt	aaagaaacac	76500
caccagcacc	agggtggagaa	ggcaggaaga	agttacacaa	ggggtagggc	aagagtgggt	76560
gctcatgtct	ataatcccag	cactgcggga	ggccgagctg	gggtgggtgac	ttgaggtcag	76620
gtgttcgaga	ccagcctggc	caacatggtg	aaagcccgtc	tctactaaaa	atacaaaaat	76680
tagccagacg	tgctcgcgtg	aacccagggg	gagaagggtg	cagtgcgcga	agatcatgcc	76740
aatgcactcc	agcctgggtg	acagagttag	actctgtctc	aaaaaaaaaa	aaaaaagtta	76800
cataggggac	agtggcagggt	gtcaagggtc	ggcagggtct	ctcctatctc	caggataaac	76860
tcatagggga	cttagatgcc	atgtgggtcc	ctaataagcc	tccacttggt	tcttgacgcc	76920
actcttatgt	gtatcatctc	atgtcaggtc	tcttctctcc	aaccacccca	gccatcccag	76980
cctggctgcc	aacccacact	cctccagccc	ctgtcacccc	ataattgggg	ccaggaggca	77040
tgggagagtc	gccatctctc	ggtgccatct	gttgcatctt	tacagataac	catggctgga	77100
tgccgcagat	cctgggggtgg	agcagccgct	gttcagagca	gtgatcaaga	cctcccccac	77160
tccacccctc	aaggaatcgg	ttttcttcca	tagccacatc	aggtgctgtg	caggaaaggag	77220
ttgaaacgag	aagccaggag	caacgagaag	gacactaaca	tttattaagc	actgcagact	77280
ctcacagcac	tccacaggaa	tcgatatatt	tatcccattt	ctgaagacca	ggcaactgaa	77340
gctcaatggt	taaggaactc	accgaagtca	ccaactgata	aaagtgatgg	aagctgggat	77400
tcaaatccaa	gctaaaactc	cttcccaagt	tactccacaa	cacagaggtt	ggggaagggt	77460
gataaaaaga	gaggggagcc	caattccact	tccacccagc	tcctgagggc	gagcttgtca	77520
gcacagctct	ctccttccca	gaataggaag	atacccatca	gaggcaagtc	ctagacacca	77580
gcagtggtaa	ctccctgccc	caaggcagct	gcagacagcc	tatggctgta	gttaactgctc	77640
ccaaagagtg	ttagaattcc	cactcccagc	ttcggggcca	ctcacacaa	gtgattgaag	77700

tggaaccag	agactetcca	caatgccctc	ctagagtaaa	tgaggctatg	taactttgtc	77760
caaatgagta	atttgaaaac	ctgggggctc	ccagctcctg	aaaagggaag	gatgtggggc	77820
cctttatatt	catactccac	tttgtgcagc	tctcccttgt	cttatgatag	ccctattaaag	77880
aaattctctt	cccagcacgt	ctccttcaaa	gagctctaga	cctgaggctg	tcagaggctt	77940
aggactctgc	ctattagtcc	cagggctctgg	atgaccagca	ggacacctgg	cattcagtga	78000
ccactggatt	agataaatga	aacagtgggc	agagtgcac	ccaatctccc	cctgaagttt	78060
gaagaggctg	agaagtgagg	ctgtccaact	gctgacctg	ctttctgtcc	acctggccac	78120
ctaacctttt	ctgggttcca	cctgcccctt	tgccatccct	ccccccagcc	caccagccc	78180
attttcaggc	atacctgggc	acgtgctgga	atagaagccc	tcgttcttca	gaatgatcaa	78240
cagggagccc	cagcccagga	gtacagcaga	gaagaagagg	ttctccagca	cagccgtgca	78300
ggccatccac	cagcgctccc	ggtacgcctg	ttgcagcgtg	ggggccatgc	tgcccccgag	78360
cctgcacaga	aacagagcgc	tgggtgaagg	gccccccagt	ggccccaggg	aagggtcctg	78420
catcatgggtg	gcacccgaga	cctctcgggc	cagccccgca	ggagccctc	atggaggccc	78480
catagagccc	tgggtctccc	agccggtgcc	aaggagctgg	ctccgcgcgc	actagcagtg	78540
ccagaggctg	acgcggcacg	gggtctccgc	tgagccacta	tcggaaacaa	ggaaggtcct	78600
gtctgcgcgc	tgcaagcttcc	tagcaggctg	ccgggttctc	tcacccaggc	cagggcgctc	78660
agggccgggc	tgctggggag	aaagtccgca	tctgccagg	tccccagagg	acagcaaggg	78720
gcagagcgcg	ctctgaagca	ccgcgggccc	atgtccggac	tctcgcgcga	ggaagagccc	78780
ctagaagctg	gcaggaagaa	gggcaagtcc	aaggctaccc	tacgacccca	tcttccagtt	78840
gccccctcaa	gacctctcct	tccctctggg	gcccggcgac	agcaagccct	ccccctttcc	78900
gtatcagggtg	acccacgacc	ctacagctctc	tcgggccaag	ccaacagctg	ccacgtggag	78960
ggagaccagc	gacgggctct	cctcgggtcc	ctcctcccc	gcgcgcccc	cactcactcc	79020
gcagggtctg	gggcaaccag	ctttgcacct	cggaaaccgc	ttgccccct	ccagccccgg	79080
gagggggctc	ggactctggc	aggaagtctg	gcggctgctg	actttataag	ggcagcggtg	79140
gcggtggggc	tggcggggcg	gtgtgtttac	caaaggagag	gaaagagccc	cagctcccc	79200
cgcccgcgcc	gctgcagcct	cgccgggagg	agagggaacg	cgggcagcgc	gggggcgggg	79260
agcgacaact	gggatgagac	cgaggaaagc	ggagaggaga	agggcaagaa	agaccagag	79320
agaggggagg	aagtaccagt	cacttctctc	agggggaact	ggtattctca	tctgtgaaac	79380
ggggcttttg	gttcaagcgc	ccgctgggac	ccgctgggac	tctggcaaac	gcgcagctct	79440
aagcagagga	agtgcagcga	gcggggaccc	gggaggaaga	gaagagtcgg	aggggtcaga	79500
gaaaagaaaa	gggaaggacg	cgcttgcgca	gatgggacac	tgtgcgcggg	gaccgcgggc	79560
gcaagtacac	gtcttctctt	gggaagcctg	gcagtgtcgg	cgggagccgg	cctcggtgtc	79620
tctcagccga	cgcatagccg	gagaccctac	gcgcgcccc	tccccgccca	cgctgctcac	79680
ctccggtcac	cggcaaatga	gcagccagca	gctgcggacg	cctccgggag	cgcaacgctt	79740
tcgcggcgcg	tcggaggtcc	cgtgggcccc	gccctgagcc	gcgccggcgc	tggggtcttc	79800
tctgcgtgca	ggacccggcc	gccacggagc	ttcagcctga	cagcccggtg	gcctcgcttc	79860
cgctgtctcc	tcggaagaa	cgggggaact	gggaacccgc	cgggcgcag	aggtctcgga	79920
agctgggctt	ggatgaagtg	gatctgcgga	gttgatagtt	gtattttac	gcgtccggag	79980
ctgcgcgccg	aggtgggggc	gggggtctcc	ttcttttccc	ctccccctag	gtcaggtttc	80040
acgcgcacgt	gactcgcccc	ctggtccccg	acactctccc	tctggcacag	ccccagcacc	80100
tacatttcca	ccctggaccc	ccatcttctc	ccccagccc	ccagactaac	atcaggcagc	80160
gccctctgtg	tccttgttca	aaacaaagt	cgattcggtc	gaagccgact	gaccgcgatt	80220
cagggccgccc	ttgggtgggg	ttttgaactg	tgacgttgga	agcagtggtt	tccgagaggc	80280
agagtggcac	gggtttcttt	ggagttagtc	agatcgaggt	ctgagtcctg	acttttttaac	80340
tgactaccct	gggttaccta	gggcaagtta	cctctctgag	cctcagcttc	ctcctcttta	80400
aatctcggtta	aaatggaacc	tacctaaact	cccaaaggaa	tcgcgattgt	gatgcaggta	80460
aatgctaag	catagcattt	ggcatagtaa	gcataatgtt	aattgttgct	gctgtcatta	80520
tttcagaaga	cctggtgac	ggatgcttcc	agatcaacaa	ttgattgact	ccaggtaaat	80580
ctctcagcct	ccctgagcct	cagtatccct	atctgtaaaa	tagactacta	tggtgtggag	80640
taatgagaag	taatctcatt	acatgtgagt	ttaatgtgt	gttaagagt	ctgctaattg	80700
atgctgagct	taataccctag	gtgatgggtt	gataggtgca	ataaacacc	atggcataca	80760
tttacctacg	taacaaacct	gcacattctg	cacatgtacc	ccagaactta	aaataaaaa	80820
aaaatttttt	taaaaaaaga	gtgatactgg	tgccaggtg	tggtggttca	tgccgtgta	80880
cccagaactt	tgggaggcca	aggcaggagg	atcgcttgag	ctcaggagtt	cgagaccaac	80940
ctggacaaca	tggtgaaacc	ccgtctctac	aaaaaagaaa	aaaaaatagc	caggcatggt	81000
ggtgtgcacc	tgcagtctca	gctaccagc	aggctgaagt	gggaggatca	ctgagctgga	81060
gagatggagg	ctgcagttag	ccaagatcat	gccactacac	tccagcctgg	gtgacagagt	81120
aagactctgt	ctcaaaaaa	aaacaagaat	gactacagaa	agctccaaga	aggcctcaga	81180
taaaaggga	ccctgaaca	gatgagccac	caagccaaga	gaggaaacta	tggtaccat	81240
agacagggca	ctttccaaaa	taaaaatact	gttattaatt	cctcaagaca	tcatgggtcc	81300
atttaaacct	catagctttt	cacagaggtg	gaaactgcag	gcttgaagct	ggagcaagg	81360
tagaggttag	atgcagagtc	aggtcggcct	ggcatttaag	tacggctcct	tccattcttc	81420
ccagaaggag	aatggcaaga	gcaaaggcct	agctgtggga	atggcacaa	gagttctcgg	81480
tggccaaagc	acatgtcagg	ctctgatggt	ttaacttctt	aaaatgcaat	actgcctccc	81540

agaacttcca	gatcaaggtc	aaactcctca	gctctacaca	gggggaccta	gagtcaactt	81600
tctaagctag	gagagtcag	gatccctttg	agaatacaaa	agacagtggg	cgcgggtggca	81660
gtggctcatg	cctgtaatcc	caacattttg	ggaggctgag	gcaggaggat	cacttgagcc	81720
caggagttca	agacctgctt	ggtcaacata	gtgagacccc	tattttctaca	aaaaattcag	81780
ctgagcatgg	tggcatgtgc	ctgtagtctc	agttaactggg	gaggctgaag	taggatgatc	81840
cctgagcctg	ggaggtccag	gaagctggag	tgagccgaca	tctcgccact	gcactccagc	81900
ctgggtgaca	gagaccctgt	ctcaaaaaaa	aaaaaaaata	gaagaaatat	gttattgatc	81960
tactcttgac	aaaaatgctt	gtgtgaacat	ggacacacac	actcatcaac	attcacattt	82020
caagggtttc	attgaccctt	tccatgagge	tctagtgggc	catggacccc	catggctgga	82080
acacttgctc	ttcctcatct	caacccacat	ttccatggag	ttggactgtc	tgctgcatga	82140
ggacacaggg	ctcattttgt	gtgttcattc	actgctgtgt	atcccagcac	ccagaacagc	82200
acctcaccta	agggggcactc	agcacatgtg	cagtgaagag	tcagtcagct	ggtttcacac	82260
ctcccagtc	ttgcacctgc	tattccttct	tgtgggaatg	acagatttcc	ttcattttctt	82320
tttttttttt	ttttgacaga	ttccagctct	gttgcccag	ttggagtaca	gtggcacgat	82380
ctcagctcac	tgcacacctt	gcctcccagg	ttcaagcaat	tctcatgcct	cagcctccca	82440
agtagctggg	attacaggtg	cacaccacca	cctgtgagct	gatatttttt	tcttttcttt	82500
tcttttttcc	tgagacagag	tctcactctg	ttgcccaggc	tggagtgcag	tggcgtgatc	82560
teggctcact	gcaagctcca	cctcccgggt	tcaagtgtat	ctcctgcctc	agcctcccaa	82620
gtagctgaga	ctacaggcgc	gcaccaccat	gcctggctaa	tttttgtatt	ttttagtaga	82680
ggcgggggtt	caccatattg	gacaggctgg	tctcgaaatc	ctgacctcgt	gatccgcca	82740
cgttggcctc	ccaaggtgct	gagattacag	gtgtgagcca	ctgcaactcg	ccattttttg	82800
tattttttta	gtagagatgg	ggtttcacca	tgttggccag	gctggtcttg	aaactctggc	82860
ctcacgtgat	ccacccacct	tgcccaccca	aagtgttggg	attacaggca	tgaaccactg	82920
cgctcagcct	cctcttctcat	ttctaattgt	ctcatccttc	acaactcage	tcaagtttca	82980
cttctctctg	gaagctctac	tctaggctgg	attcaggggc	ttgtccacat	acccacaaaa	83040
tactctgctt	acctctatgg	aagtccccac	actgatctag	aataatcagc	ttagttttct	83100
gcccccatcc	cgccccatga	gatgtacatc	ttgtgggggc	aggaaccacc	acgtggtagg	83160
tgattttgtg	gcctgctgcc	tatcacaggg	cctggcgctt	aataagcttg	cggcccaact	83220
ttgttgaata	aatgaaaaagg	gaatggtggg	aaaggaaagt	gaaaaggtag	gctaaaaatca	83280
gttttgaatt	acctctggga	ggccaaaggc	tttcagtcct	gcagggtagg	taacaggaaa	83340
ctctctggatt	ttgttttctt	ttggttttgt	ttgtttttta	tgaagggtag	cgttatcgtc	83400
aggtttttgt	gtttaattaa	tggagcatat	attggaagg	acagagacct	taaagcagtt	83460
aggagaccac	cataatagtt	cacatttttg	agccataaaa	aggaatgagg	ccaggcatgg	83520
tggctcactc	ctgtaatctt	atcacttcgg	gaggttgagg	caggcggatc	acctgaggtc	83580
aggagtttga	gaccagcctc	accaacatgg	agaaacccca	tctctactaa	aaatacaaaa	83640
ttatccaggc	gtggtggtag	atgcctgtaa	tccagctac	tcaggaggct	gaggcaggag	83700
aatagcttga	atctgggagg	cagaggttgc	ggtgagccga	gatcgtgcca	ttgcatgtca	83760
ggtacatgga	tgaagctgga	agccatcatc	ctcagcaaac	taacacagga	acagaaaacc	83820
aaacacccga	tgttctcact	cataagtagg	agctgaacat	tgaaaaacaca	tgacacagga	83880
ggggaacatc	acacactagg	gcccgttggg	gagtgggggt	tggggggtaa	ggggagggaa	83940
cttagaggac	gggacaatag	gtgcagcaaa	ccaccatgac	acacgtatac	atatgtgaca	84000
aacctgcaca	ttctgcacat	ggatcctggt	ttgtttttaag	aagaaataaa	gaaaaaacca	84060
agaagaaaca	aacaaacaaa	aataattccc	attttaaaca	ataaaaaata	ggccaggcat	84120
ggtgactcag	gtctataatc	ccaacacttt	gggaggccaa	cgcgggcaga	tctcttgagc	84180
ccaggagtgc	aaggccaagc	tgggcaacat	ggcaaaaacc	tgtctctaca	aaaaatataa	84240
aacaaacaaa	caaaatagcc	aggagtgggt	gtgcattgct	gtcatcccag	ctactcaggt	84300
ggctgaggtg	ggagaatcac	ttaagcctgg	gaggcggagg	tagcagttag	ctgagatcgt	84360
gccactgcac	tccacctgga	gcaacagagc	aagattttgt	ctctaaataa	ataaaataaa	84420
taataaaaaa	cagagaagag	gaaagacacc	tgagatatat	ttccatatct	gaatcaatag	84480
gatttatcaa	cgttctcctc	tacccccaac	actaatcctt	tccataaact	tggtctcctg	84540
acactactca	taggttaagt	ataacagcat	tatcacattg	gctgtcatgt	gggctcctgg	84600
ctagaggctg	cttcacagct	taatggacaa	gagcactgag	acagggtggg	tctaaatcct	84660
ggctctgcag	ctgattatct	gtgtgatctt	gtccaaatca	ctccatctca	tgagcctcac	84720
tcttctagtc	tgtaagtgtc	tgaataataa	agtatccaat	tcaattcatt	atttaattgaa	84780
ttattttagc	taacaaatag	ctattataaa	tatttaggct	gggcacagtg	gtcaccgctt	84840
gtaatcccag	cacttttggga	ggccaagggt	ggcagatcac	ctgagtcagg	agtttgagac	84900
cagcctgacc	aacatggtga	aaccccgctt	ctactaaaaa	tacaaaaatt	agctgggtgt	84960
ggtggtcagt	gcctgtaatc	ccagctactc	aggaggctga	ggcaggagaa	cgcttgaacc	85020
caggagacag	aggctgcagt	gagccaagat	cgtgccactg	cactctagcc	tgagcaacag	85080
agcaagactc	tgtctcaaaa	aaaaaaaaaa	aatctctgca	tgaagaatgt	acataaaatg	85140
gtgcagccat	ttcggaaaaa	agtttggcag	gtcctcaaat	agttaaacat	agagttacca	85200
ctatagccca	gcaattccac	tcctaaatat	actacaacca	agagaattga	gaatatattgt	85260
taacacaaaa	atgtgtatac	aagtatttat	agctgtatta	ttcattacag	ctaaaaagtg	85320
caaacatccc	agcagtcocat	cagctgatga	acggagaaac	aaaatgtggt	atacccatat	85380

aattgtcatat	tattttggcca	taaaaaggaa	gtactgatac	atgctacaac	atggatgaac	85440
cttgataaatg	ttattcttaag	tgaaagaaac	cagacacaaa	agaccacata	ttgtatgact	85500
gcatttatat	gaagtgccca	gaataggcaa	atccacagag	acagaaagta	gatttagtgg	85560
tgccagagac	tgaggggagg	agataatggg	aaatgtggaa	tgactgctaa	tggtatgggg	85620
tttcttcttg	gggtaaatgaa	aatgtttgtac	aattagataa	tggtgatcat	tgtaaaactt	85680
tgtgaatata	caacatgctg	aattttatac	tttattatat	tttatttttt	ttgagacaag	85740
gtctcgtctt	gtcaccagag	ctggagtgca	gtggcacgat	ctcagctcac	tgcaatctct	85800
ctgcctccca	ggctcaagca	atcctcctgc	ctcagcctcc	tgagtacctg	acactacagc	85860
atgtgctacc	atgcctggac	aatttttgca	tttttagtag	agacaggggt	tcgctatggt	85920
gcccagggctg	attttgaact	cctggactca	agtgtctccg	ccacctcagc	ctcccaagt	85980
gctaggatta	cagggtgaag	ccaccactcc	cggcctaaat	tgtattcttt	aaaagactga	86040
attgtatggt	gtgcgaatta	tatctcaatt	taaaaaaaac	aaaacaaaac	aaaaaaaaaa	86100
cctttgcgtg	tgtcaggcac	tagggattcg	atgctgaata	agacacagac	cctaccctca	86160
gagaacacag	agccccagcag	gagagagtc	cagatgaatc	aagtgttaca	tcattctatag	86220
gaagcgccat	ggaagaaaga	catggtgcca	tgagaacata	cgcttagaga	agggaaattt	86280
atctagactg	gggctcaggg	aggaatcttt	cagggtgatg	cttgtgtctc	gagttttcca	86340
tgtcagaatc	agtagaattt	atcaatcctc	cagaggagga	aacagcaaat	gaaaaatctt	86400
acaacaggag	gatgcggaga	cattccgaga	gctgatcaag	ggctgggtg	aacaaagcac	86460
ataggatgca	gagcctgtgg	tgtgaggttg	cagctggaaa	ggtaaaacac	taattacatt	86520
ggatcttctg	agacaataaa	gagtatgcaa	taatctcaaa	cgaccgaaac	tgaccttctc	86580
cctccctaac	ttgcttgctt	ccactgttgc	ccgtatcata	aaagcaccac	cctcttctac	86640
ccagtggctt	aagacacgaa	actcaagtca	tcccaggctt	tctcccacc	tcactctcca	86700
catccagcct	atcagcgagc	ttgtgggtct	taccacgtaa	agacttctca	tctccagcta	86760
ctaccatccc	ccaagcccag	atcaccatca	gctcaggcct	ggactcctgc	aacctttcta	86820
accgggtcct	cccaatccta	cccccgcaac	atgaccccaa	tagcccatca	gaatggacta	86880
atcgagatgt	agatttgatc	aggccacatc	ccttgaaagg	cttctctgtg	ccctcgggga	86940
aatgcacaaa	ctcccaatga	tggcccctga	gtcctgtgct	atctgggtct	gccctctgcc	87000
ctctgtgtct	ttgccatggg	aacctccttc	acacccatta	atactccatg	ctctctccta	87060
cctcaagttc	ttcctgggct	ggaacattct	ctgcactagc	ctagccaact	aaccttttag	87120
atcttttgtt	tgtttgtttg	tttgtttgtt	tgtttttgag	acagtcttgc	tctgttgcca	87180
ggctggagtg	caatggtgca	atctatctcg	gtcactgca	acctctgcct	gccgggttca	87240
agcaattctt	ctgccttagc	gtcctgagta	gctgagacta	taggcacctc	ccatcacgcc	87300
cggctaattt	ttgtattttc	agtggaggtg	ggttttcacc	atgttggcca	ggctgggtct	87360
gaactcctgg	cctcaaatga	ccaccctcct	cggcctccta	aagtgtctgg	attacaagca	87420
tgagccaact	tgcccaggca	acacttcaga	tcttaatgat	catttccctt	aagtgcctga	87480
cctctttag	taactagcct	gactccagca	atgaatcctt	ttgcaatgta	acctatataa	87540
catctgagtt	tccttttgat	aaaactcacc	atataattgt	tcctctgaca	gttcagaggg	87600
caagggcctt	tgcaccactt	cctcaccact	atcctctcac	cacttaacac	agaaactcac	87660
acccaccatg	cctcctgcct	gacaaattcc	taaccatcct	tcaaattctca	ctcacctatt	87720
accttctggg	aggcagtcct	ccctgagcac	caagacaatg	ggacacattc	ctttatcac	87780
cctgctgaac	atctcttttt	tgaggggagg	gtagagatga	gtgtctcact	atgctgcca	87840
ggctgacctc	aaactcctgg	cctcaagcga	tcctcctgcc	ttggcctccc	aaaatgctgg	87900
gattacaggg	atgagccact	gtacctgacc	gcaactgggt	tagnnnnnnn	nnnnnnnnnn	87960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88020
nnnnnnnnnn	nnnnnnnnnn	nnncctgggc	aacaaagacc	cttctctcac	aaaaaaaaaa	88080
aaaaaaaaaa	aaaaaaaaaa	aaattatttt	aaattagcca	gacatggtag	tgcatgcctg	88140
tagtccaagc	tacttgggag	gctgaaggga	gaggatcact	tgagcccagg	agggttaggg	88200
tgagtgagc	cgtgatcgta	ccactgtact	ccagcttggg	caacagagtg	agacctcacc	88260
cctaaaaata	aagaagaaaa	tatggcaatt	tgactgtaca	tctctaattg	gatatactct	88320
aaggatgaga	aaggaaataag	gaaggacaga	aaaaaggaaa	caaagaagta	gcaacagtat	88380
tattaagtat	tattgttatc	aagtaacatc	aataattggt	aaaccagtaa	ttatatttaa	88440
aatactatat	atgtgtatgt	acatttcat	atgcatatgt	taggaaccaa	gtttatcaga	88500
ggaagagaaa	gggtacaaa	tgtaaaatca	aggaaataaa	aatttgaata	aaaatatcag	88560
tattagcaatt	tattgttatc	tttcttataa	aaaaattata	tatatgttaa	ctctatccaa	88620
aacccaaaaa	cagtgcacac	ccaggagcaa	taaaaaacct	cagcatccag	actgtagtct	88680
ctaccatttc	caattaaaga	aaaccagggc	tagttgggaa	aaatgacaat	ttcatgtcta	88740
gggcaagaaa	cacacctagt	gaaatggacc	tgaacattta	attgtgttag	aaagtaagga	88800
aactctctag	aaataatgtg	atttcatcta	aaagacacag	attctgggct	ggtaaaagtt	88860
tcaatggcca	aagggtgagc	aatttgagca	tcaagaagaa	tcatgacaga	acagattaaa	88920
acatgtcaaa	tatattttaa	aatgaaatat	tataaaagaa	acaattagta	gccatccctg	88980
aagggtcact	gggcaccaac	tcatatttca	aactggtaaa	taaatgtgta	agccaagcat	89040
ttatttctgg	gtaacaaaaa	agtaaggaa	gtttttcttt	ctagaagaat	tctagtgtat	89100
aaaagtagaa	gatagaaata	gaaaatcatc	cttttggcca	ggggcagtg	ctgtaatccc	89160
agaacttttg	gaggccaagg	cagggtgatc	actggagatc	aggagttaga	gaccagcctg	89220

gccaacatgg	tgaaccccca	tctctactaa	aaatacaaaa	attagccagg	tagtgggtgc	89280
ctgtaatgcc	agctactcgg	gatgctgagg	caggagaatc	gcttgaacct	gggaggcgga	89340
ggctgcagta	ggctgagatt	acgccactgc	actccagcct	aggcaacaca	gtgagactct	89400
gtctcaaaaa	aaaaaaaaaga	aagaaaaaga	aatcatcctt	ttgcgatcct	aatgaaataa	89460
tgggacctagg	cattgatcat	taatggctcc	taaaatcact	aagtatatgg	ttgatgggaa	89520
actttatagt	ggatggatca	gactcgcaat	gtctaaacca	gttgatcaat	cttaacatcg	89580
taacaagaca	acagacacca	ggggctgctg	acaggagaac	agaggaaacc	catagctcta	89640
ccactgagtt	attcacggca	aaaaaaaaaa	aaaaaaaaatt	aaactgcgtt	tcctccaagc	89700
ttctaactcct	gttgtttaca	ggaaataccc	aaggaaagga	atacttttaa	atgacacatt	89760
aaaaacaacgc	caaattccaaa	atatggggaa	atgacccagt	ttcttcaaca	aataaacaag	89820
aaaaggtagg	ggggaggact	gttctagatt	ttaaaaagta	tagaagacac	agcaacccaa	89880
tacactgcat	ggaccaggca	tggatcctaa	ttggaacaaa	ccaactgtaa	aaggatgtat	89940
ttgaaatgat	tggagggaatt	tgaacagtga	ctgcacagta	gatgatatga	agaaattatt	90000
gttatTTTTT	aggtgtgatc	atgattttat	gggtgatgtt	aagtaaaaga	ggccttattt	90060
gttagagata	catgtacggg	tatagagaaa	tatttacgga	tgaatgata	cgatgtctga	90120
gatttgcttt	gaaaactcta	gcaggtgtgg	gagaagcagg	tgcatgggtg	ggggaaggga	90180
tagatgaaat	aagtatgcaa	aatgttagtc	tacttttgtc	cctcctgacc	cagcagggtta	90240
aaatacctca	gcatacctct	actcctccaa	ccaggtccaa	ggatcaggcc	aaaactccct	90300
gatgtgttaa	acagcctgac	cccttcttac	ctctctctct	ccagccactt	ccctaagatt	90360
ccccagtgtc	ctgtgcccta	gccagcccca	ctcatctgcc	cagattcctc	aatgtttcac	90420
tctctcattc	accattttga	ccccactgca	ctcctgggcc	actctccaag	gccccgcctc	90480
ttcatctcct	ccctccttac	tcactcctca	ggtcttggct	taggtgccat	tcctccagg	90540
aagccttccc	tgacaccaat	cccatcctca	cctagaacag	attatgtgcg	cttctttgtg	90600
ccccccatgg	ccccctgtgg	gtttgtctca	cggattataa	ctgcctgact	acctgccttt	90660
ctccaccctc	tagactgaga	acaccttgag	aaaaagaaca	catctatctt	gtctgtcatt	90720
gaatecctgg	tgtctggcac	catggctgac	acataactat	cactcagtga	ctagtgtttt	90780
aatgaatgaa	tgagtgcac	tagacaggg	taagaacaaa	agagaagacc	aggcatgggtg	90840
cttcacgcct	gttatcccag	cattttggga	ggctgaggcg	ggcagatcac	ccgaggtcag	90900
gagttcaaga	ccagcctgac	atgggtgaac	cccgtctcta	ctaaaaagac	aaaaattagt	90960
gggggatgg	ggcacacgcc	tgtaatccca	gctaactggg	aggctgaggg	aggagaatct	91020
cttgaaccca	ggaggtgcag	gctgcaatga	tctgagatca	caccactgca	ctccagcctg	91080
ggcaacagag	taagactcta	tcacaaaaaa	aaaaaaaaaa	aaaaaaaaaa	gagcgagaga	91140
agatgtcatg	gggtaaatga	agacctccct	tcctgggtcc	ctgaccagcc	cctgccctcc	91200
cccgcatctc	acctgtcttt	cttgcttctt	tctggtactt	ctgtttaagc	cggtccatga	91260
gcaggccatt	ccagggggca	cacagcactc	cgaactgagt	gaaggcaaa	gcatttgtgt	91320
aggtyctgac	tgacagaagg	agagagaggt	tggttgatga	gaagtttcca	aaactccctt	91380
ccaggcagg	actctcccac	cttacccctg	tctgcatgtc	cctctccccc	acaccatcag	91440
acctcctctc	gggtgtgtaca	gccctgctgg	gaggctctgt	gttcccagct	gggacatgca	91500
gatgggctac	ctcccagccc	taaccacatac	ctcgtgccat	gtccccaccg	gccatgttgg	91560
tcagcaagga	gttgagagtg	ccaatgaaga	ggtagtgcc	caactgtatc	acagacagcc	91620
acaccaggtg	ccaggcaaa	cgccgagaga	aagcgtagct	ccagaaggag	cggagttcct	91680
gcttctgccc	tgccctggg	gtctctaattg	gggagaggag	gatctgggag	tgaaattaca	91740
ggaaagtgga	caggtaggat	ggggagtgtg	gaggcttcaa	tggaaatttt	cagatccggg	91800
cccacctttc	accttggct	cccagaactc	acccggcggt	ctctgcacct	gcctcggact	91860
cacacatccc	cacctccctg	ctttgtcatg	ctggccctac	caccttggat	gacctctgt	91920
gttcttctct	attgaaatcc	gatctgtctc	tcacagcctg	gtcaatgaca	cttcttgacg	91980
taataccttc	ctgatctttc	tcagcgagaa	aggtgaaagg	aacgacaagg	agaggagaaa	92040
gtcagaagg	agaggagaat	gagtggtgat	actctgttct	aacctgtctc	tcagcacctc	92100
cctttctttt	gtaccagta	tcctgagttt	ctttgggaaa	tcttctctca	ccctaattct	92160
catggtccag	atgggacat	gaattcagtg	ttctgttctt	ctctcaagg	taaccaatga	92220
gatggttccc	cctaacaagg	cagaccggcc	atgagtttag	attaggatgg	acttaattcta	92280
aaataggctc	tacaccctgg	caagttcaat	gtcctccctt	gattttggaa	gcttcccaga	92340
acctattctc	tctttcttta	aaataaataa	ataaatacat	gttttggatc	caattgtcag	92400
atggtaaaaa	taaaaaacaaa	aaaatcaatt	ttattctgta	tattttaagat	atacaatatg	92460
aggctcatagg	atacatatag	ctactaagat	ggttactaca	gttaagcaaa	ttaacatatc	92520
catcatctca	catttctacc	tgttttgtga	caagagcagt	taaaatctac	ttgttttagga	92580
aagtcccaaa	cacaatgcag	tttgatgacc	tacagtcttc	gtgctgtgca	ttagatctct	92640
aggcctgttc	atcctgctca	tctgctcctt	tttgccttc	gacctgcate	tcctcatctc	92700
tctccacccc	cgttcttatt	tctactgtag	ctagctgcgg	tttgtgatgt	gtgtaaccaa	92760
agacgcagaa	cagagaggaa	ggaaaggga	cagtgataga	gttgggacaa	taagagaggg	92820
cggaccagg	agacctggag	aaatgggggc	actgtaccag	acttagtgca	atggcatcac	92880
agaagagggg	agaaccgagg	agtgggggga	agggaaaggca	acccatggca	ggcgggcttc	92940
aaggggtggg	gaagtgatag	gatgcgaat	agagaaaaga	gggacagaaa	agagacgaaa	93000
gccctggacc	ctccattaag	tgagagggtt	gggaagatgc	ctaaggccct	ttttctgtcc	93060

tgcccttccct	gattctgggt	ccctggggga	getctggagg	tgaggggcca	ggaaaggcac	93120
aaggagagggc	ttgggtcttg	aggagagatg	ggtttagccag	cagggctcac	cttccttcgc	93180
tgaaaggaaac	tcctttgact	gtagctccct	gttttcatgc	tcagctgttt	ccttctcttc	93240
ctttgtgggtg	ccattcccag	ggcacaggct	atggaaacaa	aagccccacc	agcaaggcca	93300
aggactgtga	gccgaacctg	agactcagac	tgaggggaaat	agcatgggtga	atccccatt	93360
ccaccgcact	ttggaatcac	cttttagcca	ctctgatgcc	caggttgag	accagaccag	93420
ttaaatacaga	atgtctggag	gtgagagcca	ggcttccttt	tctaagatct	ctatgtgaat	93480
ctagtgtattc	taataagcag	caaagttag	gaagcatgaa	aagagttagg	caggcccagg	93540
ttcaaatccc	agctctgcct	cttcctagca	acagaaagat	ggctcagact	taaccttct	93600
gagcctcatt	ttttgcatth	agaaaatgga	gataaggata	tctcagagga	ttattgtgag	93660
gatgaaatca	gagagcacat	ggggtctgac	aattagtaag	tgagcagcaa	aggaatgccc	93720
ttcctctact	ccttctggca	aatgactgca	aaaatgatca	catthcttca	cctcctctgt	93780
atthccccca	atthgaatga	gactgcagct	ctatthcccc	atgccctgaa	tctgggcccag	93840
ccttgtgaac	tgcttcagcc	aaaagaatgc	agcagaagtg	gctgtgcca	ttccaagctt	93900
aaatctcaag	aacgcttgtg	catthctgca	ctcttcaga	acctgaaat	cacggtgtga	93960
atgagcccac	gctggcttgc	tgaggatga	cagccacgtg	acccaggcat	cctgtcact	94020
ccaaacctat	gtgagtggg	ccatcctagc	atagccagcc	cccatgtaat	cctccaaatg	94080
atcagatgta	tgaatgagcc	ctgtcaaat	catctacatc	tgccctgat	cagcggaact	94140
agccagctac	ccacagactt	gtgaaaaata	ataaatgctt	aacatthtag	gctgtgagt	94200
tttgagatag	tttgttatgc	agcaatagct	aacagatgca	ctgctccagt	cctcctctc	94260
tcctgtgata	ggtttgctth	acctgttcca	ttccacccta	gggccaatga	ggggtctctg	94320
cccacaatca	ccagatagtc	cttaccata	gctgtagtgtg	gggggcagt	ggtatgggat	94380
gtgcccccg	ggcatcagga	ggaaagtgcg	tgctacatgc	caggtactgc	agacagagat	94440
gaagatgaag	gaggccctga	ggctgatgcc	tttttcataa	agaagctgca	gaaggagag	94500
gaaaaagtca	gtgtcacacc	cacgttcata	gcagcactat	tcacaatagc	caaaggatgg	94560
aagcaaaacta	agggctccatc	agcagatgaa	cagctaaaca	taatgtgatc	tatacacaca	94620
atggaatatt	attcagcctt	aaaaaaggaa	agaggcaacc	atgctggctc	acacctacaa	94680
ttccagcact	ttgggacgca	cgaggatcac	ctgagccag	ttcaagacca	gccttgacaa	94740
catagtctac	ccctcacccc	ttctctagaa	aattthtatt	taattagctg	ggtgtggtgg	94800
catacacctg	tagtgccagc	tactcaggag	gctgagtggg	aggatttctt	gagcccagaa	94860
gtttgagget	gcagtgagtc	atgactgggc	cactgcaccc	cagcctggac	aatgaaacat	94920
gaccttgcct	ccaaataaaa	aaaaaaagg	aaaggaaaga	aattctgaca	catgctgcaa	94980
catggatgaa	ctttaagagc	actatgcagg	gccaaagctca	gtggttccctg	cctgtaattc	95040
tagtgcttta	gaagaccaag	acaggaggat	tgcttgagtc	caggagcttg	agaccagcct	95100
gggaaacagc	agacacctcat	ctctactaaa	aataaataaa	taaatcagct	gggcgtgatg	95160
gtgcacgcct	gtaattccag	ctacttggga	ggctgaggtg	agaggatatg	attacatgat	95220
tacatgcctg	taatcccagt	actttgggag	gctgaggcaa	gcagatcacc	tgaggccagg	95280
agttccagac	cagcctggcc	aacctgggtga	aaccccgctc	ctactgaaaa	tacaaaaatt	95340
agtggggcat	ggtggcacgc	acctgtaatc	ccagctactc	gggtggttaa	ggcaggagaa	95400
tcgcttgaac	ccgggaggcg	gaggttgcag	tgagccaaaga	tcctgccacc	gcactccagc	95460
ctggggcaaca	gagcgagact	ctgtctcaaa	aaaaaaaaaa	ggttaagata	gtaaatttta	95520
tgttatgtat	atthttatgc	atacaaaaac	atcagcagaa	gaggcagggg	ctggaaccct	95580
gtthttctaag	gagtcctagt	acaagccatc	acctactatc	ctgtaagctg	attaggagaa	95640
cctggtacac	acatgcccc	acccaccca	agacacaccc	ggcagtagag	gagtcctcat	95700
acgaccatc	cccacagccg	gtggagcctc	ctcgtgtggc	ttcccagaga	tcttctagcc	95760
cagtgccctt	tttcccccaa	cgacagcaaa	ggcctthtgt	tcaagaaaa	ttttacacaa	95820
aaattcatct	tacaaaacac	accaatgggg	agcttgccag	tcactctcct	ctttattctc	95880
cttgggtgact	ggatgacat	caaaagagaat	ccctaagtcc	ctcaacagct	cagthttgaa	95940
accaccgacc	tagcccaacc	tcctcccat	ttacagagag	tgacgttgag	gtccagagag	96000
gtgcagtga	ttgctcaata	aattgacaga	gtaagcagca	gcaaagttag	attaaactaa	96060
gaattcctgt	tcctgctccc	tttcccttc	caactctaga	gagacaggag	agaggctggg	96120
catggtggct	catgctgtga	atthcagcac	tttgggaggt	caaggaaagg	ggattacttg	96180
aggtcaagag	ttcaagacca	gcctggccaa	catggcgaaa	cccatctct	actaaaaata	96240
caaaaaataag	ctgactgtgg	tggcacgcct	atagtcacag	ctactcagga	ggctgaggca	96300
ggagaattgc	ttaaacccac	taggcagaga	ttgcagttag	ccaagatccc	accaatgcac	96360
tcacgcttg	gagacaaagt	gagactccaa	ctcaataata	aaaaaaaaaa	aaaaagagag	96420
aggaaagaaa	gatgaggcag	ccatctgggt	tctccagggg	aaggaggagg	aaccagaaa	96480
gtgactctta	tgccaggagt	agaaaggctt	gagtgccctca	ggggtcag	ctctgcataa	96540
ccctccaaac	ctccaaagct	tatgggacta	agctagactc	atgtctgggt	ggtgactgcc	96600
agagatccct	ttctctgccc	ccataacctg	caggcagtg	caactgcctg	tgacctaa	96660
ctaagcccag	agagaagtcc	caggttggat	ggcttgagat	ccacactctt	cccttctctt	96720
cactcagcca	tctgtggtgt	gctggcttta	gtctccagc	ttgctgcctc	ataattgaag	96780
catggttgcc	acaactccag	ctatcacatc	ctcacaccac	aacattcaat	gaggaagact	96840
ttgtttttac	tctgttttca	ccttgcttca	gggaagaaaa	gtccccctga	atcttccact	96900

atatacactc	cctttatctc	attaaaaag	actggatcat	atgctgacct	ccacctatca	96960
ctagcaacgg	gtaaatggat	tgccatgggt	ggctttaatc	aatcaggatt	catccccctg	97020
gctaagcggg	tcactgccc	gataaaactg	ttcgcaatga	ataagacaga	atgggtgtgt	97080
attgacctct	aatagccttg	gcaacagttc	atccccctg	accccaacat	cagccactgg	97140
gacagctgga	caagcctctg	tgctgtcccc	tgctgtaccc	actagccact	tgccaccttc	97200
ttgtccaaac	tagaagctca	cagcagcaaa	cgccccactc	taaaggtccc	ccagcctcta	97260
cccaacactg	gcccagcac	attatgacca	ctgcacaaaa	agcttgggca	agtctgaaga	97320
aggggcttag	cgggtacaag	ctcaggctct	agaaccgaca	agcctgggtt	caagttccag	97380
tatcatggct	actagctgca	gaaccttcaa	caagcttttt	aacctcagag	actcaaatgc	97440
ttcatctgta	aaatgggggt	aacacagtac	ctacctcacc	gagttgatgg	agacaaaata	97500
tgcaggttca	caagacaagt	gtctggcata	tacaagtgcc	cagtgaatgt	aggctgttgc	97560
tataatttacc	ttaataataa	ggaagactgc	cgagggaagag	tcaaagtctc	cattgtacag	97620
agtgatgatg	gtcgaacggg	gttggccaaa	taggttccca	atctggggat	gataggacta	97680
gcctggatga	cttattttatt	catgaaacag	atacttctctg	agcaccacag	atgtggcaga	97740
ccctcctttat	acccaaactc	accctccacc	gctagagctc	ccacctcagc	ttggggccaa	97800
cccatctgag	gcagccaatt	atagaaaagg	gtctctcctt	ccctccacct	tcccgccacc	97860
ctgcccagtg	cctgggatta	gggaaggctc	ccacctgcag	gttgggtgatg	agaaacagga	97920
ttcccccaat	ggtgagcatt	ggcatggcca	ggaagagcag	cacggctgag	cctggaccat	97980
caaaagtcaga	ggtagggtgg	tgctcatagt	caacccaaac	atggagcccc	aaactctgct	98040
cccacctgct	ccaaattccc	aacaatcctg	gtatccaggc	cccaattcta	gccagcgttc	98100
cagcgtcctt	caagggtttt	taggataccg	gccaaggctt	ccccagatat	ctctgtggaa	98160
gtcttctgag	ccaccttctt	cccaaccaaa	gttgggtcctc	agtctgtggc	aggccaggaa	98220
gtctacagac	agaggcagag	ctctaagtga	agccacctct	ctcttccctc	agtaaaccac	98280
aagctgcctc	tccctttcat	ccttgacact	cctggaaaag	aagaccctgg	actcaggtcc	98340
ctggctcaac	cctctagccc	attccctaatt	tcattgggtatt	ggccttgagc	ttcaatcatc	98400
tgtttaatgg	gaacaaacagt	tcctgtctctt	cctgtctcag	gtgctatgag	aactgagtga	98460
gaaaaaggacc	atggtctttt	ctttgttcac	taaaactctga	gcacttcttt	ggtgccaggc	98520
attgtgcttg	gcactggaaa	tgcaagatga	atcagatagt	ccttgccctt	aaatagactg	98580
acatgcacac	aaatggttat	aacaggctcg	gtaagtgtga	gaccacagca	aaaaagctca	98640
agagctgggg	taggggaacc	cttgacaaat	tcttccctccc	caaacagac	ttctgcccac	98700
cattattctg	gccacaacct	atgcctgtcc	tattatttgc	taaaatgitt	taagtgtgact	98760
cactttttatc	caaaaagtat	ctatttttaa	aggacacttt	atatcactac	tgtagatgaa	98820
aacactggca	ttacttgtca	tgaatagaaa	gtaactgtca	aaataaatac	aatgaaagga	98880
aaacaatggt	attcaattgt	agctggatgc	atttgacctt	agaatgttca	aagcctaaga	98940
cctgtctcttc	ctcatcagtg	taaaatcaca	ctggccccac	atgaagacat	tctttcatga	99000
aatcagaagg	actgaaagag	aaataaaaag	ggaatagctg	ttctaccagg	tgattttgatg	99060
tttgttagtg	tagttcacgt	agtatgctg	tgccccaac	atcctcttaa	ctaccgtgct	99120
ataccttaag	aagcactgcc	taagactaat	tttagagtat	tcacacagtt	taccattcaa	99180
tttctgtctt	tataaaatgt	acatctctcc	tactactaaa	ggttgagac	tcctttcaca	99240
atagagctct	tatgggtcga	atgctttttt	caaaactgaa	aagccctata	ttatggagga	99300
agaggaggt	gtttgctcag	acgatttgca	ggcacgagtc	aaacattacc	cagccaccac	99360
ctccacattc	agttgcttaa	aaatcattta	caggctttta	gagtagatga	tgctggtttg	99420
ataagagag	tggtttgaaa	taattggttt	gaggtgctgg	gccaetctcat	gagatctgtg	99480
tgaacaaaga	cactcagcct	ctgtgtttgc	ccagcatgag	tgacagacaat	ctcatgatgc	99540
tgtcagcttt	agcatagctt	acacacacaa	gagtaatgta	ctttctttcc	taaaccacaaa	99600
attgagccac	gggtctaaca	ctaggaagga	atattgggag	gcactctctg	gccaccataa	99660
ccaaggcaat	gacagaaaga	agagtggagg	atcaggaggc	ctgcacatca	ggccacctc	99720
ccactgtctt	tctctgtggc	catggacatg	tctttgcaag	gggtcctgct	gtggcttcag	99780
tttctcctct	gtgtaatggg	tggaagggtg	gtggaaaaata	aaccagattg	gagttccaga	99840
cttaacagac	tggtgaaatt	ttaaaacaaa	gatttttgagt	acaatagggt	tgtcaacttt	99900
taccctgcta	agtaaggata	tttgcaaaat	ggtcattcat	ataatcattt	cattaaaaag	99960
agaaagagaa	cattttaaca	cataggagaa	ggatgtaaag	gttttttgtt	gttgttttgt	100020
tttttagggg	tttttgggtt	tttttgggta	gagtctcact	ttgtcaccac	ggctggagtg	100080
caatgggtgtg	atcttggctc	actgcaacct	ctgcctcctg	ggttcaggca	attctcctgc	100140
ctcagcctcc	tgagtggctg	ggattacagg	cgcgaccac	catccagcta	atttttgtat	100200
tttttttagt	agagaaggga	tttcaaccatg	ttggccaggc	tggtettgaa	ctcctgacct	100260
cagatgatcc	acccgcctca	gcctcccaaa	gtgctgggat	tacaggagtg	agccactgca	100320
cctggccaga	tgtaaaagttt	tgaataaatt	ctactctctg	aagtaatccc	tctccatcat	100380
ccttgctttt	cacattttct	caataaactg	ttttcacaga	ccagcaatag	ctcaagatcc	100440
ttccaggatt	ctttcaagct	gcagatctct	gaataacccat	gtggtctgta	tatcttgccc	100500
atagccctct	gctcacacct	gccccagcca	ccagggtgct	ctgagcttgc	atccctccc	100560
cccacctgac	agcactcacc	tgcaagagtg	aaggctatga	tgagtgtggc	ggtggtgtag	100620
aaaaatctga	aacacataac	aggaaaagca	gaatattgtc	aaggaggag	aaacctggga	100680
gaaaaaacat	gattctgctc	agccagccca	caagtgtagg	acttgaccgc	accctcagcc	100740

tgggatgcaa	cgggcactga	tgccctctgag	ccccaggctc	aaaaccaggc	gcaagaagcc	100800
gcgatgagat	tgagatgtgg	tcctgacctc	atggacagtg	cattttgtctc	attctgaggc	100860
ccaaggctag	catggaaagt	cttggacaat	gagctcagct	gacgatgtga	ttggccttggg	100920
acttagccag	gacagaatgg	gcaaaagcga	ggtcctccca	cctggaagcc	ccaacagccc	100980
aacccttgg	agaaaaggggt	tagtgccctgg	tctgcaaato	aaaggccttga	gttctaactc	101040
ctcctcactc	tgtgaccttg	ggcaaggcgc	tgctccttctc	tgggcctcag	gagccttttc	101100
tataaaaaga	aatgatcgga	ctgatctagc	tcagagtgtc	atgatttcag	gactacagtc	101160
ccaaggttat	caggctccct	tagcatttgg	gggtccttga	aggcatggag	taaaaaaa	101220
aaagcaatat	cctaaggctg	gagaagagg	aggggacaaa	ggaaggggag	gaaaggggag	101280
gtagcaggga	gccaaaggacc	aagaaggact	gaggtacagt	cattctgcat	ccaaaggcct	101340
aaattgtaa	ggactggcct	tactctggct	gtttccggaa	aggcaggccc	agccagccct	101400
cccgctctctc	tctctgacag	ccaatctcac	atgtgcctcc	ctgggagcac	ctgctctgag	101460
ctgtatcagc	ccccagcagg	ccgctgatta	ccactgagcc	tggccacaga	gcacgagatt	101520
aggatgcagc	aacacactgt	gtgtgagatc	acgtcccgaa	ccttctgact	catctgcaca	101580
ggaaaacccc	ccagctctccc	ctccagtcag	aagggaacctg	aaattccacc	agtggcaata	101640
ccaaagaaac	ttcctattag	ctaagccct	aggagtgat	tggctgttgg	ggcggggagg	101700
gggggcggtg	aggaggatga	ggatgaagcc	tgggcaacct	ggatgtgagg	ctgtgcaggg	101760
gatgaggaca	aggatccttg	gggtgaagga	agagaagagc	aatgttaggt	tttgctaatt	101820
ttgtaacccct	ggctccaagc	cagcccttac	aggaagtac	cctggcctcc	ggctcaattc	101880
agcacgtgat	aggggaagcca	catttatgca	gagcaggga	cgaggtaagg	aatggaagt	101940
ggggctgtgg	tgaagtgggc	aagtctagag	agagtccgc	tgccctgggc	tgctcctaac	102000
agctgctggg	agcgaagctgc	aggtgtgggtg	cctggcaggg	tgcccgggct	gtctgactct	102060
ggatttcaact	ccaagctagg	ctgctgcctg	aaggattcct	cttaccacc	tttgccctggg	102120
ctggcctttg	ggacttacat	ggctatgagg	cgtgccacgg	tggctttgaa	ccggtcaag	102180
atgtagccag	tggggaatgt	ctgaagtgtg	ttcatgaagg	accccagggt	gaagatgagt	102240
gagaacctct	catcttgggc	tttgagctct	ggagtagaaa	aaaggtctcc	catgcatccc	102300
agccttcctg	ccaaatgagc	acacaggctg	ggctccctcc	cacctcagac	agcttgtcgg	102360
tcgcaaacctt	gtcccttaag	ctgagttgaa	atgtggctgc	ccctaattac	ccctcaggag	102420
ctgtgctctc	acttccagga	tcaagtgggg	tcagtgctgc	tgagagctgc	tgaccttcc	102480
tctcatcata	gaaagagggg	tgggcagggg	gcagagtcct	tctgtctcct	tgccaccacg	102540
tgggagccag	acttaacttc	cttagaaaag	tcctccctgc	ccttaccagc	ctgcccgtg	102600
gcattgcccag	tcggcccagc	atctgggtcca	cacagatcct	taaagtaate	ttcattcttg	102660
aagacaaaca	ctagtgaagg	ccagccaaag	aggacgccag	caaagcccag	gcattccagc	102720
agcccagtc	gcagtgtggc	caagtgcagg	ggcaggccct	ggcccgccat	gagcagaagt	102780
ggagtggatc	ttcaaatccc	actttgtcct	cctggacgga	tcacaggcgc	cgtaagcctg	102840
gcgtttgagc	acttggaaaa	ttcctctggc	aagccaagcc	cttcccttcc	cgtagctctc	102900
tgggtgttttc	aggcctgggc	aaaaaccatc	agcgggtgat	tctctggatc	ctgtagaata	102960
aagatagagg	ctgctggaag	cggagccctg	cgggaaaggg	aaaggtagac	tagagttatt	103020
tgtgaggtgc	attaagaggg	aggatgatca	tggccgctgg	cagcaaatgt	ggggaataaa	103080
tactccaata	catcatctta	ggcactgcat	ttgagtaacc	acgtggcaag	tagagaggca	103140
ggtcttgatg	gccacctgga	gtcacagggt	agataaacga	cttaccocat	agctcccg	103200
gagcaggtgg	agaagcgagg	ctcctgcgct	cgaattctga	ataccgtccc	ctaaaataat	103260
gacagcaact	caaccagggtg	cagaggcagg	gagatttcat	acagaagaca	caaactcccc	103320
ctgccaaagt	ggtgttatct	tcagtttact	gacaatgaaa	caaaagctcc	catggatttc	103380
aggaacttgc	ccaaggtcac	agggctagtt	tagtcacgac	gcaggccatt	ctactgccag	103440
aaataccccc	aactcccatg	accctcgctc	aggactcgca	aaactgggtc	ccgccgccct	103500
tcctcgcatc	aaacttctacc	aggaaagcct	ccgggggccc	ctccccgcca	gcctccgcac	103560
cccgctcccag	cctgcggcct	gccctccccg	cagaggagcc	cgaggggcca	ggccgcgctc	103620
ggcgccccat	ggcgcccgaa	aggggaccct	tcgcccacc	cgctctgctc	ggccgggggc	103680
tctcgcgcgc	ctttccgcac	gggcccaggt	cgcatctcg	cctctcgag	cccctcccag	103740
tcctctgctc	ggcctccgccc	cctcctgccc	gcccggaagg	ggctggggca	gacctccac	103800
tctccatcac	ttccttcttc	ttttcccttg	ctcacagcct	cccgccctcc	ttttaacctc	103860
ccctcttgaa	acttctccct	ctagaacccc	ctagaacccc	agcgggtgtc	ttccctccct	103920
cctcgtctgc	tttcagcctc	ccagccccc	tgccctctgc	tcccctaacc	aagttagttg	103980
aatgtctgta	ctcgtcagg	cccacctagg	gaaaatgtca	caccagcac	ccagaggaca	104040
cacagacagc	acatgagggc	atagggacac	acacactcta	tttgtgcatt	ttgccttgac	104100
cgctgggttg	gcagggaaca	tatttttctc	atttgtcac	cagcttaacc	gtctctccca	104160
gtttcacact	cccagagctg	ccaaaaaat	cccaaccaca	gaatcaggaa	gccaagaacc	104220
aggactgagg	gcttttcaga	aacctcccc	tggaggactg	cccatattt	tactcccaa	104280
aaacccctta	gatgactccc	tgccctaccc	ccgccccca	ggttctgaaa	gagccttccc	104340
ccagactgc	attgattaac	cattcattgc	ccattttttt	attaatcaaa	gacatatata	104400
attgctcatc	ggagcttgtg	atcagcgtga	ggccttacta	agcagctgcc	ttactatcct	104460
tcagccccag	agcacgtgag	ctgacgtctt	ctttggcctg	tgtggccgtt	tccttgccaa	104520
aaagctcagtt	tggggagagc	ttcttgcgta	ttagatgcag	tctgcagact	cccaacccca	104580

gctacctgga	ccccctgagg	gcccaggaac	tccagctatt	ccaagccac	tcctcttttt	104640
tttaagagga	agaaatagag	gttacgatag	gggacagcca	gaactgagga	ttttccagct	104700
caccaccaaa	gcacaaaaga	taaaagtctg	caaccacct	agtacttga	ctgaatggag	104760
gaaggggtggc	tggggtcctg	tacccaagc	tactcactag	ttatacaacc	tgaggcaagc	104820
tctttggctg	ccccacctgt	aagacgagga	caatagtacc	tttaattatag	gaattgtcat	104880
aaaagaagta	taagatgggt	gtatgaggtc	cctgcatggc	gcaggtgcta	taggcagatt	104940
gtagggtagt	agattttcta	gtctgcagtt	atgtagacag	agccagagaa	gcagctctgg	105000
ggagggaattt	caaaggaact	tgccccaggt	cattctacaa	agctgcagta	ccttcccaac	105060
tctgaaaacgt	atgctctcat	caccccgctc	taacaaacat	ttggacatta	gagaaaacaa	105120
gtcttttctt	aaaataacat	tatttatggg	agaaaatcca	caaaaatafa	gcattcccagg	105180
acaaacaggg	cttaagatgc	aagattttct	attttactgc	aagacacaaa	gactctgaaa	105240
ttaatgcacg	ccctatcttc	tgctctggca	tacatttttag	tctcctgggg	ggatcagtaa	105300
gtgtggaagt	agcaaggagg	aaacagaaaa	aagtcaaagt	aaagagacag	atttttagaat	105360
gttaatctgc	aggagcctgc	cagaaagatc	tagctcatgg	gctatctgta	catccaggac	105420
tgaagcacgg	gacacggggc	aggctcgtcca	gggttctgtc	caccttatct	tgttacctct	105480
cttgactctt	agagcctcca	ctccacatct	cccatcaatg	ctgcagaaag	acgtggcctc	105540
cactaaacaca	agtcttactg	aactgatggg	acaggaaatt	agaatatcct	ctgaaccatt	105600
ccccatgttct	ttggttcgaa	ttccagcagc	tagaaaaggc	agatgctatt	ctgatcactc	105660
tccctgcgtgg	ctccaatgag	gattaatgag	taacatcaga	gagagaagtg	attataataa	105720
ggctctgacgg	tgcaaccgat	gtcttcaccc	ttttctcttc	gcctccttcc	tcacatctct	105780
acaccttttt	ttttttaatt	gactgattgg	ttcaacaaat	acatgtggtg	cctcaggctc	105840
tggtccaagct	gccgggattc	gtagagaaga	gattcagtgc	ctgctctcaa	ggggctcatt	105900
ctcttgtggg	agagacagac	aaagaaaccc	aagatttctg	gagtgtggga	atggctcttc	105960
aggcagatgc	tagcacagca	cattgaaagg	cacggaacct	caacaaaaa	ataacattta	106020
ggaaccagct	agagcacagg	gtgtgaaga	aagtggaaag	atttgaggcc	agcgtcgcca	106080
tctaagtgtg	ggcatttaaga	attcagccca	catcaatcaa	tcatgtccta	ttgatctcac	106140
cccttaatat	ctctcctatc	tatccgtggc	cactgctcta	tgacagacact	catcatctct	106200
cacagaggca	tcattctgct	ccaagccatc	gccattctct	tgcaagagtt	tatttccatg	106260
gttcccaactg	gatggcttca	cttaactgct	caaaacccct	ctgaggtcca	gtcaactggc	106320
tggtaaaggac	cagtcacagg	tctggggatg	ccagccatga	gacattgctt	tgagggggaag	106380
agggagcata	gaactggatc	tcctgcaccc	tactgcccaa	gtaccaatgc	tggaggtggg	106440
tttccctccc	atcatcagca	agtctggata	tccaggatcc	accctatgga	tgtttttatg	106500
gacagagtgg	gaagatggat	atgtttaggt	tagggaaaga	gggtttgcca	aagagggcag	106560
tataagttag	ctgcactcca	tcattcccct	ggcacaaaac	atggctagta	tcctctagtc	106620
ctcaagagca	ccaccttcca	atgcagtccc	tgctgttcca	cagacctctc	tcctcaaaact	106680
tcctctgaac	aaacctcagc	agggcaattg	ccactctctg	ggcagagtcc	agatatctct	106740
tcctaccctc	tgacatcact	ttctaaattt	gtatatgtag	ataaaactct	agccattcac	106800
ataaagggct	ttgatttctg	atacgccaaa	cacataaaca	aacaaacata	agctttcctt	106860
tcacaaatgg	ctcatgtaaa	ttaaaaatgt	tggttttcca	cctacggtct	tgaaaggggt	106920
ttctacagcc	tgtttttgaa	gtcagaaagc	aaaaggtaaa	tgcaaacatc	atttccactg	106980
cagagaaaaa	tctaattcct	ttgagggcag	gccaaaaata	atacaagcac	actgctatcg	107040
agccaaattac	tggtatctct	gagcttccgt	ctcctcatct	ataaaaattg	aatcgagctg	107100
tatggattaa	agataatgta	tgaaaaactgc	ccaattagta	cactattaat	aaatagcagc	107160
tactgttgtt	aacaaatatt	attgacttac	tggaaaaaca	agaggaaata	aagtcacatt	107220
tagggagaat	ttcaaagtgt	tcctaacccta	aaaaagaaat	aaattagggg	gaaaaccact	107280
aagtaatggg	tgagctcagt	ttaccttgct	taagaagtcc	caccctagag	aactgatctc	107340
tagatgacac	ccaaatgcac	tcagtacaac	cccccaagac	tgtctgggct	taaggcaggg	107400
gcttggtattg	tcctgttaagc	tggtgggaagt	ctgttcatga	gccacagtag	acaggaaggg	107460
gatggagtct	tagagctggc	tcttcagggt	atctcctagt	gtgttcaaa	cagttctcag	107520
gaggttgggg	aactactaca	tagccaaagta	aatatgaggc	ctccttgctc	tggggagacc	107580
tttctcttta	acagaggtga	atcccaaagg	aggcactgga	tgctagaggg	gggtgggggc	107640
cactctggcc	cctcagagca	gccagctcag	cttcagtggg	gctcagaggg	agcagaggat	107700
cagcctggat	cagcctccct	ttcaccatgc	agaaaacaga	gctccccac	cagaccagat	107760
actggaagca	ctggggcagg	cctaagagaa	agcagagccc	caagccccc	cacaccaggg	107820
cttatgaggg	tactgctacc	caccctccca	agccccagct	ccacttctat	gttcatcaag	107880
caactgttta	ctggtaactg	cacttcccga	tagactttgc	tagaaaggaa	tgctcctagt	107940
cactgacaat	atctaaacct	gcaactctaa	ggactaggct	ggggaacact	gtgtcaaat	108000
ggaggcacgt	cctaccctcg	agaagaaaaa	taagggaatat	caataatacc	tgctgggcac	108060
tgagcactga	ctatgtatct	gattcgaagc	gctttttgct	taatctgtca	ctgaatctca	108120
cgatagggtg	tgttatttag	atctattatc	tgggccaact	gagggcctaga	gggacgaagc	108180
aatccccca	agcatccacc	gtagcagtg	caggactggg	atagaaacct	gatgctctga	108240
ctgaaactaa	tgcttttttt	tttttttttt	tttttttttt	tgagacagca	tctcactctg	108300
tcaccaaggc	tggagtgaag	tggtgtgatc	tcagttcact	gcagcctcca	cttcccaggt	108360
tcaagtgatt	ctcctgcctc	agcctccgga	gtagctggga	ttacaggcgt	gcaccaccat	108420

```

gccagctaa tttttttgca tttttttag agatgggggt tcgcatgtt ggccagactg 108480
gtctcaaaact cctggggtca agtgttctgc ctgccttggc ccacaaagt gctaggatta 108540
caggcgtgag ccaccatgcc cagccagctg atgctcttaa tctgtgccct acccagcctt 108600
cctgggaggc ttcccaagag ctacacagag catgagttct ggaatcgggt tgatgggggt 108660
accagtattt actaatagga atgaagatgg gtaattcttt cagacagcac ccttgattaa 108720
aacaagagag tagtgctgcc tctctgtgat tctgtgtctc cctgccctgc tcacacagac 108780
accacaccca ccacacgca tgatcatgaa aagaggaaat ggatccagga gaaggagacg 108840
actcctgagt gaaaacaacg gggtttttca cattgagagc tttgcccaac accccaaaga 108900
tgaaaagagc aggaaaactgc tggggccgat tgaacactgg acttttgttg tggaaaaagg 108960
caaaagggaag ccggaagaga ctggaacagt ttccatgggt ctggaggatg gggaagtggg 109020
tagggattag ctggaggag aaggagaagc tgggggtggga ggggaacctc cacttgccag 109080
gagagcacat gtaggatggg aaccccgat gatactcaag gcatggcatt agaccagaag 109140
caagtctgtg gtgaaattag ggaaggctcc actgaggact gtagacagag cactggacaa 109200
ggaagtggga gacccagggt ccagtcctgg ctctggagcc ccctgggtgg gctgccccgg 109260
gcacctttct ctcttggggc ctccattcct acctctgtga agcagtgct gaacctctct 109320
tgccctgac ttgctgaaat gctgggactc tctccagat gctgcacggg agagggcaga aaaggcctat 109440
atggtgagtc cgccctggga gcctctgctt ggaagtgaat gtggcctgag agtgactcag 109500
aaaccacgga agttcccggt gctgatgggt tcttatagat tgtacatgca gctctcctcg 109560
tgggctgcaa aaccgcaaga tgggctgtga ccactctcaa ggaaagagcc ctatctgcaa 109620
aaagcattct ccctccagg tcttaaagca aacacagact caatccttat tccttttaag 109680
acaaaattgc ctcagggca tcagggagcc agcaggcctc aaatgtgtgc ctttctagaa 109740
ttctcaatga aagcaccctt ttgggtatta ataagacaa cagtaatgac agtcatttac 109800
tgagtgtctg ttttgggaca ggcattgggg taagagctat atgtaataata tattattatt 109860
tgatgccac agccacccca taaggaggcg gaggtaactat cattatgcca actttaaaga 109920
tgaaagaaact gaggcctcaa gagatgaagt aacttggcca agtcaactcag ccagtaaatg 109980
ggaagagata gacttccag tatccagagc ccatgttttc accattatgc tgaagtacct 110040
cttttctgt gccaatgtga tctgectcca ggaatcctgt cttgatgttc ccttccccat 110100
acagaagtcc tctctgtgtc ctcttcagcc tgatagtata tcttttcata ccattctttg 110160
gacatctctg ttatactact ccaatgggtg tccctccct accctccct gggagcttag 110220
ttgttgtgat taagtatagg ggaatgacc cactactaac aaactcataa gagactgatt 110280
gataaacctg aatatcaatt tattaatata cactgagaaa tgaaccacc cagcagatgg 110340
gaatcctaag gctgactggg cagcacaaat tctttcagga aggacaggct tttgggaaag 110400
gaaatcaata ccagaagggt ctttgttgag tacaaagtca gagggaagg agttagtgga 110460
ttgacacata ggtgaagctt gacatacctc tataaagcct ccatcctgcc aaggatcaga 110520
atatccaagg caggagacca tctgggtgtc ctctcctttg gacagtgtcg gatttttctg 110580
gtacctatga agatcttacc tttctggctg catttatcat gattgtggaa ggctttttgt 110640
ttccttgttt gcttagatta atttctgcgt atttaataga actgaaagge aatttcccat 110700
tgagaccac tgaaagggaa taatcaatac atactagtgt tgtgcccct tcagagaaat 110760
tcacttctgt gttgtcactg tatcctcatg ctctcctata atggaggagc agagatgta 110820
aaaacatgga cttggaagcc agaccgtctg ggtttgaatc ctggctctgt tacttataag 110880
ctctgcaacc ctgggcagat tacctaagtc agtttcccct tctctgaatt ggggataata 110940
tagcaccacc ctcaacatct gtcaagagga ttcaatgagg gaatacacat aaagtgtc 111000
gaacagtgtc tgccatctgg taagcagtcn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 111060
nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn nnnnnnnnnn 111120
nnnnnnnnnn taatttttgt atttttggta gagacagggt tttgccatgt tggccaagct 111180
ggtctcaaac tcccgacctc aggtgatcca cccgcctcgg cctcccaag tgctgggatt 111240
acaggcatga gccaccacgc caggcccccac acactttta aacaaccaga tttcattcat 111300
cgggaagtgc ctgtggggct ggtgtggaca tgtgggtgaa ggtggcactg ggagaagtta 111360
ggattctcca tgacctctgt tactcatatt cccacactcc tcaaatagc ctgagtctcg 111420
aggacagtct gatggctggg caaacctgc ggcaaacat tccccagccc tgccctctca 111480
accagagtcc ttccgataca gatttctggg cagctgttgt tacccgtgtc ctccatgttc 111540
ttccagagat atccatgcat gcatcggcat atgtgtataa ttattataat tatatttcat 111600
cccacaagct tttgacatca atagtagcat attattaaat tgttctgtac attattttat 111660
taacttggta tctctggtac tgctcaatat aagaacatat agacctggcc aggcacagt 111720
gctcacatct gtaatcccag cattttggga ggctgagatg ggtggatcac ttgaggtcag 111780
gagttcgaaa ccagcctggc catcatgggt aaacccccat ttctactaaa aatacaaaaa 111840
ttagccagggt gtggtggcag gcgctgttaa tcccagctac ttgggaggct gaggcaggag 111900
aattgcttga acctgggagg cagatgtttg agtaagccga gatcacgcca ctgcaactca 111960
gggtagggcaa caaagcgaga ctctgtctca ataaaaaaa aaaaggaatg tatagacctt 112020
ctttattctt tttgatggct gtaggtggat gttctaaaaa ttgtgttaacc aatctcctat 112080
tgataaatatt taagtattgt cttcagcatc atatgaaact tacaacaag gttgcattga 112140
ctatccatct gtaaatgtct ttttgacat ttctagaata attgcaggat aaactcctaa 112200
aatgagaatt tctgggtcaa agaggatat cattttacat ttaatagata tttgtcaaat 112260

```


tgtcttccaa	agtggctgta	ccaattaaca	ccccgacctg	taatgaatga	gagtgccttt	112320
tttcccacaca	cctggagag	atgaaaaatt	tatggggcca	cttgggagtg	catgggtggag	112380
gaagctgttg	gccgttatat	aacctcgtc	attaataagc	ctgggggttg	ggggggagaa	112440
agagaggtta	gttagtggtt	gcaaacatac	aattagatag	aagtaataag	ttctaattgtt	112500
cgatagcaga	ggaggggtgac	tatagttaac	aacaatgtat	tgtatatattc	aaaatagcta	112560
gaatggagga	cttaaaatat	tccaacacat	agaaataata	aatgcttgtc	tgcgggccatg	112620
ccacctgaa	tgtgccagat	cttgtttgtt	cttgggaagct	aagcagggtt	gaacctgggt	112680
agtatttggga	tggggagaaat	gataaatgct	tgaggtgata	gatatacctaa	ataacctgtc	112740
gaacattata	cattctatat	atgtaacaaa	atatcacacg	tatcccataa	atatgtacaa	112800
atataatgta	tcagtaaaaga	gagggctggg	cacgggtggct	cacatctgta	atcccagcaa	112860
tttggcaggc	caggggtggga	ggatagcttg	aggccaggag	ttcaagatca	gcctggggcaa	112920
catagcgaga	ctctgtctcc	acaaaaata	aaaaataaaa	cgaattagcc	aggcgtggtg	112980
atgcatgctt	atagtccag	ctacttggga	ggctgatgca	ggaggattgc	ttgagcccag	113040
gagtttgagg	ctgcagtgag	cctacgactg	caccactgca	ctctccagcc	taggcaacag	113100
aggaagacca	tgtttctaaa	agaaataaat	taaaataaat	aaataaaaa	aaaaagactg	113160
aaaagcagag	tggtaagaga	aaggactttg	gggctcaaca	gtactagcct	tgaaccttgg	113220
ctgttactta	cccacgtgt	gataagcaaa	tgcccttaacc	cctgtgtgcc	tcactttctt	113280
aacatataaa	atagaagtaa	aatcatatcc	cacttcaagg	gtcattataa	aaagccaata	113340
gagataatgt	atataaagct	tctggaataa	tgccctggcac	acagtaggag	tttaataact	113400
ggaaattcat	tgttgtagt	ggcagccttc	tgaatctgtg	tcctctttgt	ccactaatgg	113460
ctttgatctg	gatttggctc	aggcaagacc	tggggaaggg	cagagactga	gggcaactgg	113520
aggtatagg	tggtctgagc	ttcccagca	gagtgaggct	gggaaaggtc	tgggagacag	113580
accaggcagg	tgctgataag	accggaatgg	gaggctggag	cataaggcag	ttcagttttt	113640
cccaaagggg	ggtgtacaaa	acgatctcgt	atgactcctt	tatactgtta	atgttttcat	113700
tttatcgcgc	actgaaaaac	aaaaccaaca	tatttaatga	atgattccaa	ggggattcct	113760
gctttttacaa	aaaatgctaa	agtaggcatt	cacatgttta	aaaattgagt	tgatttaaat	113820
tttaaaatta	ctaagtcata	gtacataatg	tgtgagccac	agctatcccc	aaaatcatga	113880
tacggatata	ttaatgactg	aagttcttta	aacatcaaca	tacaatgcca	attccagaat	113940
tcagctcaaa	ttctgcaatt	acacaggctg	gggttgaaac	ccagcttttt	tgctaactgt	114000
gtaaaattag	gcaggagggc	taacctcgct	gaatctcaga	tgtctagtct	gtaaattgaa	114060
gataatgttt	gtttttatct	cacagagttg	ttgtgaagat	tcaataaaat	cacaacatgt	114120
gaggatgac	tggctgtgac	ccctgtcacc	ccactgatct	ccagagttga	ttcggctgat	114180
caggctggct	gggcagggtg	ccccctttctc	cctcaccact	ccgcattgcat	tcctcccga	114240
actgcacact	tgggtcaaga	ggaagacctt	tcctgataga	ggaggacct	tccttcagta	114300
agggatatg	agcactgttt	ctgtcctgcc	agaatctccg	aaggagctct	cagtaaaatc	114360
acaagatttt	atttgtcatg	gtagcatgag	cctgtaatcc	cagctactca	ggaggccgag	114420
ctaggaggac	tgcttgagcc	caggagtttg	agaccagcct	gcgcaacata	gtgagacctt	114480
gtctcaaaaa	aaagaaagaa	agaaagaaaa	gaataataat	agtaataaat	cacctgtgca	114540
acgtgctcac	ttctctcttt	ggaatgtagt	aagtgtacct	aataaatgtg	atcattgtaa	114600
tcacacagtg	gagcacaggc	taaagcatct	tgactttatt	ctataagcaa	taaaagagga	114660
tttgttttta	cagaactcat	tatgttgtga	aaataatttt	ccaacattaa	caaagaacat	114720
tcttcaagta	aaaggaaaac	cacccatcat	tcctcccaacc	ttcaataatt	ttcaattttg	114780
catattctcc	agactttgtc	aacatgaata	cttactttac	atggttcgcaa	tcagtgttca	114840
tgcaaatctc	tttatcctga	cttttatata	caaatatgat	gttataaaac	ggtctccatg	114900
tttctgcata	ttctttataa	ttatcatttt	gtggctgcat	aatattgcat	tgactatgtt	114960
aactgcagtt	ttcttaacca	tttactgttc	tggggaaatg	gaggataatg	ccagggtcat	115020
gcctggagct	tttttttgct	tattgcatta	tattcttaag	atcaaatccc	agcagtgaga	115080
ttagtcagtc	aaaaagtaat	aatattttca	aggctcttgt	tatattttac	tagattgttt	115140
tcagaggttt	tgacactgct	tcccagagat	gtaggaaacac	agacgtcatc	caaccttgcc	115200
agtgtctgggt	gatggtgttt	ataaaactct	gctaatttaa	taagtatgaa	atgctatcct	115260
cacacggctt	tcattttctat	ctctttgatc	attaacagg	tgaactattt	tccaagtatt	115320
tgtttactct	ctgcataacc	tccttggtga	agtagtcac	cacttctctc	acctgtttat	115380
ctgttgaggt	cttgaggctt	gtttttataa	tgtgagcgag	cacttcagag	tcaatagaca	115440
ttaatgtctt	ccagccagat	ttggccactg	aggctcctga	gcaggggaa	gcatgatcaa	115500
aactacaccc	tggcacagatt	aaatttaattg	gagaaaaatg	gctgagaggc	agagatatgt	115560
gtcactggcc	tactgtgttt	gatcctatag	tgggggcctg	aactggggca	acggcctgag	115620
tcocccacta	ccagtagcag	gaggctccat	gtgtccccc	tattagagct	tgcggcactt	115680
ccatttgccc	cacctctac	aataccccc	atacatgtac	tcactctccc	ttgcaaatct	115740
agtggcttca	accacagaa	tttaagggga	aaggaaattgt	tctgtcctgt	tcacttactg	115800
cagaaatgag	aaaagcgttg	ttcacatggg	atcacctaat	gaagggatgc	catccccaac	115860
ggtgcctata	aggaatggg	ggagggttg	agagtgtgtc	aaaatgcaac	agggaatcat	115920
cagagtctct	tgccccatga	tagagggttc	tcaaatgaag	agagtctaca	gcaactaatc	115980
tcacggccac	tctaggcagg	gcttcccaat	gcttccccc	ccccacctcc	atcctagact	116040
ttaccacttc	tgctgaacac	agatgttacc	catagcacct	tgacacatga	ttgtttgatt	116100

agcacctccc	acagtagact	gtgtttctga	taggtcagca	acatttgctg	agcacctact	116160
ctgcagggtt	gtgccaggtg	cacaaaataa	acaaagccaa	agacaacatg	gacctgaaac	116220
tcagcaagtt	cagagtcaag	tgggataagg	aggctctctt	cactggaagg	taactccaag	116280
aaacaatggg	actcaacttt	ctaaccaaga	gaactccagg	gagctaaaat	tctgacttct	116340
ggttaagact	ggtgtggagc	ttcattaaag	aagaaaagat	tcaccagagc	ttgagttcat	116400
agcctggctt	tgagctttt	aagtcatgta	acctttgatg	aagttatgtg	acctctccac	116460
ccagctgccc	ctaacacctt	gcaggggcag	ggctggagtg	caaagggagg	cactggtagc	116520
acagcctggg	aggcaccacc	ccactagtgc	aagccgggca	acctctgccc	ccaagggcatc	116580
cctagcctcc	caactgcaag	catcaatctt	gcacttgga	aggaacctca	cctttgaaat	116640
ctaggttcaa	atttagaatg	atccagctcc	ttgaagttct	atacagaaat	acagccagca	116700
gccaggcccg	gtggctcacg	cctgtaatcc	cagcactttg	ggaggctgag	gtgggtggat	116760
cacttgagga	cagaagttcg	agaccagcct	gaccaacatg	gtgaaacccc	gtttctacta	116820
aaaaatacaa	attagccagg	catgggtgta	cacgcctgta	atcccagcta	cctgggagcg	116880
tgaggcagga	gaattgcttg	aacccaggag	gcagaggttg	aagtgcagca	agatcgtgcc	116940
attgcactcc	agcctgggca	acaagagcga	aactccatct	caaaaaaaaa	caaaacacac	117000
gcagcttccc	tccacttccc	aaccacagct	ccatctcaga	caacaagggg	cctcatgtcc	117060
atgcacatga	atatccaacc	aacatgtcta	aggcccaacc	acaccctctc	caaacatctg	117120
ccctctggcc	accctttggc	catgggttca	tgactctggc	gaaaggtagt	tcagagaaag	117180
agccacaaaa	gggcccggaa	gtccacttgg	gctttttgag	attccagggt	ccaggataac	117240
ctaagtgtgg	tctagaagag	agatgcagct	tctgggaggg	acattccttg	gtcttaggga	117300
cttcttggcc	ccatggaggg	aaactggcta	gatgagggcc	aaagcagagc	cctctaaagc	117360
acagggtcca	gggaaggact	ctttttgacc	agatctaaga	gcagcactac	ctctctgagc	117420
ctgttttccc	atctgttaaga	aggggacatt	aatagactct	ccccgctaga	gttactctac	117480
atcagccagc	acacgtaagt	tcatgacatg	aagcaagggc	ttaatatata	ccccgttgta	117540
tataaataat	aggccaggcg	tgggtggctca	cacttgtaac	cccagcactt	tgaggaggcg	117600
aggaggatgg	atcaccaagg	caggagtttg	agatcagcct	ggtgaaacct	catctctacc	117660
aaaaatacaa	aaattggccg	ggtgtgtgtg	cgtgcacctg	tagtcccagc	tacttgggag	117720
gctgaggcag	gagaattgct	tgaacccggg	aggcagagtt	tgcaagttag	caagatctca	117780
ccattgctatt	ccagcctggg	tgacagagca	agactgcata	tcaaaaataa	taaatacaca	117840
catacatata	tacatacata	catacatata	tacatacata	catacaatac	atggacaggg	117900
accctaaaaa	tgagacaggg	aaagagaaaa	acatgttctg	acaaccttgc	cctttatact	117960
aattttaggt	ttcttgccctg	ttttgaaag	ggcctggaca	ggagccctgt	tccctcagg	118020
ccaggcagaa	caagggtgtg	aactcactgt	gggaagggttc	tgggtgacaa	gtgcagcccc	118080
gtccctccac	ctcccagcac	agtaggcagc	acgtgtctcc	attgactggc	tcaggagcag	118140
gcctggtgac	cagtgggaga	gctgaggagc	ccagggtggg	gtctgaagga	atccctagaa	118200
aatctgattt	tcccacaggg	cccacatcac	gtgccagagc	ctgggaaagt	ggaggcagca	118260
tggtgatctag	ctgagaggct	ccatttttgg	tagcttctag	tttgggagtc	acagagacac	118320
ctggatgata	cgaagatgta	gctttgcagg	actctctaga	acatggagtc	caagatatct	118380
ccttcaatga	tgggacactg	aggagaggtc	gtgcccaggt	gtgcccaggt	actcagccat	118440
tcggaggcag	agaccaggct	agaactcagg	acttttaatt	tggaccagga	ttccttttac	118500
cacagtgggc	agccctagca	agtggccagg	agggtggaac	tgtgaaggtc	atccgagggg	118560
tagtacacgt	gggtaggaaag	tcatactaa	gaactgaccc	ccagacctgg	ctctgccact	118620
cactccttat	gagaccacag	gtgctgggtg	cagtgggtca	cacctgcaat	cccagcactt	118680
tgggaggcca	aggcaggcag	attgcttgat	tccaggagtt	cgagaccagc	ctgggaaaca	118740
tagtgagacc	cccacctcta	ccaaaattag	ccaggcgtgg	tgggtgtctg	ctgtagtccc	118800
agctacttgg	gaggctgagg	tgggaggact	gcttgagcct	gggaggcgga	ggttgcggtg	118860
agccaggatc	atgccactgc	acaccagcct	ggatgacaga	gtgagacaga	atgacacact	118920
gtctcaaaat	aaataaataa	atgacagcag	atcatcattt	ttctttctgc	ctctagactg	118980
caatgcctat	ttctccaggt	agtcactagg	ataaaagtaa	aaataatatt	atcagcattt	119040
accaaataca	gggtcagcta	ctctgttatg	ttctttcatg	ccttggttct	tttaagcctc	119100
aaacaactct	atgagctggg	aacaagtatc	gtccttcttc	ctcccatctt	atttatttat	119160
ttatctatgt	atttatctat	ctatttcattt	atttatttat	tttgagacaa	ggtccttcta	119220
agtcaccagg	gatggcctca	aacttgagct	aggaactagt	gtcaccaccc	cccaatttct	119280
tttattgatt	gattgattga	ttgattgact	ggttaatttt	gaggcagggg	tctcgctaag	119340
tcgacagggc	tggtcttgaa	ctcctagtct	taagcaatcc	gcccgcctca	gctcccaaaa	119400
ttgctgggat	taccaaacacg	agccaccatg	cccagccctc	cccattctct	gaataggaaa	119460
actgggggtt	gaaaaggtaa	gcgacttgcc	caagggtccc	tccttagcta	gagagcttca	119520
gagccaggcg	acaaacccat	caaagcctgt	gctctgccc	attgagccac	cggacctcgt	119580
acactaacgc	ccaagtgttc	tacacagtga	aggtgacaaa	gaggtgaagg	gaagagccag	119640
ggaggttctg	tggactcact	cgggtgggtat	gccagagggg	aagggggatc	ttgggtggca	119700
cattgagagt	agctgcgctg	ttagttaagt	agaactcgga	agtccagact	catccagtct	119760
gtgccaataa	accctcctt	ctacctggtc	tcctttccaa	agccagctgt	tctccagaca	119820
atgggggtgg	cgggggcggg	tgtcctcctc	cttctcaggg	aaaatccgac	gctgagccca	119880
tctccagaga	tcttggtctc	ccgtggggct	gcagatccac	ctagagccac	cagagggcg	119940

```

gccagcactg cgcccaaggc ttgaagaccc agcacacca agcccgccca agcctccage 120000
ccagtgtcca agagtccagc cagaggccga gtcctcgatc tcaaaatgtc taactgcaga 120060
agcccaactc atgttcaggc atgatgtgtc tgattctact gggacaatca ttgccacca 120120
agaattactg ccaaaatagt aacgacatta gctacctacc accctccac acaccaacac 120180
acctcatttt accaagcact ttctcatgcc tgggggtgct ggagacttaa tgcagcctcg 120240
cgatgaccgg gtacgctcac cgtacagatg aggaaactga ggcacaagga aggggagtag 120300
attgtctagg gtcacctgga gaactctgat ctccagactc aaatttccaa ttctgcccc 120360
ctccccccaa cctaaccgg agctagggtg ggtggggaca gcaaatgttg atggggggag 120420
gtaagagggg tcagagtgtc ctacagagaa gaccaaatgc attgtggcac ctactgtaaa 120480
atgagaccag acagaccacc acccaccagc cagccaccta aaagtcttca gtgggcacct 120540
gctggaaaca cgacaatgga tgacacagat tccctgccct caaaaagctg gtagtcttag 120600
tgggggtgga gggggtgagt cagcagataa ttatgggaaa ccgtgacacc tgtataaggg 120660
gcgggtgaga acagaggggc tgcgactgcc tggagccagg gattcccga cggggcttcc 120720
ctttctcgc agctcgtccc agggaggaga gctccccccc agcttcgggg ttccgctgc 120780
cttgggggcc cggggtcccc tcccaccctc ccccgccctc agcgtctgtc tccgcatctt 120840
gacagcacac ctgagtcagc ccgcccacca actgcgtcc ggagcgagct cggcaacaa 120900
gtgatatttc gctccccggg agccagcccc actgcgtcc cggcgagct cggcaacaa 120960
accagcgac agattgtgcc gcggtcatt ccggggaagg acgccaaccc ccacctgct 121020
acccccaaac ctccctcccc gcgcgcctc ccaggccctc ccccgagcg caggccctag 121080
tcggggtggg tcctgggga acgcagggtc ctgtcctgcc tcttgaaaat agggggagcc 121140
ctgggtgagg gaagacggga gccccagaga cttttcttcc tgttcttacc tgatccgaaa 121200
acgagagggg cgggaaagga aatttagggg cacagagagg agctgggggg ccgagaaggt 121260
ccgaaaatgg aaccagcagg gggcaccga gagccagggt gccacgggc cgggagcctg 121320
ggatgaaac tggggaagag ggggagagaa agggagcgag agacaccgag acacacagag 121380
acgagagaca gagacggag gagccccgc gggagggagg gagagacgaa gacacagaga 121440
gacagtga gaagacgaag accgggcagg gaaacagacg agtagagaca gaaaaggctc 121500
gagagagagt gagggagggg gggaacagag agacagagac cacgaaatat gtagtagagt 121560
cgggggagaa aaccagagaa atcgaaatga aacgcgagaa gaacgagaga ccgtggaggg 121620
agcagagaaat gaatgggaag aataagacca acatttatca agagccgact gtatgccagg 121680
cactgcattg gacctggca cggatagga agggaggagc gcggcgcggg cagcggggcg 121740
aggggcttct gtgctcgcg gagcggcagc ccaggggggt cagcagcccc ggcaccgccc 121800
cacctgcggc tccagcagcc ccaaccocg cagcgctgcc tggccaacct acccgagcg 121860
gctcccccca gggccccgag cctatcctac gccggggcg ctcccgggac gcgccgggccc 121920
gagtcaagac ctgggtcaac cgccctcgag cctttgtagg gaagtgccta ggtgatgggt 121980
ctgctgatac cgctgtgac caggccatga agggccagag gggctccagt gagaccataa 122040
tcggccctc tttaaaaggg ggtagaggaa gttcacgca agccaacagt ctctcccca 122100
gctttgggtc ctctcctgca ccccgcggg agataagggt tccctcccg gacacatcat 122160
acatacacia aaaaacgcac acactcgac gcgcgccat ctccgacccg ctgttaaag 122220
gaggggggta catacacaca ccgggcacat atttcttcc acccatcccc aagatcgcaa 122280
gcgcaaaacc tcgcacagcc tcacgtttcc caccagctca gacatgcagc ctggcgagct 122340
ttcagcggtc caccgtgtg cacactcac tgccccccc cccgttcccc caagcccgt 122400
caaaaggtaa cgggcaagca tccgtgagtc cacctgcaca agcatcctt cgcgcacgtg 122460
cacgtcata tgactcgat cttgcacgca caaactcttg catatactat tcttatagtc 122520
gcacactggg cttgaggtct gggagtgga ggaagagtg aatcttgag ctgtcccagg 122580
ggacagaaat gctgagggt gggacactgg gcgagggac gcggctggg gcgggggag 122640
gggtgaccca gaagctcatc ttctcctgga aagttgggag gggggaacag gacaagtcca 122700
cggcgttcc ctaaaactacc gcattcccc aagaagggat ttctctagaa gagtggcgcc 122760
gcgaggaaga tcgaacacag tccctcgggt cgcttaagcg ggggggagg gggcggggtg 122820
gaggggggta gaaagcgct ccgcctcct agtggtcgag aaagggttaa gtcggcaagc 122880
cagcaaacga gggagagcc agcgagtgc ggaagagtg ggggtggtt ggaagagctt 122940
cctcgctgtc cccactctcc ctcggtagc agcctgggca caggacaga cggactgac 123000
gactctcgag cggacagcga agctagcgg gcgcgggcgc tggcgctcga cggccagccc 123060
cagccttccc cgccccgtcg cgccccgcc cgtcccgctc gggccgatgg ctctcccg 123120
ggccccgagc cggggcgggc cagggtagag cgcccgggcc cggccacgca gcccgggagc 123180
tccccggccc tccccggagc ccgcgggtc ccccggtg atccggcggg ctcaaggagc 123240
gagtgggagc gccctcccc cgctgcccc tcccccgagc atcgagacaa gatgctgccc 123300
gggtcaggcc gcctgtgca aggtaaagaa gccagcgcg ggaagagcga gggcatctc 123360
gggagagaag cagggcgctc cctctttcag ggattgagg tggggcagtt ggggaggtg 123420
ggtaacctgg ggaagggaa aagctcagcg ctggggccgc gcccccgcg ccagggtgt 123480
tctcagcagg agggcacttg gctgggagcc cgcgggcgc tgcgaggagc tctgaccca 123540
ggtgggagcg agggggcagg tggaccggcc ccggagcggg gagggaggtc caggttccc 123600
tgtccccgct ccactgctc cgggggagcg tgaggactcg ggcgggctgg ggaagcgcc 123660
actcagcaac tctcctgccc cgggtgctca gcactttctg gccacctggg aagacaggag 123720
atgtgggtag ggggctgtct ggggaggtag gaggcgaga gggaaatcca agtggccctc 123780

```

tctggtagga	gagatggagg	gcgctagaaa	gaggatagtt	ctactgattg	agtacagat	123840
aaggggtgtg	gccagagact	gggggtggg	tggggagggg	tcagggggag	agggatagga	123900
aggagaactc	aaagatggag	aaagtgtga	gggaagctca	aaggaggagg	gagatggagc	123960
gggggagggg	gagaaggaat	aaaggttaga	tgggaaaagc	gtggagggaa	gtgggaccca	124020
ggtgaagacc	aaggaagagg	gaaggagagg	aaagaccaga	tcaggggagg	gatgggaaga	124080
agactatgga	cagggaccca	gaatcctggg	atggaggtag	cgggaaagag	aatcaggact	124140
gggaccctgg	ggactggaat	ggaaaaggag	aatggaaaag	tcagaaacca	gagaaggatg	124200
gggatgtgtg	ctagagaagg	ggtatcagga	accggcgaag	agggttggag	acaggggaac	124260
atggatggga	gaggggctgg	agaggaggga	agaggaggag	gaagagaaag	gctgagagag	124320
agggactggg	gattgggggt	gctgcccagg	gatgagacaa	agaggcttct	ggtaaccact	124380
tccacgtggg	aagccctcca	ttcccaaagc	gcctgcctgc	cacatttctt	ctctcaggga	124440
gtggctgggt	ggccagatgg	gggggtgctt	gagctcaggg	ccctgggggt	ggctgtgagg	124500
gacagagggt	gaggactttg	gaaggggagt	gacagcctcc	gaggggtggc	aaacaggctg	124560
gctcctgtgc	tgccatttat	ttatccggcc	cggacgttgg	attctgcagc	cgctgccgcc	124620
accacggtgg	ctgcttattt	tgggggtgta	cattctggca	gagtgaaga	ctgtttgcag	124680
cagctctaaa	ctccgctcac	cgcgctcagt	gcctcccccag	gcccctgcgt	cactggcatc	124740
accaccacct	ccatcccaact	cctcagctcc	cacctcctca	gcccctgccc	cctcagcatc	124800
tgcccgcagg	ccccagccct	tccttgaagc	agcccgttgg	gtgtggagcc	cttgctcttc	124860
gtctggggacc	ctgtgcccct	cgttccagag	cgagaggcct	ctgctgcctt	tccaggggagc	124920
atccttttct	gggaccactc	tgcaccagcg	actctgccct	gtgggtgggt	agcctggatc	124980
ctgcccctta	cttttgggtcc	agttttcttc	tcctcaagtt	ccttcttcta	caggggcctc	125040
cggcccaaat	agtggcctgt	gggctgagaa	ctttgtttct	gagccttggt	actccaaggt	125100
ttgatagcca	gagtcctgga	cagtggctcc	tcagtgaaca	gatacttttg	gctctggaca	125160
cttcagcctt	ccgggatcaa	taccatgttc	tggcctctct	tggctccctc	ccctggtcag	125220
ttctggccat	atatcttgga	caggggtcat	ctcttcttga	ctcccacatg	taatactact	125280
tctagaacaa	ccgcaactgg	aagcctagga	ggtgaaagtt	gcagagagag	ctggagtccc	125340
ttccttgctt	tgaccctgaa	tagccaaaca	gactcagcat	tgtggctggc	ccagccctag	125400
gcacctgggt	gcaatttctc	tcctgttctt	acctcaaggg	cagtgtctca	cacattcagg	125460
cgtgggttct	gcgaggagatg	tggccacctc	ttaaagaaag	atcagagtgt	ctctctgaca	125520
gtggcttctg	gtccctcttt	tccttctctg	gttccacctt	gtactagctg	catgacctga	125580
ggccactgtg	tcattgtttct	ggggctccct	tccttcatct	gcaaatggg	ggccaacata	125640
ttgacctcca	ggggattatg	tgtgtgtgtg	tcaatgtata	aagaagttaa	cctgtacaaa	125700
tgcagtgcct	aggacaaaat	aggtgcttct	tggtttccct	ctacctgct	gtactctccc	125760
ctgcagctct	agccatcccc	tgctgacttt	agaggagggg	gtgagcagag	agggtggggg	125820
aggctgctac	aaagggtttt	cctctgtcca	tgaagttagt	gagggatgaa	atgaaggctt	125880
ctgagaaaga	caatgaaggc	gagctgtaga	gacctggtca	ggaggcctgg	ggtgctcaga	125940
aactcacact	ttccctcccc	agccctcaat	ggtgttacct	atgatgtgag	gggtcggctc	126000
taggtggcca	ccgagggtatc	cccccttcca	gctctgatac	tctgtgcate	ttgcccagat	126060
ctccacccgg	aattcacaaa	atgaaggcca	ggagtggagc	cgtggctctc	gggagagaca	126120
ggagggcctg	gcctggaggg	aaggagtggg	ggtgctgagg	aggagtgaag	acaggggggtg	126180
gggaaggggac	gtggcaagaa	agaaaagggc	acacactggg	cagggcaggg	actgagggcg	126240
ggggagagag	ggaaaggcac	agctctctag	ttcccccaacc	ccccagtcct	accacctctg	126300
ccctggagtg	ctcgtctccag	ccccagcagg	cctggggcag	tgaagcccag	agccccctcc	126360
cctccctctc	tccttgccct	cagtgaagag	cgctgcgtga	attatggatg	agctccttgg	126420
gttacagctg	ctttgcacgg	cagtggcaag	ggccagaaat	ggcaacagag	tcactgttat	126480
gcagcagctg	ttatggagga	gccccagca	ccgggtcgct	cttcagagag	cctgcaggga	126540
ccactatcat	gggctggggg	aggtagagcc	tggttggggg	agacatggga	acaagatgga	126600
aggagagtgg	ggaaagagaa	gagaagtagt	ctaagtggg	caggtgggga	gcagagagat	126660
ctaggggagag	aaagaggagt	aggcacccct	gccagctcct	gcagagttaa	ccctcaaggc	126720
cggaaggaaac	cctgatgccca	ggggaatggg	ccttgctctc	gagattgcac	atccttccct	126780
ctgtctctcc	tggggcagcg	gtcagtcctg	aggctggggg	aaagctctgt	aatcctccag	126840
gggctagcgg	ccatcagggc	tcacactctg	gtgagcttgt	ggataagggg	taggattaag	126900
ggatcagaga	aggatttggc	ttcttttggg	gtcaaagtcct	tagggaagtg	gagatcagag	126960
ggtgactctg	acaggaaggg	aagtgcctctg	gctgggcate	aagagacttt	tctggccctt	127020
tccttgccaa	cactttgctg	tgtgaccttg	ggtaagtgcg	ttgctctctc	tgagctccag	127080
tcatacactc	agtagaactg	atgcttgaac	cagaggaatc	gaggggacct	ttgcccgttt	127140
gaaatctcca	gttctaagcc	ccaaacctca	accctcatga	aacctactca	gggtccccac	127200
tgtgcttcca	cactccacct	ctgcctggtt	cagatgaggg	gtaagagaca	ttgctcctcc	127260
acccccagtg	ggtctaagaa	actcgggagg	agaaagtaat	cgtgaaacgc	cgcacggggg	127320
aggggtgaga	agggccgaga	aacgcggagg	tgggtgtgaac	gaatggaaca	gcagccgctg	127380
tgtcactgag	tattacatca	cacccagcct	acacacgcac	ggggcccggc	gctcacacac	127440
acgcggaggga	cagccagcac	gcaccgacgc	agcaccgacg	cagcgccagg	aggggcccggg	127500
gacactcacg	gtggggccca	aaagcgaggga	gcagcacact	gggagtgtgg	atcttccacc	127560
ccgcacctgt	gtgctcccc	ctctggagga	ggaacaccag	ggcagctggg	atgccagcgc	127620

cacactcggg	gcctgtcagt	cccatgcgtg	cacacctggc	tgagcagcac	tgcatattggt	127680
gagcacctgg	ctcacgccac	tacccaaaaat	cacagataca	tacacacatt	cacgcacacg	127740
gcaacctcag	gagcgtgaca	caacacacac	aaaaccacca	ctaagcaagt	gcaattttgca	127800
gccttgggag	ccccacactc	aaaatcacca	acccctcagt	ctctcccagg	gtctctgaa	127860
cccaaggagc	cccaggatgt	cagagtgcag	aaacaagtct	tcctcccctc	tgctttcaaa	127920
agcctaggac	gttgcttgaa	gcagaagggtg	ttcagtcact	gtgtgccag	ggaatgactg	127980
cctggctttg	gggtgtcagg	ctcccttttt	ccccaggcaa	aactgccaga	agaaaatccc	128040
aggagtcacc	tggaatcat	aagaaagtgt	agaggtcaag	ctagtcccg	cctagaactt	128100
tatcagctat	agtgcaggca	aaggccaggg	atgatgggag	gccttgcaac	cctattaaaa	128160
tatgagtaca	gacacctgca	ctccactctc	tagccccag	gctctctggg	cctgcttttc	128220
catcagtatc	ataataagga	tggatcatat	ccaaccttca	aaagtacttt	tggtgggaaa	128280
aaaaaaaaaa	gctttggctg	gatgcggtag	cttatgcctg	aaatcccaac	actttgggag	128340
gccaaagtgg	gaggattgtt	tgaggccagg	agtttgagac	cagactgagc	aacatagcaa	128400
gaccccatgc	ctacaatttt	tttttttttt	tttttttttt	tttgatacag	agtctcgctg	128460
tgtaacccag	gctggagtgc	agtgggtgcg	tctcggtcca	ctgcaagctc	cgccctcccg	128520
gttcacacca	ttatcgtgtc	tcagcctccc	aagttagctgg	gactacaggc	gcccgccacc	128580
atgcccggtc	aaattttttt	tttgtatttt	tagtagagac	gggttttcac	cgtgttagcc	128640
aggatggtct	cgatctcctg	acctcgtgat	ccacccgcct	cagactccca	aagtgtctgg	128700
attacaggcg	tgagccaccg	cgcccgccca	aaaattttta	aaaaattagc	tggtgtcagt	128760
ggcacggggc	tgtggtccca	gctcctcagg	aagctgaggc	aggaggattg	cttgagccca	128820
agtgatccaa	gctgcaataa	gctgtgatcg	taccactgca	ctccagcctg	ggcgtgggag	128880
caagaccctg	gtcccaaaa	aaaaaaagaa	agaagttttt	aagtaactgc	gaatgaggag	128940
agcctggggt	gtaaaatgca	gattcccagg	ctgtccccc	aggaattctg	catagtctct	129000
aggactggct	ggtggcctca	cttagagacc	cgacccttaa	ggccctccc	ggcacaaga	129060
ggctctgact	ctgcaagggc	gaaaagtaca	ggaagtaag	ggcactgggc	accagtgggc	129120
tggaagacc	agacccca	gtgagtcctt	ttcacacggg	cctcagatct	ccaaagggtc	129180
ccaagttaact	tcaggtcatt	ctccaatggg	gtgactttgc	ccccaggggg	acattttggca	129240
atgtctggag	acattttggt	tgtcacaact	ggaggcaggg	tgctgtctgg	atctagtggg	129300
tagagacacg	agatgctgct	aaatgcctta	tataggcctg	ccccacaac	gaggaaactat	129360
ccggcccaac	tgtcaatact	gaggcagaga	aaccctgacg	ttagtctttt	gacattaatc	129420
tctagacaag	gtcaaacatg	caatagttaa	aacaggaatg	aagagatgat	cattcttcaa	129480
ccaattttgca	gtgctttcta	caatggcctt	ttggcattat	tttttaatat	atgagaagcc	129540
tcagaaagtg	gaagtggcca	ggccacttga	ggctataacg	ttgtcccctg	agcccccaga	129600
catggggagca	ccagggctct	aggcctttat	ttttattttc	tattttttcc	cctgaaacag	129660
ggtcttgttg	tggtgcccag	gctggagtgc	aagggtgtga	tcgttgctca	ctacagcctc	129720
aaactcctgg	gttcaagcga	tcctcctgcc	tcagcctccc	aagtagtggg	gactacaggc	129780
acatgccacc	atgcctggct	aatttttttt	ttttctctgt	aaagacaggg	atctccctta	129840
tgttgccag	gatagtctca	aactcctggc	ctcaagcaat	cctcctgcct	tggtctccca	129900
aaagtctggg	attacagggt	tgagccacca	tattcagccg	ggtctaggcc	ttttaccaag	129960
ttggggggct	ggcccccagc	tggaactcct	gccttggaag	cccacctaag	aagttctgct	130020
tccctccccc	acagctcccg	cctcggtcct	cctcctgctg	atgctcctgg	ccctgcccct	130080
ggcgcccccc	agctgcccga	tgctctgcac	ctgctaactca	tccccgccc	ccgtgagctg	130140
ccaggcccaac	aacttctcct	ctgtgccgct	gtccctgcca	cccagcactc	agcgactctt	130200
cctgcagaa	aacctcatcc	gcacgtgcg	gccaggcaac	tttgggtcca	acctgctcac	130260
cctgtggctc	ttctccaaca	acctctccac	catctaccog	ggcactttcc	gccacttgca	130320
agccctggag	gagctggacc	tcgggtgaaa	ccggcacctg	cgctcgctgg	agcccgacac	130380
cttcacaggc	ctggagcggc	tgagtcgct	gcattttgtac	cgctgcccagc	tcagcagcct	130440
gcccggaac	atcttccgag	gcctggtcag	cctgcagtac	ctctacctcc	aggagaacag	130500
cctgctccac	ctacagggtga	gcctgccctg	ccccaccct	cagccctttt	ctggtttctt	130560
ctctctgtgg	gcccctctgc	tccccgaccc	tggtgtgcgt	ccctcctctc	tccccaggcc	130620
acccttctct	cctcagcatc	tccatttctc	tctgtctatg	tctctttttt	ctcttacatt	130680
ctccaggggg	tttacttttt	cccttctgcc	tctctacctg	tttaggtccc	ttgctgttcc	130740
tctctctctc	tctccctcta	actccacaac	cttcacctct	ctgcctctgc	ctgtctgtct	130800
gtctatccct	ttccatccat	cactgcctct	ctcaactaac	tgccctcccc	atctgtcttc	130860
tgctctctct	gtctgtctcc	cttcacacac	ccactccgca	tacaccccca	tgtctgtctg	130920
cgtgtgtgtt	tctgtctctt	tctgtgatct	cacgtgtttg	ccttcagggc	actctgcctt	130980
ccccagggtt	ccctgcccga	aaggcctttg	cagctgtttt	tctcaccac	cctcaagtct	131040
gcccacatca	cggtgaagta	gagagagaa	gcagagccac	agccactggc	atcccaacaga	131100
aaagtgcgct	tctctccaat	tcactgggca	atgggacggg	agaagccac	acccttcta	131160
gattcccat	ttccaaacct	gtcatctcaa	tgcaaggga	gaaagaaaag	ggtaaatctc	131220
tgttatgag	ctggagaatg	gatgctctga	aaatggaagg	aataaccagta	attgttatct	131280
attgttatta	ttattgatct	aattattgtt	tattgttgtt	atgctgactg	tttgacacgc	131340
aaatcatccc	actccatttc	cccaggagc	aataacacac	cctccaaacc	accctgagag	131400
aaaatcttcc	cttggtctaca	gagcctccgg	ctggaagggg	gtgaaaatat	ccaaattctg	131460

ccctctccct	acttgaacct	ggaaactgct	tcctctgcct	catccagggc	tagtgcctaa	131520
ctagtattca	atctgctagt	tggaaaatca	ggtcagtgct	gatgatgcta	atgataataa	131580
caatagccat	aacaacctaa	caaacatact	gagcaccac	tacgagctag	atgctaagaa	131640
tacagtagtg	aacagaacag	accaaaccce	ctgccttcac	agagatacca	ttcccatgag	131700
gagggaaaga	agtataatgc	acggtatatt	ggaaaaatat	gtcttatatt	attcttattg	131760
ttgcctaaat	agtacagta	atagcagtag	ccgccaccac	ttagtgggta	cacagggtca	131820
gccacagtgc	caagcacttt	ataggtatcc	actctgccat	ttacaagcgt	gtgacatttt	131880
ttttttttac	ctcctcagac	ctcagtttct	tcactctgtac	aattgggtag	caagagcacc	131940
catctcctag	ggattttgaa	agcattaaat	gcataaataa	tttgtaaagc	acttagaata	132000
gtgcttgcca	tacggtaagt	gctatataaa	tgcttggtta	aatactattt	taaaaaaaga	132060
aacgagcctt	atttaacatt	ggtttcagtg	aagtggccca	acttggaact	catcctgaag	132120
atgtgggtca	acttcaagga	ttataactaa	gtcatgagtg	agtcccagaa	attgcacctc	132180
acagtttatg	aagtgcactc	agccacctca	tctcatttct	acagcccagt	tgaggagatta	132240
ttttcacctc	cttgtaaca	atggagaagc	tgaggctggg	ggccctgaag	accctataga	132300
gatatagtca	cctccaatca	taaatctttt	caaccattgt	cgggtgtgac	ggaggcttat	132360
gtctctcac	catcatgttg	agcctcaca	caacctgggt	atagggacag	ttaggggcac	132420
tagggacatg	gaatgaatgt	tcctgaggcc	acacaccag	gaagagctgg	cgcttgaacc	132480
tcattggtctg	gtacaagg	gacagtaact	tgaggtacaa	ttgagcaggc	tcatttttga	132540
aagcacacag	tttggaactca	gcaagaccta	ggttcaaatc	ctggctccta	tatatatgac	132600
tttggaacaa	ttacttaacc	tcctcagctc	tcatttccct	catctctaaa	atggcaatca	132660
ggatagtaact	taataataat	cttttttttt	tgagacgacg	tcactctcta	tcgccaggcc	132720
tgagtgagc	tagtgcgac	tcggctcact	gcaacctctg	cctcccaggc	tcaagtgatt	132780
ttctgctctc	agcctcctga	gtaactaaga	ttacaggcat	gtgtcactac	acccagctat	132840
tttttgtatt	tttagtagag	aagggtttca	ccatgttggt	caggctgggt	ttgaactcct	132900
gacctcaggt	cttccactca	cctcggtcct	ccaaagtgtc	gggattacag	gtgtgagcca	132960
ccatgccacg	ccaataataa	tccttattta	agaagttttg	taaggattaa	aattgtaaggc	133020
atttagcaca	aggattaaaa	tgtaaggcat	ttagcacata	tgggcactat	aataataatt	133080
actactacta	ctactactaa	tactgagatc	aaatactact	acaaattgat	catgcattta	133140
atgctttcaa	aatctcctta	tcaatatata	ttagttattt	aggaggaatt	tgaggtcaga	133200
gggctgagc	ttgaatcccc	gatctactat	tttctgactt	atttaacttt	aagcagggtg	133260
ctaaccctct	ctgaacctca	cttactttat	ctgcaactgt	ggaataatga	aaataatacc	133320
ttccaccaag	aattggctgta	aataggaaac	gagttagttg	atagaaagcc	catagttcag	133380
gctggtgtgg	tggcccatgt	ctgcaatccc	agcacttcgg	gaggccaagg	tggtgtggatc	133440
acttgaagtc	aggagttcga	gaccagcctg	gccaatatgt	tgaaaccctg	tctctactaa	133500
aaatacaaaa	attaggcagg	cggggtggca	ggtgtctgta	atcccagcca	ctaggggaggc	133560
taagggcagg	gaatcacttg	aacctgggag	gtggagggtg	cagtgcagctg	agatcgtgct	133620
actactctcc	agcctgggtg	acagagcaag	actctgtctc	aaaaaaggaa	aaaaaaaaaa	133680
aaaagcccat	agttcagttg	tgaagaaatc	atgttattat	gaccccatcc	tcatttgact	133740
ctcaggccaa	caacagcaat	caggacctga	ggtcagcaaa	ggcttgggca	gaggggacct	133800
caggtggaca	ttgggtgtct	ctgaaatggg	aagtgtttgt	tctctacgcc	cctggcatga	133860
atggtaccag	gcatcatggg	cttcacacct	cttcacacct	ggcctttttt	agaggagatg	133920
gaaaacacag	cctctgcctg	tgaactgcct	ggtagggtgt	ggctgggaga	tgccacaggc	133980
aggtgaggaa	acatgggtcg	gggtgagatc	cgcagggtgc	aggtgtgacc	caagatggag	134040
ccaggcctgc	cccaaggggg	agctttggag	gaaactccac	cagaggacca	cagcttttca	134100
gaatggggaa	gggcccaggc	ctgtgccagg	tgagttcatt	catcaacaga	tatttactga	134160
gtatctacca	catgccaggc	aatgttccag	gtgccaggga	ttcaggagag	aacagaaaca	134220
gtggccctgt	tctccagag	catattccct	actcaagtgt	agccagatga	taaagacact	134280
tgttttcttt	cttttttttt	tttgagacga	agtctcgctc	tcttgctcag	gctggagtgc	134340
agtggcacga	tctcggtcga	ctgcaacctc	tgccctccag	gttcaagcga	ttctcctgcc	134400
tcagcctccc	aagttagctg	gattacaggc	atgtgctacc	atgcttggtc	aattttttgt	134460
tttttagtag	agacgggggt	tcaccatgtc	ggccaggctg	gtcttgaaact	cctgaccaca	134520
gggtgatctgc	ccacctgggc	ctcccaaagt	gttggtattac	aggtgtgagc	caccgcaccc	134580
ggcgacactt	gtttttctct	tcagtcatta	cagtggtcct	catgggtttt	gtttgttttg	134640
ttttgttttg	ttttgttttg	tgagacagtc	tcactatgtc	accagctggg	agtgcagtgg	134700
cgcaatcttg	gctcactgca	gcctcacctc	ctgggtgcaa	acaattcccc	catcttagcc	134760
ttcccagtag	ctggaaactac	agacatgtgc	cacctgtccc	agctaatttt	tctattttat	134820
agagacgggg	tttcaccatg	ttgccaggcc	tggtctcaaa	ctcctgaact	taagcaatcc	134880
acccgcctcg	gcctcccaaa	gtgctgggat	tacaggcatg	agccaccgta	cacagctggc	134940
ctgaatggtt	taaaaaatgt	ctttatgtct	aagcagatca	gatctcagtt	tgaattccag	135000
ccacacctct	aatttgctct	atggctttgt	gcaagtattt	taaccactct	gagcctcgat	135060
ggaccatctg	gtgaaatggg	gataacctgt	acctgggcga	gcagggggtg	tgaggattaa	135120
aggagatact	actgagctca	cagcccaatg	tctggtacaa	agtgcagtac	caatgaatgg	135180
tagctatcca	ttaacaccag	ggaggacacc	aactgaagct	cagcaaaaata	aaagcacagt	135240
ccaaggtcac	ccagctagta	aggaactaga	cctagaattg	gcccagggtc	gtctgactcc	135300

agagtgcagt	tgttcagagg	tctctggagt	tggaagccac	gttccactgc	atattagctg	135360
ttggacccta	ggcgagtcac	tctacttctc	tgaggctcca	tctcgtaatc	tctgaaatgg	135420
agataataat	agtatccacc	tcatagggtt	gtgacaatta	agttactata	taggatctgt	135480
gtagcacaga	gcttggcaca	tggtaaagac	tcaatcagtt	acctgcttga	caatgctgac	135540
gccgatgatg	acgatgatac	ccatcctaga	ctgatgagct	ctgtaagcgg	gggtgcctgg	135600
ccagagtagt	acactcggta	cagctctgtg	gaatgaatga	ggcacatccc	agaactcacc	135660
aattcataaa	aatcagatgc	agatgggatac	ttaaagatca	cctatcctaa	gtcccttggt	135720
tcacagatga	aaagacccag	gcccagagag	gtgcttggag	ctgcgcaagg	tcacacagcc	135780
aaagcagctca	tttgattagt	gtcagagcca	agagctggga	gtttggaggg	aggcaaggtt	135840
aaagacagga	tgctgtcagg	gaagcaggca	gggatgctgt	gttaagattc	caaatggatg	135900
cagagagctg	tgaaccggcc	agtggggagg	caagggaat	gtggtttttg	aaatggaaga	135960
ggatgacttt	agcagaggct	ctcagcccag	agggagggga	gataggaggg	ggagataggg	136020
agggggcggg	ggagggctag	ggctgtgaaa	gtcaagagct	tattaatgca	tagagaaagg	136080
ttttaacagt	ggagagagga	aggaccggat	ttgaaagcta	cattcaagga	agtggcaacg	136140
ggatttgcca	acagcttggg	tggggggagg	aggcaatgga	ccccaaggca	gaggctcaga	136200
gaagggaggg	gcaggacttt	ttgcagagaa	acaaaaggag	aggagaggag	gttagaatca	136260
agaaattctg	tgggcccata	cctggggctg	tggttcaaa	gcacctgaat	tcacctaggat	136320
ctctggaact	ttggtctact	cttctgacct	cccagggtcc	cccaaatgt	ggattacccc	136380
tgctcactct	cccccaaccc	cgggcccctt	atcgatcttc	tgaccatata	tctctgggtg	136440
tgtcctactc	ttgtctgacac	ttcataaaaa	gaggaaacccc	atthaggtgt	tttgagtggc	136500
agggtattcca	agcctacccc	ctggatgggc	ctggaagaga	acaagagcac	caggccatgg	136560
tgagtcaagg	tgaggccagg	gaggtgcaag	gagccagctg	gaggcctgag	ccaggatttg	136620
gggtggtggc	agcagggggc	ggagaatggt	ggtgtcagag	gcagccgaga	aggttgaggg	136680
ggacggatct	caatgtggcc	aagaggaggg	ctcttggcac	gctcagttcc	tgtagcgaag	136740
agggcggaag	ccagatggga	ggggggcgaga	acaggcagga	gcacagggaag	gtggaggctg	136800
tgggtgtagg	ctgggagtcg	atgccctccc	ccaacctgag	gcctccgacc	aggctcctgg	136860
gtggcaggca	tgaggagga	agcgtctccc	caggcagtg	gggagggaga	gccacagtca	136920
gggaacaggc	cccctgggtg	aactggcctg	agcagagtgg	atgctcctgt	tctgagaccc	136980
agacctcctg	gaacctgtgt	accacagtga	tgccctgcac	aagaggggag	gacctcaagg	137040
cagtgaaggc	agggagctga	agctcctgct	ccctcctctg	caagccctta	tctctttgag	137100
ccccagtgtc	ctcctctaaa	aaagttagct	gggctgatgg	gtgccaaggc	attagctccc	137160
aagtcaagctg	atcatcagaa	tcccctgggt	agctggttat	aatgcagagt	ccaggaatcc	137220
ccactggcgc	tgggcccac	acaccgcgcg	ccccccgctg	ttaattctga	accatagtcc	137280
caaggtcctt	tctgcactaa	tggtgcctga	ttaggtgact	ccctagcacc	aggcaggtgg	137340
gacagcgctc	ctaaggggag	tagtaatgca	atgtggcttc	cttctctctc	tcccctgccg	137400
cctctggggg	tggagctgat	gcccctcacc	ccaataccca	gcctagtagc	agtacttttg	137460
ttcccccagg	gagctcctct	tttaaagaaa	agggacagga	cccaattggt	actgagcccc	137520
tattgtcata	gtagccacca	tttattgatg	gttgactatg	cacctgccag	atactgtacc	137580
cttaacagca	tttatcatcc	aacctcctt	tagcctgctg	agggggttat	acataataag	137640
gaatatgtga	catactgag	aacctgagag	tccatgaggt	taaaacttgc	ctaaaataac	137700
acagctaggg	aaaaggcaag	ctggattttg	aactagggct	ctaagtgtgt	agcctgtggg	137760
cttcataatt	ggaccaaate	cctgtgtgct	gggcacgtgt	ccagcacttc	cctcatatga	137820
tctttatgtg	aacctatctc	tggaatcctc	agaacaaaac	caggaagtag	gtatactcat	137880
ccccatttta	cagatgagga	aacaggcaca	gagagatgac	tggccttgcc	aagttaagaa	137940
taatggctaa	caaaacaaaa	caaaaacaaa	aattaaaaaa	aaaaaaagaa	taatggctaa	138000
ctcatggaac	tcatagaact	ccacaaggaa	aggtgttcta	agcaccttca	tacatgctgc	138060
ttcatttaat	ctctacatta	tacagatgag	gaaactgagt	cacagatata	ctgagtgact	138120
tgcccccggt	ggcatcagtt	aatgacagat	ccaagatttg	aaatcagaaa	ggctggctcc	138180
ccagtctcca	tacttcacca	aaccagaagt	tctgaaactc	aaactgtggt	cctgccaatg	138240
gccacactgg	cttccctggg	gaacctgtag	acatggggat	tcccaggctc	caccccaaac	138300
ctcctgaatt	agaaactctg	ccccccgccc	caccccgctc	agagatccgc	aggggatcct	138360
aatacacccg	aaagtttagg	aaccactgac	ctcaccaata	ccactttttc	cacagcaaat	138420
agggttagagg	aggcagaatc	atgctatgaa	tcaaaaggte	aaccttttct	aaccttttct	138480
cttctgccac	ggtgcacccc	cttccctccc	ccggccaagg	ccccagcggg	gtctgcaccc	138540
tgcctcaggc	ccatttctct	cttctgtgcc	ccactccacc	ccaccagga	tgacttgttc	138600
gcggaacctg	ccaacctgag	ccacctcttc	ctccacggga	accgctgcg	gctgtccaca	138660
gagcacgtgt	ttcgcggcct	gggcagcctg	gaccggctgc	tgctgcacgg	gaaccggtg	138720
cagggcgctg	accgcgcggc	cttccgcggc	ctcagccgcc	tcaccatcct	ctacctgttc	138780
aacaacagct	tggcctcgct	gcccggcgag	gcgctcgccg	acctgcccct	gctcgagttc	138840
ctgcggctca	acgtcaaccc	ctgggcgtgc	gactgcggcg	cgcggccgct	ctgggcctgg	138900
ttccagcgcg	cgcgcgtgct	cagctccgac	gtgacctgcg	ccaccccccc	ggagcgccag	138960
ggccgagacc	tgcgcgcgct	cgcgcgagcc	gacttccagg	cgtgtccgcc	cgcggcacc	139020
acgcggccgg	cgagccgcgc	cgcggcgaac	agctcctcca	accacctgta	cggggtggcc	139080
gaggccgggg	cgcgcccgac	cgtacccctc	accctctacc	gagatctgce	tgccgaagac	139140

tccgccccgc	gccagggcgg	ggacgcgcct	actgaggacg	actactgggg	gggctacggg	139200
ggtagggacc	agcgagggga	gcagatgtgc	cccggcgctg	cctgccaggc	gcccccgac	139260
tcccagggcc	ctgcgctctc	ggccgggctc	cccagccctc	tgctttgcct	cctgctcctg	139320
gtgccccacc	acctctgaet	gcggtgctga	gatcgaagag	gccagtgtcc	gatccccgct	139380
tcccgctcac	ccgggggctgc	ggctccggcc	ccagtcgccc	caccttccct	ggccttgctg	139440
cctcccttcc	ccctcccaagc	tcctctcctc	cccggggagc	aggccgcctc	tccttgccctg	139500
ccccctgggc	tgctctgact	tgtggcagcc	ccaagagggc	gtgtgtgggt	gctcagccct	139560
gccccccccca	gttcttgcca	ttactcttcc	cccatcccaa	ggctgggggtg	gggcccccca	139620
ggcagccgct	gacccgcaet	cctaagggcc	cacagcggac	accagagggg	cttttgctctg	139680
cagagcgtct	tcacccagca	gagcctttgg	aagctccccc	aggagagccc	acccaggacc	139740
ctttggggga	tgctctagtc	agggccaggc	tgacctgac	ccctgcttac	cctagtcccc	139800
tcaacctcct	gacactggag	gaatactttt	ctcctaagtc	tacctggac	acttttttagg	139860
gcacctggag	agaactttcc	tctccactgt	ggccccctgcg	tggtgaagat	caaaagaagt	139920
tgtttgggaa	aaaaaattta	ttaaaaaat	ctattatttt	atctactgta	agattttgtt	139980
acttgggacc	ccgaaagcgg	gatgaggtct	cagaatgtaa	ggattgcagg	gccaggaggg	140040
ttggagaagg	ggagccgtcc	cccgcctaca	aagagcttcc	tggtggctgg	aggtgggtgtg	140100
cgctcccccg	ccatgaggag	gagctgaagc	cctgcattct	aggtgaggcg	cagtgtggca	140160
gccaagagt	gggtgctgtg	gcacctcttc	tcttcatttg	tccaggggaa	gagctgcagc	140220
caacctgag	tggtctggcg	cctgaggaac	taagcctggg	gaagacctgc	tgtctgggta	140280
acagccctct	tcacagccct	gttcccttcag	gaacaagag	cagttctcct	gcaaggaggga	140340
gtcacataca	caactcctgg	cacagacagc	cccaacatgg	ctttgggtaa	atgtgaacaa	140400
ggcactgtct	cctcagggaa	acacagcccc	atgccagagc	aaacacctta	gcaaacagag	140460
accaaggctg	ggtttccgcg	tacacttgcc	tccttggtta	agtgcccttg	tcagtgccac	140520
agcgtacaca	ctgcacaca	gcaacctgt	gggtatgtgg	tctctctctc	agctcctgtg	140580
aggtagaagc	catcagggat	gaaccaggct	agagaagcag	gtttccaaac	aggctagaag	140640
aggagccgag	gaactcgggt	gatcagaggg	acagggaatcc	caaattggga	tgctattactg	140700
gcttgaggta	caatcagaac	cttcatcttt	ctggtgtgtg	gaagagaggc	tggggactgg	140760
gaagagctca	ggctaagaag	gacttgggtt	gggatttagg	ggtgagcttc	atcagactga	140820
gcacttgagg	agaagtttgg	tagtttgaat	ttggagctaa	gaatctagct	tgggcagggt	140880
gtggtcgctt	gcacctgtaa	tcacagctaa	ttgggaggct	gacgtgggag	gatcacttga	140940
ggccaagaa	ttgagactag	cctggacaac	atatcgagac	tgagtctctt	aaaaatgttt	141000
ttttaagaat	ctagtttggg	gtgggtgtgt	atgtctcaac	gtctgtaatc	ccagcactct	141060
gggagggtga	ggtggacaga	tcacttgagg	tcaggagttc	aagaccagcc	tgcccaacat	141120
ggcagaaac	cgtctctctac	taaaaattca	aaaaaattag	ccaggcgtga	cggcgggtgc	141180
ctatagtccc	aggtactcag	gaggctgagg	cacaagaatc	actccagcct	gggtgacaga	141240
gactctgtct	aaaaaaaaaa	aaaatctagc	ttgggagggtg	ggaatagaaa	gatagagggg	141300
gcctagctga	tagggcttga	ggaagcaggc	tgaggttctg	tgattctggc	tagggagggtc	141360
aaatgatctt	gagaagaaga	gaagaaatca	gaagaaatca	gcacttaagc	ctgaggcagg	141420
tagactccgg	ttaagggtgt	ggggtgggct	gggggagagt	gagagcagct	ggtcagaaac	141480
ccaggagagc	cggagtctgg	ggtcttcag	gggcttgtgt	caggctggct	gtgaggaggt	141540
taatgggttg	gattggaggg	acagccagac	aagagctctg	gtggaggagg	ggctgctggg	141600
gcctgggcag	ggggaggggga	gctgctggta	aattagaggc	aggctgtcca	ggtcatagaa	141660
ttatcattgt	gaaatattca	tgggcccctg	gtccagatgc	tatttcagaa	cagtgaagc	141720
aagaggagtg	tgtgagcctc	aggaagaagc	ctgaagcaaa	gccactctcc	accaaccccc	141780
acccctccca	ccaccagccc	agacagaccc	acggacgccc	atcacgtgca	caccacact	141840
cccagctctt	cacacacact	cgcaccaagc	agagccatgt	agcacgtgca	agcacaccaa	141900
ccacccacgg	gtcccacaaa	caggcagggtg	tcccctaaat	tctgacatgc	acactgacat	141960
gcacacccac	tcaatcagga	cccagcagag	atcacctcca	gcgatctcac	atggcgagac	142020
ccccaaactc	tccaaacaac	ccagattcac	caccttgacc	cacacacctc	gagataggag	142080
ggatgttcaa	ggccatccag	cccaaccccc	accaatgctc	tgatggggaa	actgaggcca	142140
tagaaaggaa	gggatttgtc	tgagattcct	ctatccccctg	aaaaaagcaa	aattcattca	142200
cctccacat	tctgagtgta	ccccattct	gcattttcgt	ctgccagaca	cccagcctag	142260
ttgtaattaa	ctcctccctt	tctctaattt	cctgcactta	ttcagttacc	cagtccccca	142320
cccagccaca	gtctatccct	tccttcccat	tctccccacc	acctccctgc	tccagctact	142380
cattacctca	tgcttggaat	ataaaagaaa	actgcgataa	cctcctcgct	ggtttcctac	142440
atggaatctc	tccttccctc	ccaccagccc	ataccgtggt	gaccagatcc	atctgatcaa	142500
aatttgcata	tgttatgatg	tcactcagga	gcctgtaatg	gcttctaat	gcctataggg	142560
taaaggtaaa	acaccttagc	agagcatcaa	agatccctca	gagtctggta	ccaactgctt	142620
ttctagcctt	ttctctcaca	atctcatccc	aaaccttcac	tccagctaga	acgtttgtat	142680
catactggcc	accagttatc	atgtatgtga	aacccaccaa	ccgactttga	gtgccccct	142740
aaaatttctc	agtctctcct	gaagtaggaa	acctcttccc	cctcctcaga	tctcagactc	142800
cagagccctt	tcccaaggcc	aagactgcac	ctctctgacc	atatacaggg	gttcttcaaa	142860
gcagcagaca	gaggctcagg	ctctggctcc	ctccaagcag	acggctgccc	ccgactggcc	142920
acctgggaa	gcacagccag	gtttcagtcg	tctagaacag	agaatgagca	tctaaccgcc	142980


```

tggggagagg actaggacac cagatgataa ggtttataag cccttaagcc tctaaggttc 143040
ttacacccag agtagggggg ggacgggtct cagccctggt tccctagctg cgggctccca 143100
atthtcgatc cctaattccga gaggaactcc tctccaatga aatacagact tgggactctc 143160
aggacactgt ggaagggaaa tttcccaaca gactctgaga gtccaggagg ccagggatag 143220
accaggtggc aggcccaagg tccagctggg gtcagggttc tatatgaatt tttaatgctt 143280
ccagatagac ttgtcagatg ttctgaaaac tgagcatctc ctttcacctc tgtacatgat 143340
gcccttctcc aacccatttg cccctgcagg agggcaggcc tgggacagat attcagtggc 143400
ctctggagaa acgggttttg gacagtagaa gggtaaatga cctagttatg ttcccactag 143460
taagctgtgt gaccttgggc aagttactta acctctctga acattagagt tctgtgggtt 143520
tgtttttgtt ttgtaagctg gggacaatag tgccagccta aatcaatttg ttgtggggac 143580
tcagtgcaat agcccatggc aaagtgcact acatgcttgc tgttattatt ctctttcctc 143640
aagttctgcc tccctcttcc agcttttctt ccaaccccaa agatgtctct ggctattgct 143700
tcgaaggtag gaacttttgt tggttctccc ctttctcttc aggcccaaac tccccacctc 143760
aagatccctt ggccctttgt gaaacttcag gtgaggagggt gccagagaaa taagaaagtg 143820
tgcaaggctg gtggagttag agagagggat agatggcgaa gccctagcag aggggagggg 143880
agtgggcagt ggagagagg

```

<210> 16

<211> 215980

<212> ADN

<213> *Mus sp.*

<220>

<221> modified_base

<222> (1001)..(1100)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (2123)..(2222)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (3728)..(3827)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (5168)..(5267)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (7481)..(7580)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (8849)..(8998)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (10375)..(10474)

<223> a, t, c, g, outra ou desconhecida

<220>

<221> modified_base

<222> (12270)..(12369)

<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (13438) ..(13537)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (15902)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (15939)..(16038)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (18223)..(18322)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (20974)..(21073)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (24403)..(24502)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (27574)..(27673)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (30892)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (30901)..(31000)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (34443)..(34542)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (38205)..(38304)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (42373)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (42386)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (42393)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (42461)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (44809)..(44908)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (51380)..(51479)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (56740)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (56765)..(56864)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (62818)..(62917)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (68518)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (68534)..(68633)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (74552)..(74651)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (81446)..(81545)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (88519)..(88618)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (93791)
<223> a, t, c, g, outra ou desconhecida.

<220>
<221> modified_base
<222> (93794)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96565)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96570)..(96573)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96579)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96590)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96596)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96602)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96616)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96629)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96633)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96668)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (96715)..(96814)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (104447)..(104546)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (114521)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (114527).. (114626)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (127063)..(127162)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (139133)..(139232)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (151051)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (153242)..(153341)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (164706)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (164708)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (1164710)..(164809)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (182242)..(182341)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (192158)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (192192)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (198842)..(198941)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (199437)..(199438)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (208276)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (215974)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (215976)..(215977)
<223> a, t, c, g, outra ou desconhecida

<220>
<221> modified_base
<222> (215979)
<223> a, t, c, g, outra ou desconhecida

<400> 16

ttggggggtat	aaaccagaa	gtgggattac	tgcaccatac	aataatcctc	taacttcaag	60
caatttttcc	acaatggttg	tatcatttta	cattcccact	ggctacgaga	agggttccca	120
cttctacaca	tcttcaccac	catttctgtt	tttgtttttg	agtaacagct	gcctaatagac	180
tgtgaagtgg	tatcttatct	cagtgttgat	ttgcatttct	ctgatcatta	atgtgggaag	240
gcatcgtttc	atatgtttat	tggetgtttg	tgtatcatct	tctttggcga	tgttgattca	300
agttatttgc	ttgttttttt	aattggagtt	ttaaaaaatt	gttgttgagt	tgtgggagtt	360
cttcatttagc	tctgcataatt	aataccctga	tgaaaaatgat	taacaagtat	ttgcttccat	420
tttggggggct	tccattctgg	gctgttttta	ttcttttgat	actcttttga	ttctcaacag	480
tttaattctga	ctaaaaattca	gtttatttct	tcttttaattg	gccatgctat	tgacacatcc	540
cgtaatcact	gccaaatcca	gtcatgaaga	gtttctttca	agagatttat	agtttttagct	600
ctttaagttt	gtcatgtctg	tttcacttaa	ttttgtatag	tgtacaaaag	tctaacttca	660
ttcttttcta	tatggcttgc	tactagtata	cgaagagcta	aatttctctt	tccttgagtc	720
tcaacctctg	atgtgtagca	atcttctcag	aggaaaacat	gggtgggaagt	tccttaaaca	780
taggatgctc	catggagggtg	aaatagttca	tcctacaggg	aagcttgta	aacacaggaa	840
gtacatactc	agcagctcta	gtaagtga	gaaactgact	ggaggcacta	ggtccctcct	900
tccttacgca	tatagaagct	gtaaggattg	ggaagagata	ctgtcaggtc	agctcagctg	960
ctgcccggaa	gaagctcaga	cccactggcc	tggtcccaag	nnnnnnnnnn	nnnnnnnnnn	1020
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	1080
nnnnnnnnnn	nnnnnnnnnn	atcactcttt	actcaggcca	cctacacgct	gtttatagcc	1140
tgcttttctg	tctttggcta	tacttcctgt	ttatgtctat	gcctcccctc	tttctttttc	1200
tttctcttct	cttctcatct	catctcatct	ttcttcaggg	gggagcctgg	tctagaactc	1260
acaaagattt	gactgtctct	gtctccttgc	actaattaaa	aaatctttta	caagcatctt	1320
ttagcaattc	ttacagggaa	atcttgggaat	gttaaaactct	gattgttagc	gggctgaaga	1380
taacaatagc	tctgatgata	aattgcttgc	caggcaagtg	tgaaaatctg	agtttgatcc	1440
aaaaagccgg	gtacagaggg	caaagagtcc	ataatcctag	taggggcagg	aatcagggat	1500
gggtgggtcc	ctgggggttc	ctggtttgtc	agcgtagccc	aattgggaat	agccagggtt	1560
cagtgaacga	tgctttctgc	aagctgagag	aggtccttgt	tcaatctctg	tgacccaact	1620
ggagggagaa	gagagccagc	tctccagaag	tggtcctctc	aactttgtgc	atgcatgtcc	1680
atgttcacac	aggggaatgga	taatgcttaa	aaggaagacc	ggcagggggg	tggtaatgca	1740
cctcctttgg	tgacatgctt	tcctcttggt	catgctgctc	caggtgtggt	cggcagcacc	1800
aaaaaccagg	tgatgtttg	taatcccagt	attctctggt	cgtcagtagg	aaatgaaaag	1860
cgaggtcatc	ttcgtataga	gttagcaaac	tctaagccag	cctcggctac	atgagacttt	1920
gtctcaaaac	aaaggaaaaa	tcaaggagga	cggctcccga	gcactgtcac	ctgaagctga	1980
cctctggcct	ccacatgcat	gtgcgcaaac	acatgtcctg	cacaaacaca	cagacaccog	2040
catctgctcc	ccgacaaaag	aacctgaaac	cagtatactt	tgagaatttc	ccattcatag	2100
ttaccattgt	gtgttccttg	tgnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	2160
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	2220
nnngtgggtgc	ctttctctta	cccagctctag	aagggtctga	ggcaggggtg	atggggcact	2280

ttgaactccc	acctaggcaa	aaaacccagt	gatctctggg	ccagtgtgtt	gtttgcaagg	2340
gaataaggta	gagagcccg	gaggaagaca	ttgggggttc	tatgagtatg	tgaaggggtg	2400
cacacaccac	acacacacat	tttttttgtt	ttaaatttac	aaacattaaa	ataggctgta	2460
atgtgggtca	gtgggtagaa	aaacctgctg	tctaagcctg	gtaccgagttc	aatccctgac	2520
aagctggaag	gagacaacca	accacaactg	ctagcagcca	gaagcactgc	ttgctaacac	2580
tcaagagagc	ctggagtggg	agacactgga	tccccagcag	gcaagcctgc	aagaagatgt	2640
gccttgccca	gacaacggca	gaacaaacat	caaggctggc	agagctgtcc	aggactgttc	2700
atattaatca	tgtatagata	agagggaatg	gcacagacag	aacaattcaa	cacacggggt	2760
atgaaaggaa	aggaacaagg	cacacaaagg	acaaagaacc	tagcatacaa	gaaagcctaa	2820
gcagagagtg	gcacttccca	gaaggagatc	ataaaataga	ctgaattcat	taaaacaaga	2880
gccaaagata	aacggctcaa	aaaactcacg	gaaaacaggt	caaaataacg	tcacccatct	2940
gacagttgat	actgtcaact	taaccgtatc	tagaactcca	gcaggcacat	ctccaggcat	3000
gcccctgaag	gggtctttgg	actaggttaa	ctgacgtggg	agtgcacca	tctatggacc	3060
gaagcctcag	acagaataaa	aaaggagccag	tgagctgagc	gtcagtgctc	attgcttctg	3120
gcttctgtgc	tgtggctgca	gcgagacacg	gtgcttccctg	ctttagctgc	catgacagac	3180
cacaccccca	aacccgtgac	caaaataacc	tcctctctac	attgctttta	ccaggcattt	3240
ggtcacacca	atgagaaaagg	ttaactaata	cagcactcaa	tacttaaaaa	cataaacacc	3300
aaccttgttt	gcattgtgtg	gactttgaag	ctcacggggc	agttatgccc	aatgccagg	3360
ctgctggcta	aggggtgag	tgacacaccta	taatcccagc	tgctgtggaa	tcagcaaaag	3420
cgctacagat	ggaaggcagc	cagggcagct	gagactgact	caaaactgata	gaggtgggag	3480
gcatagagaa	aaccagatta	atagagtgtt	ccccactatg	caagaagccc	tgggtttcag	3540
gacgagagaa	ctaagaatac	agaagtctac	tgtgtagaag	cactgctagg	tcacacagaa	3600
acatcactca	agtgtctctg	gatgctacac	ggagggcgctg	tgaagtattg	cttctctgatg	3660
atctgtatct	actacagcac	tgctgtttta	gtatgcgctc	ctccactaca	gctcctcacc	3720
acaccaannn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	3840
aaacacacac	caccagttag	agaaagttaa	tcaggccgaa	tggcggtctt	cccctgtatc	3900
cagggtaccg	tcaggacggc	tcactgccac	tggcaactct	gctgaacaa	agcccgcagc	3960
caacgtgggc	ttcaggggct	ctaaacactg	caatcaaaag	ttgtgtgtgg	gggtgggggt	4020
gctgctgcta	ttcaaggatt	cccaaagctt	agatgtattc	atcatactca	caggaaaagcg	4080
tgttcaacccc	atcactcatg	agcagtcggt	accgggggtga	cctattccct	gtagaaatgg	4140
gacggatgtt	ttggaaaagt	tgacagaaaa	gttgattcat	taggcaggct	ctttgcccaa	4200
gcccctgagg	taagcaaaag	taactggcag	gagactagg	ttgccattaa	tctgagacaa	4260
gatgaaccac	ttgcccatcc	tcctgacacc	taataactaa	tgaagaagaa	atggattgag	4320
ctggcattat	taaaaacgat	agaaacagaa	gtatcaatag	tcattgtgtc	tttctcccat	4380
atgtcaaaac	aatgtgttaag	atggcatcga	acacatgcag	aaactgttta	gggaacatgc	4440
tgaaaatatg	aagtataaatt	aaaattggaa	agaaagacaa	tttgccataa	gcagctcaga	4500
gctggagaa	ggaccgaggc	agagataaca	gcaacgtgtg	gacatacggg	tctggggcag	4560
agcagtcacg	gactcagccg	gaaagggtgg	ggcagccctc	gaaggaaatt	aaggtaataa	4620
gagccacaag	gtgattggcc	caggagtgg	gccaccttca	cctcctgcct	caaagtctga	4680
aggaatgatc	ctggagtctc	ccatctattg	atatatgaaa	ttcacagtat	gttttagaac	4740
ccactgaatg	atgggtagat	taactaaaag	aaatttaagc	ggggtgtgtc	aggtctttta	4800
atcccagcac	ttgggaggca	gaggcaggtg	gatctctgtg	agttcgaggc	cagcctgggt	4860
ccaggacagc	cagagataca	tagagaaaac	ctgactcgaa	aaaaacaaa	taaaagctca	4920
tcaaaaacac	aacaacaaca	aaaaaaacaa	aaaaacaaaa	caaaacaaca	aaacaccta	4980
tagtacctgt	tggtgagttt	gagtgtgtga	gtgagtgtgt	gttagagaga	ggggcgggga	5040
aagtgtgttc	tggaaatggg	agaaagagaa	tgtgcatgtg	tgtttctggg	atgtagacaa	5100
aactacatgt	cttccatcaa	atgcaatgtt	taattatcta	tgagttgaac	catcttcatt	5160
ctgctaannn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	5220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	5280
caaacacagta	aacaaaatcc	tgtaaagata	agcctaagac	aagacacttc	ctggggctgg	5340
ggagttgctt	agaccataag	gagttcataa	tccaggcgctg	agagcccag	ttcaggctcc	5400
tgggcttcca	agtcaggagg	agaccaagga	atcaacaagt	ctcgactttg	gtctctagtc	5460
ccatgcacac	acacgcgtgt	aaatacgtag	atgttcactc	acacacagaa	gactgcacct	5520
ggctctctca	catctcagcc	aacatataaa	gctgtcatta	tcagaacatt	ctaggttcta	5580
gtttcagtc	actcttacac	agaatggcca	tcatactccg	tctacaactt	ctcctgatct	5640
acccacgtgt	cattgcttca	gtattaacaa	aaccagaaat	aaccagctgc	gtagatcctc	5700
cctgatgcc	cagtcatgtt	cttactgaga	ctactaagtc	acaaggtagc	actctggatc	5760
caaaaagcaa	tatccaattg	agagttacaa	cctataagga	ggagtttacc	ttcattatag	5820
ggcactggat	tcccaatctt	taatccaacg	tcttcagcag	atttcataac	ttccaagtc	5880
atcaaaaacaa	ctactttcct	acaaagacag	acacaagtta	gaattaagaa	ctctgcagcc	5940
tttcagatga	gttactaaga	agcttacttt	agtgtgtgtc	tggtctaaaac	tgtatccttt	6000
accaaccttt	tctcattctg	gactaacttg	agaagtatta	attcctaagt	aaatacttca	6060
cttattcttt	ccccacatct	ccaatgtttt	tgtctttaat	ttattatag	gcaattcatt	6120

tcctatctag	ttccctgatt	aaaacagtag	accttgcctgc	atgccattat	cctcatggag	6180
gcactgatac	aattttgatt	attaaataca	aaaccctaaa	acacaaaaag	atgatttttt	6240
tttaaaacaa	gatttttaaa	aaagcatgtg	ctacgcttcc	ttctgccact	aagcctacac	6300
atggctcctc	gactgaattt	ttcccctcat	tctgcttcat	ctaataatgtg	cttttcaaac	6360
ctggaattga	accagggact	tattcatgct	aggcaaatgc	tctaccatag	agctataccc	6420
ctccaactcc	catctcaaat	atcatttcca	aagacatttt	cttggtctct	tatttagatc	6480
aggtttcttt	gtcctcctgc	agctatgact	tcattccttc	agaacactcg	tcttagcttt	6540
aagttctgtg	ttaattagtg	attgttttca	ttctctctgc	tagaatgcac	tttcaataaa	6600
ggcaggtagc	cagccacagt	gcttaattaa	gcaacagccc	aacgatgtca	ttcactacat	6660
actgggacac	gatgcctaac	atcatctgca	gataaagacg	aactactggt	gtcaggagac	6720
agctaagggg	tccagggcct	gggcacgctg	agtgtgagca	ctggagtccg	gggtgccaga	6780
aacgcacata	aatgcaatat	ggatgtggca	atctacctct	aattccttct	ttaagacagt	6840
agctctccag	agcaagctgg	ctagcaagac	aagccatata	agttagctct	gggcttgacc	6900
aagacctgct	ctccagggtg	aactcccaag	caaaaggatg	atggctcaca	aatctcaggc	6960
tatcatgttc	atgtacaaaa	tgtcaaccgg	catacacaca	tgacacacac	tgaaaactgg	7020
gagaaaaata	gaagaattgc	aacccaaaaa	tgtaatgtga	ggacacataa	ttgcaggcgg	7080
ggagtggggg	gatgacagaa	ggtgaactga	gtggaccgag	ggaaagctgt	gctagcggca	7140
atgagaagaa	gggtggggca	gtctgagcaa	gggttcagca	atcaccaacg	tttactgtct	7200
gcacagctga	gctgtagaat	gctgggcttt	atcacacaga	attattcagt	atgtgctatc	7260
tttacagtaa	gcttatttcta	tcaggctatg	ctacttcaat	agaacaagcc	tgaaaaagtg	7320
gtctgtctgt	gagaacctga	caaagatgac	ctgttagaac	tgtctgccaa	gtgtggaatt	7380
ccagcactgg	ggaccaggag	ctcgagggtc	acccagatg	cagggagtta	gaggccagtc	7440
ttggcaacat	aactctatgc	ttcagaaatt	aaaaacaaaa	nnnnnnnnnn	nnnnnnnnnn	7500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	7560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	7620
taaaagttagt	agttgcaagc	atgaagaaag	acaaaccaact	tgagcttttc	ttttgtcgtg	7680
aggaggaaac	cagacaggctc	cagagagatg	gctcagagat	taagagcact	gactgctctt	7740
ccgaagggtcc	tgagtccaac	tcccagtaac	cacatgggtg	ctcacaacca	tctgtacagc	7800
tacgggtgtac	tcatacacat	taataaataa	aataaataaa	taataaataa	aatcttaaaa	7860
gaaaaaaaaa	aaaaacctaa	ccaatcagcc	aggcagtggt	gacacatgtc	tttaatccca	7920
gcacttgagg	ggcagagaca	ggttgatttc	tgagttcgag	gccagcctgt	tcttcagagt	7980
gagttccagg	acagccaggg	tgatacagag	aaacctgttc	tcaaaaaaca	aacaaacaaa	8040
caaaacaaaca	aacaaaaaag	gaggaagcca	gacaggatgc	actttatacg	tgaatggaat	8100
tgacaaaaga	caagttctat	aagtgttagg	gaaaggggga	ggacaacggg	ggttcatgtc	8160
tgtgtgtgaa	cacgtattag	aaggctctgg	gtatcctgtt	tccgacaaac	aggcactccc	8220
aatcacacag	gccactggat	gtctcaggca	gagaaagatg	tgatagattg	actttttaac	8280
aatcacagac	tgtgtggaaa	atatttgtaa	ggttgtcatt	gtcaccagg	atagagctga	8340
tggttattca	aacgaggatg	ggacaacaga	aatgggagag	agggatgtga	gaaccatttt	8400
gaaccagggt	gatttactgc	gcacgtgtat	aggttctaca	gggagtggga	tatgtagagg	8460
aggcctatgt	tcctaaacttt	ggtaatgagc	ttattacagt	tactatgcac	agcctggaag	8520
atactggaaa	aggtgcaggc	taggctagaa	aggtaactaa	tgaggggttg	acagcccctt	8580
ggatgtcagg	atgcagcaag	cctacctctg	tatgtagtca	atcccttctc	aggctatggg	8640
tctgtcagat	catccgtctc	tgtatccatt	attccagtc	catcctctga	gtggctccct	8700
cttatccagt	ttaacaaaat	gctgactgca	agctcccaag	ccaggggtc	tggctccttt	8760
actccttgtt	attgtacttt	acctgtttg	cttgggatag	agtgtgcctt	ttataaacat	8820
ttgtgaaagg	gggaatgaag	aagaaataann	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	8880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	8940
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	9000
aattcccagc	atcccatgg	cagctcacia	ctgtctgcaa	ttccagttcc	aggggattca	9060
acactcagaa	acataagtgt	aggcaatcta	cgtaacataa	aaataaataa	atgagctgga	9120
aaagaaaaca	tgtttcaaaa	tatacaagta	atggggctgg	aggagatgtc	tcaatgggta	9180
agatcattgg	gtctctcttt	ggaggttctg	ggttcaattc	ccaccacca	catgacagct	9240
cacaaactgt	tgtaaactttg	gtcctgtggg	agctgatgcc	ctcttctggt	gtgcagacat	9300
acatgtagac	aaaacacctg	catacataaa	ataagttttt	aaaaaagtta	cacatacacc	9360
cgtgtgtaat	ataacacaca	ctggcttaac	ttcctcagca	ctgactgttc	accatacggg	9420
ttcccatgag	gttttggttg	cattctatca	ccgaaaaaaa	aaaaaaaaaa	ttagaagaaa	9480
gtataacat	ataaacctct	ccctaaaata	aagttttctt	ttctaaaagt	acatccttat	9540
ttttttattt	tttttttttt	ttaagaaatg	ggaacaacag	ttctgtcac	actgtatttc	9600
tagcatgtaa	catcttgcaa	gtacttaacc	gtattctata	tcagctcaac	acacttacta	9660
ccgaagactc	aagatcacia	aaaaaaaaaa	aggaccagca	ctggataatt	aaacgtttct	9720
tttgtgttag	taagcgacct	cttccttaga	agatactaca	gtaatgctga	agaaatgaca	9780
catctactgt	aactctgttct	ctgggattcc	aacttgtttc	ctctgtact	cctcccttgg	9840
cggcaatgtt	cgtctgcac	cggctgagct	cctcgctgcc	ttgttaaaac	tccttccctg	9900
acttccgacc	tgtagttccc	gctctacagt	gcaagcagct	ggataaggaa	gogcatacct	9960

gccgtctttc	agggtgttga	cgatgaactt	gtggacctgg	cagacacagt	tgctggccag	10020
ctgccctccc	tcgaccagg	tgctcagctg	cggtggccagc	atgaacgctg	caaaagcaga	10080
gagagagggg	ctcagtcctc	aagcctttcc	ttaaacccgaa	agctcatcac	aaggagaacc	10140
attaaataca	gctgttttaa	actcctccgc	cctgcagaga	ggaagcagc	atcaatccgc	10200
cccatgtaaa	agtctgagge	tcttcctaaa	tggtatctgt	ttctcacagt	ctccaaatca	10260
tttttactgt	aattctagtt	tctggggaaa	gacctttctc	ggtcttttagc	cccgtagcta	10320
gagacaacag	gcaaataatc	cagaaaggcc	cccattttct	ttttaaagct	tctannnnnn	10380
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	10440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	10500
ttcgggtccct	cgaattttggg	tttcttctgg	gacgtggtag	catgtgactg	tcactccagt	10560
gcttggagca	gcagaggggt	caggaactcc	aggctggcat	tagctgcaga	gctggagcag	10620
gtcctggaga	acagaaactt	tggttgcagc	attaatgaac	tagaagaatt	tttttgtctt	10680
ctgttaataa	taaatacctc	cattatcttc	tcataaacag	tggtgccttt	ttattttagt	10740
ttttaaggat	caggcacaga	gactccatgc	cagactacca	ctcaaccact	gagctacacc	10800
cccaacttgc	ctttctgtcta	ttttttaaat	tgatcagtg	gccaccaaac	atggggagag	10860
gtcagggggt	tacgtggagg	aattgtttct	ctcctaccaa	gtggggccca	ggtttcaaat	10920
tcaggtgacc	tggtctggca	gcaagcacct	ttacccctaa	gccatctcat	tggtctcatc	10980
ttttaatggc	cccttcccct	gctctgaggc	aggctctccc	tatatagccc	tggtctgctt	11040
caggctcgca	ggtccaccag	tgagcaccag	gtttctgctt	gtccttacct	ccccagcact	11100
gtggttataa	gcattgtgca	ctgtgtcaaa	ctcagtcact	aagctttgcc	aagccatagc	11160
ccagcccttg	agttttactgt	ttgtctgtgt	ggtgatttgt	caaaccactt	ttgttccact	11220
gaggtatttt	gtcaagtttg	acaaaattag	ttgagtatgt	aggctctttt	ttctggaatc	11280
ttctgttata	cttttagctg	gtcttgaact	catgatcttg	cctcaacctc	acgattattg	11340
aggattattg	agatggacag	gctgtgtgac	catgctcggc	tgtgtgtttt	agcatgcatt	11400
agtcatttga	aaaacgttgg	ctcatgcac	tttacaggtc	ttccatgttt	gatatgtttt	11460
atttaatacca	acagtaattcc	agcaccagag	gctgagacag	gaggatctca	aggtcaacct	11520
agagatgcat	agcaggcggg	gccccactcg	gttaggttaa	tatcatcact	gacttcagga	11580
gaaaagtctt	aagtattggg	gactaaaaagc	aggaggatct	gaagttcaag	gtcatcttta	11640
ggaacttagc	agacttgagg	ccagcttggg	cgctgtggga	ccctgttttt	aaaccagaaa	11700
acaaattgaa	agggaaaaaaa	aaaaaagctg	agggaaagtga	atgtgagtgt	tcacatagtc	11760
ctgtttccac	aagaaaaacag	ggttactttt	ggcaacaaat	aggtgtcttc	tttgaaggct	11820
ggcatttttg	tgacttgtca	ttggagaaat	gatttaatta	agacttttct	actgagtgc	11880
tctgaagagg	ctcttttaaa	tttagtttaa	ttttatctca	ttgttagtgt	ggtgtgcttg	11940
tgacacacaga	aggcagcttt	ctagagtctt	ttcactctct	cctccacagc	tcctggagtc	12000
aaactcaggc	cttggctagg	caagctctta	ggacagtgtt	agctgtagct	tattaaagttt	12060
ttaagaattt	tataagact	ctgtttttct	ttctcaggtc	atgatacagc	aggaaaatac	12120
atccataaag	cccatcctgc	aggtcattgt	aagtaccggc	atgtgtgttt	agcataatga	12180
agatgggtca	cttatagtta	attaaacatt	ggattgggat	gaagacatgt	agttttgggt	12240
acttcccgga	aacacaaatg	cacattcttn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12360
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	12420
cttctattcc	ttcattacat	tactgttaat	agtggtaact	atgtaccaaa	gagtcaataa	12480
actcttggag	catccaaggc	agaagggaagg	ctggcaaaaa	tgtagtatga	tctgggatgg	12540
gaatgtactt	cagtttgtac	aggaggccct	tggtttcattc	catttctggc	aatgcataga	12600
cctgtaggat	ctcagcactg	gtgggggggtg	gggggtgagg	gtgaaggggc	gggaggttaa	12660
aggcagaata	gtcataaatt	caaagtctgg	gtcctggaaa	gaggactaaa	cgattaagag	12720
ctttagctgt	tcttctagag	aacctggtgt	gatccccagc	acatggtgcc	tcacgactgt	12780
ccgaaactct	gattctaggg	ggatctgaaa	accctcttct	gccctctgta	gatacagaac	12840
acacatgggt	cacatacata	catgcaacct	aaacaaacca	tatacataaa	atattttttt	12900
ttcaaaaaga	cattcaaatt	cttcctcggc	tatatagtgt	ttaccaaac	tcaaaaacaa	12960
aacaaaacaa	aacaaaacaa	agaatcatta	atgttttgcc	ttcatgtatg	tctgccacc	13020
acggacatgc	ctggtacc	gggagattaa	aagaagacat	tagctcccc	ggaatggaga	13080
taggtatgat	ctaccacttg	ggtgctggga	acctgggtcc	cctgcaaaa	cagtaaatct	13140
ttttaacccc	taagctgtct	ctcccaacgc	ctaaagatc	ttgtaacaca	gcatgatgag	13200
cactggcaag	catagcatgg	taatctgact	tcaggcgcc	agattttgag	cttaatgctt	13260
gattattaga	agtaacgtac	tagatttaat	gcctggagct	tcaagcaaca	aaatttaact	13320
aagaataaaa	ataaaaaccc	tgccagccat	gatggtaatc	ccagaacttg	agaggcagag	13380
gcaggtgatc	tctgtgtttt	gcaaggccag	ccacaatcta	catagcacgt	tcagtanann	13440
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	13500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	13560
agaaattcaa	agaaacaaat	gccaaataaa	tacacataat	gtaataaaga	gataattgtc	13620
taaaaaactc	aaggctttaa	atggttaagat	atcatattct	tggtgaaaa	gatctaatgt	13680
caaaatatat	caatttaatg	caattatgta	tattcaggag	atctctgggt	ggcttttgaa	13740
cttgatagca	ctcttataat	tcacatagaa	gaaaaaaaac	catgaaaact	gccaaacatt	13800

atagaatcac	tcacacagatg	gtattttggc	agcacataca	tcgaagggct	gtgaagatg	13860
tgtagatcat	ccacgccttg	ctagggagag	ggcgggtgtg	tgtggggggt	atagctgttt	13920
gggaaataaa	cctgtgtaatt	cctcattagtg	taaatcatag	tcagaaagct	gactagcaac	13980
ttctctctaa	aatacattcca	cctcagcagt	ctgcatttgc	aggaaacacc	tcctagcagg	14040
atctgtacgt	ggatcaaggt	agtagcatct	gcatttaatt	gacattctcc	taaatgtctt	14100
aaatttatct	tagattacct	atagtagcca	agatgatgca	aattatgtta	cactgtatta	14160
cttggggcgt	aacaagaaaa	tgtctctact	caggttccat	caggtgcagt	acttcccctg	14220
aatacttctg	tacaacacgga	tcagaagcgt	acgaagaag	ggctaaccat	atacaagcat	14280
atagtacact	aataaccatg	tacaaccata	tagtacacta	atattcagtg	cattactcaa	14340
aatgcaacaa	gatggaacaa	atccaacagc	ctgtaagctg	aaaaacaaga	taagcaaaat	14400
gtgtctggcc	tagagggccc	ggctataaat	tccaactaag	gtcagggcag	gaggattctc	14460
agtccaaggc	cagcctagac	aacttagcaa	gacctgtgtc	caaaaacaaa	agtaagagg	14520
ctgaggatat	agctcagtat	agagcatctg	cttagcatgt	gcactgacag	ccgtatcaca	14580
gagggaaaaa	aaaaataagc	aaaattgtgt	ctgtctgcac	aacaggatat	cacagcccc	14640
taccacaggg	gaacgacaga	gtaacacaa	aaaaacttag	cctgaaatc	actatggtaa	14700
ataaagaagt	gctcactgag	atcagaaaat	gcatgatctc	atttacatta	tatagaaatg	14760
agaagatcag	tgagcctcta	ggactcaaga	gatttgggat	tggcagctaa	agggtactgt	14820
gtttctttat	gggggtaaga	aaacattcta	aacttaactg	tgaagatgac	tactcaacaa	14880
gttcaagtgt	tcaaaaatca	tacttttttt	tttttttggt	ttttcaagag	agggtttctc	14940
tgtgcagtec	tggaaactcac	tctgtagacc	aggctggcct	cgaattcaga	gattcacctg	15000
cctctgcctc	ccaagtgtgt	ggattacagg	catgcgccac	catgtcccg	ctcaaaatca	15060
tacttttaaa	aattgccag	tgactcatga	atacaactag	atgcgggaga	ggacagtggc	15120
aaactcagga	taccagtgct	ttttatgtct	gctgcccaac	tgtcaatttc	ccatagttae	15180
cagagaaact	tttggtttgt	ttcatcttat	ttgttgcctt	tgttagaatc	tcaatatagt	15240
agatataca	gctggcctca	tactatatag	ctgaggacga	cttgaacct	ctaactctcc	15300
tgtttccatc	tcocaaagt	tgggattaca	ggggtgtacc	gctatgccca	cgaaacacaa	15360
agccatttga	accacacccc	agccttttca	gagaaacctg	tacaagcctt	agtgccctag	15420
catattaagg	caacaaaaga	cataatgcgt	ggctaccata	gagtgtttgc	ctaccatgtg	15480
tgaggctcta	ggctaaatgt	ccagcactta	taaaaaagag	ttaaaaacac	tcatgactca	15540
aggatgacta	tgcagtcctg	tgtacaaagc	ccgcacttca	atccccagca	ccgtgcacat	15600
cagggcaggct	ctgtagagga	cccagcttaa	ggtcatacct	aggtaaagta	gaggccttag	15660
atggctacat	tagatgagac	cctttctcat	aaacagaata	ataaatttaa	agctccttag	15720
caaacacatc	gccttcccat	cacactcaga	ataaagcact	ctactggccc	tttaaggact	15780
gcccatctgg	aagagaaacc	taagttaact	tccttgcttg	tgtcatatgt	gataacaaac	15840
tcacttggaa	tcagaaaaat	cagctcttaag	cttggtcaga	aagcttcccc	agcaaacatga	15900
tntcagagga	cataatgcag	aaagtggaca	aatgcanaan	nnnnnnnnnn	nnnnnnnnnn	15960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	16020
nnnnnnnnnn	nnnnnnnnna	tcagaggaca	tctttcagga	gttagttctt	tcctccctct	16080
atagctttca	gggatcaaac	tcaagtgtgt	actgagcgct	tagcccaggt	gcgccactgc	16140
accaggcttc	cttcttttgt	ttttatgggt	ctgaactaat	tagcaccatt	acaaacatgt	16200
tgacaatcag	caagtacctt	tctctacctg	gctagtaaga	gaagtaagtg	ccttttgtgt	16260
gtgaacgcag	tttctcttgt	gaagtgcagt	gacttgatct	tgtctcaca	gcttttttag	16320
gtccttaagt	tgtctggggt	ttatgggaaa	ggctcttggt	tttttgaaga	agattttact	16380
acaacttgat	ataatcatta	tttttaatcc	tttaaatagt	atgacttatt	ttaacagatt	16440
aatattgaa	tgtttcttca	ttcttcatata	ataaactctg	ccttaaaaaa	taactcctct	16500
aggtctcttt	ctctattttc	aaatttgggt	tatatatttg	ctgatatttg	aacatttata	16560
aaagttagga	gacacacacg	tagaaccaag	tcccatata	gctgtgcaca	tagcttcaga	16620
ttattgcctg	ataataggtc	ctgttttgtc	tctgtttttc	cacagggatg	tgttattgtg	16680
tgtgtgtata	catacatata	tatatgtatg	tatgatagta	tgtatgtatg	taatgaactc	16740
ccttttaacaa	aacaagtaact	gggctgggaa	gacagcacag	ttagttatgt	gttttaaccg	16800
acaagcatga	caaccagagt	tgagatcccc	accaaccgca	taaaaagctg	ggcatagtgg	16860
cattgacctg	tagccctggt	gctggtagaa	agctggggag	gcaggtagat	cggcagagct	16920
tactggcaac	aaactctgcc	agtaggtaag	ctctgggctc	agacatccct	tataggaaaa	16980
agatgaagg	cgaggcgcag	cggcacacac	cttctgggtg	agtgcttgag	aggcaggggc	17040
aggccagtct	ctgtgaccag	cagcctggcc	tacatgtcaa	gttgcaggag	agccagagcc	17100
accacctact	gagactgtct	cagaaataag	ttttttaaaa	aaatgagatg	aaggagctgt	17160
taagatggct	tagaaggtaa	aggcacttat	cactaagcct	gaagcccgga	gtttgacctg	17220
ggacccacaa	ctgtagaacc	aactcctcca	agttcttctc	agacctccag		

tgcacacctc	aacactcaga	aaatgaaaaa	acaggtacca	tttacacctc	cgtgttcggt	17700
ttccaaccac	tcatatgtat	gggttgtaaa	tgcttatatc	tgtagtggtc	tgtagatttg	17760
tgtatacatt	caaagttag	tcaggatcca	acgtaaactt	ggatagtagt	gggttgatgg	17820
tctgggaagc	tgctcgcagc	tgctcttttc	tcctcgtacc	ttttccctcg	tttgtttcta	17880
cgacagcagg	tcatttgctc	ctaagtgtta	gtttcccatc	ctctctcttt	tgctgatggg	17940
agccttgtag	tagtcacctg	tgctctctgt	aaaatggcct	tgccgtgtta	tttcaatatg	18000
ctatcatcct	catcttgcta	tatttcattc	aatatatgta	tattattaca	gatagattaa	18060
aattatttta	attttatgct	tatgaatggt	ttgcctaagt	atattgcacc	ttgtgtgtct	18120
agtggtccaca	gaactcagaa	gaaagtgtca	catattctgg	aactggaatt	gcaggtgggt	18180
gtaagccacc	atgtgggacc	tggaacccaa	atccaggcgc	cannnnnnnn	nnnnnnnnnn	18240
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	18300
nnnnnnnnnn	nnnnnnnnnn	nncccttggt	caaagatcct	cagtttcgac	tttgattacc	18360
cagacttctc	gtttctctca	tggaacagtt	tccccctgag	atttactagt	ggaagaaagg	18420
cactcaaaaa	gcagggagcc	ctcgtacaaa	tgagacttcc	tagctatata	attaggccga	18480
gatgcacaca	tccacagttc	ttacccttct	tcagagcctt	tgcatgttca	agtgattttc	18540
gccatgttga	acttttagaa	tggtctgttg	tggttcataa	aaagtgtagt	tggtgtcttg	18600
attgggattg	tgtaaaattt	atagattcat	ttagggagag	ctgacaaact	tacagtatta	18660
aatgttttca	tccgcggaaa	aggttgctct	gccacttact	tggtgtctct	tttatgccct	18720
taagtaaaag	tttatagttt	tctttatata	agtcttgcaac	atttccctgt	agattttatt	18780
ttgctgttaa	gggttctctg	gttggtattg	tgtataaaca	ttttctccca	attacatacc	18840
ataattgatg	gttaggagta	taataaaaag	gtagaatttt	aaaattgctg	tatgaatgac	18900
tgctttgtcta	taatttcata	tcttttcatt	tttatgccaa	tctacttgac	atgttttaggt	18960
taaatgatga	tatctgtaca	gagtaatcct	ctgatccagt	atttgcacat	cttactttct	19020
aacgtccata	gcataagata	acatcttata	ctgttgagta	catatatatt	taaggatttt	19080
accatagtct	tataatatgc	agcgtgcttt	ggttcaagac	agtgtccctg	tgcttcccaa	19140
cattaaacatt	tttttcacata	caaatacaca	ttgaccttta	tcaaattttt	aaaactatct	19200
tgagagaaat	gaccattttt	cttaatctgt	taattgtaaaa	tttttataaa	aatagtataa	19260
aataatatta	gcctacatat	ttctttctgt	ctctttttca	actcttagaa	tcagagtatg	19320
gtagtctcag	actaaaccag	gagcttccta	tctgtttctc	tggtcttaag	tcacttatat	19380
aatgtaagga	tgctgtgtat	atctgccagc	taggccttat	atacaaaaag	cacctatcac	19440
aaccttctaa	aacagtctta	ccacttagag	accatgttca	aacatatggg	cctttgaggt	19500
aattgccaca	ttcaagctat	aatattgtta	tctaagggaa	tatcttcact	tctagcagat	19560
gcctaaaaat	atctaaaggt	aaacactggt	aattgtctgt	ttgtgtgatg	ctgctcttcc	19620
tcctctctct	cctctctctc	ctcctctctc	tcctctctct	cttctctctc	ctcctgcttc	19680
tccttttctt	catctctctt	tcttttctta	tttttgaggc	atgatttcac	catgtagccc	19740
taggtaaccc	gtaacctact	atgtatgtag	accaggctag	cctctgtctc	ctgagtgtct	19800
atattaaagg	tggtgtatcac	catatccagc	aacacttgct	ttgagatggg	tagaggaaaa	19860
aaaaatatac	gtaataaaag	atggatgccca	attactaaat	tgttacttcc	agtcaaaact	19920
tgtaacctagt	ctaaggccaa	aatagggatt	ttttttctac	tttgcaagtt	ggctccatta	19980
agaggctttt	cttctctctg	tctcactaga	taggaaggag	agagaggagg	gaaggagaga	20040
aagcgggtga	ggagtgggag	gtagtgtgac	cgagaatacc	cagtaggctc	atatatttaa	20100
atatttggtc	cctagtgtat	agaactgttt	agaaagatta	ggaagcatgt	cctaggggct	20160
ttgaggtttc	aaaatttaat	gctagaccac	gtctttcaag	ggagggggcg	gtctgtctct	20220
ctctgcctgc	tgcatgcaga	gctctcagct	actactctag	tgtaagcctt	gtgtgcttcc	20280
tgccccaatg	atcataaatt	aactgttaag	aagcctccaa	ttaaatgctt	tcttttatag	20340
ttaccgtgat	catgggtgct	cttcacagaa	atagtaacct	gtggtgattt	taatatgcct	20400
ggaccaggga	gtggcacttt	taggagggaat	ggccttggtta	agagggaagt	tgctctctgt	20460
gggggtggga	atgagaccct	cgctcctaacc	atgtgagaac	cactcttctc	ctattggcct	20520
tcagatgaag	atgtagaact	ctcagatcca	cctgcaccaa	gtctgcctgg	aagctgcctt	20580
tgttccacc	ttgctgcccc	aattaaatgt	tgtacttata	agaattgttt	ttggggggct	20640
ggagagatgg	ctcagcagtt	aagagtactg	actgctcttc	cagaggtcct	gagttcaatt	20700
cccagccaag	acatggtggc	tcacaaccaa	ctgtaattggg	atctgatgcc	accttctggt	20760
gtgtcagaag	acaggacagt	atacccacat	acattaaata	aataaataaa	taaataaata	20820
aattcttttt	aaaaaagaat	tgcttttggtc	atgggtgtctg	ttcacagcag	taaaacccta	20880
acataaccct	gactaagaca	acaagtggag	aaaggtgttg	tggtgacatc	tggtatctctg	20940
gaagctcacc	tcagcatgaa	gcttgctgaa	gcgnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	21000
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	21060
nnnnnnnnnn	nnnatctaag	tacactgtac	tgcttctcaga	cacaccagaa	gagggtgtca	21120
gatctcatga	cagaggttgt	gaactcagac	ctttggaaga	gcaatcagtg	ctcttaactg	21180
ctgagcatct	ctccagccca	aaataattct	tactagtaac	atggaacaat	caagttttat	21240
tatatgatac	atattaatca	acttataagt	acatgattat	gcacatttat	catatcgtgc	21300
aaccatcact	gctgtcgttt	tgttttgttt	tgcttcttttg	aggcccggtt	tctgtgttgt	21360
tctggaactc	actctgtaga	ccaggctggt	cttgaaactca	atgatctgcc	tgctctgtcc	21420
tcccaagtgc	tgaaaaacaaa	tgtgtgcacc	accacctctg	gctatcactg	ctgtcttttt	21480

ttttttttta	acagttat	atctcgtgca	tgcattgtgtg	tataagcatg	taacgtatgc	21540
catgggtatgc	atgtggaggt	cagaggacaa	ctttcaggag	ttagttcttt	cctccactg	21600
tgggttctag	gaaccaagct	caggttgta	gacttgcattg	gcaagtgcct	ttaccacaga	21660
gccatcctgc	tggccctact	ataggtcctt	atataaaaag	atcatatgcc	gggcaaaaac	21720
caaacaaaaa	ataaacctca	aaaaacaaaa	ggaccatata	atattgtggg	ggagtggatg	21780
aagtcctgaa	cgaatgtgtt	ctgttgacat	gtctgtactt	cagacccatg	ggaattggca	21840
aagccttctt	ctggctcctgt	gaggtatgctg	atagctctgtc	taaaaaactag	agatcacagc	21900
tttctcctct	ggatgactgt	aaccccagat	tgttccctctt	cagagactgt	ccaccaagct	21960
accctgcctt	cttaagctgt	acacaatgaa	tgagctgagt	ttccagggtta	cagcacagta	22020
gacactgtcc	atcagtgaga	gcacagccta	gcctaacagt	acacatgtct	gctttcttca	22080
cgtttccaga	accaagcctt	gctggataga	gcataattgt	ctggttggct	tatttctact	22140
gataaaaagt	tttcaaggag	ggccagggtg	ggtggcacac	gcctttagtc	ccagcactcg	22200
ggaggcagag	gcaggcaaat	ttctgagttc	gatgccagcc	tggtctacaa	agtgagttcc	22260
aggacagcca	gggtctatca	gagaaaccc	gtctcaaaaa	accaaataaa	accaaataaa	22320
aacaaacaaa	caaaacaaac	aaaagccaaa	aatccaaacc	cccccaaaaa	aaaaacccaa	22380
ccaaaaccca	aaaaacaaac	gtttttgagg	tttaatttat	tgcatgtcac	22440	
agaatttcac	tgttataaaa	aatggctgaa	taatatctca	ctatccattc	acgtatttgt	22500
aggcattcat	gtgtgtagtg	gtttaaataa	aaatagcccc	cataggtctc	tacagttgaa	22560
tgcttagtca	ttgagtagca	gtactagaga	gggaattgaa	ggtgtggcct	tattggagta	22620
ggagtggcct	tgttgacagga	attgtgtcac	tttgaggtcc	cagcaacaag	gttgctctga	22680
tcacatccaa	agacattcta	ggtctatgtg	atctggctgg	aattcagaca	tgcccttaat	22740
acacaccttt	aatcccaaac	aatgaaggtta	aagttagttt	ataaaaagaa	gcacccatgt	22800
ttgaaagtga	cgtttaaatta	agagtgtatga	attagagaaa	gatctgctgt	cacagagcag	22860
agaggaaaga	gaggcagcat	aagaggagc	atggcagagg	gagagggagg	aggggttttc	22920
accagggcct	ttgtacagag	acagggttga	gagctagaac	aggtgaagac	agaacaagcc	22980
agagaatgag	agagtaggag	gagattagga	cagattgcca	atgttaatag	gctaagcaga	23040
gcatttttagt	cagaaactga	gagaagtcaa	attgaatcag	ttagcttggg	aaggagtttg	23100
agcagcaaca	gctgagttaa	actagccaac	agaatccaga	aagaactaga	aaagatgagc	23160
ttactcagca	gcaaatctca	gaggctaaaa	acatcttaga	cctagattag	actgcatgga	23220
ggctagacgc	ttccagggct	aggcctaggt	tagcagacgg	agagagtaat	aagccttgga	23280
gacaacagtt	aatacagaag	actatgtaca	gacatggata	tgaacctctc	agccacttct	23340
ccagcgtctt	gcctgtctgc	attgttagga	gtcatctagg	aaaggctaag	ggcaggcaag	23400
caactttttc	agagatgggtc	cactgttttt	tgcatggctt	ttgagaggcg	agctctgaga	23460
gggaagggttc	caagagactt	catcccagga	ttgctgctta	attacgacat	gccttttctt	23520
gtcactgtta	tttagtataa	tgactcctga	gctttagccc	atcctattgg	gcataatttc	23580
tcagatgcaa	taataagatg	aactttcaca	aattaatgct	gtttagatga	ataaatgatt	23640
ttataaaatt	cctgatttga	tttaataaat	tttaggaaga	aagctttagg	agatagttta	23700
gttggttttc	cagaagatg	taataacgtc	agaatcaaga	atagaatgtg	gctgggcagt	23760
ggtggcagat	gccttttaac	ctagcacttc	ggaggcagag	ataggcggat	ttctgagttc	23820
gaggacagcc	tggctctacag	agtgtgttcc	aggacagcca	gggtctacaca	gagaaaccc	23880
gtcttgaaaa	acaaaacaaa	aagaaaagta	agtaagggtc	gcataataaa	gaatacaatg	23940
agctttcaca	actacaccaa	gcttgggaca	aatttgtgat	aaatttggat	caaggaaaaa	24000
tattcattct	agatcagggtc	caaggatgaa	gccacaagtg	tgtgatatga	tgaacaagac	24060
catggataaa	ctgtgttttt	gagcttaaa	ataaaaacac	tgctttgaaa	ttaaactatca	24120
acattctact	gtaactttcc	ttttatctat	ttttatctat	gagataattt	tctaaagaac	24180
ttgtgtctat	aaaggtatag	aaggacagag	agaaagaaat	aaggtgtggc	atctgggctc	24240
tgctccatcc	acccaaataa	atatgtgtgt	gtgtgtatgt	atgtatgtat	gtttatctat	24300
atgtatgtat	atacatacat	gtgtaggtag	gtatatgtgt	atgtatataa	gtatgcatga	24360
acacttggga	agttgatgag	acaagtgaga	ggttggggcc	ccnnnnnnnn	nnnnnnnnnn	24420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	24540
gaaccgtgtg	cctctgcctc	taaaagtgtg	ggattaaagc	atgtaccacc	acaacccagc	24600
tagttttaa	gttttcttatt	tttttgttta	tggtgtcttt	acctgtatgt	atgtgtgtgc	24660
accatgtgga	tgcatgggtgc	ccttagagtc	cagaagaggg	tatcagatcc	cctggaaactg	24720
gagtgacaga	gggttgtgag	ctgggacttc	aacctaggac	ttctaaaaga	gcagcaggtg	24780
ctcttaatag	ctgagccctta	tctccaggcc	gtcccatgga	tttggggggc	tttgtttcat	24840
tttattttgt	tttgagacag	ggtgtgtagc	tcattgctga	atttactatg	aagccctgac	24900
tcccctcaaa	gtaaaatcc	tcctgcctct	gtctacagct	gctaggattc	gaggtcttgt	24960
accacatgct	cagcacagcc	atgattcata	acaataaaaa	aagaaagaga	gacctaaatg	25020
gccttagaga	taaaataa	attttttttt	taaagattta	tttatttatt	tattacatgt	25080
aatgcacact	gtagctgtct	tcagaccccc	cagaagaggg	agtcagatct	cattacagat	25140
ggttgtgagc	caccatgtgg	ttgtctggg	ttgaacttcg	gaccttcgga	agagcagtcg	25200
gggtgtctta	cccactgagc	catctcacca	gccccgagat	aaataaatta	taatgtatgc	25260
gtaagggtgg	atcatctcag	tctccgggaa	tcttgcctgt	tactccttcg	ctctcccttc	25320

tattcatgct	tgggtaactg	gccctggctg	attgatgaga	gctgatttcc	ccactgccct	25380
gtggcaggga	ccactgcgcc	cacagggtct	cctcaggatc	ctcagtagag	agctgcacag	25440
ctgggtggaa	gtagagggtc	gcataataaa	cacgatctca	actttatttc	tttaaataaa	25500
aattttat	aaattttata	cagctctata	taaacgaagg	aactattgaa	ggttcagcaa	25560
ggacctgcca	acggttgtca	agggtaatgg	cgatgtagt	atttttttcc	ccccttccat	25620
tttacttcca	tactttctac	attaccccac	aactggcaag	tattatttta	aatgaaagt	25680
aaatagtac	agatgacttt	gaaggaaaat	tgaatcggta	aaaagaaagc	tgagagacca	25740
cccgggaagc	ccaggctaaa	tgtaatctgg	gtcaggccctc	ccaggccctgg	ggtctcaaga	25800
tggtagctg	agggaccctg	gtgaccctct	tgggccagca	gggacgggga	ggagccggaa	25860
gctgagtacc	caaagtgtct	ctctgggtct	caagggtcctg	cacagagact	gtgtgggaat	25920
caaaggatac	aggcatgagg	actgaggcct	gacgaaccac	gctatcattc	gtcctagaac	25980
aggaggcaga	gctccaagag	tccaaccaag	aggcaggaag	ctttgggacc	cgagatgggc	26040
gattgggatta	gaaaggcatg	tttgcaaaata	ctttcaaatt	tacgatgcac	actcactgga	26100
aacccccacc	ctgggtgtcc	cttcctctgc	tcttgccaca	ccaatagct	gacatcacctg	26160
gagaaagtcc	caagaccagg	ctggctggag	ctcctgatag	gttcaccctc	cctgcagagg	26220
gccctcgaa	actagcttgc	ctgcccacac	cgccagatgt	ctgtgtcttt	ctctcttttg	26280
cctccccacc	tctgtctctc	ctccaacctc	agtggagggt	cccctgcttc	ctggggaaag	26340
tagaacttgc	cagtgtctac	tgtaatgtcg	tccctgtagg	tgtcatggtc	ccccattact	26400
gggagcaggt	atgcctcaga	tctccctcta	tctgctgccc	tttcaggctg	tctcagtttc	26460
tctctgacct	ttcctctcct	cctgaatcct	gctgtgtggc	atgcgaacag	gctcaatctc	26520
ttccatctca	aaaaacaaac	actgggaagg	tgttgagaga	cagagagcat	gggtaatggg	26580
tgcctccagct	tggctgggaa	ggggttaact	acaatgctct	actgcccagt	agggtagctg	26640
cagtgtgtcaa	ttaattgtaa	atttcaaaat	agctagtaga	gaggatttta	gatgttccca	26700
atccaacac	aaagaaatga	taaacattca	aggcagatgg	tatgctaatt	gctctgatct	26760
gatcaccgca	cattgtatac	atgtttttga	aatgtcaggc	tgtaccccat	aaatatgtac	26820
aattaccgtg	cagtgtattca	agataaaaac	tataatttta	aaaagctaaa	aacagaagga	26880
aatagctgcc	cttgaccccc	ccacccccac	aaggtccttc	ctgtttgtcc	agccacttaa	26940
tgtcagagct	tctgtgtgga	gggtgtgttt	ggtgtacaca	gacactcctt	cctccctcct	27000
tccccataag	aggagtccac	cctgtcccac	gatgccatgc	agggccacat	gcgtgatatt	27060
aaccagtaag	atgtgagcag	ggatgatacc	tgtctcttat	aacaaacgga	aaaaaaacca	27120
caccacaaaca	aaaaacaaaca	aacaaacaaa	caaaacaaaa	caggggtggg	ctgtccctgt	27180
gtcttttccc	acataaagtt	aagcacacaa	agtagccacc	atttatttat	ttgtcccttc	27240
ccccaccctc	ccccgagaca	atgtttctct	gtataacagc	cctagctgtc	ttggaaactca	27300
ttttgtagac	caggctggcc	tggaaactcac	agagacacag	agattcacct	gcctctgcct	27360
cccaaatgca	gggattaaaa	gcatgagcca	cgaactaacc	agtaccccag	agctcttgac	27420
tctagctgca	tacgtatcaa	aagatgacct	agttggccat	cactggaaag	agaggcccat	27480
tggacacgca	aactgtatat	gcctcagtac	aggggaacgc	cagggccaaa	aaaatgggaa	27540
tgggtgggta	gggaagtggg	ggggagggtg	tggnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27600
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	27660
nnnnnnnnnn	nnnggcagga	tctgtgttcc	atgtgcaaca	ctcagtgcaa	gctgtgtagt	27720
gtttgttct	gagcacctga	aggggaccaa	gcaggctgat	gccaggccca	cggttctttt	27780
ctggccccac	tgcctctccc	cacccctctg	catccccatg	atgaacatgg	ccacagatca	27840
caactactct	agctctctcc	agatccacgg	agcccatagg	tccccagatt	catctctgca	27900
gctaacaagc	tgggcagtg	caactccctc	aaggttccct	tccctgctctg	agcagcagtg	27960
tctcccacag	tgagacactc	atgtccactg	gaagataattg	tagccattaa	attcctgtgc	28020
taaaataact	agggggactt	gtcaatcact	acactcttag	ccccggactt	ctgactcata	28080
gaggggtggg	acagctcagg	gacctgcatt	ctaccaaata	gccatgtgtc	cctgatggag	28140
gaactgcccc	tggacaacct	ctgcagcaac	tgaacctctc	gtggtctcct	agttcttctg	28200
gacaggtgtg	accccatgac	ctagtgccag	gtgagagagt	gctaggggca	cactaagggg	28260
tgacaggaca	aggttggagc	tggtagatgt	ttgggccacc	aaagagaaca	ggtcagtagt	28320
aaaagccatc	atggccctgag	ccagcctgag	agtctctctc	gcagttggga	cactcttgca	28380
gtgtcctggg	gacctcttga	gggtagcatg	gtcaccaaaa	tcctacaagg	acagatcaga	28440
agtcagtgag	gtcaagggaa	cagctctagg	ttctctgtgt	ccctcacgga	cctttttttt	28500
tttttttttt	tttttaagat	ttattttatt	attatatgta	agtacactgt	agctgtcttc	28560
agacagctcc	agaagagggc	atcagatttc	gttacggatg	gttgtgagcc	accatgtggt	28620
tgtctgggatt	tgaactcagg	accttcggaa	gagcagtcgg	tgtctttaac	cactgagcca	28680
tctctccagc	cccctctcag	tctgtatgag	acagggcagc	aaaggccttg	tcccagatct	28740
gaggagagtc	atgctgaagt	ccttcctacc	ccaccccttc	cgaacccctg	aacatcagcc	28800
ccataactac	tgactccccc	acccccattc	ccttgcttcc	actgatccgg	tccctcctct	28860
ccctctggcc	ccacccatcc	tccccagccc	ccacctgatt	gtacctggtt	gtccaaacttg	28920
aaaggggagc	gcaggggagc	cttctgctgg	gcctgtctcac	tcaactggctg	tagaaatgag	28980
aaaggagatg	aagaaaaggc	ccttccccatg	gggtccccatc	ttgccaagac	ataggtgagt	29040
ccctttggct	cttcccccta	aacctctcac	ttttgagtac	ctgctggccc	gggagatcca	29100
cggcgcctcac	cgagagagac	tgttgagaaa	agggagaaca	gagaactcag	cggttctccc	29160

tetccaccct	tetggcctct	cccagatttg	ccccgcgcc	ccagcatctc	cttcagcctg	29220
actgaccact	tcccactcag	acctcagctc	tgcctcaccg	tgaaacaggg	accttgccag	29280
caggacaagc	tgagtacgag	gagcccccg	agcagtgcc	tggtcctgta	tccagaacag	29340
ggagtgttag	ttcctacctc	acgctcgaag	gccaaagcag	agactgctat	ccatgggttc	29400
cttgaccgca	ccaggtgcg	gaacctggac	tcaaaacata	gcagctgtgg	acctcactca	29460
ctctgagagg	tgggatttcc	ataagctttt	ttttccacct	gtacatttag	tcttcattct	29520
tttctgtctta	cactgtggat	cagtccctgg	ttcaaattta	aagccctcat	cttgcaagag	29580
gaccttgccg	atctcccttc	atgcctttgg	ctttaccctg	tcttggtaat	tcatggcaga	29640
agttcttccct	gctcccatgt	agatgttgag	gaccacaata	agaatctctg	taaatactga	29700
gcatgatgcc	tggcccccac	cctagcaaa	ccacctgacc	tggtgttcat	ttcatccagc	29760
ctttctcagg	ctgcccctgt	cctacccaaa	ggctctgaga	gctaactctg	gctggcagg	29820
cagccagaaa	cttctttgtt	gaccaatgaa	tgactggccc	agacaccttt	ggacttacgg	29880
gaactacaag	cctcatccca	cttctgtctc	aagttctgat	ccaggggtgt	tcggggaagc	29940
ccagctggcg	gaagggggga	ggctctcagc	ctagagagcc	ttcctttcca	tcctcagccc	30000
cctacccagg	ccttatttca	ggcaccagct	cttctaaaag	gtccttctgt	tatccctaga	30060
cctccacaac	gtgtttcaag	aaccttcagc	cagggcctca	tctccaatct	ggatatatga	30120
ttttctctgc	caagagttag	cctccagggt	ttggagtctt	agaggtttct	cctggagctg	30180
cctggaccct	tgctcctcac	caccccagga	cgctgtgaag	ctgcaggctc	cctgaataaa	30240
ttcatccaga	ccccctgcca	aggtgccaga	tgtctacttc	ctctgctgcc	caagcagcag	30300
gctgcaccac	ccctccatcc	tacctcttca	ggcttcttag	cgcagcacac	gcagcacacg	30360
gtgttctcct	ggaccagctt	gctccccacg	ctccccagct	gcagccagca	gggcttagct	30420
ctctctctcc	acaggacctt	tgctctcagc	caacccccgt	tcagcttggt	ttcagtgctg	30480
gtaaatattg	cctgtacat	gcgtttaaac	attgatattg	gggccagaag	accttttccc	30540
atcaaggcta	ccagaaccc	tgcttgagcc	tgagagaagg	gtttacagga	gcagataagt	30600
gaggaggttg	ggcctggcaa	gccttcta	gatccctcaa	cataggggat	tatccacagt	30660
cagtggagct	cagagagct	gtgtggcctg	tgtaaggcg	cagagtgggc	tccagagtca	30720
cagccaaagt	cccaccacca	ccaccaccac	caccaccacc	accaccacca	ctaccaccac	30780
caccaccacc	accaccacca	ccaccaccac	caccaccact	accaccacca	ccaccaccac	30840
caccaccacc	accaccacca	ccaccaccac	cacctcatct	accatacta	anttgagct	30900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	30960
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	tctatgaggg	ttacatttta	31020
gaacatctct	cttttttcc	tttttgagac	aatcttacta	tatttaggct	ccccttgaa	31080
gtgtgactct	tctgcctctg	cctcccaagt	gctgaggtta	caggcatgca	cagtcacatc	31140
tgctacacac	atgtcttagc	agcccttagc	agcaccagg	gtcaggaaag	cctcaactgt	31200
ccctttagct	ggcttctctt	gtgaagggtc	atgtcttctt	ccccctccta	gcattggagc	31260
ggctcttagc	ccagagcct	tccttccctc	aggtttaaaca	gcaccagttt	ttgggtggag	31320
ctcccatctc	cctctatctc	cctaagcaac	gaccttttct	gctctgactc	tcatctggca	31380
cttggaccac	aagacaaaac	tgagcctgtg	gctgtgtgtc	ctcgcacatc	attcctgtgc	31440
cccctgggag	tcaggtctag	gggaggaaga	cagggttcac	gactcagaaa	agaccactgg	31500
ctgtcctagt	gtgcctctac	ccatcctata	gcacgcacat	gctgatgtgc	ccccctcgct	31560
ccatcaccat	cctctcatgt	acacgtgcce	tcctctgcca	gacacatgca	tactaacttt	31620
ttctgacttc	ccagaaaaat	atctgatctg	agaagttagg	agtctgcat	catcagctat	31680
ggctcctaaa	attaagttag	acaatccatg	ggacatgaag	ggcaacaaag	agaagactcc	31740
tcgttccctg	ttcactctgc	ttttggcagc	accaccagca	ggaaccaaag	tggtctctcc	31800
taatccctca	tctatagcag	gtctcccggg	gggaatttta	gggacctctg	tggtctctac	31860
caggggcact	gccactcagc	tgctcaggga	gagacctctt	agaacaacaa	agaaatcaat	31920
gcagatttag	gcttcttgtt	tccttcccca	gcccctccca	tcacaggcaa	cagcctccct	31980
tggttgagcc	tcaggaggct	gatttatcag	agaggtgtct	agagaggcac	ctctggctcc	32040
tctgggttag	tagcaactga	gacaggagga	gatggtcacc	ctgggcatcc	tctaccagga	32100
agtaaatgag	atacccttgc	agatgggacc	cctgaagtgc	ctcccggggg	cggggggtgt	32160
gggtgggtgga	gtctaagtca	cagatctttg	ttaccaagtg	gttagactga	ggactgaatc	32220
tgaggtggga	aatctgatgt	gcatggggaa	acacagaggt	ccaatgctgg	ccaagagcta	32280
caagcaggga	cagggtgctag	ggggatgtct	gaatgttcca	ccccaaagcca	caggaataac	32340
ggaaatggag	actctaaagg	gcagaaaagt	aggggtgtgc	gcaggggctg	cacaggacac	32400
atgcaaggcc	ctggctgcaa	taactgggtt	ggggaggcag	tcattggcta	gccaggggca	32460
ccaggacagt	gatgccatcc	tgctccaaag	gcaggttcca	agccagattt	ctaggctcca	32520
gggggaggag	ggtccgggga	gaggggtcaa	gattctcccc	ctctgagtca	aggttggcct	32580
tcccatgtgc	cccaaatcag	gaggcacaga	aactgggatg	ttgtgtgtct	acatccaagc	32640
tgagaagaca	agtggggagc	agtacatgtg	tttcagatta	aaccagctcg	gagacaaaac	32700
tggtgtctct	cctcctccca	gagccaaagt	gccttcaagc	cacatggcag	tgaatatgag	32760
gacagtgcag	gggaggacac	ctctctctcc	actggctcaa	ggacagtttc	aaggggttca	32820
ggctgggtgg	ctcatggcta	cgccgctcac	ccccgggaca	ggttgggggt	tttccctctc	32880
gaaatcttgg	aatctgaatc	agcctgagat	acccataat	tgtaacctcc	aacaccccca	32940
gaaaggtcag	ccctgcagaa	cagaactctt	tggtccccac	ccatccccct	cagccctgga	33000

ggctgaactg	atgggagct	aagggtccaga	cagtggtg	ctcttgga	gcctgtctct	33060
ttcctttgac	tcagaccact	ccctgcccgtg	gcttacatca	ggaggtgcaa	gggctgcagg	33120
agggcagcca	gaccccaaa	accagctagg	ctaatggtg	cttattgttc	gcaagaggcc	33180
atgacctcat	ttgtctccca	gctcttttgg	taagagagaa	tgagaggaa	ctggacagag	33240
aacctagcaa	gcctcaggca	gcccactgct	ccttgctgta	aggaaccag	caccgatggt	33300
tctgaaaagc	agcgatccga	atggagtcag	gctgagctgc	aggaagctca	ccttccttgc	33360
tcactgctgg	tggaagcaac	ttcaggaaga	gcccagccta	tggaactata	gctcctccgg	33420
ggtactgctg	agtccagccc	cagagcttag	ctccctgctt	cccaccaccc	accaccacat	33480
cccttcccaa	caccattcaa	aaccccagtc	cagcctctcc	tactggtcta	cagtgcagg	33540
ctaataagagt	cctgggcttc	tgtcccccga	attctctctc	ccctctctac	tggtcacctt	33600
gggtccctaaa	ctgcaggggc	tactataacc	ctacctccac	ttccttgac	ccctcttttc	33660
tgctctctgg	ggtgcccctg	ccactccag	tccctctagc	cagggagcct	cttccatata	33720
tgctctcccc	aggctagacc	aggcgctgcc	ttacctgtgg	ttgcggcagc	ttctctcaca	33780
gctgcactc	tgaggggctc	caggaagcag	tgaggggagt	agctgcctct	caaccagcgt	33840
ccagcaggct	tcagattaca	gctactcttt	tcttaaagt	acctgactcc	atttggaatc	33900
tgtagattgca	cttatgtctg	gtgttaactt	taacccactg	ctgcccttcc	gccatgtggc	33960
tccaagacca	cacgttgggc	accctcctct	cccaccacat	ctcccttggg	tctttatctc	34020
tcttcattgg	gaccttcatt	gggacatgat	ggctaacttc	aggggcactt	ggggcagcct	34080
gggttaggtc	atgagtcctga	acttgaacat	ctgaaaggat	tggtcgagag	gcaggctgca	34140
tgagagagact	gtgagccagc	cggtatggag	atgctgggtt	cttccaggcg	cttgggtctg	34200
gctcactgca	ggtgggagca	aggtgatctt	tctcccctcc	tcacctggaa	aatgaaggaa	34260
tggtgactgta	cctgacagct	ctgaaggctc	caaaggacag	tggtgtgggg	actagagagt	34320
tggtccactg	cttatgagca	ctggctgctc	tcgcagagga	cctgagttct	gttccagct	34380
cacatagcaa	ggactcgaaa	ctgcttgga	ctccagctcc	agagaaatctg	acgctatctg	34440
ctnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34500
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	34560
catgccttta	atcctagaac	tcaggaggca	aagcagatgg	acatctgaat	ttgaggcccg	34620
catgtgtctac	atagcaagtt	ctaggctagc	caggtacatc	ataagaacct	ttctcaaaaa	34680
ataaataggg	ccagttggca	aaatttagct	tgccctccta	acacaagaac	ccaaagtcaa	34740
ccccagcaac	catgtaaaaa	gaagcaaggt	gtggtggcac	ttgcttgtag	tccagcattg	34800
tcaaggtgga	gacagacgga	tccatggggc	tcactggcca	gccagctagc	tggtctactt	34860
agtagctctc	cagccagtga	gagactgaaa	aataaataaa	taataaagg	gttaggaaga	34920
ggtaactagg	tggtcagtg	agaaaagata	cttgtctaac	aagcctagca	accctcaatt	34980
ttcagtgccc	actaaagggtg	gaaggagaga	accaactcca	aagaattgtc	tcctgacagt	35040
tttatgtgtg	ggtacacaca	cagcacaca	cacactgtgt	acatgcata	gcatacaatt	35100
aataatttaa	aatgttatgt	gtatgggtgt	tttgcttaca	tgcatatctt	tctgtgcacc	35160
acatgtgtgc	aatgctgtgt	aaggctagaa	gaggacatca	gatccctctg	gagttacaca	35220
gggttgttag	ctaccatgtg	gattctggga	acaaaacat	gggttttcca	aaagaggctc	35280
ttaatcactg	agccatctct	ccatcccctc	aatagtatat	atttctgggg	ctggagagat	35340
ggctcagagg	ttaagagtat	tgactgctct	tccagagggt	ctgagttcaa	ttcccagcaa	35400
ccacatgggtg	gctcacaaac	atctgtaatg	gaatctgatg	ccctcttctg	ctgtgtctga	35460
agacagctgc	agtgtactca	tatacataaa	gtaaatataa	aaaccttttt	tttttttgtt	35520
ttgttttttt	tggtttttcga	gacagggaga	cagggtttct	ctgtatagcc	ctggctgtcc	35580
tgaaactcgc	tctgtagacc	aggctggcct	tgaactcaga	aatccgcctg	cctctgcctc	35640
ccaagtgcgt	ggattcaggg	tttgcgccac	cgccaccacc	tccaaggctg	ctgctgcggc	35700
caccaccacc	ccaccaccac	actacctgac	tatttaactt	ttaaaggcag	ccatctcatg	35760
gaaaatgaca	cctagcattg	tccctctggc	cctacatgac	cccatgtgca	aacacatacc	35820
tgcatataca	cacataaata	cataagtaaa	cttagtctgg	ttgttttgga	aatgtgctat	35880
ggtttgggatt	gtgtccccc	aagggaataa	ttgggtccca	gagtagtact	attggaagag	35940
agtagaactg	ttagggttta	ggcctgggtg	gaagtggcca	tggttagaga	gacatgccaa	36000
ccaaggggaa	tctctggctt	catcttttcc	ctttgcttcc	aggtcctaag	acagtcaaaa	36060
ggctgcttca	ccacatgccc	agaagcaagg	ggagcagtca	tggtcgggac	cgctaagtca	36120
gctgttgatt	tccccagata	ttttagtag	taagagacag	gtgaggaacc	ccacagcaag	36180
tgtagtagta	tggtgtgtgga	ggtgccctcc	ggggacgggg	gccctcctgg	ggcaggacgt	36240
tcctcttctc	catccacctg	cactccgaga	acaggaaatg	gtgactttgg	cagagcttaa	36300
gcagagcccc	ttcatgttac	aagtatgtaa	attcataagg	accagtttct	ctccatata	36360
aacagcttca	aacaggagaa	ggaagaagca	aacattaagg	aaaagctctt	ttattgcaga	36420
ggctacactg	aagctaccgg	ccgccttctc	ggaatgtata	atcagcttcc	ctctgggggt	36480
tctgtagagc	actgagacat	taagtactac	tggtgtccag	gattctgcct	atgaagagga	36540
gggccccctg	gtccgtgttc	ctcagaacaa	agaggaaagg	ttgttaagg	tgatagtcta	36600
gcgggaaggt	gaggggagc	ggctggaggc	ctgggtggg	gctgcttctc	gccccctctt	36660
cattccactc	gaaagcagcc	ctgtgttcca	cttgggtgag	cttcaagggt	ttgccagtaa	36720
tcttgcgtga	gtcgggtgat	tcaaacacg	actgtagctc	tgtggagatt	cagagattcc	36780
attaacacca	cacacacaca	cacacacaca	cacacacaca	cactccctgt	ttgtgtaggc	36840

tgatatttcaa	gaaagcaagc	tagaagtggg	gtacctcaca	gtgacttgtg	agctatgagg	36900
cactctgtga	caggctcagt	gacctacctg	agaacttata	gccaagatgg	ctgaagccag	36960
acctggcctg	agagaatggt	ttgggctgtt	ataggacaca	tagagatata	cacacacaca	37020
acacacacac	acaccaagga	ctgagtctaa	tgggaggtgg	ttcttcattc	ccctccctcg	37080
taatgggtgc	acatgttccc	tgagccaccc	tacaaagaaa	gccacaggac	tcagttctgt	37140
cagcaagggtg	gcaggctcca	agactcagcc	ccgagcgcaa	agtggccttg	caaacatact	37200
catgtctctg	agagacttgg	taagttctgc	ttcgaagctc	agcttcagct	tggggacagt	37260
cagcacagct	tggatagtct	tcagttctcg	gtcgaatgtca	tgaatgaact	cagaggtgag	37320
gctctctctt	atcatggtca	agttctgggt	cacggtcagg	ggcaggaaga	agatgatgct	37380
catacttctt	gtcaagggtca	gctgggcaat	ctaaccacac	agagatgcgc	acaggttagt	37440
tgtgagccag	aaaaaaciaa	acaaacaaac	aaaaaaacac	caacagctgc	cttcccctct	37500
gctgtaacgg	ggccccagcc	ttgtgctccc	cagcctcagc	ctgggctgta	ggctactggg	37560
tactggcagt	ccttccatga	gtagggagtt	ttcttctcag	cctaaaaacc	acagaagttt	37620
aatgaacaca	cggttgtttg	tggttccgct	acggtttcta	ttgtgataaa	acatgactga	37680
aagcaacttg	gagaggaaaag	ggttttattt	atctgacaat	tcgcagggtg	tcttctcatc	37740
actaagggtga	ctcagggcag	gaactgaagc	ggaagccgtg	gaggaaacgt	gctttctggc	37800
ttgctccccc	tggcttctta	gcctgctttt	ttatgtctat	cagaaccact	tggccaggag	37860
tgacactgcc	cattgtgggc	tgggcccccc	cacatcaatc	actaatctag	aaaatgacc	37920
acgggttttg	ccagaggcca	gtctggtggg	ggcattttat	caattgagtt	tcacccttcc	37980
aaatgactct	aacttgtgtc	aagttgacca	cacgaatcag	ggcctggttc	ttaggagctg	38040
aagtggaaatg	tccccagag	actgctgccc	agcactgctg	accatttget	ttgtatagag	38100
cattgaacca	gaaatgaaca	ataaaatgga	tcctttgaac	agatgtgttg	atcctagggc	38160
ctgtggacac	agcgactggg	cttcccagag	ccccatgga	atcannnnnn	nnnnnnnnnn	38220
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	38280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	38340
gagagagaga	gagagagaga	ggataaaaag	ttagcccagt	ggtggtggca	cataccttta	38400
attccaacac	ttgagaggca	gaggcagggg	gagctctgtg	ggccagtttg	gtttacagag	38460
taagtttcag	aatagccagg	gctacacaga	gaaaccctgt	cttgaagaga	aacacacaca	38520
cacacacaca	cacacacaca	taagatcttt	aagaagaaaa	gaaaggatag	gaaaggatag	38580
tggggaacaa	tctgagcaga	ggaagaaatg	gggtgcgcag	gacacccacc	ctcagaggag	38640
gcccctcactg	gaggtgtctg	cacaggagaa	cacttgcact	cagcttgccc	tagggcggtca	38700
gaaactcagaa	ttcagtttca	aagcactgac	aggagcagtg	actggggacc	ccaggttgaa	38760
tccccctctt	atctaaaatg	agtaagaacc	aaaaaaacaa	aagtgttttg	gatttggaaat	38820
ctgggttatt	tgcacttaca	aaagaggtct	tggggaggaa	acccagttct	acccctggaa	38880
ttcatcagtt	ttctataccg	ctgactacac	aggggctgaa	ggtaactctca	atgttttcat	38940
aactggtgtg	gtgctacttg	ctcatgatcc	caacacttgg	aaggtaggtc	agaagttcaa	39000
gagcagctct	gactactcag	tgaattttgag	gctagcctgg	gctacatgaa	aactcataaa	39060
acaataaaaag	aaaagaaaag	gtggcgagtg	aggtggccca	tcaggtaaaa	gtgccaaaca	39120
cctcgctcga	aagcctccac	acggaaaggg	aaagccagct	tctacatgtg	gtcctctgcc	39180
ctccgcacatg	accatggctc	gtgcaacccc	acacccaccc	acccacccac	ccacatgaca	39240
taaataacttg	taatgattag	tttctgaaga	acaatatatt	cgttgatctt	gtttagggaa	39300
caaaagttcgt	gcacattgac	ctgtcgaccg	tgtagtacgg	gatccgctcc	aggaagctaa	39360
agatttttggg	atgcttcatg	ttgcagctac	tctgatgaac	agtgttgctt	tctgggccag	39420
gtaatggtgg	catatacctt	tgatcccagc	acttgggagg	cagaagcatg	tagatctctg	39480
tgagttcgag	atcagcctgg	tctacagagt	gagttccagg	atatccaagg	ctatacagaa	39540
aaaacccctgt	ctctaaaaat	cactaattta	aaaaaaaaat	cctttctaaa	cctataatac	39600
aaatgttttg	taggctgcct	taacaaagcc	caatggccat	tcagagaagg	ctcaaaagag	39660
aaagtttagg	ggactctaca	agcatcctca	ggaaggccac	agaaagcaga	gcctgggcca	39720
gtgagactttt	gcagtgggca	aggttcagct	ctttatgtag	gaagaagaga	gtcaacagtc	39780
agagtccagc	tttccataaa	acctgtgcag	ggcctctagg	caaagccctg	tgttaggggc	39840
aaaggcattt	gcagtctaag	cccgttgaca	tgagctcaat	ccttggaacc	caggtggaag	39900
gagtggtctg	actccacaaa	tttgtcctct	gatctataca	tgtagtgcag	tgcaacgcaca	39960
ctcacatata	tgtccacatg	cacacatgca	tatacatgcg	catgctgca	cccacacaca	40020
gggtcaaaaag	cagcaagaga	tgccctgtga	aaaacgtctc	attcagctctc	ccatcatcca	40080
gtgccacact	ctgagcacag	gtggtactga	tatcgttctt	gattgatcga	tcagttgatt	40140
tgagaccccc	cctcactatg	tagcccaggc	tggcctggaa	ctcacagtga	tcctcttget	40200
tctgcaagat	gagcccatca	tgccccacat	gttatgtgaag	caataccatg	ctctataaag	40260
caaacttagg	caggcaggat	ggtggactcc	tgtaattctca	ggacttgaaa	agtagaaggg	40320
agatgaggag	ttcacatcaa	tctcccgat	gcgttgagg	ctggagtggc	tggtccctgg	40380
gcgcttctgc	cagcaactga	ccaatgcaga	tgagatgct	cacagccaac	catcagactg	40440
agctcgggac	cccagtggag	gtgctggggg	gaggactgga	ggagctgaga	cgggattgca	40500
agcccatagg	aagaacaatg	tcagctggcc	aaaccaacca	gagctcccag	ggactagacc	40560
acgaaccgag	gactgcacat	gaagggatcc	atggctccag	atgcatatgc	agcagaggac	40620
agccttctct	gacagcatgg	gaggggaggc	cattggtcct	gtggagggtt	gatgccccag	40680

tgttgaggga	tgctggagcg	gtggggcagg	agtgggtggg	taggtgggga	gcaccttcac	40740
agaggcaaaa	gggatggggg	agaaggcaga	tgggatgggg	gggttggtga	ggggaagaa	40800
agaaaaaaga	tgtctctgaa	agtaaaaagt	acttgctcact	aagcatgagg	atatgagtca	40860
acccccagc	cccacagggt	ggaaggagag	aatgagttcc	cacaagttat	tttctgatct	40920
atacgtgcaa	tccatggcat	acgcagaaac	gcaaagacag	acaatgagtt	gggtgtggtg	40980
gtgcacatgt	aattccatca	ttcaggagac	agaagcagca	gagttgttgg	aaatctaagg	41040
ccaacctaaa	gacctacacc	caaagaagga	caaaactataa	ggaaaaaggt	ggtcgaccaa	41100
tgtaacattt	aagttagaaa	tctctcttca	cactgtgtag	atactgtaca	aggaagagaa	41160
aaggcagcca	catcaaaaaca	gtgtaaatca	acgagaaaac	cagaaaacaac	tcaagagaa	41220
gctgcagggg	cctgaattct	gttctcagaa	cctgcacata	gccaagagaa	tcaaaaactgt	41280
ctgtaactcc	agctccctgg	gatccaacac	ccatttcttg	cctccatcag	catcactcac	41340
aggtgtgcac	acatacacat	caataaaaaat	caaaaaccagg	gatgaagggg	tagggagggtg	41400
catgtggatc	tgaggaggagc	tgagaggcac	tggtgaata	caataaaaaa	tttgggtgat	41460
gggtgtgcac	gcctttaatc	ccagcacttg	ggaggcagag	gcaggcgaat	ttctgagttc	41520
aagaccagcc	tggtctacag	agttagttcc	aggacagcca	ggtctacaca	gagaaacct	41580
gtctcaaaa	aacaaaaacag	ccgggcggtg	gtggcacacg	cccttaatcc	cagcacttgg	41640
gaggcagagg	caggtggatt	tctgagttcg	aggccagcct	ggtctacaaa	gtgagttcca	41700
ggacagccag	ggctacacag	agaaacctgt	tcttgaaata	aataagcatt	tggtgtgtgt	41760
acagaaaaac	ccagcccagt	ttccagcaca	cacaggttga	ctcacacat	cataactcca	41820
cttccagggg	atccaatgcc	ttcttctgac	ctctgtgggc	accaggattg	catacagttc	41880
acagacatgc	acataggcaa	aacactcaca	aaataaaaata	aatctagcaa	aaaaaatctt	41940
aatataaat	ttaaagaaaa	aaataaggaa	gccgggggtg	gtgtcgcacg	cctttaatcc	42000
tagcacttgg	gagacagagg	caggcggatt	tctgagttcg	aggccagcct	ggtctacaaa	42060
agttagttcc	aggacagcca	gggctacaca	gagaacctct	gtcttgaaat	aaataaataa	42120
ataaaaaaat	aggccaagta	attcttggaa	gaatcccaag	gggacactaa	gtgtatataa	42180
agggcttcca	tagggctagg	aatgaggtct	agcgagagca	acttcgctgg	tgtagtaaa	42240
tccctcagct	gcattgtggt	cctttaatct	aggctctccc	gaagcagagg	cagaaggatt	42300
tctgtgagtt	caaggccagc	ctgggtgtaca	tagctagttc	caggacagaa	agggcgatat	42360
aatagaaaaa	tcntaccctag	agcccnccca	aaaaaagggg	agacctgaga	ccagagagat	42420
gactcagttg	ctaagagcat	tgactgctct	tccagaagtc	ntgagttcaa	ttcccagcta	42480
aaaattttatt	taaatgttta	ttacttgtat	tattatttaa	atttaataaa	ataagtaaat	42540
gggagcctag	gtttgagttc	ccaaatcacc	aagaaaaaat	gttatcattg	ctaataatca	42600
aatataagag	ataagaactt	cttttttaag	aatctcttatt	tatttttatgt	atgtaagaac	42660
actgtagctg	tcttcagaca	caccagaaga	gggcatttga	tccatttaca	gatgggtgtg	42720
agccaccatt	tagttgctgg	gaattgacct	caggacctct	agaagagcag	tctgtgctct	42780
taagctatga	gccatctcta	cagctcttat	caggttgata	aaatttaatc	tcgtggagcg	42840
ctgagaccaa	gaactaaagc	tgggagattg	aaaaatgcag	accaccaagg	ccctgctcat	42900
ttctccagtt	ctgatcagct	cccgtaccag	gggtctaac	aggcctgtgt	ctgcttccct	42960
gagtagacca	gaggccccat	ctaaacagcg	tgccctgcag	agctcctctc	tctaggttga	43020
cagatgggaa	tttcagacca	atgtcatctt	ccaggacatc	aacacagcag	ccaaatttat	43080
tggtgctgtg	gctgccacag	ttggtgtggc	aggatcaggg	gctggcattg	gcacagtgtc	43140
tgattattgg	acttcttaaat	aaccagtctc	tcaagcagca	gctcttctcc	tatgccatgc	43200
tgggggtttg	cctgtctgag	gccatgggac	tcttctgttt	gatggtcgcc	ttctcatccc	43260
tcttcgcat	gtgaggctcc	ctggggtcac	ccagccgtcc	ctgctgcctt	gactccatgc	43320
cagtcctggt	gctggagttc	actgagattt	accattaaac	agcaacggtt	ctctaaataa	43380
ctattaatta	attaattaat	cacgtgacaa	ccccagcgtc	catatgggtg	tggaataatga	43440
ggaactctac	ccatcataca	tggcgactat	gaagaacaat	gtgacagaaa	atgctaacat	43500
catgtgtgac	cgcattgcac	agccctgact	gctaaaagtg	gacaagcccg	aagcgaaagc	43560
ccaatgttct	acttcttaaat	gcatgcacca	aacgcctccc	acaggaccag	aggtgcagct	43620
ctgatagggt	ccttgccctg	catgcatgaa	gctgtggaca	cgaggcatta	ttcgcaagaa	43680
cattctagct	gtctggaggg	ccctcaatcc	actgtgttcg	cgctgttcca	gcaccagtgc	43740
ctcctggggc	tgacacctga	aaaggggact	gcttaagagg	gctcctacca	agcctactgc	43800
cacagatgca	tgatgggaaa	gccttcttga	agcaactggc	tgccaaaggc	tctggacaag	43860
agatcaccct	ctactggaaa	ggtggtttca	gtctaggttc	tgtgggattc	caggaaatta	43920
gacaacactg	gcagtccaac	agacagacga	tctaaacttc	caaggcacag	ctggtagaac	43980
ttgctgcgga	accagacaac	aaggtagcag	ctactcccat	acaacataca	aaaaagcaga	44040
gagagagtca	gagacagaga	cacacagaga	gagacagaga	gagagtctaa	agagagttag	44100
gggtctcagga	ctgagggtat	agtctactgt	agagcatttc	cccagcatat	acaagacctc	44160
ggattcaatc	tgaacacagg	aaaaaagggg	gggggggact	tcgattatct	caaattctcc	44220
tttttgtgac	acacccctaa	agtcaactgc	tacttccctc	accgccatga	agtaagagac	44280
tggttgcgct	tatgtctaca	cagtctcgcc	tccacttccc	tcctccccct	tgcttctgtg	44340
ctcatctcct	ctgaaaccac	tgacagcaag	gacttgtgtt	gactgccaca	cggaaactct	44400
cctcagtagc	agcgagcaga	gcagagctct	gtcttctcgg	agcttcttct	ctcttctcgc	44460
cattttctcc	cacccttaag	tacctatctt	tctctgtctc	tgcttgttga	tccttggacc	44520

cttttccctt	ctatgaacaa	aatatctcct	taaaggatct	cttctagtct	aggggtcccc	44580
cgccccact	gtggagaaaa	cccagggcct	tgacacatgct	cagcaggagc	tccatccagt	44640
ctctagctcc	atgacttaaa	gcacatctgt	gctgtcaaat	atacacttcc	agcccttacc	44700
aaaatattca	gtcaactcct	tgccattcaa	aatggatgac	ctcaaagcca	gagtcagcgg	44760
tgctatgact	cccagatcca	tccacttggt	agcccaggaa	tgaactcann	nnnnnnnnnn	44820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	44880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnng	ttactgggat	ttgaaactcag	gacctttgga	44940
agagcagcca	gtgttcttaa	ctgctgagcc	atctctccag	ccaccaccac	caccaccacc	45000
accaccacca	ccaccaccac	cacccacca	ccaccccgcc	ccacctgctc	attcctgatt	45060
gggttggttag	tttagtctgt	gagacaggag	ctgtcccttt	tctatagtgg	aaggtgaata	45120
agaaactcct	gaaagtgaag	gcctacaaaa	cagccacact	tatttggttg	aaaatactgt	45180
aaatgtgaca	tgtaaataca	tgctaaaata	attcgttaag	tcagtgaaca	accttaaaac	45240
acagtctgta	gcctgaatta	cagacacgac	acgagccatg	acagaggctg	aaataagacg	45300
cctttgcgaag	gagaagggca	gaagcttcca	tccttgctag	caaatctttg	ttccaagctt	45360
tatcagattt	tattgctttc	tctttctgtt	tttctttatc	atatttggtt	atttggtggg	45420
ggaaaagccta	atcttccatg	cccatgtatg	tgagcacttc	agcatatgtg	tgtgaacacc	45480
aacagcacac	gtgtatgagc	accacagcat	aggtgtgtga	gtaccacagc	acatgtgtgt	45540
gagtagcgata	gcacgtgtat	agagtccaga	ggagaactga	gagagtccgt	cttttccctcc	45600
tactgtgtag	ctctcagggg	tggaacttgg	gctcagcctt	gggtggcaagc	tcctttatcc	45660
acagagtcac	gctgccagcc	cagctttctc	tttttctctc	tggtatgtct	atccactctg	45720
ttcaaggcta	actcactgac	tctgagttat	cagaactgct	tgtagagaga	ggagtaactt	45780
tggaacatct	tgctggtagg	aacaccatcc	ccactcggtc	tggtatgacg	aggggaaaaa	45840
aagcatcacc	aaaggatttc	accacctcaa	ccagcaaaata	tttacctcct	atacatggat	45900
aggtgggggtg	ggtaggcctt	gtgatttatc	gttaggatct	catgggagtg	attacagctg	45960
gtctactcca	tgacaaaaat	ggtagcgggtg	gctgacaaaa	aaagaaacagc	tacacctggc	46020
tctagtttttc	tttctttctt	ttttcttttt	cttttaccct	acggtactaa	ggattgaacc	46080
caggaatgca	agagctctgc	caagtgcgct	acattcccag	atctgttttt	ccattttctt	46140
ctttcctttt	agattttatt	tcgattttatt	tgtctatggt	tgtgtgtatg	tgtgtatgct	46200
tatgtgtacg	tatcagtatg	ccatgggtat	acagaaacct	gagaaggcca	gaagagagtg	46260
tctgggttac	aggagtttcg	agctgtcctg	tgggtttctc	aatgcagcag	caccaccacc	46320
aatcaccccc	acccccaccc	agccttcgag	ttcaattctc	atcatcacia	aaacacacac	46380
acagacaagg	gcctgcaaga	tggtcagca	ggtaacgaag	ctcgtgtcat	aagcctgaga	46440
acctgagttc	actgtctgga	acccgcgtaa	agggggaaag	gaagaatcaa	ctctatgatg	46500
ttgtcctctg	ctctcccat	gtgtgccatg	gaatgaacag	ccctcccaaa	cacacatcaa	46560
gaataaataa	aactaaaaat	agcttagtaa	cttttatggt	gaaagtgggt	tttacctgag	46620
ggggcaacaa	taacaccgag	agtagaaagg	caagcatgta	tgtcactgaa	cagcattgaa	46680
gaaaaaacaa	acacatttcc	tgtacatcgt	tctgggagtc	tgagttaggg	tttctatgct	46740
gggataaaaa	caccctgaca	aaaaattaacc	tggggaggaa	gctgtttatt	tcagctttta	46800
tgtctacaac	atgacctgtc	acccaggga	gtcagggcag	aaattcaaac	aggtcagaag	46860
cctggaggta	gaagctgata	cagaagccat	ggagctgctt	ctggcttgct	ctagcacaca	46920
ggagtactag	cctaggggct	gtactgcccc	cagtgggcta	ggctctccca	cagcaatcat	46980
aaatttagaa	aatgcactac	aggtttgcac	acaggccaat	ctggtagggc	cattttctca	47040
attgaggttc	cttcttccaa	aaggacttta	gcttgcat	tggtgacata	aaaactagcc	47100
agcatattgg	gatttatagat	attctcataa	aaaaaagaca	tttagattcc	cacataacac	47160
catattcaga	aattaaactca	atgtgaacca	gaagctctga	aagtaagagt	taaaactatg	47220
aaaaattctt	acaacctacc	atacaaaaaa	tctgatgccc	tcttctggag	tgtctgaaga	47280
cagctacagt	gtacacacat	ataaataaat	aaataaatat	ttaaaaaat	atatgaaaaa	47340
tcaggctgtg	gagatggctc	agtgggtaag	agcaccggac	tgctctttcg	aaagtccaga	47400
gttgcaaatcc	cagcaaccac	atggtggctc	agaaccatcc	gtaataagat	ctgactccct	47460
cttctggagt	gtctgaagac	agctacagtg	tacttacata	taataaataa	ataaatctta	47520
aaaaaaaaaa	aaactatgaa	gaactatgaa	ctacaagaag	tcaggaatag	ggctgggggt	47580
gtaacccaac	agaaaaacac	ttgcctggcc	tgctgttggt	ctctagcacc	accaacgtag	47640
aaagagaaca	gcagaggatg	agggcatcct	gacttgagtc	aagtgaaca	tgataatcct	47700
cgagacacca	aatcaccaat	gataaaaagag	atcaacaagt	tgggctttat	ctgaataaag	47760
agctgtgtcg	ttaaatacca	cgcagggaag	gaagaggagc	tgagtctggt	aacacaggcc	47820
tgaaatccaa	gctactgggt	ggactgaggg	aggacaacag	ctagctcaag	gcccacctgg	47880
acgccagagt	taactcagag	agcagcttgg	gtagctttaa	tgagactctg	cccaggccag	47940
tgacacaggag	agatggctca	gtgggttaaga	gcattactcc	tcttgcaag	gacctgagtt	48000
caattccag	caccacgtg	ggcaattaca	atcatccata	actttagt	caggggatcc	48060
aatgcccttt	tcacagtacc	aggcattgtac	acagtgcaat	tacatacata	catgcatgca	48120
tgcatacata	cacaggcaaa	acttacataa	aataactaagc	agataaatct	taaaagaagc	48180
cgggcgtgg	ggcgcatgct	tttaatccca	gcacttggga	ggcagaggca	ggtgtatttc	48240
tgagttcgag	gtcagcatgg	tctacagagt	gagttccagg	acagccagga	ctacacagag	48300
aaacctgtc	tggaagaaaa	taaaaaaaa	aaagaaaaa	atcttaaaag	aaaaggagag	48360

gactggagag	atgggtccac	agttaagaac	acttgttctg	aggtctacag	agtgaagtcc	48420
aggacagcca	ggactatata	gagaaaccct	gtttcgaaaa	acaaaaacca	aaacaacaac	48480
aacaacaaca	acaaaaccac	ttgttcttac	agaggacttt	ggtttgattc	tcagaatcca	48540
catgatgggt	cacaaccatc	agttgcaggg	atccaaggtc	ctgtcttctg	tgggcaccag	48600
gcataatgtg	ggtgtacata	catgtatata	ctcatatata	taaaataaaa	agtttttaaa	48660
aggaggctgg	gtttgttagc	cagaggtaga	ggtaaaaaga	ctctagcttg	tttaagtgtg	48720
acatgaaaaa	aaaaagacat	ttagattcct	gcatcacacc	atatccaaaa	attaactcaa	48780
tgtgaatcat	aagctctgaa	agtaagaata	agcctagtat	gcactgtaag	gctctgggtt	48840
cactccccag	cactgcacaa	gatcatgaaa	ccagaaatgc	agatcctctg	aaccacagca	48900
tgggaatgta	actcagccga	tgcagtgtct	acctgtctgt	tacagagcac	aggataaatt	48960
gatttgtgtg	gtgcatacct	ataagctcac	tacgtggaaa	gtagaggcag	gacgaccaa	49020
gggtcagtg	catccttgg	cacatagaga	atttgaggcc	agtctgggtc	gctgggtctat	49080
ttggaatgct	gtctcaataa	ataaaagaaa	gaaagaaaaa	gaaaagaaga	agtcctatga	49140
ttgtcttaac	ctctgacctc	tgtgttcatc	aagtctcctc	ctcaggaaact	cactggctcat	49200
cttgtgaaaa	cctaccccg	agtctctgtt	cagaggaccc	aggtctccagc	tgtgggttac	49260
acataggatt	tttatactag	aaaaataaaa	tgaataagta	tgtatttttt	aaaaagggtc	49320
agagctggat	atggtgtgtg	ctagtatatag	catccagaac	tgagacagga	tagccatgag	49380
gttgagaaca	gctagactat	acgggtctcaa	caaacaaaag	taagggatct	gagtagatga	49440
ggtttttaatt	tttttctttg	tgtttgttac	ctaactgtga	tggttgtttt	gaatacatgc	49500
atgtctgtgt	atcacttgtg	tgcctgaaac	ccaagggaagc	cagaggaggg	catcgggtcc	49560
cccggaagta	ttattacaga	aggttgtgtg	cagccatgtg	ggtgtctggga	atcaaatctg	49620
aaagagccac	ctcgggctgg	agagatggct	cagtgtgtta	gagcactcaa	tggctgtctc	49680
tcagaggttt	tggagatcaa	atcccagcaa	ctacatgttg	gctcacaacc	atatgtaatg	49740
ggtatccgatg	ccctcttctg	gtgtgtctga	agacagctaa	agtgtactca	aataaataga	49800
tcaaaaaaga	aaaaagaaac	agccacctct	ccactctccc	tttttaaaat	cctcttgctt	49860
ctgtccctta	atgttaataa	cacaggtata	tgatactatg	ccttgtttat	gaatagaaaa	49920
tacacgtgct	aaagcaagtg	tgaaccttaa	atacattatg	ctgagtaaaa	ggagtgaagt	49980
gcacacaaga	cttttctgct	caagagtatc	tgtatgaagt	attgaacatg	tgaactctga	50040
aatcgggagc	tgaaggaagt	atggggagtt	ctaattggcta	caacatttct	ttttggaatg	50100
atgaggatgt	cttagaactc	aaaaatgggtg	ataactcagc	atatatacta	aaactcattg	50160
aattgtacac	tttaaatgaa	tgcataaaaa	cttgtctcag	taatgtgggt	tagaagatgt	50220
acagacatgg	gtgtgtgtgt	gttaaaacat	ttcttggcat	ggcaataaaa	atacagtttt	50280
agccaggtgt	ttgtggctca	aaaaataatg	ataataacaa	taataaaaaat	aatgaaaaa	50340
gaggctggag	agatggctca	gcgggttaaga	acactgactg	ctcttccaga	ggtcctgagt	50400
tcagttccca	gtaaccacat	ggtggttcac	agacatctgt	aatgggatct	gatgcctctt	50460
tctgatgtgt	gtctggaacc	agctcacagt	aaagtcattg	caaggacttt	acaatagtga	50520
coatgataac	attgaagcta	gacttgcctc	tactgctgag	tgtgtctgct	ggctctttct	50580
aaggagtaat	gttagctttt	tgtcctaaat	ttgtttcctt	cctttcctct	ctccctctgc	50640
tgtttttctt	taccctctt	ttactttgct	ttcccctctc	atctcctctc	ttaacagagt	50700
tgtcctatgc	agcccaaatg	ccatcttctt	gcctcagcct	ccccagtgtt	gaaaaatact	50760
ctttccacag	gttatgttag	gagactggag	tctgtctcag	cggggagggg	gcctgggtca	50820
agttctgagc	tcaattctct	ttctttcttt	ctttctctct	ttctttcttt	ctttctctct	50880
ttctttcttt	ctttctttct	ttctttcttt	ctttctttct	ttctttcttt	taagacaggg	50940
ttctttgtta	taccctggct	gtcctggaac	tcactttgta	gctggcctgg	aattcagaaa	51000
cttgctgccc	tctgctcccc	aagtgtctgg	attaaaggtg	tgacccacca	ctgcccagcc	51060
ctgggctcaa	ttcttaacat	tgtggagaga	aaagatttgt	agctgttctg	gccacctgga	51120
attactttgt	ttctgatctt	ttgctgcagt	caaatccttc	tcacccatct	ttcctcgtca	51180
ggctataata	tagactctcc	ttgcaatact	tggaaatgct	ctacagtcat	ctacatcctc	51240
agtcctgtct	ctatattttt	tactaagctt	ccttctaagg	tctttatttg	tttatgattt	51300
acacagaaca	tttttttttc	ttgtctatag	catgcttag	agtgatcggt	gccagataga	51360
ggaaagagaa	atgagagaan	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	51420
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	51480
cagctactga	ttcctctctc	tccctcctcc	ttcctcctcc	ctccccagcc	tcagtctctg	51540
ctcatcttgg	acttctgcgc	atgtcctcag	cccagacctt	ctgctcttgc	ttctcctctc	51600
cccagcagcc	ccccagttct	cttctgaaa	cttctgaggt	actctccatc	acctcctttg	51660
gctcctgtct	tgattgggtg	cacctgctgg	ataggcttgc	tcctgactcc	actgttctgt	51720
tctcaattag	ggaccctcac	cctctgatat	accacacatt	tcctagtgt	ctccacctcc	51780
cacccccacc	ctatacgcac	atacacactt	agctgcacat	ggatcctaca	ccagggactt	51840
cttacccttc	taatcctccc	caccggacac	tgcccagggg	cactggggct	ccagagggct	51900
attgccacac	ggacacacag	gagatctcat	caaggagatg	tgccctacccc	agagggttagc	51960
tctcaccatt	cacaagcaca	ccacttctgc	ctccagcttc	tactctctcg	caggaagttag	52020
ccagccgggt	gccaagtatc	cccactacac	tcctcaaaaat	tctcagacac	tgccagcctc	52080
cagctgtcag	cctggcccg	gctggcgggc	gcctgtcctt	ggcatagcga	ctagggtgta	52140
attagaacc	cgctagctcc	ctaattgcca	gttctgagct	gtccttgtta	ccggtgctcc	52200

gaggcacaca	tagaggaaaa	ggctgagagc	tgagccaggc	tggcatggag	gtagccctag	52260
tagacctaga	gaggactggc	atgtggccag	ggaccaaacg	tggcacagag	agggctcagt	52320
gcaatctgcc	ccgtgggtgc	ctcccagcca	catccatttg	cccagaactg	tgacgtcaaa	52380
ccagcccggc	ccattcattc	tttatctcag	tggcataaaa	atcactacaa	aaactttaca	52440
aaagagtctt	gggagctaaa	gggtcccttc	cttgccctcag	tccccaagat	tcctggcagg	52500
ggaggacaag	agagagaaga	aggaggaaga	ctcctggcag	tgttggcatc	tccaaatacc	52560
agaggggtga	cttgggtgac	aggacacagg	ttggggacct	gaatgtcttc	agcaagggac	52620
actcttgtag	ggtaggtcag	cctccaacca	tgaagtataa	caccaaggcc	agtctaagct	52680
tgggagacca	acacttgctc	ctccttttcc	caccaggggt	gtctggaata	tgtctaaaga	52740
tggcctctcc	agcctctgct	tacaaatgtg	gagggacctt	aagttaggga	cttgcctaac	52800
ctacctctag	ccaaaaactgt	gtccacaagt	gccagcccac	aaaagatcac	cccctgagcc	52860
ccttgggaag	aaatgaagat	tcccatgccc	tgccctcctc	caggccccac	cccactgct	52920
gcaagagaac	agcttctaca	ctgggtgagg	tccttccggg	cccaccctat	cccacaaagc	52980
tggttagaaa	gagtcacagg	agctgagagg	ctgatccagg	tggggactca	ggatgctgct	53040
gcccaggggc	cctctcact	tgggggagct	gaactggggg	tagtcttctc	ccatgcgggg	53100
tgcaagtctt	aagtcaggac	caaaggctct	gcctccatgg	aagtcagctt	tgtcattctg	53160
gcctatgagc	ctgttgctcag	gggaatctcg	ctgttctctg	agctggggca	gcgcgtggg	53220
gtttaggttc	ctcacactgc	ccacaaagag	gggcacgcct	atggtgtcct	ccatgatgaa	53280
gaagaggaag	ggtcggttca	cagtgaaggga	ggagagggac	attcagattca	tggctacgct	53340
ggtagctgcy	gctgcctcca	caccagcctc	gctgagctcc	atggtagact	gatgttgac	53400
gctagacacc	accagattct	gctcagagat	cccacgaagg	tctgggccct	ggaacaattc	53460
ctgcaggcct	cccagaaca	gcagatgact	ggtcagtgct	gccccaggcc	tatgtggatc	53520
tgtctagcat	cctggctaaa	gggaacactt	gaacccagcg	ggtgattgga	atctgttaga	53580
cctcagctca	gacaacactt	ctagaaacct	tttttttttt	tttttttttt	ttttaaatca	53640
ggatctgcgc	taggtacagg	acagaaagtc	tagaggagca	tatcaaagtc	tcccatccag	53700
gaagcagggc	cacctctggc	tcaggcacac	tggcagctcc	cgtactctgc	ccagaccacc	53760
taggggcacc	ctatccccaa	gctccttacc	cagttggctg	aggggtggca	ccagggtccag	53820
ctgctgttgc	agatggagtt	taggcagcca	cacctgtgtg	ggcctctctc	gcagcgaggg	53880
atggtacaga	gtaatccagg	tcagggttgc	tagtacctcg	gacacgttcc	actcaaaata	53940
agtgggcata	acgaccacaa	agctcatggt	gttcttaaa	gggaaatgag	ccacctacag	54000
ataagaaaag	gagagaaact	gaggaccaga	cagcacctgg	acctgtctgg	agctctggcc	54060
aaaattactc	ctgtactttt	gagacaagag	ccagaaatcc	agggttagca	tgctttcact	54120
taactgggtg	agtggataaa	taccacttac	ccctttgcaa	ggtgacatgg	gaccaaatga	54180
gataatgctt	ttacacctct	ctgtgtgcac	acataagcat	atatgtttgt	atcgggtgtg	54240
gtgtgtttgc	tcatgggtat	atggagtcag	aagtaggtaa	acatcagtcg	tcttctctca	54300
ttgtctccca	cttttttttt	tttttttttt	gtgttgccat	ctttttgttg	ttgttatctc	54360
aagacaggct	ttctctgtgt	agccctggct	gtcctggaac	tcactctgta	aatcaggctg	54420
gcctcgaact	tgcagagacc	caactgcctc	tgccctcctg	gtgctgggat	ctaagatgtg	54480
tgtactctca	catagctccc	tctttttttg	acacagggtc	tcattggatcc	caagctggct	54540
ttgaaatgac	tgtttggggc	tggagagatg	gctcagcggc	taagaacact	gactgctctt	54600
ccaaggttcc	tgagttcaaa	tcccagcaac	cacatgggtg	ctcacaacca	tccgttaacaa	54660
gatctgactc	cctctctctg	agtgctctga	gacagctaga	gtgtacttac	atgtaataaa	54720
taaatataat	tttttttaaa	agagaaagaa	atgatggcta	catacttctc	tctcgtctct	54780
ctgcccacag	tgctgggatt	acagagctgt	acaacaagcc	caagtgttgt	gtgttttaga	54840
catgctaata	tatcccaggc	tgtcctcaga	ctctctatgt	aattcagaac	gaccttgaa	54900
ttcttttaag	gtttattttt	atcttatgtg	tatgggtatt	ttgcctgagc	atttgtctgt	54960
gtacogtgct	cttgagctac	cctcacagtc	cagaggaggg	caccatttcc	ccctgaactg	55020
gttgtagctc	gcatgggtgg	tgttggaat	caaaccctgg	tcctctgcaa	gagaagccag	55080
taagtactct	taactgctga	gccacttctc	caccttgagc	ttttcttctc	cctatctcga	55140
tctaaaagta	ctagggatgg	cggatgtgcg	ttcatgtgcc	tggtttatgt	gttgctaagg	55200
gttgaacaaa	gggcttttgg	catgccaggc	aagcactcaa	caactgagct	acacatcccc	55260
acagactttg	actcttctag	tagtaggttc	tccactacag	cctgagttct	ctatctgctg	55320
tcagcaagct	gtacaaacaa	gctatggggc	ttcctgtcct	tgccctctcag	ttctctccgc	55380
aggtggggct	actggccttc	aaaatgaccc	atagaggagc	cacagcaaac	agtaggaagc	55440
ttgcccctcg	tctttcaccc	tctcccagag	agtcagctat	aattcgagtt	tttttttctc	55500
ctctctctct	ttaaacagga	tctgggttat	tggccctaac	tatcttcaac	ttcagttctc	55560
ctgcttcaac	cttctgagtg	ctgggattat	ggtgtaagcc	accacactca	gctcacacaa	55620
cctttttttt	tttttttttt	tttaagaat	ccatgcagtt	aggacagcat	ggaaatgacc	55680
aggctcaggc	ctccctgggt	accagcataa	tgccctgcag	cgggtcctct	gccagtgggg	55740
ggatggaaag	atggagccag	aggatctttc	ctctctgaac	ctcaatgtcc	cacagtggag	55800
cactcatgtc	cactggggaga	tactgtagta	ttcaaggaag	aagcaacagg	aaggtgagag	55860
ctaaagtgag	ctgagcaggc	tcgtatcctc	tcaccacggg	ctacagagaa	gtctggctgc	55920
ccctccaca	tggctcctcc	ctgcagaact	ggcaatgctg	ggcccggtct	gccagtcaca	55980
actaaccaac	agaatggatg	agcatgtgtg	gtgccacaca	cctgggaccc	cagcactcag	56040

acagctgggg	cagaagggtc	atgagtccaa	agcgaacttg	tgtaacattg	tcagaccctc	56100
gaacaaacaa	aactagcccg	tcctgtttatc	tcagccacag	atgatggggc	caaggatcag	56160
tactctagcc	aaggagtcac	ggttaggcta	gaagcaagg	aagccttagc	tgagacagct	56220
tggcacggag	cttcatccaa	tcagaatgtt	cagagcaata	agctttgaaa	cccgaactcc	56280
atctatgaag	cactgtgtgg	gaactcctct	cttcccttac	gagcagggcc	ctggctctct	56340
tgggctccgc	taaaacccca	gcacagagaa	cagttacctg	gcacgtgaca	aaaactcaat	56400
atattttctt	tgaggagatg	aacctcaaa	aagctgtgtc	ctggatagac	acagcataat	56460
aaacccttca	ggagctacct	accagggac	cagactttac	ctcccagtac	caggcctcgt	56520
ttgccagcca	aaggcaaaag	ccagactgac	ctgtatctca	ggttgctcca	gcaggaacca	56580
tcgaagagga	tatgacaccg	cgtgcatcat	gtccaccgac	actgtgaacc	gctcatccag	56640
gtggaagaaa	tctttctggg	tgaggctcgg	gtcaaaactg	gtcctccaga	aacctgcagc	56700
caggcagagg	gcaggagcca	tgtaacataa	aatcagcctn	ctgcctgtct	tgctagaac	56760
ctatnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	56820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	56880
ggaaaaagga	atcttaaat	agaagatgcc	ttgataagat	tggcatgtag	gtatgtctca	56940
ctaattgattg	atgttgaaag	tcacgagggg	tgtgtgcacc	ctgggcagat	ggcctggggg	57000
atataaaac	acaggctgaa	caaacacaaa	agcagtagtc	ctcaatggct	tctgctttag	57060
ttctgtctc	aggttcttac	cttgacttcc	ctcagtgaag	gcattgtcaca	tgagagttgt	57120
aagaggaat	aaacccttcc	ctcccacat	agtttttggg	tatgatgtta	tatgtcaaca	57180
acagaaacta	taactaatat	agttgggttt	cttttttgt	ttgttttgt	ttgttttgag	57240
acaggggttc	tctgtatggc	cctggctgtc	ctggaaactca	ctttgtagac	caggctggcc	57300
tcgaactcag	aaatccacct	gcctctgcct	ctgcctccca	agtgggtggg	ttaaaggcat	57360
gcgccaccat	tgcttgctg	gttttcttt	tttttaata	catttataat	gcattttaga	57420
tttaaaaaa	aaaatggcca	tggcatataa	tataaaaga	agtgtctaca	aatcaccatg	57480
tgcccttgcc	ataaattatg	taaaaatttc	catatggaca	tcagtctcaa	gcttacaatc	57540
tcagactca	tgagcctgag	gcagaggcag	gaggatggg	agctcaaggc	cagcttagtc	57600
taataacaa	gatcctgtcc	aaataataac	aacagtaata	atttcataca	tagaactaga	57660
aggggccact	gcaaagacag	tatgacaaaa	ccactggccc	tgccataatg	tattttaaat	57720
aactgtccct	ctctctgtaa	ttttcagttt	ctaattttta	cataactacc	atgtattctt	57780
tttgtaattt	taattagttt	tttaataata	gaaacaagct	aagtgtcaag	aatattttca	57840
tatgaacatt	ttcaaggcac	ttgatacata	cctcagattt	gocctccagg	tgagcagtac	57900
caattacgtg	ccaccagcaa	tgttagcttc	cttttttccc	taocatctga	ttctgtttca	57960
gtctattcgt	agttctgac	ttgttatata	cctttttatt	gtttccctgg	gttccaacac	58020
ctcccagttg	agtgttctca	ttgaatttca	ttagcagctg	tttcattaat	ggcacagaag	58080
aaggattcag	gtgttaacta	ggatagactt	tgacaaagaa	ctatgagaac	atatcttatt	58140
atctttgcat	aaattctttt	taatcaaaag	tctctctctg	ctctctctct	ttcccatctc	58200
agggagtagg	tctggccact	gatgagtgct	caggccacag	tacaggtgtg	cgtggttctg	58260
tccctgtggg	aagggcacat	ctgtgttgta	acaggattcc	tgtcttaaca	agccttgctc	58320
aggctctaag	tgttcttgag	ctagctaaact	gcccttggct	ttcccttgat	taccagataa	58380
ctattcactc	ttctcatttt	gcagagcact	taccaggtag	ctatgtcctg	gaagtacgaa	58440
tgagtccttc	tattgttttt	cttttaacta	aatcccat	gaaatgcgcc	agggacactt	58500
caatccaagg	tacacttttg	ctaaagaatc	actcattttt	atatgcaaaa	tgccacctat	58560
taactgcagc	tgatatggta	catacatatt	ctctcttctc	attatccact	aataggtgac	58620
taatgcgaaa	tattgagtaa	tttttaaaaa	tcaatactca	attttttaga	aataattaga	58680
gagacattca	actctgacac	cagcacccta	ctcagttcct	gagccttcc	ctgccggagg	58740
agaatctata	aataactcac	gaagctgaca	ttactcactg	tgttgacagc	atttttttct	58800
gagaaaaattt	tagcaactgt	tctaataag	cctgccaagt	atcagtagtt	gagaatgcaa	58860
gtcaactttt	aattatgcag	acgttgatta	ttcagacgac	aaattgttgg	tgctgcacg	58920
gctccttctc	gctgcctacc	tttaaccgtt	ctcagtgctc	attagcacat	gttccagaag	58980
gtaggctttg	gagggcgga	caggcactca	aaccagctaa	gcacttagag	aagctctgat	59040
gaaagatgtt	aatgcagttt	gtagaattat	tgactaaaat	tgagtcattt	ggattccctg	59100
tgaattgtat	ttacatgccc	tgctcctgtc	ccccatagca	acagataata	ggattgtctg	59160
cagagagaca	acatagttct	tatatttaat	tttttcttt	gtcgaacatt	ttcacatgat	59220
ggttcgtggt	gtttcctttg	ttcattacat	ttgtatccag	actagttact	tctgataagc	59280
ggttagttag	gattcctggc	acgcggacag	tgacaccaca	gttgtctgat	cgtttccac	59340
ttttttacaa	aaacgtttgc	ctttaagagt	cagtggtttg	cacatttcac	ccagattatt	59400
ggaaatatta	tttccctcct	gcttaaacgc	aagctgtgat	cataatttaa	gcctttctag	59460
gtagccgatc	ttacatgtat	catacctatt	tctggcatat	gtttgtctat	tacaaagacc	59520
tcgtaggtat	gcagtttagaa	gcctctagtt	aaatgaaatg	ttgcgtgtgt	gatgaacctg	59580
gagtggggat	ggccttttgt	gtgcccgaag	gctgttgtgt	ttcacacagt	tgttttctgc	59640
ctcctctggt	ctatcactat	cctgccactg	ccagaaaacc	ctgctgtgtg	ttcccccgct	59700
ggaggatctc	tgcttctgaa	cttctttggc	ctgagaaaact	ccataaccaa	atcagtttagc	59760
attttgttta	aagagcaggt	aggctgttag	agcttgggtc	ttacatgtct	cccaggtcca	59820
cttgccagcg	ccttgaccac	tgtaaacctt	tgtaaaccaa	ctcatctttt	gctgcctggt	59880

tttttggggg	tttttttggt	tttgtttaag	ccaagatcag	ttatatggcc	caggctgagc	59940
ctctcttccc	agcctetcaa	atgttagaat	tacaagcatg	catccctcag	catacctttc	60000
ctttgtcttt	tttaaaatag	agttttgcca	tagcaacaga	aatctaaccat	aactaagcat	60060
agccgtgcac	atgtgtatag	gaactcacat	atgttggaat	ggaagtctcat	agagaccggc	60120
atcactgcct	agagggccct	ttcttccctc	ctgtcagttg	tcgtgctagc	tgactgtact	60180
acaaaagagg	ttgtctgagg	cataagacta	ccttcaataa	aacatgcaca	gacagtttgc	60240
ttctcttgaga	tttcagagca	tgcactacct	tcaataaaaa	atggacagac	ggtttgccta	60300
cctgagactg	cagagcagtt	ttcaaaaaat	ttagacaaag	ggtagagatg	agaaggctgc	60360
ggggttttgc	acacacttaa	ggtgcgtaag	taaaaaaact	gagctacact	gacaggatgc	60420
togttctagt	agccaaacca	agagcagttg	aaccaagcga	cctgagactc	aaacatcctg	60480
gggagataat	cttaggagtg	ctatgcttct	gggtcctaca	agtattatga	aactgtctag	60540
aaagaccccc	actgtgtaat	cctttttgat	tatttttttt	ataaattcta	gtcttggggg	60600
tttgagtggc	acacagacat	aatggttagg	cttcgggtgtg	tgctcattca	ccttgcttcc	60660
tggggaccag	agtttgcgat	gagtcactgtt	ccatctgatt	tcctgtcggat	ccggctgcag	60720
agccatgact	cagatgggct	tcaggcccgac	ctgctcagtt	catcttctgg	ggaatagatg	60780
acaaggacgg	gacaaatgtc	ctgacgcaca	ttctcctctg	ttcttgcact	tccagggtct	60840
aacgagagca	tcattaccaa	cagcaggcgag	atacgccttg	ccaacaggcat	cttccctggt	60900
gtcagccctc	tggaacacct	ctgcaggccc	aaacccagtg	tgtcctctac	tggcactgtc	60960
gccaccgtgc	gggcagcaca	gaggatcgca	aaaggacagg	agattctgca	ctgttatggt	61020
gagccagcct	ttctttccac	taccctgctg	tgccctcacac	ctcacatgaa	aaggataaag	61080
ggacaggaat	cagcagatat	gggcccagtg	cctctactca	tcctctgagt	cttctctgga	61140
aaggggcaatg	catccttggg	ccaataaaaa	aggttctctg	gctgtataaa	aaaagccogt	61200
tggggcgagt	gagccatata	cctccatgcc	ttgtagacag	ccatccttga	aaatgagcga	61260
ggagcacttt	cttggcttct	ttcttcctgc	cccagcagct	tggaaacgta	tccactttca	61320
cccggtgtttt	gttgtttttt	ctgatgatgat	agggcagagt	acccaacctc	atataggcta	61380
ggctagtgtc	tatcatctgag	ccagagcccc	aaccagcac	caccatgcca	gtcacgtgat	61440
gactaggcca	gcccctcggt	agagtaggca	ttgactctct	tggtgtgact	aggaactgtg	61500
ggtaatctct	ctccagggcc	tcacgagagc	cggatggggc	ttgtcagagt	gcagcagagg	61560
ctgagttctc	agtaacttct	tgactgcgcg	tgtggggcct	gtgcagctga	gacactgaga	61620
gcagctgcag	ctccagatg	ggaagccctc	tgttgtaaga	cttgccagagc	gctcatgcag	61680
gtaaatctct	gctgttccca	ggggcagggc	tccagctaaa	gggtgtcagt	cgcaggaga	61740
accattctctg	cttcccttct	tgtaactcct	ccctacatgt	cggccggtcc	tgcagaaaac	61800
acaggttgta	tttctctaata	ttttccctat	aagtgcacaca	aaatcttaaa	ttcacaaaag	61860
ggaccaaana	aaaaaaaaaa	aaaaaagccc	tagaaattta	cttgctcaaa	taagtcatca	61920
aaagttgtgc	atcaggccta	gcacttgggt	actggttaacc	ctgacactca	ggagggtgag	61980
gaagaaggat	ctcaagtcgg	agccagctgt	caagtgcacac	cccacttaag	agatcaccat	62040
ttcaaaggagc	tatttcagag	atggtttaat	ctggggaccc	agatttgtga	ttttctgtct	62100
gttcaattcc	atctctctgt	ctgtggcctca	tcgacacacac	tctgtagtaa	ctgtgggaaa	62160
atccgaccca	catagttttc	cctcagcctt	tgaccagagag	ggaagagcca	cagtggagag	62220
catgagagca	gacccttggg	tgctaactgc	aggtaatggt	gtgacacatg	gagtcttcaa	62280
cattcatgcc	ccaatgcana	atggtctcca	caccagagca	tggcaattct	attagaaata	62340
agtaaaatgga	attggctgtg	ttgaaaaatt	taaaagccaga	ggtaacagaat	gaagccttcc	62400
ccagcatggt	ttgttttgtt	ttgtgtttta	ggcagcgctc	ctctgtgtag	ccttggtctc	62460
tgccctctgc	tacctctccc	aggtgtgcca	ccatgctggg	cctaagcgcc	ctgtgcatta	62520
gtgctccctc	gatcctgtct	actctttgaga	cagctcttct	tctactctgt	atccccagat	62580
aacctagagt	tcactctaga	gcccaggctg	gcctcaaat	tgagatctct	gtgtccagc	62640
ttctcaaatg	cagtgatatt	tacaggccta	cacctggcct	tcctgatag	attcctagta	62700
agatgattat	cctttgagcc	atatctctct	tctgtcttct	cctctcttcc	tgcagggtgt	62760
atctagaatt	tatttctaaag	ctgactggcc	tcagaattgc	catccttctg	cctttagnnn	62820
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62880
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	62940
tgcccttaaga	gttcttttgt	tatgggaatt	ttatgggaat	ttccacagaa	ccctccagc	63000
ggagttcttg	ctgttgtttt	ttaatcttta	ttctctcttc	gtgtgtgtgt	gtgtgtgtgt	63060
gtgtgtatgc	gcgcgtgtct	aacttgcaaa	attgcaaaat	tcagtctcct	ctttccacc	63120
tgtaggtctc	ggggatcaga	ctctgttagg	cttgggtgta	gggtcctttc	tgagccatct	63180
tacaggccccc	ccatggacaa	ctttttcttt	gaaaacctgt	ttctggcttg	gggtgtgatg	63240
ctcacacctg						

cctttaaccc	cagcactcag	gaggcagagg	caggcagatt	tctgaattcg	aaggcagcct	63780
ggtctacaga	gtgagttcca	ggacagccag	ggctatacag	agaaaccctg	tctcaaaaaa	63840
aaaaaaaaa	aaggacttta	aattgggctg	gagagatgga	ttaaaagcat	tggtctctct	63900
tcccagagg	cctgggttca	attcccagca	ctcaaatggt	ggctcacaa	tgtctatata	63960
aacgcaatct	aacaccctct	tcaggcatgc	aggttacatg	tagacaaaac	atccatatac	64020
ataaaataca	taagtaaatg	agtcttttaa	tgtatactag	aagctgggtg	gtgggtgcatg	64080
cctttaatcc	cagcacttgg	gaggcagagg	cagggtgatc	tctgagttgg	aggccagcct	64140
ggtctgagta	aatagagcct	tgtacttcta	cttatcacta	cagttacatt	ttataacttt	64200
ggggcccttt	gtttccattt	tcactgtttt	gcttaaccac	tggggcctga	agctttttgtg	64260
ctgacacttt	tgttcgctaa	tcacagcgca	accaatgggc	tctacactcc	atcaccatca	64320
acacaaacaa	aacaaaacac	aacactacgg	atcctggcat	ggtggaacat	ctttagcccc	64380
agtacgtggg	cttgagtcca	aggccggcct	ggtctacata	gcaagtctta	ggatagtagg	64440
gtagtctttt	aaaacaaaa	actattttat	ttatgaacaa	aacatgtaaa	gaaagaaaaa	64500
aaactgcaaa	tttatctatg	aatgaagtct	aagtaatact	tcaatatggt	aaatagcttt	64560
ctaaaaatatt	tttatttaaa	gaaaactcag	caaattattc	aaacaacctt	ataaacgttc	64620
gttataaaag	taaagaatta	tttgcaattg	ccttaagggt	ccaaggtggc	agcctcttaa	64680
aattcagaac	aatccaagct	tcacattcca	gttcaacatt	tctacagccc	taacgtattc	64740
aaatacctcc	attctgacaa	ctgtttccccc	tcttcttttc	tcttaagctg	cttagatgtc	64800
tgtccaggcc	ttttcatgat	tttagtcatt	cacacaacta	gcaaacatta	tctagggact	64860
aaaacttgcc	atatactggg	atatcacctc	aaagggggac	tgaaagtagc	tcgaggctac	64920
agtctctaca	atctcctgaa	tgaaatacaa	agtagctaata	atttacccaa	taaacatgta	64980
cacctgtgat	gattgtctagc	tgtactagca	gaagctaaac	actaaatcta	gaaactcagt	65040
cctccaacta	gccccctgct	cggcttcage	ctcattttta	caaaccaagg	aaagagtttg	65100
gaatgttgcc	caaagccata	cataagtga	caaaaaggag	ttggagcttc	caaatgcatg	65160
gatttgggct	agttactttg	ccaaccaact	cagtaacaac	tgagctgaac	aggaacactg	65220
tggtagcaaa	agaaactgga	actatcaatg	gcctctagag	caaaaatata	tttaaaaaga	65280
aaaaaacaaa	caaggcctgg	caaggagact	gtgagaagag	tgtgtctgact	gaaattgact	65340
agttcagcca	acaaaagact	attccagggc	tggtgagatg	gctcagtggt	taagagcacc	65400
cgactgctct	tcggaaggct	aggagtcca	atcccagcaa	ccacatgggt	gctcacaccc	65460
atccgttaaca	agactctgact	ccctcttctg	gagtgtatga	agacagctac	agtgtactta	65520
catataatca	ataaataaat	ctttaaaaaa	aaaaaagact	attccagtgg	ggatggaaaa	65580
gttaagtgtg	gagttaaaa	atacttcaac	tggtgatgga	ctaggtgtcc	agagtccggg	65640
aaaaggtagc	tctgtggtag	agggtccctg	tgtgtaagcc	cagctacctg	agctcaatcc	65700
acagaatcca	cagcggagtg	ggaagagaaa	caacgtccca	gagttgtcct	ctggcatccg	65760
acgcacattc	gccatcccca	agatgtcata	catatgtgta	catactacac	actggcgcac	65820
gcgcacacac	actctttttt	aaaattcaga	cttagaggga	cataaaggat	ttgctctgat	65880
atatgttcaa	ttgaaaaatga	ctttgaagat	agagggcaga	tcgaagggaag	ctcagcagga	65940
aagaattaat	aatcatgcagg	tgaaaggcta	taaaactagtc	tcgagagggc	cttggtctga	66000
caaaaaaatc	tatgggggtt	gccggtaaaa	taaggaaaaa	gttgtcaaca	tgaaacacag	66060
aacactagca	agagaggagt	gttagcagaa	agaagccaac	aagctcaaac	aattaggctc	66120
gctgaaaaat	tttaaaatgt	cttctgattt	ggctactggg	aagccactgg	tgacttcggt	66180
cagcgttttc	tctctcgtga	ccagagagat	gtctagtagc	aataatgagt	taggaggatg	66240
taaaagaagt	aaaacagccg	aaaacaagtc	caaaaagtgt	gggtgtagtg	agaaagggag	66300
gaaacagagg	ccgcccgaaga	tagacagcgg	catgtttatt	tgtcttgttt	tcttagatgt	66360
aaacaaacta	aaaaaactcg	tgagttcttc	tgccagtacc	gggttgccct	cagcatcctc	66420
tgatggctct	agagaccccg	ggatgctccc	ccgcggccgt	ataatttcc	ccctgacgct	66480
ctcccgatcg	acagcggctc	cctcccgggg	tcctctttgc	accgctccaa	ggccgcgctg	66540
ctagggccat	cgagcccgct	cagggtcgct	tccttaacct	gatggccccc	tcgctcaggt	66600
gtcccaccat	ggctgcaccg	ctaactcccg	cgtcgcgct	cttgacccgc	ctgagcttct	66660
ctgccggggt	cccgcgggct	gctcaacgat	tggttagagc	aactgtgcgt	gccgatccgc	66720
ccccagcggt	agcgcgggtg	gagggcgggg	cctagacgcc	gatagccacc	gcattggcta	66780
ccgcgcggca	ggcagagcac	gtgactcttc	cgaggccggg	ttcgaggcct	agtggcgggg	66840
tggcgggacg	tgagggcggg	gcctgggttc	gcagtgcgcc	tgtgtcagcg	cgggtgctact	66900
gagttgttcc	cccgccagct	gtcggaaact	tgcccgccea	gtcctttggc	ggacagacag	66960
aatggcaacc	caggggaacag	tcggagctct	cccctggtaa	ctgctgtctaa	ataatagtaa	67020
agcagtgaac	tgggtacttc	ttcacgcagt	gcgtgcccgg	cgccgggtgc	aggcccagag	67080
cttggcactg	tggtgataaac	aaggtaaatc	agactcagtc	tcgcgccctc	tgagttccac	67140
ctgagagttg	tgccgcgaag	gaaccagcc	tcaaggatgg	tagacgcgat	atgggccaca	67200
gtgctgggag	tcagagtggt	gggtcaaaaa	tcaatcaggc	tttcgagagg	cgatgcgggt	67260
tgaaactgag	taaaagtgtg	gtagaaattt	gtcaggtgga	ttccagttag	gatagtgatg	67320
ttcctaaaaa	cccaaatggc	ctatgcaaaa	gtattggaga	gcctggcggt	ctggctgggt	67380
ctgatctgtt	tgtaatccca	gcctttggga	tgtagaagca	gcaaaagtgc	aaggtcaccc	67440
ttaaacccgt	tgagttcgag	gtcaacctga	actaaatgag	accctgaaaa	atcaaaattt	67500
gggacccagg	cgtggtggca	ttcgaggtaa	aagcaggcag	atctctgagt	tcgaagccag	67560

ccaggctaac	ataagatccg	gtctcaaaaa	aaaaaagtaa	taaaaataaa	aagggagaga	67620
ggctatatga	actgaaagaa	agacctggag	atcaaaacag	aaaactgagc	cgtctaagaa	67680
atgaaatat	ttaacttcat	agttgctgga	gtaagaagtc	tggaaaactt	tgggcaacta	67740
aggtaaacag	gtctagaaag	actggaatag	tagccatcta	ctgggtatfff	gatctctgtt	67800
tgtacaacca	caacctacta	tagtttctca	aacagttcca	aagaatatgt	ctgggtgaat	67860
tggtagcaca	ccacagatta	actctccttc	agcatatcaa	cagctataga	aaacccaga	67920
agaaatgatt	ttggttgcgt	gtcacttggt	aggatgaaat	ctcgattttc	tagaaactatg	67980
cattaataga	aaactgaatc	ttcatgttct	gactttacag	agctgcggca	gcatggatct	68040
accggtggat	gaatggaggt	cctacctact	taagaagtgg	gcttcaactcc	cgaagtctgt	68100
gcaggacaca	atctctacag	cagagacttt	gagcgacatc	ttccttccct	cttcttccct	68160
tcttcagtaa	gtgaatggaa	acttcaggga	aatttttggtc	tggaaaatgt	tctgccttgt	68220
catttggtct	gaatatctct	tttttatagg	agagagtagc	tttatattct	ttatagtatg	68280
gggcatttag	cagtacttgt	tggttttcac	gtttctccct	agtctgtgat	tactagaatg	68340
ggtaggcact	aactgccttc	ctcttttggtc	atgtgttata	cttaagggaat	gtagtatctt	68400
gctgtcgtcc	cagtgtgtgc	actcatagga	tctgggtcag	gttgtgtagc	tgcccctaga	68460
agctcattca	gcctaatgg	ggagaaaagaa	ccctggcact	tggttagtgg	agaccanana	68520
cttctcaagt	tctnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	68580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnntttagtt	68640
tccaagtgcc	actttactgc	aatgtgtcac	cacatccaga	gttctgtgtt	tgtttatttg	68700
tctgtttttg	agacagggtt	tctctgtgta	gccctggctg	tcctggaaact	cactctgtatg	68760
accaggctac	cctcaaattc	actgagatct	gcctgcctct	gcctccagag	tgctgctgta	68820
cactaccacc	acccatccag	ctttctatat	tggttttccct	atggcgtttt	aaaatagcca	68880
ttatagctgt	gtttatcatc	taaagtcctg	gtcccaaaag	gagatgagag	gggctgctaa	68940
ggtgaaaagg	attacaacag	cttatcaatt	ctggttcaaaa	attaaacctc	agagtgggat	69000
tagctgtctc	tttcattaga	attgctatca	gaattcactc	aggccttggt	tgcgtgtgtt	69060
attgaagagg	tctcttctct	atcaggtcag	tgaactcctta	gctcaagtac	atgcaatatg	69120
cagtattgat	aactgttctc	gcttaggaat	aaaaatagaa	ctgacttcca	cagggaaatg	69180
atgtgtctga	ctgtagcaac	gaatcttgca	caaaactctgt	cagcagggac	cagctagtct	69240
ctcgcctgca	ggacgttcag	caacaggtct	gcactggccca	gaagcttctc	cgaaccggta	69300
aaccagtgga	gattggctcc	ctcgccttag	gcctcagccc	ttcccttggt	tattttggta	69360
tcaccttgcc	ttactgagca	gtcctcaata	aatgactgag	gacttgaatt	taattatccc	69420
agcaccagcc	acaagatggc	tatgtagggc	agttagacca	gactgtgacc	agctgttact	69480
ctgggtgcct	tgaaagtctt	cctgatgggt	taagctgtgt	ctgctgcgcc	agatagttct	69540
agcagctcga	gcaccagaaa	ggctgtctga	cttccatggg	ctttgtgtgg	ctccagaggt	69600
ccaatgccat	catctgatcc	ccagcttaag	gacctaaagt	ccgagaaggt	tgctctgccc	69660
tcagcagcag	cagcaagtcc	tgagtgtctg	ctgggctcgt	gggtgtgactc	aggagttagag	69720
ctcggtagct	agcctgagct	gagagctgag	agaaagaaa	gactcctctc	ttttcagaaa	69780
gggatttgca	gaactcgatg	ttagaccctg	acatggtagg	aatctgtttt	gactatttcta	69840
tcttagatga	tgaaagtggac	cttttagccta	gagtcagaaa	aactaatgat	tacaggagga	69900
atgtagagtt	ggttgttaaa	tgttggttgg	aaaatggatg	ttagaagccc	agggtaaatg	69960
tgagggaagcc	tcacttaaca	cctcttttac	tgaaaagagaa	aacataagca	accaacagct	70020
tccttggaat	gcccggctgt	tgactccgtg	agataaagag	gcattttcac	tttgacctaa	70080
ccgatatgaga	ccttgcaacg	tgtctctctg	tgtccaggac	tagatctgta	tctgttgtga	70140
ggcatttttc	ctttgaatcc	atagagcaag	ccattcagca	gttggttgcgg	tgccgggagg	70200
ctgctgagag	cttcttgtca	gcagagcaca	ccgtactggg	ggaaattgaa	gatggcctgg	70260
cccaggccca	tgctacctta	ggtatgctac	cttaggtata	gccggagttc	tccttccctg	70320
ccgtgtgttc	agtgcggccc	ttgccttgct	tgtttggttc	tctcttgcca	tctgaattga	70380
cgctcttctc	cctcccatcc	tgcatctctt	gccccagag	ccttaggcta	atgggtgttc	70440
ttttccgga	tgagacattt	ctcttctcac	agggaactgg	ctaaaagtctg	ctgcccatgt	70500
acagaagagt	ctccaggtgg	ttgaaactcg	ccatgggcca	tccagtgttg	aaattggcca	70560
tgagctcttc	aaactggccc	aagtccattt	caatgggtag	gcctttcttt	ttcctagtgt	70620
ttggccaggg	cacacagtgc	tctgtgtttt	cctaggtgct	tctgtgtatg	gctttttgct	70680
acagtgtctt	aaagcatggt	gaaactcttt	tatttctctt	ttaggacaga	tatttgccct	70740
ctgcttcaact	gatagacttt	aagctttgaa	ttccttccctg	aggatgtgga	gaaagccatt	70800
aggtctgcat	ggagcttccc	agggaggatt	tggaggcagc	ctcaccgcc	tctagcattc	70860
ctgtctgctt	aatcacacct	cccttggtcg	cctcagtcoc	tgctctctca	actccagggc	70920
tcggcccttt	ccctgggttg	cctcttatcc	cttttaagc	agtggttttc	aactagaagg	70980
gattgcaaat	ggcattttggc	agtgtttaga	gacagttttg	attgttatgg	ctgccagcat	71040
ctagttaaagg	ctaaaacctac	agtgacacgg	accgcctcca	cagtggagag	accaagtta	71100
gctatgtgaa	ggctgagaat	ccctgctttg	gagattaaaa	aagggaagctg	agggaaaccac	71160
tcagttggaa	gcacccttgg	tggcatgcac	aaggccctgg	ttctgtccct	agctctgcac	71220
aaaaaataga	atacaaggaa	gagtaaccct	aatgagctgg	tcctcaccct	agtggtccac	71280
tgaggtcact	tgaaagggaag	tctagcccca	atttagtatt	ttttgtggct	gccatacctc	71340
cagccttgat	caaactctcat	ggtatacatt	ggtaaagaaa	agggttttaa	acatagacct	71400

gatactcgga	catggaaaca	gtatgtttgg	tcagagagag	cgaaggacct	gatagacgag	71460
ggcaatatca	gagagagggc	atcagtcggg	ttagacacga	gcattccaca	gtgagcagct	71520
ctggataaag	ttttataaat	gctggttaag	gttttgatt	tgcccaattt	tgtagcagatc	71580
ccagagtcta	tcacaaacat	acacagtttt	ctcaaatctg	ctttgcagta	tgcccgtaga	71640
tgtctcttat	ctataacttc	agatggtaag	accctgaggg	cagaggaaact	cagacccttt	71700
gtgccccctg	taagaccctg	gggatgacag	tggacccgac	tttgtgttct	ctgcacagaa	71760
aggagtcac	ttctgttgag	actaagggaag	ggaactgaca	agcttccctt	tctgggttca	71820
ggttggcagt	gcctgaagct	ctgagtcca	tctggaaggc	agaaaggatc	ctgttggtgc	71880
actgtggccc	tgagagttag	gaggtccggg	agctccggga	aatgaggtcc	tgcttactgg	71940
actcgtcatt	cgatccctgtg	gggcccctgg	tgtagagcaa	tcatcctcac	cctcaagaag	72000
gagctctggt	gatgactgag	atgttctggt	ggcttgagc	tctcatcaga	gaggacggga	72060
ccttcccacc	tgacctgagc	ctagtgtctg	gcacagagag	cacttgaaaa	cagattgaga	72120
cactcacctg	ccatgctggc	tgctgcttgc	aagagctaac	tgccctctga	tggaaacccc	72180
atgcccagaa	aagactaaat	ccagtatcta	aaggctgctt	taaaggggtg	tcactgcagc	72240
cgggcttggt	ggcacacgcc	tttaatccca	gcactcggga	ggcaggcgga	tttctgagtt	72300
caaggccagc	ctgggtctaca	aagtgaattc	taggacagcc	agggctacag	agaaacccctg	72360
tcttgaaaaa	ccaaaaaaat	aaaaaataaa	aataagtaaa	aaataataaa	ataaaataat	72420
aaaggggtgt	cactgatctg	caggcagctc	atgctagcct	aggcttttgg	ctcgatttca	72480
tctcactaaa	cgatgaatct	gtttccctgg	aacattccta	tggtttctag	tagtaatgaa	72540
gtgctgtgtt	ccactccagt	gagaacttca	attcttagtc	ttgtattata	attgaaaaat	72600
aatatatagc	aagaaatcag	tatgactgct	tacctcaaga	gacatacaat	tccacttaca	72660
atatcctgct	tccttaaat	tttcattaa	actgggtgata	tataatttgt	gaatggagaa	72720
ataaatacgt	cttactgttg	gcagtttctt	cctgggatgg	caactctgta	ttggtttctt	72780
accagtgtcc	taattcttca	tcagtggctt	tcattgagtg	ttcttggcac	tcactgtcca	72840
agcactgagc	caaggtcaac	ctgtagcatg	acttcatagc	acaggccctcc	ttgttagcac	72900
acctgaaagc	agaccactct	ggctgtttca	cttgacagaca	gaatcttact	ctgtaagcca	72960
gtctagcctc	aaacaaacac	ctcctgcctc	agccttccaa	gttctaggtt	tataggaaaa	73020
ggccaccttg	cccagcttga	gactgcttct	tactgccatg	tctcttcagg	ctcacacatg	73080
aagtccaggg	cactccaggg	ggagccgtga	gtctgtctgc	agggcactcc	agggggagcc	73140
atgagtctgt	ctgcagggca	ctccaggagg	agccgtgagt	ctgtctcgag	ggcactccag	73200
gggaagccgt	gagtctgtct	gcaggggcact	ccagggggag	ccatgagtct	gtctgcaggg	73260
cactccaggg	ggagccatga	gtctgtctgc	aaggcattcc	aagagcagcc	atgggcgtca	73320
ctcattggtg	gactgtgagg	ctacatctcc	agatgccccg	agtgtctgtg	ttgtgagcac	73380
tgctgtcat	ggtttccaac	tgagacagag	ggaaggactt	tgcccttttc	cctaaggatg	73440
ggtagtaata	gtccagacca	caagggacag	atagctatgg	ggttttctga	ctcatcotta	73500
gtacattatt	gctgatgacc	agtttgtttg	gatgagttag	tgggaaagaa	gacccaagtc	73560
catacactct	gctttttaga	acttgctcat	cctagccatg	cccaaggagc	agccgttgac	73620
gtcatgggca	ttacagtgag	gaaataaaca	gtcctgaagg	tgccctggcag	cagcttttca	73680
agaagctggg	gttaaaaagc	agtattcaaa	catctgcgga	ctgggaaactg	ggcagcattt	73740
gagtctcctg	ctgtctgtta	atttaccctg	acaaggaggt	gacttgaaag	gtttgttttg	73800
tttggggtag	agctttttca	ggaaaaaagt	ttagtccctc	agacaactct	atagttattc	73860
tagtccaaac	ctatgccttg	tgttttattc	ctaaaagccc	tgctcacactt	tgtaaaatag	73920
gtgctcttcc	tcaaaggata	tatttaacgt	tttatataac	agggccttatt	ctgtgcatgg	73980
aaagcttttt	tagatgcttt	gtaagatggc	tcagtgttga	agagcatgta	ctgctcttct	74040
ggaagtcctg	ggtttgattc	tcagcagcta	acaccagctg	ttattccagt	tcctgggatc	74100
tgatgccctc	ttctggccta	tgtagcact	gcagtgtcgt	agtgcacaga	caaatgcagg	74160
caaagcactc	atacataaaa	ctaaattcaa	aaaactcttt	cattgtctca	tgtagacctag	74220
cttgagaata	ctgtgtctta	tattataatc	tagtatgagc	cagccacggg	agcaacacac	74280
ctattatctc	agcactcaga	agattgagac	tagatggtca	agagctagag	tctgggttac	74340
aaaacacctg	tctcaaaagt	aaaagggctg	aaaaagtgtc	tcagcagcta	agagcacaca	74400
ctgcttctcc	agagggcctc	atttcagttc	ctaataccca	caccgagtga	ctcaaccacc	74460
tgtaactcca	ggtccatgag	atccaacacc	tctggctcgtc	tgcataagct	cctacactca	74520
attatacaga	gagagagaga	gagagagaga	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	74580
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	74640
nnnnnnnnnn	ntctagagtg	tttcaggttt	ttttgttttt	tttttttttt	gagacaaggt	74700
ctctctatta	tgctgcctgg	aactttctat	gtagaccagg	ctggactcaa	acttatagtg	74760
atccactact	tctgcctctc	agtaactgta	ttgaaggcat	gtgtcaccac	acccacttac	74820
ttcaagatct	tagatttcca	aagaagccgt	agcctagaaa	aggttaataa	gtactgattt	74880
aaaacagaaa	gaaatcaggt	acacttagag	ctgtagaatg	tcagcatgtg	acatttgtga	74940
caagttgtca	aaactttgct	cttaattcta	aagagagaag	ctgtcaaaag	acttgaactg	75000
gggctgtagc	caacttggtc	gagcccttgc	atgaagctgt	gtgttactc	cccagcactg	75060
tggggtttga	attgatttga	accagtaga	ttcgtatatt	tgaatgttta	cctcatgggg	75120
aatgacatat	tacaaggtgt	ggcctgtgtg	gaggaattgt	caatttgggg	gtgagctttg	75180
aggtctctct	gctcaagctc	tgcccaggtt	agaaaggag	cctcctcctg	gctgtctaca	75240

gaggacatag	tctectggct	gccttcagat	caagatgtag	aactcttggc	tcctccagca	75300
ccaagtctgc	ctgcacaatg	ccatgcttcc	taccatgatg	ataatgaact	gaacctctga	75360
aactgtaagc	cagccccaat	taaagtgttg	tctttataag	agttgccttg	gtcatggtgt	75420
ctcttcataa	caataaaagc	ctaactaaaa	cacattcctg	ctgggcagtg	gtgggtgcacg	75480
cctttaatcc	cagcaacttg	gaggcagagg	caggaggatt	tctgagttcg	aggccagcct	75540
ggtctacaga	gtgagttcca	gaacagccag	ggctacacag	agaaaccctg	tctcaaaaaa	75600
aaaacaaaaa	caaacaagca	aacaaatgcc	agcatttggg	aggtagagtt	aagaagattg	75660
ggagtacaaa	gtcgtctcag	ctagtatggt	tgaggccagc	atggaccaca	tgagacggtc	75720
tcaaaacgaa	agaaacgaat	gaatagataa	acatttgagt	gtccagtttt	ttcctttctt	75780
tcttgctttg	tttttgccgg	tgctgaggat	taaaccagg	accttggtca	tactaggcaa	75840
gcattctcca	ctgagggaaca	ccctggcgag	tgcttagtct	gtctgtctgc	ctgcctgcct	75900
gcctgcctgc	ctgcttggtt	tgtgtatgag	tggtaacctg	catgtctgtc	tgtataccac	75960
agacatgcct	ggtatctgca	gaggccagaa	gaggatgttg	gatcgccctg	aactgggatt	76020
acaaatgggt	gtaagctgcc	atgtagggtat	tcagaattga	acctgggtct	ctgaaagagc	76080
agccagtgtc	cttgttggtg	gttttatttg	gggcaggagg	tagttatttg	gttgggtggg	76140
tggttggttg	gttgggtggg	tttcttgaga	cagggtttct	ctatgtagcc	ttggctgtcc	76200
tggaacttgc	tctgtaggct	caaactcaga	gatctgcctg	ccctgcctc	ccgagtgtcg	76260
ggataaagtc	atgtgccacc	aaactccagc	aagcagccag	tactcttaac	cactgagcca	76320
tcattccagc	ccttctttgt	gttttgagat	ggtcacaaag	tacaactcag	actgagctct	76380
tgatcacctc	ccctcagcct	cttgactgct	gggggttaca	ggtgtgtcac	tgctcctcat	76440
tctgagtgtc	agatcttgaa	aaaccttct	cgtgaccttg	atccttaaaa	caaacctctg	76500
gagaatgagt	tctgataaact	atttctcact	cctcttcaag	aaaaggaaag	ccagagaaag	76560
gggggggggc	aagccccaga	aacattgata	acttgcccaa	agttacacag	caaaattcag	76620
acagcctgca	catctcagtg	gccatctgtg	ccatatccac	cctgcccctc	tctgacctcc	76680
ccacctccat	ccctacagac	cttgagttg	agatcagagt	ccaagccgta	tcgttaagtg	76740
gccttaggat	ctgacatcat	ggggactctc	acggtcctgt	cctcgtccaa	atgaaaatcc	76800
tgagggtcg	tctttctcga	gtcaaaactg	gttaccact	gccctggaag	gaaacagatg	76860
gagcatcctg	agccactgtc	cccagaaagg	ccacaggtcc	acgctgtgcg	tcactgtggc	76920
aaagcaacct	gagctgtcag	cagcaagaac	acaggagccg	ctgggtccca	gcatgtgtgg	76980
cacagaccat	aggtcccatg	caccacgggt	tctggctatc	ctcctgtagt	aaactcagaa	77040
ataagtgggt	gttctctctc	tgacttggat	caccacgctc	cttctgttta	aagtggcctt	77100
taatatgtcg	gtgtgtggtg	cgtgcctgct	ctcctgtccc	ctggggactt	ggagttagaa	77160
gccccagggc	tttctctcaa	gttaggatcc	actcttgcta	ctactccata	agatggtcac	77220
aaagcaacgt	aaaatggaaa	ttaatcaaac	catcctgcc	acaagaataa	aacagatctc	77280
aggggaggcc	tgtggaaggg	tctcctgagg	ccttaccact	gtctagaagg	aagttgacag	77340
cagttcttga	ctgggggtgc	gactccagga	gttggggggt	gctctgagag	caggacagca	77400
tgtattgtag	agtgtctggg	agggagctgt	gttatcctta	ccgtgaagat	gaggacacgg	77460
gctcatgggg	gcagagccag	gattaaacct	ggtctgattc	aaaaagccag	agatctgtgc	77520
ccagcccccac	gcagccattt	cactggtcaa	ctaattcaga	aacacttggt	ctgatatgct	77580
catatgctac	aggtcactgt	gccttcagat	ctccctctgg	cctggtaact	gcattcaggt	77640
tcaccaccat	caccacacac	acacacacac	acacacacac	acactcggcc	agagacaagt	77700
ggggaagccc	tcacccttga	agtaagccac	gccaaaggaga	aggatgctga	gggcactggg	77760
catttccctc	gtgggacggg	caatcttccc	tttcatctgg	gcctgcaccc	agttgttaat	77820
ctcctgaagg	tctaactcgag	ggttgcccg	gaggatccgg	ggcctggtcc	cataggactt	77880
ctccagaggg	gcaacaaagc	tggatttgac	tcgaagttct	tgagaggaaa	cagatcaaa	77940
atgagagctg	aatcagcacc	ctcaatttga	aagcatgcc	gacccagct	tcctgctcag	78000
catcttctct	tgaactgtcg	gggcatctgc	cggcttgccc	agaccctggc	tagggaacag	78060
tggattccac	cgtttgcatt	ccccgtccca	ggccctcctg	ctgtctccca	gagcccactt	78120
cctctttctg	ttcctctgtg	gtctcaactg	ctctttctct	cccaccagtg	ccaggcctcg	78180
cctgagcaca	cacagcctat	tgttttagaca	tcattggaagc	atacagacaa	cccaggccaa	78240
tgaagcaact	tcacgccagg	cataatgggg	cgtgcctgcc	cttcagaagc	agaggcagct	78300
ttatgagttg	ggggaccagc	tgagactcta	tagactttga	gaggtgggg	gggggtgggg	78360
ctactgactc	ctctcaaa	caattctgga	agcactcttg	aggttcttct	caggggcagt	78420
aacagaggca	aggagctcct	tgtagggtgt	gtggatgtca	gggttggtga	tcagggtcgt	78480
gtagagagcc	cgtgtaatga	cagactctgt	tcgatgttca	gctcctgcc	gagagaaaag	78540
gatgccaaag	ttcataactg	cccgtagggc	ccgcatcagg	atagggacgt	tagacatcaa	78600
tccttttctc	ctctgagagc	ccgaggaggc	cgatattgca	gatgttttag	ctggacaaga	78660
tcttcagggc	gtggaaagaa	ataatgaccg	ccttgctagg	aagagctcta	agacagggca	78720
aggttatcag	agctacagag	agaagagtgg	gatgtggtcc	tgaagttctc	ccatcgtaac	78780
ctaccctgtt	ctgagagagg	gccagctctg	ctcacggcag	ctgtaccctc	agaacctggc	78840
taaatgacta	aaacacgata	ggaggccact	taaggaaacca	aggtcgagtg	ccacttacaa	78900
agtggttaggg	attgtgtgtg	tggccccac	cgcccttctc	ctgttctctc	gacggcgcca	78960
gcattgaaac	tctgagtggt	ggaaattcag	gtccacctgc	agccttcttc	agttgacact	79020
caccacagaga	aagggcagag	agggccgtgg	ccacgctgag	tgagagacag	aggacgttgc	79080

ccgttgggct	ggcaetggat	ctcaggcggt	acagatcgta	gccgaagtgt	gagacagctg	79140
ctgccagctt	gttcacaggg	accttgaaga	aggggtccctc	ctcctccacg	ggctcgccccg	79200
tgctgtccgg	gactggggag	ccctgggtta	gaatacaagg	accagtaggg	aggcacagtg	79260
agtacatcac	ctcctgggtg	ggttgggtcct	ctagtccctg	gggccatgag	tctgaggtca	79320
gaatgagtg	gtgctctctg	actccacaac	ctgtgtgctg	ggaggtgggg	agtgggaagg	79380
gcaacacaaa	agggtcttgc	agacctgaac	tgtgtgtctga	gaacctgaag	cctggcccac	79440
tttaaaataa	aacttgtagg	gctggggaga	tagcacagta	gataaagtac	cagcatgcaa	79500
gttcaaggac	ctgggttcag	tccccagagc	tgggcacggg	ggtgcatgct	tataatccca	79560
acactgggga	ggcagagatg	ggcaggtcct	ggggctcatt	ggccaatcag	cctgaactaa	79620
tcacggtatt	ccatctcagt	gaggggtcct	gtttcagagg	gcctgaggaa	tgactctggg	79680
ttgactacta	gcctactctg	tgtctgttct	tgtctgtctg	tctgtctgtc	tgtctgtctc	79740
cacccctctc	tgtccctctc	cctctgcagg	gaacttcctc	accaccacca	acccccaaag	79800
aaaccacccc	tcagaccagt	cttcctctatt	cagcttgctg	gctggctccta	gtctgcctag	79860
gttctgtctg	gacgcctctc	ctgtctttct	tgacaagcca	tccccctctg	ctagacccca	79920
gaggaatttg	tcgttttctg	acctgttttc	agtgtcagcc	tcttccctat	gagactttct	79980
gcttttttgt	tttgttccca	gggttttgga	tcaaagctgg	gctcttaaat	acgttaggca	80040
aatgcttgge	cacccagctg	tacctccctg	ccctgttctg	tttcgggttg	gaggactttt	80100
tttttagttt	tctgttttgt	ttttgttttt	gattttttgt	tgttttgttg	tttgttttga	80160
gacaggattt	cgctatgtga	ctctagctgt	cctgggactc	actatgtaga	ccaggctggc	80220
cttagattca	gagatccacc	tgctctctgc	tcctaagtgc	tggtgatttt	agattttaat	80280
ctgtacctac	caacctcaaa	ggaagtgtcc	atggatagag	ttcagtacta	catcatgtgt	80340
gtaacatgtg	tgagggcctg	ggcttcaccc	ccaacagaga	agggggagtt	agtaggtgag	80400
gaataagtga	ctggctagtg	gcaagacagt	attgtctaa	gtcactaagc	cttaagccac	80460
acttaagccc	cacaatccag	gtctaataat	cccatctgcc	ttgtccttgt	gtgacatgac	80520
ccacccctca	cttccctcct	atagtggcag	ctcctctgga	tcctgaaagg	agaggggaaga	80580
tattcttgtc	tcgatgttaa	agtaaccaag	gcctagaaga	gtgaaggcca	aagccccccc	80640
tggtatccagg	gctgcctccc	tgcaactgtc	cttctgtctg	cccacctacc	cacctactctg	80700
acctcagagc	tgtctggggc	gttctggctg	ctgccgtgoc	cgagcagggc	tcagctccag	80760
aggagtagca	caggggcctg	catcccggaa	ctacaagaga	aacaagagag	caaacgactc	80820
ccctcaccca	cacctccccc	tgccactgca	cattgcacac	tgcacagggg	caagagttag	80880
gcagagttag	cccttccccc	ccctccagct	ctcagcccca	agtggaccct	tgacttgagg	80940
tcttccctcc	ccgaectgccc	cctgcacttc	tccttgagct	gtgcccccat	gttggttcct	81000
atcaggagac	ccaccttccc	atctaagctc	cagcacaggg	aagaccagc	agcaggctct	81060
tcaggcccca	agacaatgct	ctagcacaaa	cacacaccaa	ggcttttccg	tgagggcaca	81120
cggccagctc	ctttggttag	atttggaacc	ctgtctcagg	atgggagcag	agcccaggtc	81180
atagacttac	agaacatctg	gtctgggtcct	gctatccacc	aatagtctct	tgaccaaagc	81240
ctatgtttaa	gacacacaca	cactttcttc	taggtaggtt	cttgtgtatg	tagcccaggc	81300
tagccttgaa	gttgacagta	ccctacttca	tttgccctct	gagtactaca	atgccaggtg	81360
tgagccatca	tgcctgactt	gactcccttc	ttctatttca	agagcaattc	ttagttaaga	81420
gggtatgaac	cagggccaca	ctgccnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	81480
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	81540
nnnnntctta	tttttagctt	gtttgttttt	cttacttgag	acctggggag	gggaggtgta	81600
tgtgtcttcc	atttgcttct	tctacctaata	aaagtttctt	ggtgatgttt	gggggggggg	81660
gagggggtag	gactcaagaa	gggattctcc	cataagctgt	tccgtttggg	tagtactatg	81720
taagggaagtt	acaggtgggc	agagctcggc	tctgectgac	tggcgtgctc	tgaggtaaa	81780
gtgagatggt	gcaagatttg	ggccctcagg	agttggctct	gttggccctg	taccttctgg	81840
tctgtgggta	aggatgacca	gtaggtgaga	gatgagggaa	ccagaacaga	aggtgaaagt	81900
tagtggggcg	gagcccca	ctagtccagg	gggggtaaac	tagatgactt	tctggaaccc	81960
caaggggctc	ggagactagt	ggtgttgagg	aagacctcta	atgtgttgta	aggcctctga	82020
actcagtagc	cgaacttgat	gccagaaagc	cccaaactgc	taaaccaca	caggagcggg	82080
acgccatccg	ttccatggct	tcacccgagg	tggcccatag	gctgcgcca	tcaatgagca	82140
gccgagagat	aggggcgtgg	acaagccagg	aaaagttaca	gcacgctgga	aagataatac	82200
aggccaggaa	gccccaggca	cagcagggtg	gaaaagctag	atcccgattc	tgccggaggg	82260
ggggcccttc	gaggtcccg	gcaccgggtg	ccaggatcag	agaaaactga	tgaaacctag	82320
ctgacctgcc	cagacctgg	catcctgggg	actccttggt	gctggcgctt	ccttcacggc	82380
gtttcgggga	ctgcactggg	ggctgcagct	gctgccacc	ccgaaatctg	ttcgggaccg	82440
ctggatgtgg	cggaacattt	tcgtttcgtc	gatacacagc	ctactctctg	gagtaggggc	82500
gctggctcgg	tgccgaactt	ggggactgac	aaagcactga	ggggcggggg	tggaagagag	82560
ggcctggga	actgaagttg	gaaccttttg	gaatggaact	ggtttgggtt	gtggatgggt	82620
gggagtaccc	agtgggagaa	tggtactagg	tctgggagaa	attgacctta	gctctttgtc	82680
ttctccaggg	tgtggcagtt	tcctcaaatg	gtcaccgacc	caattaatga	tcacccaccg	82740
tgggcacggg	tcctagttag	agtgtcagtg	ggtgagtgtg	cagaaaaggc	tgaatcgga	82800
aaggccttgt	tggaccggga	attctagggt	cctcccccct	ctttggaatg	gagcagatgt	82860
tgctggaggt	ttgctgtgag	gaattaaggga	cctgagaaaa	gtgggacttg	agatatctag	82920

gctgtgcatc	agctctgagc	gaggagcctc	atagtcttct	ccggtgcctt	caggttattt	82980
cgctgcagat	ggagttgata	tgctgtggaa	ccagacattg	gcccaggcct	gggaccttct	83040
ctgtcaccat	ttggcggtaa	gactctgaag	ggagaggcca	ggtagtaagg	gagcatgtcc	83100
aactcaagg	cccaacctct	ctcttcagtg	ttctgtcctc	tgacttttcc	acaaagcccc	83160
ctgaaaaact	atcctctcag	acttggttg	agttggagg	aggttttgac	tggtctagcca	83220
ctcctgggca	ctgcccagg	agtttggttc	tccccacaaa	cctccagctg	atcataaaaa	83280
aaaaaaaaaa	aaagccagg	atgaaagcta	gggtatgcta	tgcaaatagt	gtggcttggt	83340
gtaagagaa	ctctgttcca	gggctgtcca	tgccccctag	ataagggtca	gcagaaaggt	83400
caggattgga	ggcagtccta	aaaaatgctt	gggtaatata	aagtgaataa	ataaaaaata	83460
aataaatact	aatttttaaa	aagctgatac	ctggaaggat	gaggcagaqa	gtagaaaaaa	83520
catgcgtggg	tgtccctagg	ataaggagct	gggacttggt	gggcacaggt	catgcaaagc	83580
ctgaaccttg	aaccttgect	gcaggtagtg	agctgcctca	gcaccgctgt	tgtgtctggc	83640
cactatgtgg	gcttctctat	gggtatccctg	cttctggagc	tgaaactccat	ctgtttgoat	83700
ctacgggaagc	tactgtctgt	ctccccataag	gccccatcct	tggtccttcag	agtaagcagt	83760
tggtggcagcc	gggcccacct	ggctcctcttc	cgccttctgc	ctctgggatg	gatgagctctg	83820
tggttgtccc	ggcagcacta	ccagctgtct	cttgcctctgg	ttctgtcttg	tgtggctggg	83880
ctgggtcaccg	tggtgagcat	aagcatctcc	acagggatcc	gaattctgac	caaggatatc	83940
ttgcagtctc	agccctaccc	gtttatcctc	atgcacaagg	aaaccaagac	acgtgagcct	84000
gttgccagga	acacttccac	tctcagctctg	aaaggtgtgg	aagttttctc	ttctgtcagc	84060
ccccaggagg	gtggggctgg	gaagaggaga	tggtagccca	ctgcatagtc	tactatgtag	84120
caaggactag	actgtatcat	cagagagaga	gagagagaga	gagagagaga	gagagagaga	84180
gagagagaga	gaacattgta	tgagatctcc	attacagtca	ggaaatcagg	agatctaaat	84240
aactttaaaa	gtcccacagt	ctttacatat	tcttaaaatt	tcaatctctt	taaaaatctc	84300
atctcttttt	aaattcaaa	tctttttaca	attaaagtc	tcaactgtgg	gctccactaa	84360
aacagtttct	tcttcaaga	gggaaaatat	cagggcacag	tcacaatcaa	aagcaaaagt	84420
caatctccaa	ccgtccaatg	tctgggatac	aactcacgat	cttctgggct	cctccaaggg	84480
cttgggtcac	ttctccagcc	aggccctttg	tagcacacgc	gtcatcctct	aggctccaga	84540
tacctgtact	ccactgtctg	tgctgtctct	gggtgtctct	tcatggtact	ggcatctcca	84600
aaacgctgca	tgaccctctc	agtcctgggc	cttcaagaga	gaagactaga	gcctggcaaa	84660
gtggcacatg	ctgataatgc	tagcacttgg	gaatgacaag	cagaaggatc	agaagttcaa	84720
ggccagcctg	ggctacaaga	gactctgttt	caacaacaa	acaaacaaac	aaaccaaa	84780
agagaaagaa	aaaactggac	atgacagccg	gaacattatc	tgacattcat	aaggtcctga	84840
gttcaatgcc	aagttggcag	tgccctagttt	gataagggtc	tagccactct	ggtaatacca	84900
tggtgctgct	aacaccttac	ccagcaactt	gctgtagac	tctgccttcc	agcaaaaggg	84960
aggagcttcc	ctgaggagag	aacattgaac	cctattgtat	atgaataaat	tgctgtgcaa	85020
atgatttcat	cagtctcttg	tgaatgtgat	tgctttgagt	catttttctt	ggctccagt	85080
ttactcctgg	ctgcaagtgt	gtgtggagtt	gtggaagctt	tgagttggga	gggtttcctg	85140
ttaaagtttc	tctggctctt	ttcttctctc	ccggtttttg	ttttgtttgc	ctggtggggg	85200
tctctgtgtg	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	tagaagttgg	85260
cggtgggtgg	agggggctgg	agagatggct	cagcgggtta	gagcgccaac	tgctcttcca	85320
aaggtcctga	ggttcaaatcc	caacaaccac	atggtggctc	acaaccatcc	gtaacaaaaa	85380
aatctgtatg	cctcttctgg	agtgcttgaa	aacagctaca	gtgtgcttac	atataataaa	85440
taataataata	ttaaaaaaa	aagaagttgg	catggatgat	gtagtgaaga	ctggcattag	85500
atatctctgg	atccccctgc	ctctacctct	tagacactgt	gagtatggaa	gtgtaccacc	85560
gcaccaggcc	aggctaggac	attctctgat	ctacaaatac	ctagagtatt	attcctctat	85620
gatcagaaaa	cagacccagg	gggcccacaga	aatgtcttag	taggtaaaaa	cacttgcttt	85680
caggcctgat	aacctgcggt	tttttgtttg	ttctgggggg	cgggagaggc	tggtctggctg	85740
gctggcctgg	aatttcacaga	gatccacctg	cctctgcctc	ctgagtgtca	ggtaccagga	85800
tcaagggtgt	gtgceaccac	acttgcccta	actgcctgag	tttgagcatc	agtactcaca	85860
tggtactgag	gatagaatag	actctcacca	gctcttctga	cttccacatg	tgccctgcag	85920
catgggctct	ccttcccaa	aggaaaaata	aatgtaagaa	ttaaaaaaa	aaaaaaaaag	85980
caaacccagg	tcttgtgtga	tggtcagca	tcaaagctac	ctcccggcac	agctgaccac	86040
ctggtgataa	cttatagcct	tgttatgctc	tcctttgacc	tccacgggca	tgctgtacac	86100
gtatgtgtgc	ccacacaaac	acacaatcaa	gaaataaatg	cagccaggcg	aggtggcaca	86160
cccctttaat	cccagcactt	gggaggcaga	ggcaggtgga	attctgagtt	cgaggccaac	86220
ctggtctaca	aagtgagttc	caggacagcc	agagctacac	agagaaaccc	tgtttcgaaa	86280
taaccaaaaa	aaaaccactt	taaatattat	ttttattttg	ttttgtttat	cctggaactt	86340
ggctctcaga	ccaggctggc	cttgaactca	cagagatcca	actgcttctg	cttcccaagc	86400
acatttaaagg	atgtccacc	actgcctggc	taaaagattta	ttttttcttt	ctttttgttt	86460
tggtttgttt	tggtttttct	aaaaaatttt	tttaaaaaga	accatccctc	ctagcactca	86520
ggagactctg	aagtcagggc	cagccagggtc	tactgagtga	gctctagggc	agccagggct	86580
ccacaaagaa	accctctctc	aacaaacaaa	caaaagagaa	cagacccaac	cagacctgag	86640
gacacacact	tgtaatctaa	gcccttgaga	ggctgagaag	ttcaaggcta	gccacaagt	86700
tgtggtgcat	tcaagagcag	cctgggtggg	ctacagaaaa	agaaagagg	agagagagaa	86760

tggtaaatga	agatgactct	ggaaaagtga	aactcaagag	aaagccctc	agatttgctt	86820
aagacgagtt	gaggggtggag	aaccgcca	gcggacgagc	cagacagaga	ctgccaacaa	86880
agttcaatcg	gttcagggtac	attacttcca	aaacgccatt	gccacatcag	gatgcttcaa	86940
tcagccaaac	caacgcagcg	actattgact	tctgcatttc	agagacttcc	gtctctgtcc	87000
agggcaatgt	cacttttagct	ttcctttgca	gaaaggaaaa	gtccctgcct	ctgatgtggt	87060
agatcctcac	acaccttctg	ccagatccag	acactgggtat	gactcagcct	cggggagctc	87120
tatctacaga	gataagggta	caaggcggtg	gtgtttaaag	tatgtgttta	aaagtacaaa	87180
tgtagagtc	ctggaaaggg	ctccctgccc	tcaccatcac	cgaagcac	aaccttaggg	87240
taatatctga	cattcctgga	aatgtatgta	tgtattcatt	atgtagccct	gactgtcctg	87300
gaatggggta	taaaccagga	tggcttcaca	tctcagagac	ccatttgcct	ctgcctccca	87360
agaactaaga	ttagaggcat	gcactaccat	acttggctca	tgatttactt	aactttattt	87420
tatgttcacg	aatgtttagcc	tgcatgtatg	tgtgtgcacc	atgtgcatgc	ctggtgcccc	87480
agaggccaga	agaaggtggt	ggttggattt	cctggagatg	aagtcccaa	caactgtaag	87540
cagtcacatg	tgtgtgctgg	agatgaaact	tgtttcatcc	acaagagcag	tatgtgtctc	87600
taactgtgga	ggcatatctc	cagcctcaga	tttccagtt	aatgtttgct	ttcgcaacca	87660
ggcccatctg	ggcatgcgct	ggagacctcc	tttaccgcct	tgagcctcat	tggccaattg	87720
tggctgggag	acttgcagat	cccaagtgtg	acaagagaa	aataaactgg	tgtgctatga	87780
atcaccctct	tctctgtagc	cattggctga	gcatactttg	cctcaacctt	ccgcccctcc	87840
ttccccta	cctaatactt	tgcctctcc	aatgtgtctc	ctccccgcga	gtaatccagt	87900
ggtcgctggg	gctctagaga	gatggggggg	gggggagcaa	cgggtacagc	ttaaggcagc	87960
tgagcagaa	cttttttgc	gtatattgag	tcttaaaaa	tcatataaac	tttgtgttct	88020
gtttctaaat	ataaccccat	ctgtttcaac	acaaaatgca	acaacaaaat	gtttcaaat	88080
gctatttggg	ataattaaaa	aatttcaata	cttgatttaa	aaatgcttta	acttttttaa	88140
taaattttaa	atgtttattat	ttttaaaaa	ttacaagt	aaaaaaaaga	aagatagaaa	88200
tcacataatg	aaattaacca	tacgcaagt	aggctcggtg	cactggtaca	cagttacagt	88260
agcatttggg	aatgggagcc	atggcgcttc	acattgaatt	ttatacttcc	tttatagaat	88320
attttttatg	cacctatcta	ctactgataa	caaaacaccc	atgagagagt	tagaattaga	88380
catcaattag	ctttgatcct	ctgtcataac	tctgttccac	tccctgcctt	agtcctacct	88440
catccctgtc	ctcttttcta	catcttatac	tgaatccaca	cactcagttg	tttacacaaa	88500
cacatacatc	actgtccann	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88560
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	88620
tgccctggctc	tgttttgc	ttgttgtttt	tgtttgtgtg	tgtgtttaga	tttgggttgg	88680
tttgggttga	ttttgttttt	tttttagaga	tcttactatg	tagctcaggg	tgtccttgaa	88740
ctcacagaga	tctctgtctc	ctgccttcca	agtgtctgga	ttaaagggtat	acaccacctt	88800
acctggcccc	tttcatctat	ctatctatct	atctatctat	ctatctatct	atctatctat	88860
tatctatcta	tctaaaattt	atctgtgtgt	gtctgtgtgt	acattcccca	gagcctgtgt	88920
ggtagtcatt	aagtaaccct	cagaagtgtg	ctttctctaa	tccttggatc	aaacttgaat	88980
tgttaggctt	ggttagcaagc	atgtttaccc	actgagccat	ttatgacccc	atggcccagc	89040
atcttccatg	ggttctgggg	acacaaatgt	gtactctgat	gtttacagga	caagcgctta	89100
accaaccaag	tcaattttccc	agccccatcc	tgaactccat	taagtgttct	ttcccccaac	89160
ccaggaccaa	atctagagga	gtgtccatgc	ccaacaaaca	ctctgccaa	cctctcccc	89220
tactgctctt	ctcccttccc	ttccttcat	tcttctgtcc	ttcttttctt	tctttttgaa	89280
acaggctctt	tctctgcac	ccaagctagc	cttgaacttg	tgatgtagct	caggctggct	89340
ttgaactcac	agctgtcctc	ctacttcagc	ttcccaaa	ctgggattat	agacctatgc	89400
taccacacct	ggctcatttt	tcaataaat	aaaaagaaaa	tcaaaaagtt	cctagaacag	89460
tcacaggatt	cacaaaaact	ttggaaggag	actaaaaatg	gattttttaa	aaatgcttga	89520
agcacaaga	gttggtgaaa	gaagagagaa	gaggaaaagt	tagcttagta	ggtagaagtc	89580
aatcaagcct	cacaccctga	gttcaattcc	tgaccctatg	gtagaaggag	aagagcaatg	89640
cgggaaacat	tatcctctga	cctccagacc	cgctgtggca	cgtgcatgca	cacacacaag	89700
caggagccct	tggagggaag	tcctagaaat	gaatcttact	gaagcagggtc	tgccaggccc	89760
tgtgctcagc	cattttat	ttcctttgtg	tacccgacac	gcttccatcc	tcaaagtgtg	89820
gagtcctgaga	ggaagtactc	actgtgtccc	cagtgtgctt	ctgtcttacc	ctgggtccat	89880
tagatggggt	cacttagtgg	tagccttgg	gtggagaaag	agaacacagg	tccgagtagc	89940
cagtagacct	gagtccttat	atctgcaaa	ggtgttgggg	cataatcaaa	tctccccccc	90000
ttcccggtgt	cctgatacca	ggttgattta	agtgtatgtg	catggtcctt	ccaagtcttg	90060
acacatcatc	caactccaag	tggctttctc	atttttcctt	gccagtagcc	tcttgggtgag	90120
gaaatggctg	aggaaaacag	agttgcagaa	agacagggcc	atggcctggc	tgcaggcttt	90180
ctctgagtct	gaagagggtc	agcgactctg	agaaaatga	ctatttctga	gtgagagggtg	90240
ccaaagaagg	aacacggcag	agggagagcc	cccagggaga	tggagacaga	agccggagag	90300
ggaccctgtg	cgaggctgga	ggggagggaag	aggggggagg	agtgtgaccc	actgtcatct	90360
gttggggcaga	gaggggctac	attcatctgc	agtatgtgtg	agaggggaca	gagagtgtatg	90420
gtaacaggaa	aaatttgggg	ttgagggggg	cagcctgtag	ggctgggccc	cagcagtgtga	90480
cagctaggta	gagtagacag	taactcccag	aattctctgg	ctccactaaa	tccctgttcc	90540
gctccgtgca	gagtagaaacc	cacacagggt	ggatttcagt	ctcctttgca	ccccctcca	90600

ccccccctcc	acccccagct	ctgggtcacag	ccagtcagag	ttggggggtgg	ggcagatcctt	90660
gtaaaagagg	ctgggtgagga	catcggaag	tctgtacct	ccactagcaa	agtgccagac	90720
gtccctgtac	acttttaaatg	cctcagataa	aacagtgaga	gactctcctg	gtggcaggca	90780
agatgatggg	tcagggaacct	cagcgccctct	gaggctcaga	caccaggata	aagaataaaa	90840
acaccacgga	gacccctgtg	acccctcgt	gcagagggag	aatgccaatg	tggccagct	90900
agctgctgga	gcgcaagcct	caaggcctgt	gctagtattg	agtctactgc	tgctgccttg	90960
tcccaataaa	cccccttccg	cccaggatta	gtggacacgc	cttgctcaag	ccgagtcctt	91020
gatctgccac	cttatcacac	acatacaaaa	atcccttgag	gatggttacc	atcctgggac	91080
aaagcctcat	tctctgctta	acccaagtga	cacctatag	gcagatccct	gtgtccttct	91140
cctgatgata	acaaccccttg	caatccaata	gaggggaact	cgggtttctg	tcagcttctt	91200
ttatgtctgat	agaaatgtac	tctgcatgtg	gggagcctgc	cttgctcacc	ctgagacccc	91260
atggggctgg	ctggggcttt	gcacatcatt	gggactcaga	gatgttgact	acatgaacgt	91320
cccacacttg	gttgcacaa	gcagaatgac	aggatgttat	gcctgggtgtg	tgagtgtgtg	91380
tgtgtctctg	tgtgtgtaaa	acgcctctct	ctggagccct	cctgtctgtc	tgccctctgt	91440
tcaatggctg	ccaattgttc	ctttctcttt	ccaaggacct	ctgtatgggt	gtgtccttca	91500
ttcagtgccct	ttcctctgtg	ggtttgcct	gctagcccc	tgtcactgag	aaagtcttct	91560
gtctgtcctt	gggttgtctg	gctagaacac	agacatcatt	gtcttttttt	tttttttttt	91620
ttttttttta	agatttattt	atatgtaagt	acactgtage	tgtcttcaga	caactccagaa	91680
gagggagtca	gatctcgtta	ggatgggttg	gagccacat	gtggttgctg	ggatttgaac	91740
tccagacctt	cggaaagagca	gtcgggtgct	cttactcact	gagccatctc	accagccccg	91800
acatcatgtg	cttgcccacg	actgctctcc	agaatgggtg	gcaggaggat	gtgaccccc	91860
accccagggc	accgggacac	aacatcttct	acacgtgtag	gtcttgtgca	ctggctttgc	91920
tttctctctt	caagcaggtc	tcccaggaaa	tggcacttac	agagattgaa	gagtttaata	91980
catgtctcgc	tgcctctctt	ttcgggaacc	ccccagaggg	agcagcagaa	accagggctg	92040
gcaggggctc	taagctgcct	gggcaaggga	gcagggggta	gcatggagcc	ttagccaatt	92100
tggaagcac	tgtgacccaa	gcacattttg	cagcagtaat	gtcaaatctt	gccgttcagg	92160
catgccattg	atgtgcacgc	tgccacacag	aaaccagtga	cacaaaggca	cagccttctc	92220
caccctcctg	gtgcttagga	actaacggct	ctaagagaa	atgagagctg	aaaggagaga	92280
gacggggggc	ggccacagca	gcgcaggctg	gcactgcgtg	ttggaggagg	ctgacccact	92340
tctcgtagag	gtaagggggc	cactgaaatg	tcacttaaat	tagccaccac	tcccaacact	92400
agatctcctt	tgtccccata	cctcagcccc	acgtctcttt	ctttttttct	tctttttctt	92460
ctcctctggg	gcagcctcaa	gcccagcacc	cacttttttag	agctgtaaac	caccctggtc	92520
ctagaagccc	tcttacgtta	ggggatgaca	ggaggtagag	atcaggaagg	agggagggag	92580
gggaggaggga	aaggaaaaag	gaggggagag	agggaaagag	atcgagagag	catgcattca	92640
tcacaaagag	ccctcttttc	tggctttttg	actgcactgt	gagttattta	gccaaacata	92700
gatgtttatg	tattttttta	gaacccgtat	ttattaacag	cctgaaaggga	gagagacgga	92760
gatttatata	ggaagtgcag	tgagttaagg	ggggcaatta	agagagcaga	aagagatacg	92820
gaacacagac	ttgtaaaggg	tttgttaaca	tccaatcaaa	ggtgcttcag	gtattttcca	92880
aggaagcaga	aggtaaaaaa	aaaaaaaat	tgtccatta	gaagctgaca	ctgagtgag	92940
caatggccca	ggcggaaact	ctgcttga	gaaggtgaga	agggagggac	acagaccagg	93000
atccgatgag	ccagagtgtg	gccatagctg	ggtcatgagg	cccagggttg	gaaggacccc	93060
actaaagtgt	gcactggcct	ttccttgaca	aaggatgcac	ctatagctag	gcgtgggtgg	93120
aagtgttgtt	tattctagta	cttaggaggc	tgaggcagga	ggatcaccat	gagtgtatgc	93180
ccagcctgga	ctgcatagca	acacccagtt	tcaaaataac	aacaaaagga	agtgggggtg	93240
gggaggggcaa	catttggaat	gtaataaat	aaaacatttt	ttttaaaaaa	agaaaggggc	93300
tagtgagtta	gttcagcggt	taagagcgct	gactgctctt	ccgaagggtc	tgagttcaaa	93360
tcccaacaac	cacatgggtg	ctcacaacca	tccataagga	gatctacgcc	ctcttctggt	93420
gtgtttaaag	tcagctacaa	tgtacttaca	tataataata	aataaattct	ggagtgaggg	93480
ggccagagca	agtagaggtc	ctgagtttaa	ttcccagcaa	ccacatgatg	gctcacaacc	93540
atctgtacaa	ttacagtgca	ctcatataca	taaaaataat	aaataaatct	ttaaaaaaag	93600
agaagagagg	gtggggcagg	ggaggggaaga	agaagaaagg	taagaagcta	aataaaaggc	93660
acagagatga	gcttcatgtg	gaaacacagg	cctgtagtcc	tggcactcag	gtgggggttg	93720
ggggggctac	agtgaagata	tcatgagttc	aaggtcaact	tgggtgagac	cttgtctcaa	93780
aaaatacata	ngcnaaaaaa	aaaaaaaaaa	acatagccag	gcagatgggt	atacatttat	93840
agtcctcagca	cttagaggac	tgaggcaggg	cagaaagaaa	aggaattcaa	gatcaggctg	93900
agctgtatgc	agtcctgatc	ctatccccct	ccccccccc	ccagagacag	acagacagac	93960
agacagacag	agagaaacac	aaagaaaggg	gccttcagat	ggctcagcaa	ttaaaggcgc	94020
ttgtatttca	gaccccatga	cctgagctca	aagcctggga	cccaaggtag	aaggcaagag	94080
ccaactccac	agagctgttc	tatgatctct	atatgaatgc	tggggcatgt	gcctacacta	94140
tggtgtgcac	acatgcacag	attagaaaaa	gaggagggaag	aaaaacataa	gattgtttca	94200
agaaagaaa	ggctggcttc	ttccacgtca	gtgtgagagg	agggctctggc	ccctttgtag	94260
ccaggtcctt	cccagttccag	tgggggctga	actgaggcag	cggaggaggc	aataacggag	94320
ctttcccaac	gcagtgtcca	gcaaaactca	ctctacagcc	tgtcctgatc	cacagagaag	94380
ccttctctggc	tccctcacca	atgcgggggc	attggtcccc	aggctcctgg	gccccccccc	94440

acacctgtgg	agtgtctaggt	gatttgcata	tgttgggcaa	catttgccca	cgtgggggttc	94500
ttggctcttt	ggtaatatag	atgcctagca	ggaggggcga	gcttggagg	gggagtcctg	94560
gggttgcctg	tggctccctg	cagctggggg	gtctggccag	ctgaagaagg	agccatggca	94620
cgcaaatggg	agagcatgga	acagaggctg	tggatgctaa	gcaatatggg	aggcagtcct	94680
agcttggaag	cagcagggtg	ctgggaacgg	gectgtggcc	caggcagatt	tccagtggag	94740
actccagttt	tttggcacia	ggaacaagct	ggctgagccc	aagaggcaag	tggtgataat	94800
gaaacccgca	gtttaggaac	agcgggttaag	ggtgccatgg	gagcccatgt	gctcatgaag	94860
aggctggggg	gtgaagaaga	gccccatgca	ggaagccaca	catccctcgc	agttccaggc	94920
agaggcagag	tccctgagtg	gggtccctcg	ggctccctcc	tacctaacca	gtctcccgcc	94980
acccagcaaa	acaaaatccc	atccataatt	tgaggtttat	agagacctca	aaggctgagc	95040
tactgtgtgc	cactaaccat	cagcctaacc	ctccccact	gtcttctcta	gctgccccct	95100
tttcttctga	gactgtgata	gtggcgggga	cggttggga	gtgtgtgtga	agccctctcc	95160
gactctccaa	ccccagctga	gccccctgtt	ctgcagctca	gtaacacagt	aacacaggct	95220
cagttctaca	ctggttgaga	acactcacgg	ctctctcagc	tccttagaga	gcctgttttc	95280
tcattttcct	gtcccaaaag	cctagacaat	ggctggtcca	tttctaagct	tatctgagga	95340
tgccaggggc	caccccatgt	ctccactagg	ctggcaatgt	tctctgtcac	tgtagtacag	95400
aagactgcct	ggttggagggt	gagataagga	aagggtgtgt	ctccctggg	gttccacac	95460
agtgtctgag	ggaaaatggc	agaatgggct	gggaggtaac	tctgttgcta	gagtacttgc	95520
ctagcatgtg	caagggaagg	cctgggttcc	atccccagca	ctacagaacc	caggctgtgt	95580
ggttctatgct	ggtattctca	acattcagga	ggtacagtca	ggaagagcag	aagttcaagg	95640
ccatcctcag	ctacatagct	agcttgagac	cagcctgggc	tatgtgagac	tttgtctcca	95700
acaaaacaac	acaaagcagc	agaaggccaa	ctggcaagag	gagtattacg	taaaagtaaa	95760
ccatctcaaa	aagcaagtag	catgtatctt	ctttcatttt	tttttacatt	ctataaaggc	95820
ctatcaagtc	atgtatatat	gcattgtatgt	ttgtatgata	tgaagtagag	gggctgggata	95880
gatggctcag	cagttgagag	cacttggatg	ctctttcaaa	gaacctgggt	tcaattccta	95940
gcacccacat	ggcagctcac	aactgtctgt	aattccagtc	tcaggggatc	tggcacccct	96000
acacagatat	ccacgcacat	aaacaccaat	gcacataaaa	taataaattt	ttaaaaaaag	96060
aaattggaag	taaaactctc	taaggagaca	aaagggactg	aggggaagtg	ggaggggcat	96120
gaagggggag	ggcataggtg	tgtggtgtgt	ttaacatgca	gaatacacct	ctataaaagc	96180
tttgggggttc	attatgcaat	gtatacatgt	gtgggtgcaa	gatgtaagct	gtgcatatgt	96240
gtgggggcca	aaggctctct	cctcaatccc	tctctgcctt	attttcattt	aaattataat	96300
tattactatt	agtgtgtgtg	gtgatgtgtg	tgggtgtgtt	aagccctcac	ggcaatcaga	96360
ggatgtctgt	ggtctgagga	tgtctctcta	ccatgttctg	gtgggttctg	tggatggaac	96420
tctgtgtatc	aggtttgcaa	agctagtgtc	tttatctgcc	gagccacctt	gctggccttc	96480
aaccttattt	tttgcattga	acatggaact	tcctgagttg	cctggacagc	aagtcacca	96540
gacctctctg	ttcctgcctc	cccentgtcn	nnntcacang	aggacacacn	gcttantggg	96600
tnccggattt	gctgcncacc	tcctccgcnc	ccnagcctcc	tgctccccc	cccctcgccc	96660
ccgctggmcc	ctcccccccc	cccccccccc	cccccccccc	ctcccccccc	ccccnnnnnn	96720
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	96780
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	96840
agctctaact	gctgttaaat	tcacactcct	tcacatcccc	acgctaggac	tctaaggagg	96900
caccagcgaag	gcccaggtcc	agcttgactt	agagcaaaag	atcctccccc	ctccacacaa	96960
tggaaacgga	cggaaagggg	catggaagca	gaaccagaca	acagcagcct	agccaagccc	97020
aggactctgc	tccttccccc	catgcctgcc	gtgcaactgg	ggaggcaaa	ccccagccgg	97080
tgttttctga	ccgcttagcg	gaagacaagg	ggagcctgtg	attatgattt	ctgctgattt	97140
gcaatgaaac	actaatgcag	tgggcttttc	attaagccag	atttattcaa	tctaaagatt	97200
ttatttctct	tatgtagaaa	gtgcatcttt	atatgttgtt	ggaggagcag	agatgtgata	97260
aaaagaaatt	tctcttatga	actaatagca	ctgatacata	gtggtagcta	tgctagggcc	97320
tctctctctc	tctctctctc	tgtctcctgt	gcattgtgtg	gtgtgtgtgt	gtgtgtgtgt	97380
atgaatgcac	acaaagtgcg	cccccccat	attatttctt	ctgtgggac	tccagactca	97440
gcaaatgggt	gtgactggga	agtctggcca	tgcaattctt	gccttttctc	ttgccagccc	97500
aatccctttg	cattcaaacc	cgggctgctt	gctgtggcca	gccctttcac	ctggagtcct	97560
tctctctctc	tacctgtctt	cccatccttt	gcagacaatt	atcctcaata	actagccaat	97620
tacccttaag	gacaattata	ctcttccatc	agcaaacacg	ggtgttcttt	ccttgagtct	97680
tttgatgaag	tcgatattaa	agagatgctt	tattttacata	aagtcaaata	gctccctttt	97740
agaagggttt	gggttcgatg	tcaaagtttt	aaaatcttaa	ctagaggatg	ggtgtagagg	97800
gcttttggct	agggttagaaa	agagatggag	atacttatto	tgatgtgtgt	ttaaaaggta	97860
ggatgccag	agaagggtga	aggatggggg	agggagggtc	cctcctcaag	ctaataaatc	97920
taaaagcagg	gatgagctgg	gcgttaggag	tggaaaccagt	cagaagtgtc	tgcctttgac	97980
tgaccacagc	tctctgcctc	ccctcccca	gtctctctgt	gaaccgccag	cattaggagc	98040
taatcgcttc	agaaagccag	attggaatgt	gttgctcacc	ctccactgct	cagaaaaacct	98100
ttattccagg	caaggactga	cccaaaccca	tcattggcatc	tgccaatcag	gaggccaaag	98160
gtgcccggcag	ggcgggacct	agctgtgcag	aaacagctcc	gttatggcgc	gcagaaaaag	98220
ctgggggggaa	aggctacagt	tttatctctt	ggcagatggc	ttctctcttt	gatgctttgg	98280

gccttacctg	ttactgcctg	cacttgactt	gacctaggca	aaaatagcag	cgagatacag	98340
gttctcgaag	ttagaaggaa	aaaaaaaaag	ccccaaccca	caacacaccc	cggaagtgtg	98400
cccccgctgt	gtttctaaag	agctgttttc	ttcccaagct	ctacagcgtg	gtgggtctaa	98460
tcggaaattt	ctttttaatc	atagcaggag	tcccaattag	cgtgttgggt	aatctttcaa	98520
gtagagtggg	agttccgtgg	ccacagagag	cagaggcaat	attcagcata	aagccctaga	98580
gaaagagggtg	ttgtgggcct	gtgcacacat	gtgtgtcaac	gcacatgtgg	cttgtggagg	98640
ctggcttccc	actctcaaga	tgagggtgtg	gcacccagg	ccttttgatt	ctcaaaagctt	98700
tattaggacc	acagggactg	tgtgtgtgga	gggtgtgtgc	tcacagtgc	gaaacccaaa	98760
cctggcttct	ccaggagccc	acatgccaac	aaacaggctg	cacactcttg	ctagtacatc	98820
ccctaaaggt	atggggatga	gggaccaaag	gctttgcaag	acagcaggca	cagagtctctg	98880
ggacgctcct	gtaccccca	ctcagccgcc	accagggccc	agctctgac	tggttgacc	98940
tactttcttc	tgttgtgtgt	tttggaaagt	ctgatgtcaa	tcagaatttc	agcagagtgg	99000
ctgagtga	aaaaagagga	gaggaggaa	aagggggggg	ggacgggacg	ggccgaggcc	99060
aacaggaaag	ggcaggcaac	aagacaatga	ccacaaggct	cctgtaacta	cactaactgc	99120
ttactcttcc	tgaccccca	ggcttagcca	atatagctga	gacccagtct	tggtgtctgtg	99180
gcttcaggct	aagtaaacag	ggaagagttg	gacatgggtc	tccattctct	ctcctcatcc	99240
aacaagggga	ggaggcagtg	gccaggcagc	catgcccacc	gatgccatcc	ttctgggagg	99300
agccagacat	ttcaggcacc	tctccttccc	tgggtgccta	gaggtgctgt	gtctgcaccc	99360
atctgccatg	cctgccatct	gagagaggcc	actgggactt	ggtagagagg	ttctccacac	99420
atgctggcct	ggaggaaatt	ggtctttagg	gacactgaag	gcagtctcct	ctgttcagtg	99480
gctccttgga	aaaccacgtg	acagagctcg	catgacaact	tgccggctct	caactcccat	99540
ttcttagctgc	ctcaagcact	gtaaggttta	ggagagcccc	agatgtaagt	atggatggga	99600
agaccctcca	gggagtcatt	gcctaccctt	ctgaactcta	acatgggtcca	gcttttccat	99660
tcacaatttg	aggagacgcc	agacctggca	ggggagcaag	cctttgttcc	tgacccattt	99720
gcaaacccca	gccactgagg	aacttgcata	caagaaactg	cctctggggc	tctcctggac	99780
tgagccctgc	ctcccagggg	acaactgggc	aacagatcct	tccaggtggc	tgacgtgaca	99840
gatccatgct	tttatgacat	agaaaggcct	cagtctcagg	atttcacaca	ctgtatttcc	99900
ctcatcctgg	ggaccaggga	aggcgagcat	cttctgctcc	ccccaaacaa	gtgtgggaat	99960
gattaaaaatc	atgttttttt	tctgtctcat	gaactcatac	agttttcaga	taccgaggag	100020
acaaagccct	cctgtgctga	aattagaccc	cgaaaaatag	gttagctgac	aattacttgt	100080
ttctaagtgg	agtgtgatgt	agtggcagga	gcgcaggatg	ggctgccagg	gctgcagtct	100140
cccccccccc	aaacttactg	tctcttaacc	tctcgagtcc	ctgggtttct	tgccgggatg	100200
ataattctcc	ccatctccct	cctctgggtg	gctgggtgga	agcgtaatga	atcaacgctt	100260
gaagcacgct	gaagaggcca	gactcgggat	gccatgtaag	tacacagcat	cgccagccac	100320
ctctcaagtc	tacacggagc	tgatttatct	acctcccggt	aaagagacaa	caatcatcat	100380
atttacactt	catgccgcag	cttctctcgt	ggcacggcag	cacccctctc	ctctccgctg	100440
ctgaggactc	catcaagcac	gctgccttgc	caggatgaca	gcagcccact	ctcagcctct	100500
ccttgccctc	cttacagatc	atgacctcct	gccccgtgag	gtctgtcacc	cgaaaaacac	100560
ggtacaccgg	gggctgcagc	ctctctatgg	gggaggctga	ggaaatgaat	tccgtaggta	100620
aaaggcttcc	taggaaatca	gacgctgcta	gtaatttaag	agcgaagcat	aggtgcgtga	100680
aaggtaaatg	gatgttatct	aaatgttgcg	tcattttaag	agtgtcctgg	tgcttcagtt	100740
ccttgttacc	atgcagggct	gtggacgggt	ggcaattagg	ctggcacggg	tagagctcac	100800
ctgctgagct	gagggagggt	ggggacacac	cttccggtaa	ttgctgctgg	gcagctctgg	100860
gtctccccac	ccccgcccc	gccctcactc	cccccccccc	acttctttcc	tgacagctct	100920
ttcatttgca	gcagcttaca	gggcttgttg	cccttaacca	gaaaatcacg	ttggaaagaa	100980
tataagaaaa	agaggaatga	aagagaaagc	cagaaaagtt	catattaggt	tcggatctgc	101040
ggccaaacct	ggccgagaga	atccatgacg	gtccgcgcgc	atataaccct	gtggcaacag	101100
ggcccgccac	aacaggggcc	gccacaagag	cttcttgagt	tgccacctgc	caggagacag	101160
gatgaatgaa	tggtatcatct	gtccttagag	cacaagccag	gcctgattct	ccaatattga	101220
tggtgtgagg	agatgtcaac	agagggtccc	taaagaatga	tgcttctatt	tccatgctaa	101280
tcttggggcg	tcagcttcag	tcggaacagc	cggaccgtta	ccttagctct	gctgttctcc	101340
tgtctgtaac	ccgcagaggg	aagggcgggg	tcacccagca	ttgccactcc	ccccccctc	101400
acgtggtcca	gacccctctt	gggttgatct	gctcctgaaa	aacagtgttg	gctcaagttt	101460
gcctctgaag	gtatgtcacc	gctggctcag	ccagcttacc	tccccggtgc	tttcaagatc	101520
aaaacaccca	aacgaaagaa	aaactttgtt	tcaagagcag	agtgtggtgc	caactctgat	101580
caaagtgttt	ttcagcatga	caactcactg	cccgtgacaa	ccagtacttg	gctgtttgtg	101640
ctcagagtga	gatgaggagg	gaagtggatg	acaacagctg	tatccaggto	caaacagagt	101700
agattcacgg	ctggcagaaa	atggctgaga	gccttgggct	gcacccctcc	ccccctcctg	101760
cctctctctc	ttttcaagg	ggtttttgga	aatgtccttc	ctgtgggttg	tggtgccttt	101820
ccatgtagga	cctggggcct	gtgcagatgg	ccctgtgttc	ctggtgctgc	tggtgagatg	101880
tgacagagtg	ataggaaccc	aggcactaaa	cacacaatgt	ggttgtatct	gactagaagc	101940
aaggcaagag	caggaggcat	ttgagggtaa	aggagtgtaa	ggactgtgta	aagagatgag	102000
ggttctatct	gggaggcagg	agtcccaatg	ccagcaataa	caatggactc	tcctggtcga	102060
cccaaccaga	gagaattcaa	gatggcagag	ggacaggctg	tctgagtttc	ctatggctgc	102120

accgataaat	gggcataaagc	agagtagagg	aaaaccacag	acagaaattc	atgccattga	102180
gactagaaat	ctagctcaag	gttggtgtgtg	gcagggttgg	ttcctgggtg	ttcaaccttt	102240
tcacactgtg	acatgatgct	gtggtctgca	gatgtgttgg	gctgcatcca	tagctacctt	102300
gggacacatt	catggaccgc	aggttacaca	tgctatttaa	aaactccaag	ggaagggtta	102360
gagaaatggc	ctggtagtta	agtatgcttg	ctgatcttcc	agaagacctg	agctctgttc	102420
ctagcagcca	tgttgggcag	cttacaacta	actatgactt	ctgagctcca	aagctctctt	102480
ctaatacata	catacatata	tacatacata	catacatata	tacatacata	cgtacacaca	102540
cacacacaca	cacacacaca	cacacacact	ttaaagaaaa	aaattctggg	ttggagaggt	102600
ggctcagcaa	ttaaagacac	tgactgctct	tacagaggtg	ctgagttcaa	ctctcaacca	102660
catggtgggt	cacaaccatc	tgtaatggga	tctgatgccc	tcttctgggtg	tgcgctgaag	102720
acagatacaa	tgtactcata	tacattaaat	aaataaataa	gaaagaaaaga	aagaaagaaa	102780
gaaagaaaaga	aagaaagtgt	aaacgaggaa	aattcctaata	taaaaaagaa	agaaagaaa	102840
aaagaaagaa	agaaagaaa	aaggaaagga	attctgaggg	agaatctgcc	ccttttctta	102900
acttccaggg	ctataggcaa	cctgtggcct	ggggaagctg	tagacaacct	gtggcctggg	102960
gaagctgtag	acaacctgtg	gcctggggaa	gctgtagaca	acctgtggcc	tggggaagct	103020
gtagacaacc	tgtggcctgg	ggaagctgta	gacaacctgt	ggcctgtggc	agcatcatgt	103080
caacgcctca	ccctctgtgc	ccaatttctt	gttctctaag	gacacatgcc	atcaaatgca	103140
taggacactc	tacatcaaga	tgatcttgtc	tcaagatgtt	taacaaaatt	acatctgcaa	103200
agacctatct	ttacatgtga	ggctactcca	cagggtctag	acataatttt	gaggagccac	103260
catccaactc	actatgtgac	agagtcactt	agagatttgt	gtccaggaca	gactggctgt	103320
atctgctctg	agagtccctt	gcctgcccgt	gggaactccc	cagtgggtcct	taagggccct	103380
gaggactttg	gatctgcaaa	gccacatctt	ccaaaacctt	tttctctttt	tggagagcta	103440
ctctaccctg	aaaccttttt	ctctgagggtg	gcttttagag	aggcagggtc	cagcaggcca	103500
ctgtgcccac	aagaagtcct	ggggagaagg	gaccgaagg	ccagtgtctg	actatcgctg	103560
agactgagaa	acttgtgtct	cacctaaaat	cggtggctgc	aaggaccaag	caggctctat	103620
aaatgtctta	ctgctctttt	tcttcttctt	ccgctccatc	ttactcctca	tttttgtttg	103680
tttgtgtgtt	tgtttgtttt	cttctgagat	gtagcccagg	ctggccttca	gctcactatg	103740
taactaaagg	tgtctttaa	cttctgagat	tttcttccct	ccacttccag	agtctctggg	103800
cagggtgtgtg	ccaccgtacc	ccagctttat	ttgagactat	gattcaggct	ccatacttca	103860
tgcatattag	gtaagcatgc	taccaacttg	gctatattcc	cagcctttct	ttctttcttc	103920
tttgagacaa	tgtctttttt	tttttaatta	tatgagtaca	ctgtatctgt	tttcagacac	103980
accagaagaa	ggcattggat	cctatttagag	atggttgtga	gccaccatgt	ggttgttggg	104040
atttgaactc	aggacctctg	gaagagcagt	cagtgccttt	aaccgctgag	ccatctcgcc	104100
agtccttgag	acaatgtctt	gctatatggc	acataatggc	ctcaactca	gaatccttc	104160
gtctcagcct	cctaataact	gggattacat	gtgagccatg	gtgtttggct	tctagccttt	104220
cttccctccc	tttcccttcc	cttttccctt	cccttccctt	tttccctttc	cttcccttcc	104280
cttcccttcc	cttcccttcc	cttcccttcc	cctcccttcc	cttcccttcc	cttcccttcc	104340
cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	cttcccttcc	104400
taactctggc	tgtcttggaa	cttgccttgt	agaccaggct	ggcttgnnnn	nnnnnnnnnn	104460
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	104520
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnaaga	agaagaagaa	gaagaagaag	acaacaacga	104580
cgacaacacc	ggcgccgctg	cctccactgc	catccactgc	agacaggact	caaattccaga	104640
ccaattttta	aaaccagtg	ttcaagcccg	tacactgaag	tagtagtccc	acttgggatt	104700
atagctcctt	actttgtttt	gctttgacgt	tttgtgacat	ggtgtgatgt	agtcttggct	104760
gtcctagaac	tcaatgtgta	aatttaggct	gccttgaact	tgcttctgcc	tcctgctggg	104820
atgatagact	gatggtgtaa	aactccactt	aggaggcaga	ggtgggagga	tcagaaattc	104880
aaagtcatcc	ttggctatgt	tgtgagtttg	aggaccaacc	ttggctacat	gatatcctat	104940
ctcaaaaaga	aataaatgta	ttgccgggca	tggtggcaca	cgctttaa	cccagcactt	105000
gggaggcaga	ggcaggcaga	ttttctgagt	tcgaagccag	cctggtctac	agagtgaagt	105060
ccaggacagc	cagggtctaca	cagagaaacc	ctgtctccaa	aaacaaaaaa	caacaaaaaa	105120
agtgaacccc	aacagtactg	ccggacagtc	tggtgtcttt	cctaagtctc	ctttcaactc	105180
tgtttaacca	ggtgtaccca	caaggtgtgt	gagcagctct	ataccagag	gtgatacgg	105240
tgtttgaatg	agagaaaagt	ttcccatcag	ctcgggtgtg	tgaatactcg	gtccccaggt	105300
ggcagttattg	cttggagagg	tgttggggag	gtgtagcctt	ccggtggaga	tgggcctttg	105360
gagtttaaa	gcttctccac	gaagaccact	gggctgatct	tgttaccaac	agaattggat	105420
tggcctgctc	ttggctccgg	cctcagttta	tctaaaattt	acatgttacc	tgatcaaaaa	105480
ctgtttcctc	ccccaccctt	ctccctgtct	gtattccctg	ccctctagt	gtgctggctg	105540
tatactacac	tggtgatctt	gactgtattt	cagtttacct	cttgtctct	ctctgctgac	105600
tctagagatg	tcctttcatg	gctgggacct	ggctcagaaa	tttctaaggc	actgagcctt	105660
cctccatctg	aaactttagg	aacttcttgg	cttaaggtg	tatttctgac	ttagtatgca	105720
atgagacctt	ggagcctgca	ctttgtttaa	caccctgggt	gggtgtgggtg	gtggtgtgtg	105780
tggtgtgggc	agtagtagaa	ctctgatgaa	cagttagtta	ttcaaggccc	atctagaaaa	105840
agaaaggctt	tggtgtcaca	tggtataaac	tcagtggcga	gagacgtgct	tcctatgaat	105900
aataatgatg	acgatgaaaa	taattctctg	tgactgttct	ccccacttcc	ctctctctca	105960

cctagctct	tatctaccga	atccctgcac	aagcaccat	ggggtttaca	gaatctgggg	106020
cggaacgtta	gtcacttccc	ttcgctact	tcagtattgt	gtttccagaa	gtaccatttt	106080
tggctagta	ctgaggaaaa	cggcagctgc	ctgtgggcca	ccagcccatg	ccaagttagg	106140
tcagcaagaa	agaagctgac	agcaaatgtg	ccaactgttg	gtctgtctga	tttctactgt	106200
gctaagtgg	ttcaagaagt	ttcttcttaa	ccccctacaa	gaaccacaa	attttattat	106260
ctacactgtt	ttgtagatga	agaaaacacc	attccgaagc	tactgccag	tcagcactgg	106320
aactggaatt	tgtttggcta	attcagtggt	tctcaaccag	ggtggatttg	gactccccag	106380
gggatatatt	gggacacttc	tggtcgctac	aattgggac	atgtgctact	ggcacctagg	106440
gtagaggcca	gggtgggtact	gaccttccta	ctatgaacag	ggcagccatg	tacataaatt	106500
ctctcgatca	aaacatcaac	agtgtctagg	ttgagaaatc	tcaggctgaa	accctgtcat	106560
ttggccttga	ggtgggtggg	aggaggttag	agagtggga	aaaatcagaa	gggccaccac	106620
agaggcctcg	agtgaggagg	gaacagggct	cctatgctag	ggataatgga	gaataggcca	106680
gcttgttgaa	acttttcttt	cttccaagct	tggttagagc	cctgtctaat	ttcccccaac	106740
tctgccaaagt	cagtcccggg	acttcgcact	aagtttgtcc	tgagtgacc	ctgactccag	106800
cttggagctg	gggcaacaca	ttacctggg	tcttccagg	agtgggttaa	aagtcacaga	106860
taagtggcat	ctagaagtt	aaaagtggg	tgagtgggta	aaggcaggag	gaccccatca	106920
atcagctgac	ctaggaagga	agagagcaac	tgaagcagca	aagagctggt	ccaggagact	106980
ggattctgac	ccactaggct	tatcttccac	agcctttctt	gtttaggctt	gggctcagtt	107040
tcctcgcatc	atccgcagga	gccctggag	caccacatt	cagccggccg	ccaggacagg	107100
ctccccagca	gtggcctccc	actaactga	cagtgtgtac	aggaaatata	tccccattca	107160
atctcctcag	aagtctgaat	aagtaaggga	cagatgttgg	ggagaggcgt	cactcttggg	107220
ttgatgaaga	aaagatcatg	agaagcatat	attttaccog	ctatggttgg	ggttccatta	107280
ccacccatgg	cggtgttgag	ggggaagggc	agaaaaaagg	agatggagaa	ggacagacac	107340
gtagaagga	ggtatgtggt	agggtctgatt	cgtccctggg	ggtcgaaaca	tcagctgtta	107400
ctggggcgga	aggcaggaag	tctccttaaa	gacacaaat	tctgaacgtt	gaactcagga	107460
tttgaagcaa	ggccagcagt	caccttagtg	gagcccatat	ttaatttaac	acagaagcgg	107520
ctatctcagg	cttccctctc	atttctttgt	ctcagatgct	ctcacttaga	aagcttagat	107580
gctcttagaa	atgactcaaa	agtcaagaac	cccaggcca	aagtttctct	ttgggggtgg	107640
ggaggagtg	aagaggtggt	cccagctctg	tccctttaa	taagcaattc	agcagctttt	107700
gccaagtcac	tgggttcatt	tgggtttttg	cccatcccc	gcctttcaga	ctctgattgg	107760
cccttaggga	aggagccgcc	tcttcattgg	tctccacctt	tgaatcaact	tccctaaagta	107820
ggcctgagtc	agagaagcgt	ttcggagggc	gggactgaat	gggtgttaat	cttagaacccg	107880
ggtttcttgt	tgatactact	ttggttaaga	tcttccctta	attttttaa	agacgcttcc	107940
tctctaaaag	tgagggcgaa	tcctttgtta	agaacgtgcc	ccttgagaag	ccgtgggctc	108000
ttcagcgact	aagacgagac	atcactaga	aaagatttca	ctaaaccac	gagggataga	108060
ctagacctcc	agtgaagatt	gggcctgtgc	gggtgacatt	tgtccctata	ccccgaagac	108120
ctcgagctag	ctctccagtg	aagactgggg	ccgtgcgagt	gacagtggtc	cctatacccc	108180
gaaaaaaaa	aagtctctatt	tgtggaaaaa	aaaaaagact	tgggtgttcc	tgtgtcatcg	108240
gtggctggct	tccatcttta	gttctactca	ctcctgttgc	ttcgctgtct	ccaccttcgc	108300
ttagctcagg	cctcctgtga	atcagttttg	aggctaaaag	aagttccaag	aaggaggggc	108360
tgtagccctt	taaggacttc	cccgcgaccg	agtcagagat	cagtttaaaa	atgccaaact	108420
acagagcgcg	ctgcattctg	gggaagctgag	tgtcaccgta	agaacttcat	tgaccgggat	108480
gcactgcaaa	aatacacgcc	tatacttctt	tctgtctctt	aaactgtagt	ttgacgtaaa	108540
gctggtctaa	gcaagtcgcc	tagggccgag	gttagccaca	ccttttcagc	cattggccag	108600
ttggttagtt	ggtaggcggt	gcttagagaa	gctcctccag	gcaagggggt	ggcctccttg	108660
ccaatcagag	cccagacgcc	tgaatgggcg	ggagtaagca	gaggtgctgg	cgcctccag	108720
tgggtgtggt	cacgttgcgc	agcaatgggc	ggtgattggc	cctgggtggt	tcatctcgag	108780
ctcgtgcgtc	acgacgcgcc	cagctgatcg	gagactggag	ccggtgtgtg	ctgggctgtg	108840
ggaagagaca	gagcgtcgcg	ccgtgcggac	aggtcgagct	gattttgctc	ctctgtccac	108900
agcaaccccc	gcacccagca	tcaggtgggt	gtgatctggg	gacccggtca	tcccgggggg	108960
aaccgcggtg	accgggtgat	ggggaaagta	gggtcctgac	ggccacaccc	tgcctctctg	109020
ggggaggggg	gagggggcg	cggggacagg	ggcgctcttg	ggagaggagc	ctggactctc	109080
ccgagtagtg	tgtctggacg	tttaaagaga	gagtcctgga	caggagtcgt	ggcagaaggt	109140
ttggagaagt	aactggggag	gaatatgaga	ggccagaggg	ccggggcggt	ctaaccocga	109200
cgcccttttg	tttgaggatg	cccgagctga	ccatttagcc	tagggaggat	ctggacgagc	109260
gaggggtgcg	gaggtgcatt	gcctctaccc	gcgctgactg	ggtcaggggc	agttcaagtc	109320
cctggcaggg	aaggggtcgc	tgggcggtcc	ggccctcctc	ctcgttccct	cccggggatg	109380
ttatgtaagg	ggggagggga	aaggagttag	gggcggtggt	gaggaggcct	tatgcaacc	109440
aaaggttagg	gtttcacccg	gggttggcg	gaggttgggg	ggggcgga	ggaggagtgc	109500
ctggaaactc	tacccgcacc	ccccctccca	gcctaaactg	ctgtcttgga	cagagagaag	109560
gtcacctttg	cacctccccc	ctagtatgtc	cggtagagag	gcccctagcc	cggtcttgcc	109620
ctgactgcct	gggaagccgg	ctggctgggt	ggggcgccct	ggttagtcat	cgctgggctc	109680
cctctctccc	cacctcctgg	ccaactcttg	gcccctcccc	acggcctccg	gttaggctaa	109740
cgttcccaac	tccctctggc	cctagtttca	gtctccaaact	catttggcct	gtcacccctg	109800

ctgttagagt	aggctagaag	ctgtcatggt	gccagagagt	tgatggagca	gctggtcaga	109860
gggtcagtg	cctgggccc	ccccgcccc	cagccaaggg	cacctgcttg	gcacaaactc	109920
tcagcagcca	gtgaacctg	tggcctgaac	agagctatcc	tgggcagaga	gaagtggaca	109980
gagactgatc	acctaggaga	aggaagatcc	gacaaagttt	atacttccca	agaggctttt	110040
ggaatttgaa	agttgcccac	cctagtgtaa	tctttccact	ctctgaaaat	agaaatccca	110100
aggcaagtc	tccttggccc	ttctatctgg	cagtggccat	gtccttggac	tgactgtgca	110160
gaaccaccct	ctcgggctcc	cagccctcta	gcctgccacg	ccccagccc	cctccctgag	110220
ccatgctgta	gggccccgc	ttttactgct	gattcatgct	ttggaactgt	ggggcgggg	110280
cttggaaact	ggaacaaagt	tcagacgtgg	aggggcccgc	agacagcctg	gaattcatac	110340
cagatgtacc	cggaaatgtg	aagcggaaatg	cctggcatct	ctagtctga	ggaagctgcc	110400
cagccaccct	accatacct	cctcccccct	ctgcttttgg	tcagctgtcc	tccctcagac	110460
tcctgagagc	ccttgcctgac	cttccaaactc	tagtgcccct	cccatttcta	accctacaca	110520
aacctcctt	gctgctgaat	tccctaagaa	caagtcatth	gagttgatca	cagagctcat	110580
atthctgaag	tacattttt	tttttaactt	gggacttggg	ttctacccc	tgccttttga	110640
atgccgaaga	tgctgggctc	cttagcaggt	tgccaagagt	tgcagctcc	tagtctgtaa	110700
aggggcacaa	agcaagtgc	tttagaagcc	tcttgcctct	tattcaagaa	ccccctatta	110760
gaaggtactg	aaagtcagct	agagccaggt	ttggatggcc	tctgggtcgc	tggccctgtc	110820
accagcttt	cctgtttttt	tttttctcc	ccttcccttt	aggaaacctgt	gcctccaca	110880
ccctcacctg	gctgagccgc	agtagttctt	cagtggcaag	ctttatgtcc	tgaccagct	110940
aaagctgcc	gttgaagaac	tgctgcctc	tgcctcggc	ttcgtggagg	aaagaggaga	111000
gcagcagctt	tgcctatcat	ccggaaggtg	acagaactgg	ggtgggaagg	tctggacagc	111060
tggggtgatg	gccttatggg	agggaaaccc	tggtcctctg	gggagccctt	acccccactg	111120
gccagtgaa	agatttaggt	taaaggcact	gtctataaat	tggggaaatg	gtgactccac	111180
ctccccaa	ttagttgatg	tctgtgtggc	agtgggaaga	aatagaagga	aaagtctgtc	111240
tgthtactga	gacttccctg	taggcctgcc	tttcttatct	tcatcatcac	catgccaca	111300
cacacacaca	cacacacaca	cacacacaca	catacacaca	catacacaca	cacacacaca	111360
cacacacttt	cctttccatg	aggtccaaaa	gtaaatgtac	tcaggaaagg	ggacattgaa	111420
actcogttct	aagtatcat	ttgtgtatth	actttttttg	tttatthgtt	tgattgactt	111480
tcgagacagg	gtttctctgt	atagccctgg	ctgtcctgga	actcactttg	tagactaggc	111540
tggcctcgaa	ctcagaaatc	tgctgcctc	tgctcccaa	gtgctgggat	taaaggcgtg	111600
tggccaccac	gccccgctgt	atttaccatt	ctttatttat	ttttagtctg	gcccagattt	111660
tgggttttag	ggtacttacc	cttaccactg	tggatttttc	cacctgtata	atggggaaatc	111720
ccatagataa	gtaggcagga	gggcattaaa	agtccaccag	tgttgactca	gagcctgggc	111780
tcttcttctt	ctcgtggatg	gaaacgaaac	agctcttcac	atgaactgtt	gtccttcccc	111840
caccccttga	ctactacccc	agctcagggg	gattaggatg	gaaggaaagg	ctatggttaa	111900
gtcccaggca	agctcgtggg	aggctagtcc	tctactggct	tctcaccatg	catgggtggg	111960
ccaaggcttt	ccttccacct	aaagcaaaac	tgtagctctt	ggttgggttc	tagcaaccac	112020
tgcattttat	tttctgcctt	tgctttccag	gatagtgaga	ctctgctcaa	tactgtgcag	112080
gcaagaaatt	gtcaggggag	atgggttgta	tgatatgagt	cccttctgct	gcctctagct	112140
cctgattcat	tctcacgtat	gggcttggtc	tctgattgtg	gttcaccttt	ggcccagctc	112200
tcttaacaga	agatgggttc	agggggtaca	ggaggctggt	tgthgtatth	gacaggagga	112260
ggagtthctag	cctgttcccc	atttgtgaga	aactgaaagt	cataggggag	actagatcat	112320
ctaataccagc	ccactgcag	tctaagctga	gggataggat	gtgtaaggga	ctgtagcaga	112380
cgggctgggg	aggctgagtc	ggctcacaca	ttgcgacaaa	gattgccctt	ccctcgacct	112440
cgcttgcttt	ctttctcctt	cccttccctg	gcacacagtgt	gtccctccag	cactgggtac	112500
atggctctgc	tgctctcatc	caacatggag	cctcagaggt	gagaaagggc	agcctggaag	112560
caacagaggg	aggcacaaga	cagtggagga	cctggcctgg	aaccacaagg	gcctatccgg	112620
acatttggtca	gagaggcacg	tagaagcctg	gagaacacca	ggaaagagag	cagccagcca	112680
gcctcagtga	aagacacgtg	cttccagcca	tctcctctca	ggacctgcct	tcctgggaga	112740
tgaagggcct	ccaggaagta	tgggtcccatc	tctaccctgc	agthttctata	aacagcctca	112800
aggagcatga	gccacctctg	aaaggaaata	cacagcaaat	tcaaaaagag	attcaaatgt	112860
gtaacactgt	gggaaaaaat	atctatgact	ggggttgtag	ctcagttggt	aggthttgctt	112920
aacatgcacc	aagccctggt	tctgtctctt	gcattgcata	aaactgaaca	ggttgggcca	112980
ggtctgcaat	cccggcactc	tggaggtggt	ggcaaaggag	cctacattca	aggtaatcct	113040
ctgctataca	atgatttctg	agccagcctg	ggctatatga	gactgtctca	aaaaataaaa	113100
caaaataaaa	taaagcattg	gttagtaatt	caaagaaagc	agatgtggct	gaaaccgttt	113160
tcctgatcca	taatacaaca	agcaaatgaa	agccagaaga	aggctcctgt	gccttgtgtg	113220
tggcagtagc	aaccattgtg	agagatgcct	ttggacctgg	tagthttgctg	tcttagaaat	113280
gtatcctaaa	ataaggattt	ggttataaaa	tgthcatctc	agggthgtaa	tagagaaaaa	113340
tggaaacgag	ctgthttgtt	ggaagtccat	tcctthtctg	ctgtcatgaa	aatgtatagc	113400
tagggccttg	ctaagtaaat	tatttccatc	tgttggtggt	gttctgtgca	gccatccaaa	113460
gtctttacaga	agaaaaattg	agtggaaata	taaatattga	atactaaaaa	gattataaaa	113520
gtatgagtht	gtgactgtht	ttaaaatatg	aacacatact	tgtaatatat	ttthttaaaa	113580
accatccaat	tgagtggaaa	tataaatact	gaatactaaa	aagattatga	aaagtatgag	113640

tttgtgactg	tttttaaaat	atgaatgcat	acttgtaata	tatttttttaa	aaaaaactg	113700
aaagtggatt	caaaatgtta	agaatgggtg	tttttgtatg	gtgggatagt	acaattgtga	113760
attttccctt	tgttttttct	gtctttctaa	tttttaataa	ttgtgcattg	ctttcatatg	113820
ttaaaataaaa	tacaaaagac	aaataaatgt	tttaaaattt	ttactctttt	atgagtgttt	113880
tgctgtgag	tggcaggaat	ccaacattgt	cttctgaaag	cagctagcgc	taacttctga	113940
gccgtctctt	cactccctct	gtaattttta	aaaaatata	ttgtatgtta	tattatgtgt	114000
ctttgtgcac	cagagtgtgg	gtgcacattc	tgagaggcc	agaagagggc	atcagattcc	114060
ctggagctgc	acgtgttttg	gatcttctga	tgtggatgct	cagaatcgca	ctcaggctct	114120
ctagaaagac	agcaaatgct	cctagccact	aaagccatct	ctccctctag	tcctcattgt	114180
catgttttaga	ttttggagaa	tttgcttgta	ggaggatggg	ctacaçcaag	tgccagggtga	114240
aagaaaatgt	ttgcttgagg	tacctattgc	ttcttgagtg	tgtgtgtgca	tgcttgtgtg	114300
tgtgtgtgta	tgtgtgtgtg	tacactggag	ctaggaatca	tatccagggg	ccttttcaag	114360
ctccaccaca	ctaaagtcaa	ttctatgaac	ttcatataat	gtctgaatcc	acctactctc	114420
tacacacagg	aagcatttct	ctgactttct	gactgtcagc	cagctaagga	ggtgtggctt	114480
agaataagaa	agaaggga	tgctcaaaa	ctgtcactct	ntggggnnnn	nnnnnnnnnn	114540
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	114600
nnnnnnnnnn	nnnnnnnnnn	nnnnnncccc	tattttacca	agcattgcag	gataaataag	114660
aacatagctt	aaagaattct	agagaaaaa	aaaaaagtcc	actaatgttc	aatgttttaa	114720
ttaattttta	atgggaaaaa	gtaatatcaa	agaaatccat	gtctacggca	tattaatgtg	114780
gttaatgcaa	acaggacaga	ttatcttgcc	caaattaaac	cacagtaaca	cgttttgaa	114840
taaaagtcat	ctattttata	gagatgaaat	gtgaattagc	gtgcgcgcgc	aggtgcacac	114900
atgcacacat	acaaacagag	tctttttatt	cttaagccta	ccaaataact	ctttaagcag	114960
taaataattt	taagctctaa	aatttaaaaa	atagtgaac	cccttcaggc	tatgacagaa	115020
tgctgctttt	gccatttctt	ctgataaaa	tcccaagg	tgtcataatc	tgtatccttc	115080
ttagaaaaat	aaggagcaca	tcctatgagc	tgttgacctg	agttttacac	ccaagtca	115140
cgtcagcaca	cagcaatgtc	ctggaccatt	tgtgaggagc	cccacgctgg	tcctgagcaa	115200
caactcactt	ggactgttac	ggggtgagc	tggecatggg	caccactttg	gactgtgtcc	115260
cactgaagg	ctgcagcagg	ccagttccgg	caggttttcc	aatggctttt	gaggcaccat	115320
aagccttagc	tgagtgga	cctatgcaa	cacacaggaa	aagggcagtc	attcccgttt	115380
tatttctcga	agtgaatgta	caggcaccat	ggcacgcac	tgagggtcag	atgataatct	115440
caaaaaaatc	caacttttgt	gagacaggct	cccttcattg	tcttcccacc	acacttgcca	115500
ggttaggggc	acaagctccc	aaggagtatc	tgctctccta	agggcactgc	catcagagac	115560
actctgacta	ctacacctgg	cttttaggtg	ggttctgagg	atttgaacgc	agatcatcat	115620
ggttgtatag	caaattttta	ttcaccaagc	cagacaaccg	ttttttaact	ttttttaaaa	115680
aaagattttat	ttttatgagt	acagtgtggc	tgtcttcaga	cacaccagaa	gaagacatca	115740
gatcccat	cagatggttg	tgagccacca	tgtggttgc	gggaattgaa	ctcaggacct	115800
ttggaagagc	agtcagtgtc	cttaaccgct	gagccatctc	tccagccctt	atttaacttt	115860
taaaagttta	aaattaatc	ctattatgca	cgtgccacag	ggaagtgtaa	cataagcaca	115920
tgtgtttgaa	gtatatatgg	ggatgtagtt	atgtgacggc	cagaggacaa	ctctgtgaag	115980
ttgattcttt	cctttcactg	tgggtcctcg	gagcgtcaat	agggctagca	ggcattttac	116040
cagctgagct	atcttgtggg	ccaccaaat	tatttttaaa	aagctagaag	ttgggttaaag	116100
agaaggaatt	catgtaatgg	aaatagaaaa	cagagtttga	agaggagga	gaggaggtga	116160
gtatgataatg	gtaagaaaaa	ccaagaaaaa	acctctacaa	gtaaatctag	togtgttcca	116220
cagatcacta	tgttcaagga	tggtgtccca	gaaaagggtc	ctaaaacaaa	gcccaggaag	116280
taatcaaaat	gatcccatcc	agacctgcta	cagagaagct	ggggggaagg	agggagaggc	116340
cctgtaacct	caagtcccg	ggcagcacag	cagagctgtg	cccaatgatc	tccgtgatac	116400
aggacaacag	gatattctgag	aaggagctcc	agtgagcgtg	tagcagaagc	cagagccctc	116460
gaacaaaacc	aacaacttca	ccatgatgaa	catcatgaag	atgtgtggac	aaaagggtgc	116520
actctgggac	acagtgtcac	acactacagc	ttccacctgg	gacaaaaagt	cacacactac	116580
agcttccacc	tgggacacaa	tgtcacacac	tacagcttcc	acagtcatag	tatttttctt	116640
ctagcggaga	ggctgcaagg	gtggagggca	ggtacaagg	aacagggtg	agtgggatcg	116700
gggtgcatga	tgtgaaactt	acaaagaacc	aataagttag	gaaagagaga	tgaagacata	116760
gaagccatgg	ccaacaaccc	cagcactctg	gaggcaggca	ggagggtctc	catgccaaag	116820
gtaatgaatg	ctatgtctaa	agagatgata	aagtcctcatg	agagttacaa	agtcagagga	116880
gaaacttgga	cagaaaagcc	aaatgtaaca	aatgcattgt	aggtggggaa	aggtggggat	116940
ttaaaagaaa	tttcacaaaa	aggcacagaa	tgtaaaggta	aagcgagcga	ctaccagata	117000
aagcatcaac	taagaggagt	caggaaacagt	gtatttttaac	tacctttata	aaaaaaatag	117060
tcgggctggg	gagatggctc	agcggttaag	agcaccgact	gctcttccaa	tgggtcctgag	117120
ttcaaattccc	aacaaccaca	tgggtggctca	caaccatcca	taacaaaatc	tgatacgggtc	117180
ttctggagtg	tctgaaggca	gctacaagtg	tacttacata	taataagtaa	ataaataaat	117240
aagtattaaa	aaatttgcaa	taatttggaat	tattctttgc	aatctacta	tgaatggagt	117300
ttttgttctt	gttgtttgtt	ttccgagcct	gtttcatgta	atccaaactg	gcctggaacc	117360
tgctaaatag	tggaagatga	ccttgaactt	ctgattcccc	tgtctccacc	tcccaagtgc	117420
tggcctaggc	tactctgcct	gactcaggct	cagtgaattt	tcaaacactg	cttcaacctt	117480

gacccctaac	ttccatctct	tgggtcccat	cttaccatt	cccagactcc	cattctgtgg	117540
ctggggcttg	ctggctggtg	gtgtcgctgc	agaggctggg	gaagggactg	cttgctgctg	117600
ctgctgctgt	tgctgctgct	gctgctgccc	atatcctgga	gatcaaaaca	ggccaattca	117660
gtgcaaaccc	agtcaagaga	tttcaacacc	agcagtaaat	accttaggaa	aacccacctt	117720
tggctcttag	aggagcagac	tgcaccacgc	tgcaacccct	gtgtctgcaa	agggctcact	117780
tttttgtctg	cgcatagtaa	agactggttc	catttccctt	tcaactgttt	agaattgaca	117840
gcttttcagtt	ttaatgaggt	tctccctgaa	gccttgccca	tttctccttc	aagaacctgc	117900
agtaaaagtc	ctgggctcaa	tctgtgatac	taccactatt	tccatagcaa	caaatcgatg	117960
tcatcaacag	gacgggtgtg	tcgtggcagg	atccattaga	agagaaagca	gggtggttaag	118020
gaaaacctaa	aagcagttca	attgtctctc	aagtgtcttg	gcttcaaagg	aaaaggagac	118080
gtttacaaa	ggttctaccc	tttcccccaa	acaaagcaac	cttattttgc	aactgacagc	118140
tgggtgaaaa	tatgtattga	aagagtctca	ggtaccgtga	agtgttagtg	aactgtcaca	118200
gcaggggagag	atgagaacag	agttcagaag	cagccgtgtc	aaggcaacag	ggagataaaa	118260
aggagggggag	ccgagcacgc	ttctaaagca	gcaaataaga	cgcgccaccc	tgctggagtg	118320
tgtattttgcc	cttcaccttt	tatacaaatg	tttaccaggt	tgaagaatgt	taacattgta	118380
aattgcctac	cactattttt	aaaataacga	gtttatgggg	tttgatttta	cttctgtgat	118440
tggctctgaa	ggctgaagcc	agcctcggtt	aggctgtcca	cctggagtg	ttgtagatga	118500
caaaggcagc	taaaaaaaa	aaaatgtccc	agagctcctg	aaacactaaa	actggtgcgc	118560
acaggcgagga	agtctccctt	gccactgagg	ctctcccttc	tccaactgta	agttcttaact	118620
cctgctgggc	ctcctgaggt	cccaactcac	cgggcatgct	acatgcccct	ggaattcatt	118680
tttggcaaaa	ttactcttaa	gtagctcaag	gaatgagacg	taattgtgtt	gctggagaca	118740
aatacattca	gcacaaagtt	gagaagatta	aatagaattg	attatgcttt	agttcatcct	118800
acagagagaa	aaagtgtgac	acctaattac	ataggagagg	ggcccaggct	actgacaatg	118860
agcccttcct	ttacagctca	gccttcccat	tatcactcaa	ctcagacacc	cagcccagtt	118920
cagtctatgt	gaagcatttt	taaaatcagc	aagagaaagg	tgagtgtctc	agtagtacac	118980
tgaaggaaag	tagaaatgag	cacacagtga	cctgtgcat	aagacacagt	ttaaaagggtg	119040
acctaattcc	cagcgaagtt	cttgctctct	ttaaaagaa	tgtgttattt	gctgtgtatgt	119100
gcacctcagc	aaggtgcttg	cagctatgtg	gagggcaga	accatcctga	gggaatcggtg	119160
tctccttcca	ccagagtcct	agggatggca	caaaggcacc	gggcttgaca	gatgccttta	119220
cctgctgagc	cctttcccca	gcctctcctt	ttctacagtc	tccaaattac	ctgaggtagg	119280
agcctcattc	tcagaacctc	ttctgtctgc	catgtctgac	tgaactggat	atggctcactc	119340
gactcaattt	ctataaaaag	ataactgagg	agccggcttg	gttggtagag	aaaacataaaa	119400
gacccgagtt	ttgttcccac	atggggcggtg	cgtgctagag	aggtgaggac	ggactaggca	119460
gggagcagag	gtaggcagat	cccaggcctc	aatagccagg	cagcccaacc	tgagctctaa	119520
gttcagttaag	ggagcctgta	accacagaag	acactgtgcc	gtgtaaccac	ggcaagcggc	119580
ccgctgcagt	ggcatctgct	catagccaca	gctaccagg	aggctgagcc	acaagactca	119640
ttgaagctgg	aggtcaaggc	cagcataggt	accataggta	gacccccatc	tcaaaagtta	119700
acagtaaacca	tatatattac	tacaattaaa	aaaacaaggc	cgggtgtggt	ggcgcatacc	119760
tttaatccca	gcactcggga	ggcagaggca	ggcggatttc	tgagttcgag	gccagcctgg	119820
tctacaaagt	gagttccagg	acagccaggg	ctacacagag	aaacctgtc	tccgaaaaacc	119880
aaaaaaaaaa	aaaaaaaaaa	aaaaaaccaa	aaacccaaca	catgtacctc	ttctaacata	119940
aaattttcat	tttttataaa	aattacagtt	ataattttta	gttgaacata	atacaatgat	120000
aagtctccaa	cttgataatc	tgaggctggg	ggtgtagctc	atttagtaga	gtacctgcct	120060
agcagtcgct	aagtccgggt	agtcctctgc	ctgtatccca	gtgcttgggg	aacaggcaga	120120
aagaggacca	gaagtccaag	gtgctcctcc	tcttcaggta	atgaggagtc	tgagccagcc	120180
tgggatgcgt	gagacacacc	agtaacagca	actactgcct	ggagtgtctc	ctctggacca	120240
ggaacaagag	ttaagtgcgt	tactcccgta	actgtctgca	caatgatgga	gacagtacgt	120300
catctttaac	atctttggct	gagaagagaa	aacctgggtat	tcctcagctc	gttctggcta	120360
agttcatgta	cttcatcatc	atgcagttaa	tttgtaaaca	accaatcctg	gcagcaataa	120420
ctctaattat	atagaacata	agatgtggta	attaggaaaa	gctactaatc	cacttaatatg	120480
agtaaccttt	atctcttgca	aatctggtac	aagagacagt	cccaaatcaa	atgatggcaa	120540
gaattccaga	gcattgttaa	tagcagcaat	tgcccttcaa	ttaacgcatt	gtaacgcagc	120600
agctgcccac	aagacctcaa	atcaatcagt	ctatagctaa	ggaaaaatct	ttctaaagcc	120660
aaaaccattc	tacaaagcag	taacgtaggc	tccgtttata	ataacctgtt	ttggggccacc	120720
tgcaaatcaa	gctatcccag	gaagccagat	cgtaattctt	aggctctgct	ggctacacac	120780
tgggtcccaag	ccatgaggga	actagattac	agcaggctcc	gccctcggtg	acctgtctcat	120840
agctatcatt	cttacagtct	attatggcaa	gtgagctctg	ggcagagaaa	aattcacaaa	120900
caaacccacc	acattcccaa	gcaagcattt	tcttaacaaa	cacaaagaat	aaataaatag	120960
agcctgccat	ggtagtgcac	acctgtaatc	tcagcatgtg	ggaggcagag	gcaggcagat	121020
ctctgccagg	agttccatgc	cagcctggtc	tagacagtgt	caagatcaac	aacgctatat	121080
ggtagggccc	tgtctcaact	cccaaacacc	tgaaaacaag	taagacgata	tggatcaatg	121140
caatatttct	cagttcctac	tgacagaaaa	tggacacaa	taggctgggt	ataattccaa	121200
tatgaataat	aagtatatta	tacagtacct	agtttaagtc	tgagcaagat	attatcgcca	121260
acagcacaga	tacaggacac	acacacagct	gcaagcgtga	aggattttaa	ggcaccatc	121320

cctacagtat	atgcagcatt	gactgtctag	ttttatttcg	cacactttga	atcatccatc	121380
catttttgc	caatatggca	gcagtaataa	aatgtatatt	tgcattttga	tgcatgggtg	121440
aattcttact	agcctgggct	gtgtttgctc	actaactcca	gtggacttct	gacgtaagag	121500
gcgctggact	agcttcagca	ggatgtaaat	ctaactttgg	ttctcggcct	cccctgaagc	121560
ctttgctgtg	gtgaaagggtg	ctgtttctga	agccacgaca	gtcccatggg	ggtttgagta	121620
acaatactcc	tggtctggta	acaatgccaa	aaaaataccaa	aaaaaccaa	acaaaaacca	121680
aaaaaagcac	agagctcaca	tctgagccaa	aaaaaccgac	actccctatt	ttttgaagaa	121740
ctcacagaaa	tcaagaagaa	aaacaagcaa	acaaacagca	cataacaggc	taacaacaac	121800
aacaagtcca	cttcagaggc	caaaaaacca	aaggcaaaag	ggctcttaaa	catttacaag	121860
gacagacctc	actcagaaca	caagctatag	ccatgatact	gctcttcatc	catcaattct	121920
taaagacaca	gaagatgggc	tgtctggcatg	actgggaaga	gaccagtgtg	ctcactcatt	121980
gctgggtac	gctgtaacaa	gcagaggaga	atgtctaaact	gtgtctgcca	atcacaggca	122040
cattccagct	taacccagca	atcacttctg	gcaaaaaaag	cccactctgc	ccgcacactg	122100
gtgtaccagc	gaggtaatct	actgcggtct	ctgggtgatc	ttttttctct	cttgcaagtgc	122160
ttgggttcaa	attaggtaca	agcacacttt	tatagtccaga	gcacggagac	aaaaggatct	122220
aaccacagag	ttggttaaat	aaacaacaga	acagccacac	cagaactgct	ggggggcggg	122280
gggagggggg	ggagggaaag	gggagaagct	agaatcccag	cattcaggat	gcagagggtg	122340
gtaaacaaagt	ttcagggtca	gtcgggggtat	aaggtaagac	tttatttcac	aaaaataaat	122400
aattttttta	aagaggacag	aaaaaagaca	gcgtagagaa	ctagctatct	tttatgtaaa	122460
ccatgtgggg	taaggagca	gcatttgcat	ttggttactt	ggcgaatagc	cctggtagga	122520
taaagaaacc	ttaaactgtg	ttatctataa	agggcagcag	tggtgtcaca	atgcagtggg	122580
taggagccca	cccattgcac	gccatagtct	tgattacaata	gtcacgcagc	tctactcaaa	122640
aattaaaaca	aacattaatg	cttggttaag	aaacagggct	gcagagatga	tggtctcagt	122700
gttaagagca	cagcgcccat	atggaggctc	tcagccatct	gcttctccaa	ttccaggggg	122760
agtcaggccc	ctcttttggc	cttcaagagc	actgcatgca	catggtaagc	ttacctacat	122820
gcccgcaaac	attcaaaaga	aaatacaaac	tgctataaaa	cccatatatg	accatcttaa	122880
gattctttca	tttttttgag	acaggggtttc	tctgtgtagc	cctggctggt	ccagaacttg	122940
ctctgcagac	caggctggcc	tccaaccag	agatctgcct	gcctctgcct	ccacagtgtc	123000
ggaattaaag	gtatttaaca	cacacattat	acatatcttc	cttcttcttc	ttcttttttt	123060
aaagcgtttt	gttgtttttt	gttaagtctc	aattaaaaaa	ctacatagtt	ttataggcaa	123120
acataattaa	aaatgccaat	gtgaaataaa	taatatatac	atatataaca	ttctgtaata	123180
gattcactca	cacaacttat	atacttaaat	acaattttca	caataatgaa	aagccttgaa	123240
atgaagactt	ctggatacat	tagaaacgta	ccctgaaaat	cgcaaatgac	ggttttcatt	123300
tctttgtgtc	agacattagt	gtgagtgtct	aaacttgcac	aaaggctctc	ttctctatca	123360
cttctaccct	attgcagtgg	ttctcgctaa	acctggaaac	tggggctcct	gttctttggc	123420
tggactacaa	ggcagcaagt	cccagcaatc	ctcctgtctc	acctttcttg	gaaccaatgt	123480
tataagtgtg	tgtgggcact	agccttgtta	catggctgct	gggactagaa	ctctggtctt	123540
caatattatg	catcaagagc	tcttaactgc	taagccatct	ttctaccctg	attagaattt	123600
cttgaagcaa	aagaaactca	cagatgggtca	gagtttacac	acacacacac	acacacacac	123660
acatacacac	acacacacac	tcacacacca	aggcttagtg	accactgtga	aaagggaagt	123720
gcggtgagga	actgtaaaaa	taagggtgtca	ggaaagctct	gcacaaaatg	gtgtcctctg	123780
gacaagccag	ggcctctgcc	ctcatcagct	cttagtaact	atgggtgcct	acagcaaac	123840
atgccaggga	ccactctaac	atggagctgg	gatgggctcg	agaggccctg	ttattaaagg	123900
agaagctgtg	gagagtgtat	ttgatggatt	ctaaagttag	gggaatcgat	ttccttttat	123960
gatattgggtc	agccatgctc	tagtgagtgg	ccccacaccc	acccatgagt	atgtgtggag	124020
agcacacacc	ggacctggca	agtcaataaa	acaaaaacaa	aaacaaacaa	ataaacgttc	124080
tgggtccacca	tagtggctgc	tcgtgggtgg	gagctgtccc	gcacttatgg	caggaagaca	124140
gtggctgcac	agaggacaaa	agtctctgga	actgatcaac	tctagagtct	gcttgttatg	124200
agaactggga	agtaccgcgt	gggacagaag	cagactctga	aggtgatcag	gacagagatc	124260
acagaaggag	agactgggta	tcggaggaaa	tctgaacat	aactcgacgc	atactgggtc	124320
aaaactgggtc	ccatcactac	aacagcagta	attgaattgg	gcacaacatt	cagaaaacag	124380
aaaaagacta	cagagtacgc	accctggcta	tcatacaacc	aggcgattct	ggcactattg	124440
gaagcaagcc	agactggaga	aaaggaaaca	aaaagtattt	taacaaaact	tcccagagca	124500
tggttaaaaa	aaaaaaaaga	aagccagaca	tgggggtgca	cgtttttaac	cctagcactc	124560
aggaggcaga	ggcaggtgag	gcaggaggat	ccatgagttc	gaggccagcc	tggtctgtac	124620
agtgggttcc	aggaaagcca	ggcaacaaag	aaaccctgtc	tcaaaaatca	ctgactgggg	124680
agagagggaag	tggatccaag	agcaagagag	agagagcaga	gagagtgggg	gtttggatgt	124740
cagcgttatt	aaatgacagc	agaaaagatg	ggccccacca	atgacatccc	agaaatggca	124800
aagatgaaaa	aataaacaca	agttctaaat	atcatttttaa	taatggctgt	gtgtctgctg	124860
gctcattctt	gttatccaaa	caaaactaaag	cagggtgggc	cggtatttac	ggccagccag	124920
gactagagtg	agacactgct	ttaaaaaagc	aataagtgca	cgctaaccat	taatcagcgt	124980
aactggagtt	tgagggaggg	agtgggcccg	ggaagcctgt	gcctataaac	ccaacctgcy	125040
agggcctgaag	cccgaagggt	gaactgagaa	catctcaaga	caaagcacag	gcacaatctc	125100
ttacaaacag	tttaaacaca	ataccagcaa	taaattgtca	gctttatgac	agataggctg	125160

acaggcatatc	cacaaagatc	gggaagaaga	aacgggattc	ttgcacaaca	ttttacaaat	125220
cgacacagct	gagcctagt	acaggccgtg	ataagctcaa	aacatgcaca	ctgcaaacac	125280
cacagcatgg	ccagagtgac	aggggtcacag	caatacagca	acagagacag	taagaaacat	125340
tggaaaaagg	aagaaggcaa	agagcaccac	gagagaaaag	ttaacccgcg	attccatatg	125400
agggcccagc	atccctctcc	cacctgacag	agaaaaccag	gaggacgcag	ctgaaacact	125460
gtacaactat	aaatactcct	tcctagtgtg	gacagagttt	accaaagggt	atcgtaatct	125520
gaagcacaac	cataacactg	taagttacca	gaagtttaga	cacgttgaga	atttaataga	125580
agttgagaca	tcaaacacac	gtttgatcct	aggggaatta	aattattggt	aacagaaaag	125640
ctcttttaca	agtctccaca	tttataaacc	aaatacactt	gtatcagaat	tgttgtctga	125700
gccagaggcg	gcggcacaa	ggtggctgag	gcagcaggat	caagtctgga	gaagcccctg	125760
ggtccaggga	gcactagtga	attcaagtaa	acattaaaca	ttaaataaga	aagacggatt	125820
ctacaggagc	ctagacaccc	tgtaagggtc	gtactactaa	gataccaaaa	ccaaactgac	125880
tcagggggct	gatgagatgg	ctcagtggtc	aagggcatgc	actgctcttg	cagaggacct	125940
gagttcagtt	cccagaaacc	gtatcaggca	gttcgcccac	ctgtaattga	ccagagaccc	126000
taaacatttt	tatctctcaa	gcacagacac	acataaacat	aattaaaaat	aaatcttaaa	126060
aaaaaatctc	tcgtggacag	acagcaaaaa	atactgaacg	aaatctaate	aggtagactc	126120
agcaaaagca	atccaaatat	gcgaaaagga	taatacaatg	caaatccaga	cttatcccag	126180
aaaccccagg	tcgcttttagc	atccaaaaaa	ttcaatcatc	ataattcacc	gtattagcta	126240
accgaaacag	aaaaggcatt	tgattattaa	taaatacagg	gaaagcattt	gacaacattg	126300
accatttca	cttaataaaaa	atgtttacca	acaggagaaa	gaagtataat	ctcctcaacc	126360
cgatgaacag	gaacacaaac	tggctgggca	cagtggcaca	tcctgaaacc	ccagcaccca	126420
gtagctggcc	agactgggct	ggcatgagaa	ccaagggtaa	gaggcaccca	aacacttaat	126480
gatccaaagt	gctgcctccc	cccccccctc	gtgtcactta	aatccgtaat	taaatctctg	126540
ggaggagggg	tttgacacag	ggtctcactc	gtgtggcacag	gctagcctga	atggagtcaa	126600
tgctggcctc	aaaactcttc	tgctcagttt	ccagagttag	gggaaacaat	catgagccac	126660
ctcaccagct	tttaatgccca	tttaacagta	caagggttaa	cgctgcctga	aaagaacaag	126720
acaaagaaag	atcacacaga	caaacatcac	acaggtgctg	tttgagacta	ggtctctagt	126780
gatgtaggct	ggtctcaaac	ttgcacaaga	aaagaatact	tttgccatca	agatcaaggt	126840
tgctatcacc	catggtggca	tttcgggaagc	agaggcagga	agatcagtag	tacaagggtc	126900
tcctcagcta	ccatgggtgt	gaggccagcc	aaggcagcac	gtgagccagc	tgtggaagca	126960
cacatctgta	accccaccac	tcaggaggct	gaagcaggaa	gttcaagcca	cacaagagcc	127020
acttaaaaa	ataaaaaaac	ttggagagag	ttggagagag	ggnnnnnnnn	nnnnnnnnnn	127080
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	127140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	127200
tgctcatctg	aagggttgaa	atcactggag	aatggcta	gacaaaaact	cttaaggaga	127260
atttttattaa	aaaacaaaac	tccttctgga	ggagtactgt	gctcccttca	gcattaaacta	127320
ctgtgcccag	gaaccactg	ccagcagaca	cccaacttca	aggataatca	agcttgctac	127380
gtacatctgg	tgcttacgca	gacaccacac	acacttctct	tatgctgggc	atcacctcta	127440
gcttacttat	aatatcaaac	ataagtacca	tgtgagcgct	acagggcctg	gccaatgaac	127500
agcgatggga	aaaataacct	gcacatggct	agtaatgggc	aatccactcc	ccagacactt	127560
ctaacctcag	ctgcagactt	gtgggtactg	agaacggagc	tgactaagtc	aattcacaga	127620
agcagagccc	tgtgctgtg	ccagggaacg	gaggacgctg	ctcaatgggc	agagcttcag	127680
ggcgggaaga	gatggaaagt	ggagctagg	cccgatgacg	tgaatatggc	taatgccata	127740
ttgtgtgtgt	gtgtgtgtgt	gtgtgtgtat	gtatgtgtat	gtgtgtgtgt	tgtatgtgtg	127800
tgcatgtgtg	tgtgtgtgtg	tgtatgtatg	tgtgtgtatc	caaaatgggt	ataaaaaatcc	127860
catacaatgg	ctatgcttat	gtgtctttta	ccacaagtaa	aaaatttttaa	gtaatcttag	127920
gaacacattt	ctaaaaattg	gaatacttgt	tggggaagct	atctgagccc	cagcatgaat	127980
caggaaatgg	gtaggagagg	ggccttagtg	ggccttagtg	ttaacacaca	catggtggcc	128040
ctttcctggc	accaactagg	tcactcacaa	ccaggatctg	acagcctctc	ctggcctcct	128100
caggcaccag	gcaggaaagc	ggccccacat	ccctgcaggt	aaaacatttg	tactgaaact	128160
aaataaaaa	ggagcagaag	ctgggcctgg	aagtccttta	cctccagcac	cggggaagcag	128220
gtggatattt	tgtgagttcc	gggtctacat	agtaaaaaact	tgtcgccaag	taaaaacaaa	128280
caaaaactgg	gggctggtga	gatggctcag	tgggtaagag	caccgcagctg	ctcttccaaa	128340
ggtccaaagt	tcaaatccca	gcaaccacat	ggtggctcac	aaccatccgc	aacaagatcc	128400
tcttctggag	tgtctgaaga	cagcaacagt	gtacttacat	atattataaa	ataaatcttt	128460
aaaaaaaata	ccttttaaaa	aacaaaaaaa	aaactcccac	taaaataatt	ggaaaagtcc	128520
agggaaagcc	acactcgatg	gcgcaagcgt	gtgatcacag	tattctgtag	gtaaggcaag	128580
aggatcacag	cgggacggag	gccagcctca	gctacacagg	ccggtctggg	ctacagtgtg	128640
agaccccggt	ctcaaaaacaa	aacaaaaaaa	aaaaagcggt	cacattatca	tatactcaag	128700
ggcacaaaac	accttctctt	tcccaatcct	tgaatgctat	cgaatttggt	tacttttttt	128760
tggttaatttt	ttgtttgctt	ttcaaggcag	ggtttctctg	tgtagccctg	gctatccttg	128820
aactcactct	gtagaccagg	ctggcctcta	actcacagag	atgtaattct	ttttcaagc	128880
tggaatttga	tttgggggtg	tgtgggtgat	accacaggga	cacttgggcc	ccaaacaaa	128940
ccaaagaaaa	aacaccccc	cccaaaagta	tgaacatcaa	ctgtatataa	aactaacagt	129000

tcatataagc	taagtagcct	gcaggtacat	ggttatggaa	acagctttat	catacagact	129060
tctttcagta	gatgccat	gagaaaaaa	aaaaaacaaa	acattctttg	gattttcaac	129120
atgtaaacga	aaatatgtta	aatcttaaaa	aacctttaaag	cagacgccac	tgctctttgg	129180
ggtcagggaa	gcacggctgc	aggcccccag	agcagcacat	tccctgaggg	aagccttgtg	129240
cgctcctacc	agcaccaga	acagccaact	ataataagct	cataaaaaat	cctgaaaggc	129300
tgcccaagcc	tgaaaatctg	atatgaaac	agagggcaag	acagataaaa	ggcaaaactat	129360
actttaaaag	tcaattccag	ctatctgtgt	ggaaggactt	ctcgggacgg	actatctgct	129420
aagaccttgt	gggtgattta	aacagcgttg	agggcgagaa	aagaaaggaa	agtgcaggaa	129480
acaatggcaa	aagctcctcc	tccgtggcct	ctacacgcac	ccaagcgata	tggaggagg	129540
agggaggac	actatacaca	ggtaccaatc	cccacgaact	agaaaacaca	gcttttacta	129600
actagtctat	tttttttttt	accttagcta	gtgtctttct	tatgtttggc	ataatttctg	129660
ctctgcatta	aataaaccta	gtgtataaga	aggcaaatga	gtaaaggtaa	agtcagttaa	129720
tgttcatttg	ttttgactgt	gtggtgtgca	tgccaggcgc	attcgtggag	gtcagaggac	129780
aaccgcattg	agcaggtcct	ctccttccac	caccccgagg	gtcgcagggg	tcccagggat	129840
ggaactcggg	tggtcagatt	tgcttgccaa	gggcttcact	cactgaggcc	ccccagggt	129900
cccaattcac	ctgtttactc	taatgtatat	atttttgaga	cagggctctg	ctatgtagtc	129960
caggctagct	ttgaactcac	tgtggacact	agctggcctg	gaggtctgag	tgaacctcct	130020
gtttccaccc	aagtgcctgg	acaacagaca	gcaaccatca	aggcaaatg	aagtcttccc	130080
ttcacacca	agtggcttcc	tatgaccttc	tcctgacaac	cccaaacagc	aagtgcccg	130140
atatgcactt	cggtttctct	ttttttactt	caagctttga	ctccactttc	ctaaaagggtg	130200
tttactaggc	tgaagcacac	ttcaggagac	accctgtagc	tctggagtga	ccggaaacac	130260
accagctctg	atgcataaac	gattggggatg	gggcaaaagt	ggcttgtgct	ggcttgtgct	130320
tctgtgcaat	ggcattccct	gacgcattac	ctcacaccag	atacatacag	aaaagacaaa	130380
ggttaagtcc	cactcagtcg	gcgcacccac	cacacccaaa	gcaagcattc	aacagctcag	130440
tggtcataagg	ccagcatctat	cacagcggcg	acaaacacaa	agccaacctt	ctctttgagt	130500
tccccaaggt	acatagtgtt	ctgtccaaag	atggcctctg	cagactcctg	gaaacacgtc	130560
accactgggt	tctcttga	atctgcaata	tttgcttaaa	aaacacaaact	ctattatttc	130620
agcgaataaa	taacaaaagg	aagacacccc	aaggacggag	gaagattatc	tttccctact	130680
taaaagtaact	tccaagggag	aaatgacatt	gccgccacac	acagcccccg	actcttttgt	130740
taacgggtcc	cccttatgag	gaagtattca	cttctgggtc	ccccgtgttt	cacttttagt	130800
acagtatcta	atgaattaca	ggagccaact	aataacctta	tcataaagtg	ggctttgtgt	130860
taggtgatcc	tgcaccaagc	taggctatgc	aaatgctctg	agcacactta	ggggaggctg	130920
ggcgtgtcta	gccctgtcac	atgcaggtaa	tgctacgcc	gtcttacgcc	acagggaagaa	130980
gcttaaaatg	tggtcgaaac	caatctgtcta	gaggtgtgag	tttagacctc	gggagacggg	131040
ccaagtaaac	ttccgcagat	gtcggagcat	cagacagagc	tgtccctccc	tgatgcaaaa	131100
ggcttcgcac	ggcaagtttt	taatgagcct	ccatggtaat	agtggtttcc	tctctctcct	131160
ctcccttat	caatatgctt	gacatctttt	agttttttga	gacagcaa	aacgtagccc	131220
actctgccc	catacttact	gggcagccac	agctggccct	gaactcctcc	tgcgtctcct	131280
ttaccccgta	acaagtgtcg	gggttacaga	catatgccgc	catgttcagt	gtaagtga	131340
gtgcccagc	agggcaaa	cttcccttatt	tacaaagcag	cagccaagcc	ctgcagccca	131400
ggcctatctg	atttccctcag	cacaccccca	agggctctcac	cgatggcagt	cagtcctatga	131460
acaccgtagg	cttctcacaa	tccacactac	tcaccgataa	gatcatgcgg	tatttgaaat	131520
tgggaaatcc	ccggctcacac	ttcttcacagc	ggtacaaccc	attctgtctg	tcaatcactt	131580
tcttattgca	gtcctgggtt	gggcaggcct	ggtacataca	gttctctttg	cggagaaaca	131640
ccaccgtgc	cacagtgtcg	aaatagtccg	cctgcaggcg	aaaggaagac	cgccatcagc	131700
aagcaacaca	ggtctggaac	aggcaactca	aagctgctct	tccttgcaag	tggtgagcgc	131760
gtgtacatcc	tagctccac	gctcacgagc	gatatgcaga	atcactaact	ctggtttcag	131820
aaatcacacg	tgctacacgc	agaacccaaa	gaagtaacaa	accggcactc	acgcacctca	131880
ctttttccat	tttgagagac	gcggctcact	agctagcctt	gaactcagag	tttggtctgc	131940
ctctgtctct	atagagtgc	tggctaccac	acctgggcac	tttgttttcg	agacagggtt	132000
tctccggaac	tcaactctgta	gaccaggctg	gcctcaaa	cagaaatcca	cctgcctctg	132060
cctcccaagc	gctaggatta	aaggcatgtg	ccatcagcgc	ctgacccaat	tctttttatt	132120
tatagttatt	atttgttaca	tatgtgtatc	agtgtgtatg	acgtccat	gtatatgact	132180
gtgtgcagtg	tgcacatgtg	tgcaagtcag	aggacaactc	tcaggagtca	gttctctcct	132240
cctactgtgg	cgtttgggga	actcaggttc	caagatagca	ggaaaagtgc	ctttaacagc	132300
tgagttatct	cgacactgac	atgtaatgat	tcagtgtgca	cacagtgggt	ttgcatgtag	132360
tcatataaac	cacagcgtgc	atgtgcaggc	taggaaacaa	cctgtgggag	ttggtgatct	132420
ccctcacca	tgtgagtgc	ggagaggaac	tcaggctgtc	aggctcgggtg	gcagtgcctt	132480
tactcactga	gttcccttgc	tggccaagca	ttatttataa	gatgggtgatg	tctaccctta	132540
tcttttaggga	tcagtaagtt	tttccccaag	acaagccaga	aaaacctttt	aggcccaatg	132600
agccattcag	tcagttctta	ctgccactcc	tcaacctgtc	tgtggggcag	gacaaagccac	132660
agacaacctg	aagacggaag	gtgagccaat	aaaatcttac	tgacagtaac	agccagccag	132720
ctcacaggct	tgacaggcaa	ttcttggact	gaatgcgttt	aagagaatgc	agaattaccc	132780
atgactaaga	tcttctaaat	ggaaaaatgt	ctggttaagt	aaccagcaa	ggagctaagt	132840

cacgcaagcg	gtggatacct	gctttgcctc	tgaaccctgc	acaggttttg	gtttatgttc	132900
aatcatgtca	agtacctaca	aaatccagat	ctagcctgaa	ttcaaagata	ggctacctac	132960
cctgcccccg	gacccccacc	cggggtctca	ctgtgtagtc	ctggctgtcg	tagcgtcttc	133020
tatacagacc	agctgttata	aaattcagag	accacttgc	ctcccaagtg	ttgggattaa	133080
aggcatttgc	cactatgcct	ggctctcact	ggttgtacca	gaggagcaaa	acaatgtggc	133140
catttaaaga	gacgaagcta	gaaaccagtt	tagacagctt	tggggctgtg	ggtgtagctc	133200
agtgggaagag	cttgcttagc	atgcacaagc	tgttggtttt	aaccctcagc	gtgacagaac	133260
cgaatacata	agaaggctta	gaggaggggg	tgggtgggga	gagatgggtt	agaggttaag	133320
agctgggtgc	ttaatttccc	agtccccaca	tgggtgctca	caacatccat	aactgcagtt	133380
ccaggggatc	tgatgtcttc	ttctgacctc	cttggggaca	tgcgactcat	ttggctcaca	133440
tgaggatgc	ctcggggcac	tatgctggct	caggggtgaa	ggtgcttgc	gccaaagctgg	133500
gtggatttga	gtttggctcc	tgggaccac	aaggaagagt	ttgacttcta	tacactgagg	133560
tggaaatgc	atgctctacc	cgcaaattaa	aaacttaaaa	tttaaagagg	aagctgtaga	133620
gaaatagctt	ttaggaggat	gcctaaggaa	cttctctcgc	ttttcagggt	agattcagac	133680
tcaaagccca	atttaaagt	ttgagtgcgt	tcacgtgttc	tgtatgcca	gttctggctg	133740
tccgttgtct	gtcttttagt	taagagagca	actgggtgag	aagtaactga	gagtctagcc	133800
gatgtttaat	tctcaagatg	tctgtgtata	gcattataag	ttgctgtgga	tgacagtgat	133860
ggatcgagca	cacagtgaag	tgagacagtg	aggaagaaaa	cctatatatt	atagcaggac	133920
aataatggag	cagtgggcag	cgatatattt	ggaatataat	agaaataaat	gattcaata	133980
tccaagagaa	cgaggttag	accaaggtag	gaagaggatc	actgggctgg	agagatggct	134040
cagtggttaa	gagcactgac	tgctcttctg	aaggtcatga	gttcaaatcc	cagcaaccac	134100
gtgtgtggctc	acaaccatcc	gtaataatat	ctgatgccct	cttctggagt	gtctgaagac	134160
ggctacagtg	tacttacata	taataaataa	ataaatcttt	aattaaaaaa	aaaaaaaaaa	134220
aggaagagga	ccatctagca	cacattgaaa	tggaaagatg	aactgaaatg	caataaccga	134280
gccagttgtg	ggcatggcag	caggaaattac	agtctgttca	aaaccagca	cggacctgag	134340
gtacatttat	gtccctctac	actggccctg	actgaaaagc	agcaacgtgg	taccctggag	134400
caggccctgag	gtcccaggtg	aggacacagc	cctgacctgg	cactaggaaa	caggtgtaaa	134460
aaactcaagcc	ctgcgcgttc	tctgggtgtc	tggagcgagg	ggtgcgaggt	acctgtcttc	134520
cctggccccag	gttctcagat	ttagcctcat	gcaaaagtgt	ccagttgggt	ttgccccctc	134580
cgccccctcc	actcctgtgg	tcagagatgg	aaacaccatc	taaggcttgt	ccttctgagt	134640
caaaccctagg	ggagaaagaa	acaaacgtac	tgctgacagg	aacttcacat	ccttctcaaa	134700
tgtgctgtga	atgcaccagc	cccgagctgc	gccccctcgg	ctctcaccta	ggtctcagca	134760
caactagtcc	actcaagaaa	tgcattgccg	ttgctccttc	catctcgtgt	ccttcaatgc	134820
ccctcctcct	ccaccgcgaa	ggtcagacaa	actgccaaaag	taccacagac	cctttcccta	134880
aaactggctc	acattgacag	cagccatgaa	caaaggcagc	aacagcaagc	cacactggca	134940
ccctccccc	caccctcggc	ctcaagcctt	caccccaatc	accaaaatga	accagatgaa	135000
aaggagacatt	tttgtttcat	gagtctgcca	aaaaatgttt	catggggagaa	aatcttttaa	135060
gcccaatcag	ttatagaact	caatcagaat	gcattatctg	ctgtgttaga	gatttgcatt	135120
ctcagtgaga	atttgtttat	ggaaatatca	agggttaaat	catttatatg	aaattattat	135180
taaaattgac	tatacttctt	gacctctat	tggctaatct	gaaggaaaag	aggccaggga	135240
gaaggtagca	aggacaggct	agtggcagag	gacagtgagc	ctgagatgaa	aacacctcaa	135300
ctacacgatt	ggtggtaact	aaagctgccc	ccacacagga	ctgactccct	aaggagactc	135360
ccaacagagt	agcgggtggc	caaggcaagc	agtcacatcc	gtcagaggac	aaggtccata	135420
aatccgtctc	accactacta	gcgcgtcaag	tttaacttca	cacaggctac	ggcaaccaaa	135480
gacggctcat	caggagtga	cagagccagg	agacaggcag	cctctcttat	agtaaaaggt	135540
ccactcttag	agccaagatc	taaagatcgc	taatccttgc	ctgggtaggg	gggtataact	135600
gaacttgtag	atcaatgaca	gtattttgtt	cgggggacag	acatgcctgt	ttgtgatttc	135660
tgtgatcaca	gttactgaca	gtgaaatatg	aacaccttaa	tgtagacaca	caattaacct	135720
ggggttggtc	tcagagtctg	ggagaaaaac	caaccaacaa	accagaccaa	tgccatggct	135780
gtctctggct	agacttctgt	cctacaatgc	aacgctagct	atcagcacag	ttacacagcc	135840
agaccacac	cactgaggat	gccttaagtg	acagacaccc	cagatggatg	ctctaagtag	135900
taaatattgt	tattggttta	catcagactc	aaaacaaaaa	tgaggagcgt	gaaatatgag	135960
ctgtttttct	tgagtctctt	aaattctacg	acacagctgg	aaaccacaca	tgcccccttc	136020
ggtactacag	ctattcaatt	tccttaagct	tgggctagg	aaagtatttt	ccttgaccac	136080
ctcaatcctt	ctacacaaca	ccctcaggat	gaagtgtctt	tccaaggcag	agtttaatgc	136140
cttaaaaaag	tagatatggt	gtggaggttc	taacaagctc	tcctcatctc	ctgagtgcct	136200
tgggaatgac	aatgtacagg	caacaaagag	tccattgttc	tcactgtact	ttaccgagtc	136260
cagtgacaac	acaacagaag	atgtgaccag	ctccagagac	tcagcagagt	aagaaagtag	136320
attctacccc	catcatcagc	gcccacagtg	caaactgaag	atgctctcag	cactgttctt	136380
cagcgcgagc	ctgttaaaag	cttgccgtac	ctggcgggat	caaggtcagc	acttagattc	136440
taattgcttt	ctctggctct	ctatgttcat	caggggatcg	tgggtttgtt	ggtttaagag	136500
ggctttcagg	ccaccatctc	agaagacata	atgcctcgaa	gaagtatagg	aaccactcca	136560
ctcaatagca	aaggcttggc	aaaacagctt	ggcagcagtc	cacatgctga	cagtgtccag	136620
ctctggttta	ccagagcccc	agcacagata	acccctgagc	tgagttgcac	aaaacctacc	136680

agccacgaag	cttatagggc	tctgggatgt	caggattcac	aatgacagt	ctggatgaga	136740
ggaccgagag	gctccgtcca	ccgaagtcag	agactcgggc	tcctttgatg	gccatcacgg	136800
gctgccgaga	gccgtcaaac	ttgtcagcct	gcaaatcaaa	cgcacgcaca	catcactgac	136860
aagaagaatg	tacaggatgg	ttttatggag	aagagtcagt	caatgtctgt	ggactctaag	136920
acagacctcc	cactcgggaa	ggagcattgc	actacaagaa	gctgcaataa	ccgatcatct	136980
cacacagcga	aggctctcaa	atacttactg	gatagcacia	ctccctgagc	taaagcccca	137040
ccctcaggac	tcgggtcagt	gcaaggactc	acatcttctc	cccacagagt	tgtggtcacc	137100
accttccctg	acatgtccat	caaatagata	tttctcttag	caacttctct	gttggtcgac	137160
ttcactgtga	ttttaatcga	atcttcatag	ctcttgacga	ttccaatgat	gtctgaaaca	137220
aagaacactt	tgtaagagct	cccatcaagg	ctcctcttta	aaacacggca	gggacgaaag	137280
gcaagcacgc	tagtcacgca	tcaccccaaa	cactggagac	aaaaatcacc	actctttgcc	137340
ctcaaccctg	gacgcacctc	ctagtgcgtc	tttagccttg	ctctctaggt	caccgatccc	137400
tgtgaaatca	aactgaactg	tgggttaagt	atggccatct	tcacagggaa	ggacagaagt	137460
ctcattattg	aaggctcatct	catagtcatt	tttaacagcg	gagaactggt	tgtagcgat	137520
cttcaggggc	ccctttgaga	agtaatacac	ctgcaaaaga	ggccaagtca	gggcaagcta	137580
ctcagtcctc	caaagcccg	cccaaatgc	agccttccaa	actgtgtcta	tggtctctac	137640
agttcagcct	gtcagctctg	actcaacacc	gtcttcttac	ctgtgtcact	tcaataaggg	137700
gaaagaactt	gtccacttgc	tcattgaaag	cagtagctct	gatttcaccc	tcgagaagca	137760
agggagagca	cattagaact	gcgtgtctag	gccctcgctc	tcaacagcga	gagcaggcca	137820
ttttgacagt	tagacaccat	cctcggagca	acagtccaag	tgtaaaagaga	ccctcacagg	137880
gctgcagaca	ttgcagaatc	ccacatatat	agttacattt	tcaagctatc	agctctggga	137940
aaacaaaagc	agctcagcca	cccgagccta	cagttgttaag	ttaacacata	aaacgagggg	138000
gcctcgagat	gatctggaag	acaaaggtgt	ttcctgtgcc	aacctggcca	ccctagttta	138060
atcccggaac	cctcgggatg	gagaagacaa	tcgccccacc	aagttattct	ccagcctttg	138120
cacacgtgtg	catgcataca	cacgaaaata	attcttacat	cttatcaaaa	cattttttag	138180
aggaatagct	taatatccg	ataatgaaac	ctaattattct	gggtcccag	tgatgggctg	138240
gaagcccaca	ggaaggcgtg	ctgaactcca	atgattcagc	actcttcctc	acgcccgcaa	138300
cctaagtgtg	ggcttacttc	acagtcagga	gctgggtgcca	aaaatgttcc	aacaagtcat	138360
gctttcaagc	tgactgcact	gttacttttc	tgtaggcaa	tatgttacag	atggataaga	138420
aaacttaacta	aatggctaaa	accttaccca	gacaagctgc	aggaaattag	tttctgatac	138480
tgaactgtcc	atcacagtta	agatatctgc	tgtgttacac	cgtgacaaga	ggaagaagag	138540
ggggagagag	ccccaaagct	ttgcccttcc	cggtatcatg	gcttatttta	cgtgtgtggg	138600
tgttttacca	gcatgtatgt	gtgagcatca	ctgagtgtct	gatgcctgag	gtgaccagga	138660
gaggtactgg	atcactgaca	cgggagtcac	agatggttat	gtgctactct	gttggtgctg	138720
agaattgagc	ccaagtcttc	tcaaagaaca	gcgagtgttc	ttaaatgctg	agcagtctct	138780
ctggcccttg	cattttgtct	taagtaaagc	ctagttagta	taaatgatca	gaagccctcc	138840
ccgcagacta	gttaaagctg	agtagctgct	gtccttctct	ctggaggcaa	acccgccctg	138900
ctcggcaggt	atcaccaccg	ggcttcaact	gtgcccttag	caatgtaatt	agagcgctgc	138960
agtetctgca	gacggagact	atttacagtc	caccaatctg	tattgtctaca	gacgcgtccc	139020
tttagggaga	cttttatctc	tgtctgaccc	tgtgccccag	atagcatcaa	aggtctccaa	139080
cagaaggaaa	gccacaatga	ccgccatgtt	cctcggagct	gagccccacc	cgnnnnnnnn	139140
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	139200
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	139260
atctctatat	atgaatgagt	atcctgtagt	tatcttcaga	cacaccagaa	gagggcatca	139320
aatcccatct	acaggtgggt	gtgagccatc	atgtaattgc	tggaatttaa	actcaggacc	139380
tctagaagag	caatcagtgc	tcttaacccc	tgagccccct	cccttttttt	tttttttttt	139440
ggcagcttat	ttttaattgt	tttaaatcac	gtatatgttt	ctgtatgtag	ctatgttcac	139500
ataagcgag	gcactcatag	aggctagaga	tgtaagatgc	ctttggagct	ggacttaaac	139560
atggttgatg	gccatctgat	atgagcacca	agtggaactg	ggtcctctgg	accgccttgc	139620
tgccctccac	ctagaattac	agtatatgca	attagctgct	cagccacccc	ctcagtcctc	139680
cttagtactt	tcaacatagt	gtaaatcgga	gagcaactct	tggtgcaaa	attctgtgtg	139740
ttctatcatt	ctctacaccc	aggcttcatt	gggacgtgtg	atgccagcga	tgattcattg	139800
gcgtgaagta	tgagtttagg	tcagacaatc	agtgctctct	gctctgagtg	gggtctccctg	139860
ccctgtgtgc	tctgagacag	gggaggatcc	tcctgttgga	tcacatagtg	catagagccc	139920
tggtgtgcac	gcgctctcaa	ccattttgtg	ccatttcaaa	tacagtgaat	tcagagctct	139980
ccttgaaata	acagatttct	ccattttgtt	tgttcccttc	tttcgtcaca	aagacctgag	140040
gatgagatgt	ttttgaagga	actttctagt	agttactcgg	tggaaaagga	caatgatgct	140100
ccctcttctc	acagagaaga	aggaaaacag	aaattccaag	aaaaggagta	cacagatgct	140160
gcagtgtgtg	actctaaggt	aacgtgcgtc	caaccagcag	ttgaacaagg	cggaagatag	140220
aggcaaatgt	cagactcata	accagtccat	tgttttgctt	tggttagttg	gggttggttg	140280
tttgtttgtt	tgtttgttct	ttcttcttta	ttagtccgat	tggatttggg	aaaaatgtag	140340
tagaattcta	ttgttgatta	tattcacaaa	caaaagactg	tttatatgtc	ctgaattttg	140400
gcaagattct	tcccatattc	cttcagaccc	atagcagcag	caagcttagt	ggcccttttc	140460
caggtattct	cactgagctg	tacagcattt	gtctgaagaa	ctgatgaaca	ttttagctta	140520

tgctcctatg	ttagtaaaata	gctaggtttt	tagctagcca	ccttgctagc	ttacaactag	140580
aatgcctctg	gctgagtttg	gtggcacttg	cctttagttc	cagcactcag	gaggcagaag	140640
caggtgaatc	tctgagttca	aggcagcctg	gtctacataa	caagttccag	cctggccagg	140700
gatattgagt	aagactctgt	ctcaataaag	taaacagcca	gaatgactta	aatattgcta	140760
aaaaaagaag	aaagaagaag	agaaaagaag	aaagaagaag	agaaaagaag	aaagaagaag	140820
agaaagaaga	aaatgagtgg	atgtgccata	gcctagtcaa	actagatttt	cttggttata	140880
tgtaacata	ctactattaa	caacaacaaa	tactttacac	taaccatgac	aagctatat	140940
tttaaattat	tattttatat	gtatggtgtg	ggaagctgtt	gcagagtggc	agttggctac	141000
tgctggccac	cacacataca	taggcagtga	aggttctttt	gccaagacaa	gttaaccaat	141060
cagatgtgag	acacgcctct	cctaggcccta	tgtaagcagc	accagttctg	ggctcagggt	141120
ctcttcgcct	ctacaatcaa	gctctcccaa	taaacgtgtg	cagaaggatc	ctggtgcagc	141180
gtcgtttctt	ctggccagtt	gagcgcgcac	aagagtatgg	gttttcacct	acatttatgt	141240
gaactgcgat	tgcaactggg	acccagggag	cccagaagag	ggcatcatat	cccctggaac	141300
tagagtacac	ggttatgtat	gggttctagg	aatcaaaccc	atgtcctctg	gaagaacagc	141360
cagtgttctg	gatttaactg	ctgagtcac	tctccagccc	caccatgatg	gatttaactg	141420
ctgagtcac	tctccagccc	caccaagctg	gatttaactg	ctgagtcac	tctccagccc	141480
caccacgctg	gatttaactg	ctgagtcac	tctccagccc	caccacgctg	gatttaactg	141540
ctgagtcac	tctccagccc	caccacgctg	gatttggaag	cagaactgag	gctttgaaca	141600
ctgtgttatt	ctaactcttg	cttcccttga	accctgagaa	aattccttct	attggccttt	141660
cagagatctg	tatgggctta	aaacaagaat	atgtccaact	cttgatctct	gattttatat	141720
attaaaagaa	tagagtgcag	ctaggagttt	ggcacactcc	tatgttccaa	gcttcaggaa	141780
gggaagaagg	tcaggagttg	aaggctagtc	agtcttagct	ttgtgacaag	tttgaggcta	141840
acgtgagctc	tatgagacct	tgcttaggag	aggagaaaga	aaggggtaaa	gtgaggtgtg	141900
gtccacgct	accacttgcc	taacacgcac	gacgcccaag	gtttccgttc	ccagccctgg	141960
gggtcagagt	attgtgcaca	ccctgtatct	ttaaaagtca	atagtcaacc	cttggaaata	142020
actttccgaa	aaatattaaa	gctacatcat	ccactctaca	aacttgtaaa	agcccttctt	142080
tgtaatggtg	taagaagata	tttggtctca	gttttatggc	ctttgcatcc	taaatgttac	142140
cccatggaaa	gttgcttaat	gaaaacaggt	aaaaataaga	caggagggac	tgcaaaagat	142200
ctcaggctca	cagagtctta	ctgcccttgc	agagacctg	gcaggcgatg	ggctctcccta	142260
ctgcagggaa	tccaaaaccc	tcttttggcc	tctgcaggca	accacattaa	cacatgcaca	142320
catacacata	attttaaaaa	taaaataaat	cttttaaaat	gagctctaga	acgagtttga	142380
catcagctca	ggctacacaa	gaccttgtct	caagaagaaa	gaaatgaagg	tttgctgtgg	142440
atggagaggg	agatgcactt	cctattccat	gaagctgcta	ttttggtgat	tatggtactt	142500
tgcaatttta	tagagagctg	ctattttctt	tcttttaaa	taatgcttgt	tgttttgact	142560
gtaaaagtaa	tgaagtaccg	tttggaatat	atagagaagt	ataaagagta	aaaaaaaag	142620
tcataaccag	tgatgtacct	tttaacattt	tgcttctatc	cggccagcca	gagtttttcc	142680
tgtagaagt	tggttttctt	ttacaaaatt	gggataatgc	tgcaacttact	gtttttgtagc	142740
ccactcttct	ccttccagat	ttattgtacc	cattttctca	cagtattaaa	ttttcagctc	142800
caagtaattt	tccatagcta	actctgtgtt	ctgttacaga	gaagaaatgt	acttaattta	142860
agatcttaata	ttggcataat	atttggcatt	atgatgctat	aataagcatc	tttctgtata	142920
aatattttta	tgtagcagca	gtgttttgtt	tctcaggaaa	ggactagtta	acatccattc	142980
gggtcaagcc	tctgagtgc	gctgaattgt	gcaacactgg	ttattattac	atactattat	143040
tgaaagaatg	tggaatgct	catttcttaa	ataggaaata	tgctcatata	gttcaatttc	143100
tttggttagta	ttatacattt	atttaggggg	gtgtggtttc	ccctagttaa	gtctcacttc	143160
aagagttgca	aaaatatatg	gtgtggtttc	ttttcctgtc	gcataatgtg	gtgaataaaa	143220
cagccccata	atctaccttg	ttaaatatat	tgatatatag	aactgctctt	gcagagaaca	143280
aacagccctg	tatggtggcg	ttaacctcgg	gttacttgag	agggatgcag	tgccctctggc	143340
tgactttggc	tcgaaagcc	ttcctggaat	tccagctcca	ggttttaate	ttaggaaact	143400
atccttgggg	acgtcactca	tgtagcacaca	tacacatttg	atgcaaaggt	ctggatgaca	143460
caagtgggcc	aatgagatgg	ctccatgtat	aaaggcagtt	ccaacccctg	cgagttgtcc	143520
tgagttcagt	cttcagattc	tgcaagataa	caggagagga	cacatgcaca	gaagtccag	143580
tctgacctca	gtacacatgt	catggtacgt	gtgtagtatg	ccagcctggt	ctacaaaacg	143640
cactcgggag	gcagagggcag	gaggtatctt	gagttttagg	acaggggttg	gggatttagc	143700
atttacagtt	atataaagaa	actctgtttt	gaaaaacaaa	ggtcagtcct	caactctgga	143760
tcagtggtag	agcgttggcc	tagcaagcgc	aaggccctgg	ggaggagagg	gaaagggaag	143820
agagagagag	agagagggag	agggaaggag	agaggagagg	aagaaagaaa	gggaaagaaa	143880
aaggaaaggaa	ggaagggaag	aaggaaaggaa	gaaagaaaga	aagggggaag	aagaaagaaa	143940
gaaagaaaga	aagagagaaa	gaaagaaagg	gaaagaaaga	tatatataat	tttaagactt	144000
gaaagacaaa	gcaaagcaaa	actaaataaa	atacatataa	taaaactctt	ggcttctaaa	144060
gagggcatag	atcagtgtaa	taatgtctta	ctagcatgcg	attatatttc	tagttgtcag	144120
cctgacaccc	taccgcaaaa	tatttgcctc	gtcttgctaa	actatcccac	tcctctgtca	144180
actactgtgt	actttcacta	gcaacataat	gagaatgttc	tggttgaaaca	agtatttcag	144240
agtaactctgt	tcttggtttt	atttttcttt	gccaaattga	gtttcttate	cttttatgat	144300
ctagcctaga	acacacagag	aactctcttt	atgaactcaa			144360

ctccagaggt	ttgtttttgt	gggtttatta	gtgttttttg	tttgggttgg	tggtttgttg	144420
tttgggttgg	tggtttgggt	tgctttactt	tttctttatt	atttttttat	ttctttatct	144480
ttttgttttt	aatttaatgg	actggttcca	tgtggccaag	gatagcctca	acttttgtagc	144540
agaaactggc	tttgaacttc	tggtcttccct	tcatctacct	cccaagtgat	gggattaagg	144600
cacgtgccac	cacatctaac	aatatctggg	tttctttatt	ggagtgtgaa	agggattcct	144660
ccagcattac	tttgactctt	catagtttct	tccagagtta	ttacactttc	atttgtttaca	144720
ttaaaggttt	gatccagggc	tgagagatg	gctcagtggc	taagagcacc	aactgtctct	144780
ccagaggtcc	tgagttcaat	tcccagcaac	cacatgggtg	ctcacaatcg	tctgtaatgg	144840
gatctgatgc	ccctctctgg	tggtctgtaa	gacagctaca	gtgtaatcat	ataaataaaa	144900
taaataattc	tttaaaaaaa	aagagtttga	tccattttaca	ctggacttct	tgaggcagca	144960
ggatcataaa	ttcaagggtga	gcctgggtga	actggcagaa	gtggcagaag	ctgtgtctca	145020
cactgtctga	ttcattttct	cattgatctt	cagaggtttg	ctaacgggaa	gtaagtggaa	145080
cagaaggttc	agtattcttt	ttttcccaat	tctattcagt	ctttagtagt	agatccctca	145140
ttatctgaga	tgccagagtcc	cttttatctc	tgtgaccatc	tcgtgtgttt	tcagggagtg	145200
ctctcattcaa	ggcctaacac	tgaggacatt	tcactgtgct	atgccaatcg	ctctgcagcg	145260
tctttccatc	tggttcagta	tgaagtgaat	attgaagaac	ctggtgtcct	gcctgtggct	145320
gcagtggaaa	atgagctcct	ctctgttctt	ctgcacacat	tgaatcaaac	tagcttgcaa	145380
acactgacat	ccaccagac	ccattctctc	ttctgactca	tgccacctct	cataggtgac	145440
cacaacaact	atgtagttag	caagaagtag	ctatgtcatt	gtccacagtg	catggatttg	145500
ttccaatagg	ccagcacttc	tgtgtccata	tcagctagat	gtgctgtgta	tagtatttta	145560
gattccaaaa	tgtgtccaga	tattacctcc	ttcgtttgtt	tcttctaaat	aagccaggca	145620
caaagacttg	aaagatggct	tgatggctct	tccagaggct	gggtttgatt	cccagcccca	145680
acatagcagc	tcacaatagg	ctataacgtc	atttccaagg	ggtctgactt	cctgttctgg	145740
cctctacagg	cagaagagcc	agacatacat	gcaggcaaaa	cacataaaca	taattgaaag	145800
aagatatttaa	ataatagccc	acgcttgagc	ttattcctct	gatgatacag	ctctcctgaa	145860
gtcatcatgg	gcagtgtaaa	agtaaagggt	ccccgccctg	cccaggggca	tcagtggagt	145920
agctgactgt	gagtggtctt	cttctcatcc	cctaactgct	gcaacatcat	caaactgtgga	145980
gcattatata	cgtggatttt	aagttaggaa	atgacaaaag	ttagatctat	ggccaggcac	146040
tagtgacgcg	ctttaatccc	agcactcagg	aagcagaagt	aggtggatct	ctgtcagttt	146100
gagtttacag	agagtgtcta	ggcagccagg	gctatataga	gaaattctat	ctggaaagaa	146160
taaacaaatc	agatctgtat	ttcaggaaga	tgccatagag	ctgtttgaca	tgtgtgatag	146220
aggtccttaa	ggacaggaaa	gtattccaca	tgtgtgcatt	ttacagaaat	tggtttataga	146280
ctggagttaa	ggactgtttg	aatagttgaa	tgttcacact	caagtgttct	caaatacaat	146340
agaagaacaa	gaatctatct	gggcatggtg	atgcacaact	gtattcctaa	catgtagaag	146400
actgaggcat	gctatgtgtt	tgtggctaac	ctgggctaca	tagtcaatat	tggaacgtca	146460
gagctgctac	taaaaacctaa	aaaacaaaaa	tctaataatc	tggtatttta	tattttcctt	146520
tttttaaaaa	aagagtgggc	agaatgtctc	tgattttgtt	cagatggcca	cagaacctag	146580
aaaaactgct	ctgctgctg	ctgctgcata	gcacacagct	aatatttgac	tatacgtata	146640
taaattttgt	tgtatacttt	agctgtgctg	tcaacttttg	aaaaaaagta	ttccagttta	146700
tcatttttaa	ttggcactgt	acagaaatta	acagccatat	tagtctagac	acattaaact	146760
tcatttttcc	atttatacag	aagcaatgta	ctgtatttaa	tattcagctc	tatctacagg	146820
ggtttgatta	cagaaactat	caaagtattc	tctaataatg	gaaaaaagat	tcaagaatct	146880
gactgtagat	ccaaaggaca	agtggagaaa	aacttaggaa	gaattttccc	tttatccctc	146940
ccctatatgt	atcatctctt	ttacttctaa	taatagtggc	cattttattga	acatacccag	147000
gagttccttt	catcacttta	gatataataa	ttatctctac	ctaaaaatgac	ctgttgatga	147060
gtcatctctc	tttcagatga	gaaacaaaga	cattgaaaaa	tctaacttgc	ctgcataaga	147120
tcacacctag	cctcttactc	actccatgaa	tattcctttt	ttttttttcc	tttgagacag	147180
aatctcacta	tgtagtcttg	gttgccttag	aactcaatat	atagaccagg	ctagcctcaa	147240
actcacagag	atctgatagc	ctctgacctg	cgagtgcctg	ggttaaatgg	atgtgtcacc	147300
aagcccagca	aaattttacct	tcttaatttt	ctaagacgtt	tctcctctta	aaaaatggaa	147360
ctattgagcc	agtcatgggtg	agacaggctt	aatctttaat	ctcagcactt	aggaggcaaa	147420
gacaggccta	tggttctggg	gcagcctgat	ctatagagag	agttctatgg	gttaaagttt	147480
aggggttaag	ttttgagaca	aaactttgtg	tcaaaaacaa	acaaaacaa	ccagactgct	147540
taataagaca	aatcagacat	aatattataa	acaagtatta	gtgtcactca	attaaaaagt	147600
cactcaggag	gtgagatcaa	tcccagcat	aatgagggga	ggaggggaga	gaagggaatg	147660
aatgggaggg	gaggaggaag	gaagagagaa	aaaggaaaga	tctcaaagca	gaacacagga	147720
tgtaatttaa	ggcctaagct	ctcgactgaa	gttgtccact	tttaatgacc	cttttcatgc	147780
tcattgtttt	ctgtctctgg	tacatagtga	agagtggaaa	ccaaggggtca	tcctggaatt	147840
tccttttgtt	ttcaggcatg	tcttaaagac	atagtggaa	cagggtatgca	tggttatcct	147900
gaaagactgc	agcccaagat	gatgggtgct	aagacagaat	gcctggtgaa	cctggggaga	147960
ctccaggagg	caagacagac	catcagtgat	ctcgaaagca	gcctcactgc	caagcccaac	148020
ctggtgcttt	cctcttacca	gattctgcaa	aggaatgtcc	agcatctgaa	aataaagatc	148080
caagaaaagg	agactctccc	agaacccatc	cctgcagctc	tcaccaatgc	cttcgaggat	148140
atagccctgg	gggaagagaa	cacacagatt	tctggggcct	ccctctctgt	cagcttatgc	148200

acacacccctt	tgaaggccg	ccatctagtt	gccacaaaag	acattctccc	aggagaactg	148260
ctggtgaagg	aagatgcttt	tgtaagtgtc	cttatcccag	gagaaatgcc	acgacctcat	148320
cattgccttg	agaacaagtg	ggataccaga	gttaccagtg	gagacctcta	ctgtcaccga	148380
tgcttgaagc	acactttggc	cacagtacct	tgtggcagct	gcagctatgc	caagtattgc	148440
agccagggaat	gtatgcagca	ggcatgggac	ctctaccata	gcacagagtg	ttctcttggg	148500
gggctgtctc	tcacactcgg	ggctctctgc	catgttgccc	tgagaatgac	tcttttagcc	148560
agatttgaag	atgttgatag	agttgtaagg	atgctttgtg	acgaggttgg	tagcacagac	148620
acctgtttac	ctgaaagcaa	gaatctggtc	aaggcatttg	attacacaag	tcagggaag	148680
agtgaagaga	agagcaagat	aggtgaaccc	ccaattcctg	gatgcaatgt	caatggaaag	148740
tatggaagta	attataatgc	tatcttcagc	cttttgcccc	atactgaaaa	gcatagcccc	148800
gaacacagat	tcactgtgac	catcagtgtc	tccgcactgt	gcagacaact	caaagctgac	148860
agcgtgcagg	cccaaacctt	aaagtccctt	aagctgaaag	cagtgaaccc	agggctgtgt	148920
gcagatttga	ctgtttgggg	agcagccatg	ctgcgacaca	tgctacagct	gcagtgtaat	148980
gcccaggcaa	taacatccat	atgtcacaca	ggtaagtcag	aaatggtttt	tacttacatt	149040
attggtattt	caagagctaa	tgtttaagga	gaaaaaacact	ataaaggaaag	cctggcatca	149100
aataaatcag	tgacctaata	ggaaacacaca	gccgtcttat	atatcattat	gctattgaga	149160
agctttgagc	acattttctgt	gaacccagag	cttgggaggt	ggagatagga	tgattaggag	149220
tcctaagacag	cttttagctat	acagcacggt	tgaggtcagc	ctgaactaca	tgagaacttg	149280
tctcttaaaaa	acttgagcca	gagccaggtg	gtggtggcac	atgcatttaa	ttctagtact	149340
caagaggcaa	agtcaggcag	atccctgaat	ccagcctcgt	ctatatagtg	agatccccc	149400
caggctacat	agtaagatcc	tgtttcaaat	aaataaatat	aacaaaaaca	gcaataataa	149460
caatagcaac	aaattaat	tttagatgta	tttatttatt	ttatgtatga	gtacaccatt	149520
gcttttttca	gacacaccag	aagagggcat	tggtatccat	tacagatggt	tgtgagccac	149580
catgtatgtg	gttgcctggg	attgaaactca	acacctctgg	aagagcagtc	ggtgctctta	149640
gcacatgagc	gtctctccca	gtccattaat	taaaaaattta	aaactagagt	atttttaaac	149700
atttattcat	tttgtgtgtg	gtatacatat	tataatacag	gttcataagt	caattctctt	149760
ctaccatgtg	tgtcttggag	atcaaaactca	ggttcttagg	catgggagca	agtatttact	149820
tcctgaacca	tctccctagc	catttctagt	attcttttct	tttgtcttga	aagatttatt	149880
tatttatatgt	aagtacactg	tagctgcctt	caaataccgg	aaaggggaat	caggtcttgt	149940
tgagatgat	tgtgagtcac	catgtgggtg	ctgggatttg	aaactcaggcc	ctccagaaga	150000
gcagtcagtg	ctcttaactg	ctgagctatc	tccagcccca	tttttagtat	tcttattaag	150060
tggtttccat	tttatccaaa	gatgccttta	agggcctggg	aagatggctc	agtgggcatt	150120
gaacttggtg	tgtgagcatg	aagaccagag	ttcagatccc	tagcaccag	gcagatgctg	150180
aatgatgggt	gcctgcctga	gattccagga	caacggagac	agacaggggc	cctagctaac	150240
catactacac	actagctgag	ctgtgtgctc	aagagagcag	ccctggctta	ctgtgcagga	150300
tgagagtgta	tcactccac	aggcaagcac	acacctgagc	acacagacat	gcacaaaagga	150360
aagaaaagtg	cttttaaggt	ggtggtggtg	ttgttgtttg	ggggtctttt	gttttgtttt	150420
tttctccccc	tccctcattg	tgatggcaca	ttcctttaat	ctcacatctg	ggacaaaagag	150480
gccggaggat	ctttgtgaac	tggaaggtcag	cctgttctac	atagcaagcc	catttcagcc	150540
aggacgacat	agataatacc	tgtctcaaac	agacaaaaat	tatttatttt	atatatttga	150600
atggttttgc	tgcatgtatg	tctatgcaca	ttatgtctgg	tgcccatgaa	agccagaaga	150660
gggcatcaga	tctctcagaa	ctggaatttc	agacacttat	caagtaactgc	ctgagtgcta	150720
ggaatcaaac	caaggtcttc	tggaagagca	gcaagtagtc	tttttttttt	ttaatatttt	150780
tttattacat	attttcctca	attacatttc	caatgctatc	ccaaaagtcc	ccataacct	150840
ccccccccc	ccccgagcag	caagtattct	tcattgctgg	gccatctccc	catctccttt	150900
tctagttaat	taagctgaaa	gggagggagg	tagatgttgc	ccaaacttag	gattttattga	150960
cagattaata	ctctgttagc	ctaactacac	tatagaagct	tattcttttag	actttcacat	151020
tacactgtcc	agattttgcc	atcctttttg	ngtgtatatg	tctacagatc	ttaattcagc	151080
tgccaattta	tacagtgttt	ataggtattc	tttgtgacgt	ggatctttta	cccatcttaa	151140
agcagtagga	tttgaaagct	gacatttatg	tggcctatgg	tccgtttaa	tcacatttca	151200
agttagtctc	tgtggtacac	attttggggt	ctatctgcgg	ttccgcatct	cacacttttc	151260
cctctcaggg	tgtccagaag	ctgctgcaca	ctgggctgga	aggatgaagt	ggagtccaga	151320
gtgagtggaa	ttctgcagca	tcccggtcca	gctgggagtg	aatgctgggg	tcaggaggag	151380
atgggtgaga	gggccttctc	caagggcctt	cttagtggtta	cagctctagg	caaaggcctt	151440
ctctgacaat	cttagcctgt	gcatagtttt	ttattcgaga	tgagcttgta	tgcatacact	151500
ttattggcag	taaatcagag	gttatccact	cttagggaag	gagataggaa	tacccaaggt	151560
ggacagaggt	cattggctga	aggataacgt	actgagatgc	tcattagcac	gggaggcat	151620
cccagggaatc	tcagggtgctt	gctgactggg	tttcttaggg	ggttgagagg	gtagcagtg	151680
tttcaaccaag	gttatgtatg	gcagagggta	taggggttct	aggtcccca	gacaaagaag	151740
gagaaggaga	agccctgctg	ataaggggaag	tccccatttt	tgagccactt	cagaaggcta	151800
tcaagacatg	gatagacttt	gtccttaatt	agcagggccc	aacaagtgtc	tgttttcttt	151860
tctgccttca	ttggctcttt	gagccactgc	tacaaaatat	ccaaatctgg	gccaggaagc	151920
tggtctcggt	agtaaaaggtg	tttgctctta	agcctgaagg	cctgagtttt	cacttgattt	151980
ggtttggttt	ttgttttttt	gagacaaggt	ttctcagcat	agccctgggt	attctggaac	152040

tcactctgta	gatcaggctc	aacttgaatt	cagagatctg	cctgcttcta	catcccagat	152100
gcttagatta	aagttgtgcg	ccaacactgc	ccacctaaaa	aaaatatgag	gggctggtga	152160
gatggctcag	tgggtaagag	cacccgactg	ctcttccgaa	ggtcggaagt	tcaaatccca	152220
gcaaccacat	gggtggctcac	aaccacctgt	gatgagatct	gatgccctct	tctggtgcat	152280
ctgaagacag	ctacgggtgt	acttacatat	aataataaat	aaatctttaa	aaaaaaaaaa	152340
aagaacacta	ttcacaaagt	acaacaacag	tgtactagga	aacttttaaa	tagcccatca	152400
ttcattctag	gggaaaaatc	tttttaaact	ttagtgtgta	agaaagagag	aggggctgta	152460
gaaaggccac	agcacgtgta	tccagggttag	gagtcaactt	ttcagaagcg	gagtcctccc	152520
ttctacctgt	ttttgaggca	gtctcttggt	tctgccttac	actttgtaca	tgaacttcaa	152580
gatggttggt	ctatttctgc	ctctcatggt	gccctatgca	tgctgagctt	actgatgcca	152640
gccaccacat	aagcatggca	ccagcactga	gccaagggca	ggcatctggc	tcacatggag	152700
agttaccat	caagccatct	tgctagcccc	agaaaatgta	tttttgacag	gtgtggtggt	152760
gcacatattt	aatcccagca	ctcaggaggc	agaggcaggc	agatctctgt	gtctgaagcc	152820
agcccagttt	acaaatcaag	tcccagaata	gccaaggcta	catagagaaa	ccctgttttg	152880
aaaaacaaac	atccttctgt	gttccagtca	caatgactgc	tgtacaata	atatgaggat	152940
ttgggctgt	caaaaatcac	aagtagcaag	gtgtaatggg	caggccttta	gtctcagcac	153000
ttgggaggca	gaggcaggag	gatctctgtg	agtttagcac	agccagggct	gttacacaga	153060
gaaacctgt	ctcaaaaaaa	ccaagcaaaa	atagaattac	aagttaacca	gggtattggt	153120
gttaatatga	aatagcagaa	ctcaagatag	ctaataaag	aaaggattct	attttataaa	153180
ggagacattc	catactgaaa	tatatgcaga	gcctgatgct	tgccatgcaa	ctgaactaca	153240
cnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	153300
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	153360
aaccgtcagg	ctcaggaggca	ggatggcagt	gatgtgtttt	ggggcacaa	gcccccttcc	153420
attaaagcac	caaactctgt	actaaaaaca	gccctctgct	ctcgatctgc	agagatacga	153480
gagaggcaca	catagcttcc	gggcccttgc	ccctctgctc	cccggtcagt	tcctggactc	153540
ctgggagaag	cttgaagctc	aagaactggc	cgctgttgcc	tttgtgctga	cgccagaacc	153600
agcagctatt	cagcagctgc	ggctcttggg	caagcttgga	agtagctcgt	tccttctcct	153660
tctgcaagg	aaagcaagg	agatggcact	gagctgggta	gccaggatgt	gggaagtaaa	153720
ttgtgtctat	gtgggtgagg	gaagtgccgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	153780
gtgtgtgtgt	gtgtgacaga	gacagaaagg	gagagtgcac	gttcttctgt	gttctctgtg	153840
gtcagtgctc	cttttggtct	ctcctgccaa	aaagcatgct	gtttctcag	agtcaagctt	153900
ggccaggctc	gcccacccca	ccaaagccca	ggatctgcca	ccccattaag	aatgaggggg	153960
taggaaataa	aaaacagggt	tcttgtccaa	gagaattttt	attattattt	tttctctctt	154020
aatggattc	agcatttctt	actcctcagt	atcctctctg	gtagggaatt	cagggtctgt	154080
tatgcagagc	acaagagact	gtgcttgcca	gaaacccctg	ggccaacagc	cctattccct	154140
gactggggtt	gcctgcaggc	tgcttctctg	gctgacccct	gagtctggcc	ctctgacctc	154200
tgcccgtcct	ggggctatcc	aggaggagca	ggaccactgt	gatacagggt	tcctgagagt	154260
gctactgaaa	caccctcatc	tgctaaggcc	actagatttt	tttccactgc	ccagactact	154320
cagggccctc	agtaggtcac	tgcccaaaga	gcctagatgt	taaaactgca	gctaaagcct	154380
ctcctgaggg	cagagctcag	agcctccctg	gcctgcacaa	agtgtctaag	aggacattgt	154440
ccatccaacc	acttggaact	acactgtaga	cgctgccttg	tctgccagct	tgaaaacccc	154500
agaaacctct	tcccagctc	cgctagctt	gcctccagca	ccctgacatc	cacttctctc	154560
gctaattgtc	gcagcttcta	cagtagggtt	gacggggtgc	tcgggggtgc	cagacaggcg	154620
tgtgttgcat	ataaaaaaga	ggatgatgtt	taagtatcta	agaatgttgg	tccctgaag	154680
tgattcttgc	tgcttctctt	ccttcccagc	tcttcccact	gacacctttg	ccccagcaa	154740
gcccagggac	tattgttgct	ggctgggggt	ctgaacaaaa	ttgccgcagt	tttttgtttt	154800
tttttttctt	gctagaagtt	accgatacag	tccttaattg	agcaaatata	ggttcccgtg	154860
tagtttataa	acatgataag	acatacagtt	tggtaggag	tggttgagg	gacctgtgca	154920
tttatataat	tatatatata	tatatattat	gtgggtgtgt	gggcagagt	aggatatata	154980
taagtggaca	taggtataaa	actgcacctt	ctgtgtgact	ctcattgcga	gtacagttct	155040
aaatgtcatc	cactggcgat	ctctcccttg	gtgattgggt	cttggaaccc	agcagggtctg	155100
ccgggggctg	cctgagacgt	cagggaatgag	aggcattacc	cccatggata	gggactgagg	155160
tgggcatagg	gttgagacag	gcaggttaac	taagtgatct	cagacaagag	ccaagagtgc	155220
tctgagattg	ctggttgccc	cagctggctc	tgggagagcc	ttgttctgag	tcctgtcct	155280
tccaaaaacca	gcagggtcct	tcagcccttc	tctccaaatg	accaggcttc	cgagagagcc	155340
agcttcttca	aggggcgcat	gtcccgacac	cactaatgac	tcactttgcg	tgcccttgag	155400
cactgtgctg	gagtgagatac	ggtccagagg	cgctcgggtc	ggacagccga	gtgagacgtg	155460
atacccttcc	cgtctacggc	tgtacatttt	gggtctataa	tccaccagga	aggggtccaga	155520
cggtggctga	aggcttcagc	agccttttcc	tgaaacccag	cagatcttcc	acttaggaaa	155580
aaaaaaaaga	agaaagaaag	aaaaattctg	ttccttgctg	gacatgtggc	gacatgtggc	155640
agtgtctggg	gacggagccc	tggtgcctca	gaggagagtg	actgccagcc	ccagtatcca	155700
ggccaggagt	ggggccccaa	ggggcgtgcc	tgaggatgcc	tgctgcaccc	cactgggggc	155760
acggatgggg	gtcctgcgag	cacacttgcc	cttctcttgc	ggctgtgcgg	tgggcatgat	155820
gtcaaaagctg	aacttgtgct	gatagtgcgg	ggcatagtcc	tgcatgtcgg	taagctcttt	155880

cccagagctc	accttagaga	tctgggtccg	gttccctgtg	ctggtgcagt	tcttgccctgc	155940
cttcttgtaa	cctgacctgg	agccaggcgg	atggccatgt	gggtggcctt	tgtccctgga	156000
ggcccatg	gacggatggt	gtcccttgcg	ggcagccctg	tcagaggtgg	taagcgtgtg	156060
agacttgatc	tggtgaggag	acactggtcc	tgtgcagttc	cggaggtcct	ccaccctcag	156120
cagcttcaga	tcctggcctt	gccgcagctc	gggggtcgcg	caggggacag	cagagctaga	156180
gccacggaac	cttcgcagcc	attccacag	ggaacgtgcc	cggcagccac	agtcccaagc	156240
attcccatgt	aggcgaagga	actccaaggc	caccaggggg	gccagacagt	caccctgcag	156300
ctcagttagg	ctgttggtga	agagaaagag	ggtggttagc	ctgtggaggt	catggaagac	156360
cttgttggtg	acccactgta	gctggttctc	atgcagcagc	aaccggtcca	ggttcaccag	156420
gccccggaag	atgccttgcc	ccaggctcca	tagcttggtt	ccatggagaa	acaagtgaact	156480
gagattgacc	aggtccacaa	agatgtcatc	ttggagggtac	tcgatatggt	tgtcctgcaa	156540
gtagagatac	tgaggctgt	gcaggccacc	aaagatgcct	gcgggcaggg	cgctcagtc	156600
acacttatag	aggtagagg	cgtgaagctt	caccaggcct	tggaggtct	cggtgccag	156660
cgcttcgcagc	tgtcgggtgt	ctccaaggtc	tagctcctcc	agatgcacaa	agccctcgaa	156720
ggtgttgga	gcaatgaaag	tgatgttggt	ggagttagat	cagaggtgta	ccatggcggg	156780
gctgaagtgg	ccttgctgga	ggaaggtgat	gcgattgttc	tcaggaaga	tgcgctcact	156840
gtctcttggg	atgcctccg	ggatggcagc	aaagtgtgt	gcctggcagc	tgacagtcac	156900
gggcgagg	tagcacacac	agtctcagag	acaaccacca	cccagaggta	gctctccagc	156960
gagcagcaac	agcagcaatt	ccacacagca	ccctgggtgg	gagagacaga	acagcagtga	157020
ggggctgccc	agaggaggtg	gagatagatg	gggaacagag	ggtggaatgg	gggactggca	157080
aatgactctg	ttggctcaca	gaggttctgt	cctctgtatt	gtatgcaggg	gtcccttgga	157140
cagggcattt	ggggccaagg	cccacattat	ctcctcacct	ctttagctct	gtccctaaag	157200
tctctaattc	catccgaaca	cttcttcaga	ctgtcagccc	caccgcaggg	tggaaacacac	157260
ttgtgaacac	aggcgagtcg	gcctccggtc	ctgggtccgg	ctctgccact	cgctcactgt	157320
tagctgcctt	agcaaggaat	gactctaaca	aagcaaatct	ggagtcctga	atgatcactt	157380
tatttaataa	attcctcaaa	ataaagaaag	cattgagtcc	atggtaacca	agcatgcctc	157440
aataagcgcc	tctttcacac	tgtggtacaa	aaacttcaaa	cctaacaactc	ctccatggct	157500
gtttccatcat	gagttaaaca	cagttcacag	ggctgtgtgt	acagacaagg	cacatttctg	157560
tgaggggctg	tgagtgcacac	cagggtgtac	gcacgaggtc	tcccttgggg	ttcacagtac	157620
tgccagatcg	aggctgcacg	ctcctctccc	ccattcacac	ccccccctcc	tgtgctggag	157680
attgccaggc	tgtggtgta	aaaccgggcc	ttgcctcttg	actgtccaga	gcatttctct	157740
tgtagcttcc	ctctaattgg	gcattaatta	ggcattcgtt	aatggatcct	taaaataatt	157800
atthtgggat	gtgtccagcc	tgtggtcggg	taataggcct	atgctcatta	tggagccgc	157860
ctcattatgg	acgattgtca	ttacctgcct	ttttccaggg	tcacagctgc	ccccagtgge	157920
ccagcaggcg	cgtcaggaa	atggggacag	gctccaggcc	tacggggcgc	caccctgaag	157980
ggccaggcag	ccacgacct	tgtgcctca	gttgccctct	tgcctctct	tttccagctt	158040
gttcagctgg	gactcttggg	agagccaggg	cccctggggg	aatatgagct	gagctgaact	158100
ttcttctgct	tagctgtgct	cagagcaagt	ggaaggagca	gggaccttct	gaccaggtct	158160
cccacttggg	gtcccagcc	cagggactgc	ccaggccccc	gcagagttagg	ttgccacctt	158220
gacttctgac	gccccccccc	cattcccaac	agaaacagca	tcgtaagttg	acagctccca	158280
gcgtttggga	ggttatgggt	cccagagagt	ggcagctgct	tctcgtccct	gtaatcaccc	158340
ggcttcagct	agaatgttct	tagcacataa	aaatcatcgc	atataattta	gtttttgcat	158400
aattgggttc	agttgtgatt	tcagagcaat	tatgacctca	gcagcagggg	tggcacagca	158460
taggaacctt	ttctgggccc	ggccatgccc	tgcaggccct	ctgcagtgtc	ttctgcccac	158520
cgcccttag	acagcatgca	ggctataacc	atctctgcct	caatttctct	ctccaaaacg	158580
tgtctagatg	ttactctgtc	gatcttctct	ctcagcatcc	tgggtgtggc	ctccagcctg	158640
ccagcctctg	ctctagggat	cctgtgtctg	agagggaggc	acagtcggag	ggaggggagc	158700
ctgcctgtg	ccaccagcac	tcacactggc	tgacagttcc	acagaccac	agctccaacc	158760
tccttctgtg	gcttgacccc	tctcttcag	gaaggccatt	cttgccagaa	cctttcccaa	158820
cggtcccttg	ggaagacctg	gactctaggt	tcaaggacat	tcatgatgct	tgcctccacat	158880
tttatgctgg	atgagacaca	gcagagcctt	cttcaactgg	gggtcctgtg	aaaatgaaag	158940
cttttcttcc	ccggcctgca	gctgcaggca	ggtagggttt	gcagtggtgt	tatcactaat	159000
accattcgac	attttgtacag	ctcatcggag	tttacagagg	gcttttgttg	taccctcaac	159060
ttccctctgt	ttctccacc	tactgtggct	gctctgtctc	tgtgcacccc	aaaagaatct	159120
ggaagtccct	tggggagatt	acccccctta	cataggggcc	tccaaggaat	acaggctaca	159180
gctctattta	ggaaaaaaa	aatcaaaact	aaccaaacc	caggtgtgga	cttagtaacc	159240
agtttataaa	cataccgtgg	caagtggagg	aggcaggcgg	cagaacggca	tgaaggtagg	159300
actgggggtt	tccctctaaa	agggcaatgg	ggggacaaag	ggactctagc	caagacctga	159360
tgagaaacca	ggctcagttc	ccctgttatc	tcaaggctat	catcactagg	aggcctaagg	159420
caatggacca	caaggacctt	gtccttgtag	gacagtcact	tcctgcagtg	aagtgtctct	159480
ctggaagcat	actaatagga	tgtaggctca	ggacagctgg	tctctgtcct	ttagtatatt	159540
tcacatgccc	agggatgtta	accttccaag	cctccatctc	ttctaattgg	gggggggtgt	159600
tgaggggctc	agccactctg	catccatgtc	tttgaaagcc	agtggattta	ctccaggacc	159660
ctgagcaagc	tgtcctagtc	agccctggcc	acttctggac	tccttgcctg	agtcagtagg	159720

tgccaatcct	aggattgtca	ccagcagggt	tcttcctagg	gaggcaagca	ctgtatcacc	159780
atggcgccct	ctatgcccc	tctatgaggc	ccttgggagc	ccccccccc	tgattgctcg	159840
attaatgtac	caacaatgag	gatggagcct	ttgccatgca	ttttaacatt	gcaaatatagc	159900
aggaatccaa	gtctctgtgg	aggggcccctg	cacctcttct	gccagactca	tcaagcgctc	159960
cctgggcagg	gctgccttct	acttgagggg	gcggaaggga	gaagacccag	ttccactctc	160020
cttccctctc	aggaggtgcc	cttcatcggt	ttctgcttcg	ttactctcaa	gcctccggcc	160080
tcccacgcac	gtgagctccc	aaggggctct	acagcctccg	tcattccttc	ttccattcat	160140
acttgcccc	ggctctagga	gagccatgga	agacagtgtg	ggaagggtct	gacaatgagc	160200
atcatgcccc	atttgcatat	gcgggtggca	tacctgtgtg	ggtaccaggp	gagtataggg	160260
gaaattaaga	gaggggccta	aggaagcct	ctgctatccc	tggtctacca	gtcagcattg	160320
cttggctcact	gatccctctt	gtaacaccag	cccttctgca	acctgccaga	gtttttgacc	160380
tttgaactag	ggctgagaag	ggtctgctct	gttcagctgc	cttggctggg	aggggaatct	160440
gctcagacct	cagcacacac	tcaacagaag	gcatgcaagc	aaggagacta	gcagtggcct	160500
tgggtcagct	ggcaagcccc	aaactcttcc	tgccaagctg	agcatgaaaa	gccacctcac	160560
catgggtccca	tgggaccaga	cctggtagga	taggtggcaa	ggctaaggca	gcggaatagc	160620
atgtgcaaag	gcactggggg	gggaaagggc	ctgtgcttct	cacccctctt	aatggtgcag	160680
agcctccaag	gaatactgta	acctcagctc	agctgggtct	gggtggccag	agagcttggc	160740
accagaacca	gcatacaacg	ggcctgtctg	ctaaacccag	acctcacaag	ccagtttagt	160800
agggggccctg	tagcacccctg	gccaccagaa	ctaaccagga	agatctgacg	ctgggaatat	160860
gtctttaaatg	aaaagccctt	ccggaagcca	catttgccaca	gaagaaaatg	aggtgccacc	160920
agcatcagtg	ggctgggtgc	agctggagaa	cacagcaggg	ggacaggctc	taccaagcta	160980
ccctggccttc	aggctggggc	tctagccagc	tccctgatgc	ctggagtagg	taaagcagcc	161040
tcgaaatggg	ctgggtcagc	tttttcaggc	tccaaagggt	caggacagct	gctgcagctt	161100
agcacccaaag	gggggtgccc	ctctacccct	aagtagaggc	atccccatgg	cccttgggca	161160
gggtcagtg	tctctctgaa	gctttgtagg	ctgcttcttg	gccatgtagc	caatctctct	161220
ggccttcagt	cctccctccc	tgccccagc	ctggccagct	gctcttctct	gagcaatcga	161280
tgttaaccga	atgctctctt	gctgtgggga	tgccggcctc	agggcaggcc	agctgcactc	161340
ctggggctgc	tggcgcccca	ggccacttgg	cacttgtgct	acttgtgttc	taaacacagg	161400
ctccttctctg	tgtcggccct	gaagacaaga	aagctggcca	gggaacagct	gggctcccat	161460
ctcagcctcc	actgtctgtc	agagcggccg	gcagcctcct	atccatggct	gtgagttaga	161520
cagggtctgtg	gagccagagg	cctaagttag	atcctggctg	ctccttttaa	tgcttgccag	161580
agcctcggtt	tcttcaacctg	caaaatgggg	cagtcattgg	aagctcagcc	agtcctccag	161640
cccacagata	ctggtaccaca	cctgcccctc	ccacatctcc	atctgtctaa	ctgaaacct	161700
ccaaaccaag	ctctctctct	cctctccggc	ctcctccgtt	cacaaactct	ccatcttggg	161760
ttaagatggt	ccctttacat	cagctgcctg	ggaggaacac	ctcagaatca	ccgtgggtca	161820
ctctggggaa	attctgctgg	ctagatttta	gaatgtatcc	agtatctatc	cacttatcat	161880
actcgctatt	gctaccattc	accagtagc	ctcctggatg	ccctcccccc	ccccccgcaa	161940
gcccctgcctc	ctcacccttc	aacaggaaact	agggtagtcc	agggaaagtg	agggaaagtg	162000
agtcaggaag	ggctgctcct	tagtctgcat	cctccagagt	gccatcttaa	ctagaagctg	162060
cccagggtct	ttctccagc	ccagaggccc	tgctctctcc	tggtacaca	gacctgtctg	162120
cgttccctcc	cacatgccag	ccctcagcct	ccctcctgcc	tttgtccatg	ctgttccatc	162180
tacctggacc	ggttccagct	gtgtctgcag	ggctgcttcc	cagagaagcc	actcttgagc	162240
agcgatttgc	aaggagcttc	tttctgggga	cattttactc	tacagcacgc	gacctcttgg	162300
acagcacgac	ataagtcact	tgtttccctt	atgtccgctg	ccactttgtt	ctgatgagtg	162360
tgtcccttgc	acacagtagg	tgctcattaa	cagttctggg	ggagggaaat	accttctcaa	162420
gtctctggag	aattgaatga	tgacactcag	gaagccctag	gctcaagcct	ggggcctgga	162480
tcatagtagg	tgctcgataa	atgttggttg	taattagtcc	tgggagactc	agagccttca	162540
ggagaacaga	cacctgaact	tggttcacat	caagactcct	tagggccatc	aagggaagtga	162600
ccgtttgttg	gatgagcact	ctgagaagga	cccaatacca	gtcctttcct	gggcaagggg	162660
aaatagactg	ggactgggga	gtttcccaag	gtatgggatt	ttcagcacta	aactggaata	162720
atgctaaaga	aaaaaaaaag	ttagtcaact	taaaaggggc	caggacccaa	acctttcaca	162780
cagaaatgtc	tgggtttatg	aagagaggaa	gccagatagc	gtggggcaca	tctttaatcc	162840
aggtgcttgg	gaggcaaaaga	cagatggagc	tctgtgagtt	tgaggccagc	ctattctaca	162900
aagtgagttc	taggacagcc	aaggctacat	agagaaaccc	acttgacttg	ccaccocaaat	162960
taaaaatctt	agtgggggag	cagtcaagga	ggaaacacac	acacacacac	acacacacac	163020
acacacacac	acgttagtat	aatatcatac	tatggctctg	tgctgcagct	ccaggaatga	163080
gggtggaact	cagagtgtta	gtgtgtgcta	gtggatatatt	gagctctgta	tttatgtgca	163140
tgtctgtgta	gatgtgtacc	tgaggtgttt	atgtgtacac	aggtgttggc	ctgttgcata	163200
tggtgggag	catggttgtg	tttgctaggc	atttgcatgt	gtctgagttc	atgcacataa	163260
actcacatct	acctctggag	actgagagtg	acaacccagg	gcccttttat	cctgcagcac	163320
cccaggccca	gcaccccgac	ccagcatccc	agggccagca	ctccagaccc	agcacccagc	163380
gccagcatc	ccacgcccag	catcccaggc	ccagcacccc	agacccagca	tcccaggccc	163440
agcacccag	accagcacc	ccaggcccag	catcccaggc	ccagcacccc	gacccagcat	163500
cccaggccca	gcacccagc	tccaaagcact	caatgcccag	caccccgacc	cagcatccca	163560

ggccccagcac	cccaggccca	gcattccaagg	cccagcatcc	cagggacagc	accccaggcc	163620
cagcatccca	ggccccagcat	cccagggaca	gcaccccagg	cccagcatcc	caggtccagc	163680
atccccaggga	cagcacccca	ggcccagtat	cccagggaca	gcaccccagg	cccagtatcc	163740
cagggacagc	accccaggcc	cagtatccca	gggacagcac	cccaggcaca	gtatcccaca	163800
tgagggcagc	acatactgaa	gatagggaat	gtctctgagg	cctcttatct	tggtccttac	163860
cctcattgct	ttcagcacct	gctctcctca	cactcggaat	caaacaccct	gtgcagggtc	163920
tcccagtaacc	aggattcccc	tcagctgagg	aatgggtagc	taccattttg	gcttttgtct	163980
gtctggggtt	ggcagcccca	tgctaattgg	actgacagtt	tctcctgaga	gcaatttggg	164040
cagcacatcc	tgcccattag	gcctaacctt	gcctgcaggg	gtgtgctgtg	ggggcaggga	164100
tgagagctac	cctgtatagc	tctgtattga	ggcactcccc	caagctatga	cccagtcagg	164160
tgggagtcac	ttcacctagg	caactccaga	tgggcacaaa	aatctctcca	ataagggtag	164220
gtatgggaat	aggtaaggag	agcatagtga	gcctggctgg	gcacctgaga	cctgagcagc	164280
ctgcacggga	gatttgtgtca	ctgtgggtcc	agactgccaa	gacatcttgg	ctttcacccc	164340
aaactcaggt	ggctccagaat	ccagagctct	taagagagca	gatgctgaga	ggcacttaac	164400
ccagggtctaa	gaccttctct	tggacggttt	cttggctttc	tactctgtcc	tctgtcccag	164460
tctgtcatcc	ccatctgtgc	ctaaccgtcc	tctgtggaaa	acatgaggcg	tatgagctct	164520
ctacttctcc	cagcatccca	tgcccgcacc	ccagctcact	gtgtgcccct	atgttactca	164580
aatcttctgc	taggtttgag	ggcccagggt	tgaggctgtg	ggtttccctc	catctgtccc	164640
tcccttttac	caccaccact	aatcctcttc	ctcctcttcc	tctcttccct	cctcctcctc	164700
ttctctnctn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	164760
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	164820
aaattgaact	tacagaggaa	acacaaatga	gctgtaatga	cagaaaaagg	atactatttc	164880
ccttgtgatc	agaaagctat	agattaaacc	ccagcccttt	ttgcccttgt	ttttcatatt	164940
tagtggttct	ggggactgaa	cccaagggaag	gacagctgct	tggaatgttg	gagacctctc	165000
aggcattgct	tatgctaagc	tttctggaat	tttctggaat	catatggtag	tatctattca	165060
tcttaagatg	cctataacca	tggttgaggc	aattctctaa	gagctctgtc	gaacttctta	165120
agcatgtgct	caaaaaacaca	cataaagatt	gtttactgaa	gcagtatttg	caacagttaa	165180
acaaaaccac	acaacaacaa	aaagtactta	atagctcatc	agtaggaaaa	agaacaacaa	165240
attgggacac	atttatacta	ttctcgttct	atcacacgac	tattaagaag	attaagggtg	165300
gaggatgaaa	agcccctgca	gaacaacgca	cacaatgaaa	ttgaatagag	aataaagaaa	165360
aaaatatata	gtgtatttat	gtaaaacata	cataaacaca	taaaaagaat	agaaaaactc	165420
ttgactgtat	atctttgaaa	agaaccagca	gagtgctatg	gtttataaat	gaaaagtctc	165480
cccaaaatcc	ctgagttgga	ggttcaatgc	agagtccata	ctcttaggaa	gtaactaaat	165540
caggaatgct	ctgatctaat	caatggattg	accaactgat	agattaataa	tctgaaggca	165600
acacttgacg	gaggcagaaa	caaggtaggg	cctaactgga	ggaagtttgt	caccagaggc	165660
atgtccttgg	agaggctacc	tcatccttga	tccttccctt	tctccgcttc	ctggctgtaa	165720
aagtatttat	ttgactgatg	taactgtcat	cttgagggtc	tactggctcc	atcagctaac	165780
ctaggcctag	cctgtgaagc	ttctagcttc	catacaatct	aatccaagcc	tagaatgttc	165840
cagcctttag	gaactgtctg	tgagatcacc	gtttcctgtt	ctttctgaac	tctagctggc	165900
tgattcagtc	cccctgttcc	gggctcaaac	tcctctcccc	gatgatttta	ttcacaatct	165960
gtcttttctc	ttggcctctg	aattgtctct	cttggctcca	aactaactct	agcaatcttt	166020
tctaactctc	tgtctccttc	acactctctt	gcttgttctg	tctttactgt	gtctagtttg	166080
ttctctcttc	catccttctc	tgtaaaagct	tcctgggttaa	cctgcctcct	cctccccctc	166140
tgtgcgctc	tactctccct	ctcagctcta	ctgcaactgt	ctccacagct	ctcctgtatc	166200
ctgtgctgca	ctctcttctc	cggtagccac	tgtgtctccc	ttacgtagct	tccttttctc	166260
ctctcttctc	ctgagggttg	ggcagatcct	atcctgtcaa	acctttctct	gattcttcac	166320
tttgtctgct	actcaattag	acatcacttt	caagcaggag	tgccctctct	acaaaccaac	166380
tttaccttca	ttgtttcaaa	ttaaaaggta	gtactaaggg	tgtgtctctt	tttcagccag	166440
tgagagtaaa	gatgtgtgct	aataaggctg	agccaactct	agctagaaat	agtttctttt	166500
tctccataaa	taacagaatc	ttagggttca	caatacgate	aaatatcctg	agacagctgg	166560
ctgccatgtg	gtgaagagaa	tgcaagaggt	taaggatgtg	ctcttgaggt	ttcagatggg	166620
aataagcatt	ttcttgggag	ttggattaga	gtcattcctg	tgacactgtg	acaaaggact	166680
cgactacatt	ttgccatgcc	ttgagactgt	ggagggtgga	gcgatggatt	aatacactgg	166740
agtgaatttt	aaggcagcca	aggctgtggt	tggactgtta	ctagctacgt	ttagctggat	166800
ttatatatta	aattgggaac	aaaaaagcag	agtagaaagg	acagtttttg	cagaaggagg	166860
tttcttcata	tacaacttag	tttttatagt	tagcttttta	tggtgctatg	accaaataac	166920
cttatggaaa	ctgaagaata	aaaattttta	ctgaactctt	cttcacccca	gaaccgcacc	166980
cctccatctc	agagattggt	cccgaacac	tcctgaactc	ttcaccocag	aatgctttcc	167040
tgaactcttc	accctagagt	tcgaacccct	ccaactaaaa	actgttccaa	gaacattttt	167100
gagataaggg	cctcctaata	caacctcaaa	atgaacccgg	tacattgcca	ataaatagga	167160
catgaccctc	tagttacgta	gattcccttg	gcagaacccc	ttgtcccttg	acagaacccc	167220
ctagtgtagt	aaactgttac	tttccctgcc	cagctctccc	cccttgagtt	ttactatata	167280
agcctatgaa	aaatttggct	ggctcgtgat	tctcctctac	accactaggt	gcattgagtt	167340
cgacccacga	gctctgggtc	atgttccatg	tgtttctctg	ctgtgtgtct	attaaatctt	167400

gccttctaca	ttttgagtac	ggtctcagtg	tcttcttggg	tccgcggctg	tcccggggct	167460
tgagtgcctt	agtgaagggtc	tcccttcggg	ggtctttcat	tttgggtgcat	tggccgggaa	167520
acagcgcgac	cacccagagg	tcctagaccc	acttagaggt	aagggttctt	gttctgtttt	167580
ggtctgatgt	ttgtgttctg	tttctaagtt	tggtgcgac	gcagtttcgg	ttttgcggat	167640
gctcagtgag	accgcgtctc	gagagggaa	gcgggttgga	taaggataga	cgtgtccagg	167700
tgtccaccgt	ccgtttcacc	tgggagacgt	cccaggaaaa	acaggggagg	accagggacg	167760
cctgggtggac	ccctttggag	gccaaagagac	catttggggt	tgcgagatcg	tgggtttgag	167820
tcccacctcg	tgcccagttg	cgagatcggt	ggttcgagtc	ccacctcgcg	ttttgttgcg	167880
agaccgtggg	ttcaagtcct	acctcgcggt	tgggtcacgag	atcgtgggtt	cgagtcctac	167940
ctcgtgcaga	gggtctcaat	cgcccgccct	tagaaaaggcc	atctgattct	ttgagttgct	168000
tgtgggtcgac	gcagagtcgc	cgccgtttct	ggtttctttt	ttgtcttagt	ctcgtgtccg	168060
ctcttgttgt	gtctactggt	tttctagaaa	tgggacaatc	tgtgtccact	ccctttctct	168120
tgactctgga	gcatttggag	gaggtgcggg	tcagagccca	caaccagtcg	gtggaagtca	168180
gaaaggggtcc	gtggcagacc	ttttgcgcct	ccgagtggtc	aacgtttaga	gtaggctggc	168240
cacctgaggg	tgcttttgac	ttgtcactaa	tcgctgccgt	caggcgaatt	gtttttcagg	168300
aggaaggggg	tcacctgat	cagatccctt	acattgtgac	ctggcagaat	ctcgtccaat	168360
tcccacctcc	gtgggtcaag	ccttggaacc	caaatctctt	gaaactgacg	gtcgcgggtg	168420
cccagtcctg	tgccagccga	aagtcgtggc	catcagcacc	cccccaagtc	tatccagaga	168480
ttgacgacct	cctctggata	gactcccaac	ctccccctta	ccccctggcc	caacagccac	168540
ctgcagctgc	cccaccacag	ggaccaatag	cgagaggggc	tcaggggacg	gcgggggaga	168600
ctcggagtcg	ccgaggccga	agccccgggg	aggaaggggg	gccagactca	acagttgcct	168660
tgccactcag	agcacatgtg	agagggccag	caccaggacc	taatgatctc	attcctttac	168720
agtactggcc	tttttctctt	tctgatttat	ataattgaaa	aactaaccac	cctcccttct	168780
cagagaaccc	ctctggactt	actgggctcc	ttgagtcact	tatgttctcc	catcaaccca	168840
ctctggatga	ttgtcagcag	ccttttcagg	ttctttttac	aacagaagaa	agaaaaagaa	168900
tccctcatga	ggcagagaaa	aatgttctgg	gagaggacgg	cacacccact	gccctcccta	168960
acctcgtgga	cgaggctttc	cccttgaacc	gcccccaact	ggactacaac	accgcggaag	169020
gtagggggag	cctccttgtc	tatcgccgga	ctctagtggc	aggtctcaga	ggagccgcta	169080
gacggccccc	caatttggct	aaggttaagag	aggtcttgca	ggggcagact	gaaccaacct	169140
cagtccttct	tgagcgtcta	atggaggcat	ataggagata	cacccctttt	gaacccctgt	169200
cagaggggca	gagagccgct	gtagccatgg	ccttcattgg	tcagtccgct	cccagacatta	169260
agaaaaagct	gcaaaaggct	gaggggctcc	aagatcatac	gctccaagat	ttagtaaaa	169320
aagcagaaaa	agtctatcat	aagagggaaa	cagaagaaga	gagggcaggag	agagagaa	169380
aagaaataga	ggagagggaa	aatagacggg	atcgccgtca	ggagagaaat	ctgagtaaaa	169440
ttttggccgc	agttgtgaat	gatagacagt	caggaaaaag	taaaataggg	ctcctgggca	169500
acagggcagt	gaaaccgcaa	ggtggcaaaa	agataccact	ggaaaaagac	caatgcgcct	169560
attgcaaaga	gaaaggacac	tggtgctagag	attgccctaa	aaagcgggag	cgatccaaag	169620
tcctaaacct	agaagatgat	taggggaagtc	ggggctcaga	ccccctccct	gagcctaggg	169680
taactttgtc	cgtggagggg	actcccgta	acttctgat	agacaccgga	gcagaaactt	169740
cagtactcac	taacccctta	ggcaagctag	gctccaaaaa	gacctgggta	attggagcca	169800
ctggtagtaa	atttttaccc	tggacgacca	aacgagctct	tcagatagac	aaaaatatag	169860
tgaccacact	ctttctgtgt	atacctgagt	gccctgctcc	cctcttgggg	cgcgatctgc	169920
taaccaaaact	aaaggctcaa	gtccaattta	cttcagaagg	cccacaagta	agctggggaa	169980
agggccctgt	tgccctgcctt	gtcctcaaca	cagaaaaaga	gtaccggttg	catgaagaac	170040
aacccaaaaa	tcaggtctct	tcaggttggt	taactgcgtt	ccccaatgtc	tgggcagaac	170100
aagcaggaat	gggttgggt	aaacaagtgc	ctccggttgt	ggtagaactt	aaagctgatg	170160
ccacccccat	ttcggtaaaa	caatacccca	tgagcaagg	agctagaaaa	ggcatccggc	170220
ctcatatcca	gaggttgctg	ggccaaggag	ttttagtggc	ctgtcagtc	ccctggaata	170280
caccacttct	gcgggttcaa	aaaccaggga	ccaatgacta	tcgcccggta	caagacctcc	170340
gggaggttaa	caaaaagggtc	ctggacattc	acccacagtc	cccgaaaccg	tacaatttat	170400
taagctctct	cccacctgag	agaacatggt	atacagtcct	agacttaaaa	gatgccttet	170460
tttgctgtgc	tttgacccct	aagagtcagc	tcctgtttgc	ttttaaatgg	agggaccag	170520
agggcggaca	gactgggtcaa	ctaacttgg	ctaggctacc	acaggggttc	aaaaattccc	170580
ccacctgttt	tgacgaggcc	ctccatcggt	atcttgccgc	ttttcgcgct	cgaaaccttc	170640
agcttacctt	actacagtat	gtagatgac	tcttgggtgc	ggcggcctcg	aaggagctgt	170700
gtcaccagg	aactgagagg	ctcctcacag	aactgagtc	cttgggggtat	cgagtttcgg	170760
ctaaaaaggc	acaaatctgt	caaactgagg	taaccttctt	gggttatacc	ctccgagggg	170820
gcaaaaagtg	gctcacagag	gcccggaaaa	agactgttat	gatgatccca	tcgccaacta	170880
ccccacggca	ggtagctgag	tttctgggga	ctgctggctt	ttgtagactc	tggattccag	170940
gctttgcaac	cctagcagca	cctctatatc	ctttgactaa	ggaaggggtt	cctttcaagt	171000
ggaaagaaga	acaccaaaga	gcttttgagg	ctatcaagtc	gtctctaagt	actgccccca	171060
cgctagcatt	accagacttg	actaagcctt	tcgtctcata	tgtggacgag	agagcgggtg	171120
tagccagggg	agtatgaca	caagcactgg	gacctgaaa	aagacctgta	gcctatttgt	171180
caaaaaaatt	agatcctggt	gctagtggat	ggccacatg	tctgaaagct	attgcagcag	171240

tagccctgct	gatcaaagat	gctgacaaac	tgacaatggg	acagcaggtg	accgttgtag	171300
ccctctcatgc	cttagaaagt	atcgtgcgac	agccacctga	cagataagat	gacaaatgcc	171360
cgaatgacac	actatcagag	cctgctgcta	aatgagcggtg	taacctttgc	gccccctgcc	171420
atcctcaaac	cagctaccct	tctccctcta	acaaatgatt	ccgtcccgagt	acatcaatgt	171480
atggacatcc	tcgctgaaga	aactgggacc	agaagtgacc	tgactgacca	accctggcct	171540
agagctccca	gttgggtacac	ggacggcagc	agtttcctga	tagaggggaa	gcaaaaggct	171600
ggagctgctg	tggtagacgg	gaaaaaggta	atttgggcaa	gcgctttgcc	tgaaggaaaca	171660
tcggcacaaa	aggctgaact	tatagcgctt	atacaagccc	tcggagaggc	taaaggtaag	171720
atcgtttaata	tctacactga	cagccgatata	gcttttgcta	ccgcacacat	ccatggggcc	171780
atctacaggc	agcgagggtc	attgacctcg	gctggtaaag	acattaaaaa	caaagaaaaa	171840
attctggccc	tgttagaagc	catacatgca	cctaaaaagg	tagccatcat	ccactgcccc	171900
ggccacccaa	aaaggagaaa	acttgggtgg	caaggggcaac	cgaatggcag	acttagtggc	171960
aaaacaagt	gctcaagggtg	ccatgatctt	aactgaaaaa	ggtgatccgc	ccaaaagccc	172020
tgaggatggg	aggtataaca	taaaagagct	atggtagacc	agtgatcccc	tcccatactt	172080
tttttgaaag	aaaaatagaa	ttaactcccc	aagaaggaaat	aaaatttgta	aaaggactac	172140
accaattcac	ccacctggga	gttgaaaaaa	tgatgagact	aattaaaaat	tcccgatacc	172200
aagccccaa	cctgaagtca	gtggctcaa	agattataga	ctcctgcaaa	ccatgtgcat	172260
tcactaatgc	aactaaagcc	tacagagaac	ctggaaagag	acaacgggga	gacctccctg	172320
gagtgtatg	ggaggtagac	tttactgaag	ttaaaccctga	aatgtatggt	aacaagtatc	172380
tgtagtagt	tgtagacacc	ttttcaggat	gggttgaggc	atttcccact	aaaacagaga	172440
ctgcccagat	tgtggccaag	aaagtccttg	aagaaatcct	gccaaagattt	gaaatcccta	172500
aggtaatcgg	gtccgacaat	ggaccagcct	ttgttgccca	ggtaagtacg	ggcttgccca	172560
ctcagttggg	aaattacact	gtgcttaccg	ccctcaaaagc	tcaggacagg	tcaggacagg	172620
tagagaagat	aaataggacc	ctttgactaa	attagccatt	gagaccggca	gagaccggca	172680
gaaaagactg	ggtggctctc	cttccctctg	cgctcaaaac	cccctggctg	tttcgggctc	172740
actccttttg	aagttctgta	tggaggacct	cccccttaa	tggaagctgg	tggaacatta	172800
gtttccgact	ctgaccctgt	cttaccctcc	tctttgctta	ttcattttaa	ggccctaaaa	172860
gtgattagga	ccagatttg	ggaccaactg	aaagcagcct	ataccccagg	gaccaccgca	172920
gtaccccacg	ggttcccgagt	tgagagacaa	gtccttggtca	gacggcatcg	aaccggtagc	172980
cttgagccac	ggtggaagg	accctatttg	gtgttactga	caacccttac	tgcggtaaaa	173040
gttgacggaa	tcgcctcctg	gatccacgcc	tcccacgtca	agagggccgc	cagtcaagat	173100
gaagaaaacc	acgacgacaa	ttggacagct	gcagtcactg	acaatcctct	taagcttcgt	173160
ctgcgcgcga	ggcgccactc	tagacctagg	gaaccttaac	cctcatgctc	caattcaaca	173220
gtcctgggag	gtgcttaatg	aaaaggaaaa	cattgtatgg	gcaaccactg	cagtccatcc	173280
cctctggatt	tggtggcctg	atctcacgcc	tgacatctgt	aagtttagcg	caggatcccc	173340
caattgggac	ctctcagatc	atactgactc	tagcaaccca	ccccctgagg	agcggtgtgt	173400
cccaaattgg	ataggagaca	catatgggtg	ttcgggcgag	ttctacogag	ctaactcttag	173460
agctgcacat	ttttatgttt	gccctggctc	gggtcagagc	aaaaggcttc	aacaaaaatg	173520
cgggggggca	tcagattact	tttgtggtaa	atggacatgt	gaaacgacag	gagatgctta	173580
ctggaagccc	tcctctaaat	gggacctaat	cacggtaaaa	cgaggtagtg	gcta tgataa	173640
gtcaaacgaa	ggagaaagaa	accctataaa	atatcaagag	agtgggtgctg	cttttaaaaa	173700
cagagcaccc	tcaggaccat	gcaaagataa	atactgtaac	cccctacgta	taaggttcac	173760
cgagaacgga	aaacaacacc	gtctaagttg	gcttaaaagg	aataggtggg	gttgccgagt	173820
atacattcca	ctaagagatc	ctgggttcat	tttccagatc	agattgacag	tgagagaccc	173880
ggagaaaacc	ctcgtagggc	ccaacaaggt	ccttataaaa	cagggggccc	ccagtctgtac	173940
tggtctcccc	aaagggtccc	actgtaccag	ctccaccaac	tccacagccc	aacacagtgg	174000
taccctccct	aggaactaat	actctcctca	taaagcctac	cttggcttcc	ccaccgcccc	174060
taggaacaga	ggaccgtctg	gtcagtcctag	tccaaggagc	tttttttagtt	ctaaaatagaa	174120
ctaaccctaa	tatgactcaa	tcatgctggt	tatgctatgc	ctctagcccc	ccttattata	174180
aaggaatagc	tcagatcagg	acttataata	ctacttcaga	tcattctcaa	tgcctttggg	174240
gaaaaaacag	aaagttagct	ctagcagcag	tttcaggaag	agggtcttgt	ctgggcccgg	174300
tacctcagga	taaagggcac	ctctgtaatc	agaccagaaa	catccagtct	agcaaaagcg	174360
gtcagtatct	ggtgcctccc	ctagacacag	tgtgggcttg	caataccggt	ctcactcctt	174420
gtgtgtctat	gtctgttttt	aatagttcca	aagattttctg	catttttggtt	cagcttattc	174480
ccagactctt	gtatcatgat	aatagttcct	tttttagataa	atttgaacat	cgggtccgct	174540
gaaaaagaga	acccgttacc	ttaacttttg	cagttctatt	aggattggga	gtagcagctg	174600
gagtaggtac	aggaaccgct	gccttaatta	agaccccccc	aatactatga	agaactacgt	174660
gcagttatgg	atattgatct	tagaactata	gaacagtcta	taaccaaatt	agaagaatct	174720
ttaaacttccc	tgtccgaagt	ggtgctgcaa	aatagaagg	aattagactt	attattcctt	174780
aaaaaaagag	gactctgtgc	tgccctaaaa	gaagaatggt	gtttttatgt	tgaccattca	174840
ggagtaatat	aagattctat	ggctaaactt	agagaacgcc	tagatatacg	taaaagagaa	174900
agaaaaagcc	aacaaagatg	gtttgaaagc	tggttttaata	agtcctcttg	gctcaccact	174960
ctcctctcca	ctatagcagg	acctttaatt	acacttatgc	ttttgcttac	ttttgggccc	175020
tgcaccttta	ataagtttag	agctttttat	agaaaaagga	taaacgcagt	ccaggttatg	175080

gtactaaggc	aacaatatcg	ggccttcag	gaggttgaaa	actcgcctcta	agattagagc	175140
tatctcctaa	aagaagtggg	gaatgaagaa	taaaaatttt	tactgaactc	ttcttcaccc	175200
cagaaccgga	ccctcccat	ctagagattg	ttcccggaac	actcctgaac	tcttcacccc	175260
agaatgcatt	cctgaactcc	tcaccctaga	gttcgaaccc	tcccaactaa	aaactgttcc	175320
tagaacattt	ttgagataag	ggcctcctaa	aacaaccgca	aatgaaccg	ggtacattgc	175380
caataatag	gacatgacct	cttagttacg	tagattccct	tggcagaacc	ccttgtcccc	175440
tgacagaacc	ccctagtgat	gtaaacttgt	actttccctg	cccagctcto	cccccttgag	175500
ttttactata	taagcctgta	aaaaatttgg	ctggctcgtc	attctcctct	acaccactag	175560
gtgcatgagt	ttcgacccca	gagctctggt	ctatgttcca	tgtgctttgt	tgctgtgttt	175620
ctattaaatc	ttgccttcta	cattttgagt	acggctcag	tgtcttcttg	ggtccgcggc	175680
tgtcccgggg	cctgagtgct	tgagtgaggg	tctcccttcg	ggggtcttcc	aaaactactt	175740
cagaggaaaa	atgtattctg	cctcatgggt	tcaggggggt	tccctcagca	aattcagggg	175800
agacaagatg	gaacagctca	acctgctggc	aggagggtgt	gggaaaggac	aagtgttcat	175860
tgtgtgggtg	acaggaaaca	gagagctgcc	tacagtctta	caggcctacc	accactgacc	175920
tacctctgtc	cgtcaggccc	tacatcttaa	aggatctaca	gtttatttaa	agaacactac	175980
cagataggaa	ccaagtatca	aaccaccagt	ttgtagggga	taaaaatata	aggaacacat	176040
ctcaatagga	gtgtgttcca	ggatgtggac	aaggagaaca	cagttgttta	aaagcttaac	176100
gctggccagg	agagctgcac	acctttaatt	ccatcactcg	taagagggaa	gcagggtcat	176160
ctctgtgagt	tcaaggcaag	cctgggctat	acaattctag	attagccaga	gctacatcgt	176220
aggagcctgt	ttcaaaacaa	acaaaaccaa	accataaaaa	agcatttctg	aggctttggg	176280
tttaattcccc	atgacctcaa	atagccaaac	agctctcctc	agtcacaaac	aaactgcata	176340
attggagcta	gtgagatggc	tcaacatatg	aaagtccctc	ccaaaaatat	tgacaactgt	176400
agcttatctc	gtgggacaca	cataatggga	gaggaccaat	ttctacaagt	taccctctga	176460
cctccacaca	tatgcctccc	acaaataaga	aaatatatat	aataaaaaaga	aagaagtcta	176520
cagctgcaca	tggtcatgca	tgcttataat	ccagcactcc	agaggctgag	gcaggaggat	176580
tattagtttg	agatcgcata	gcaagcagta	ggctagacag	ggctacatag	tgtaaacctg	176640
ccttaaaaaca	caaaaatcaa	ttaagcaaca	ataacagtaa	caaccacaa	aaaaacccaa	176700
aagagtactt	tgtagtaagg	acaataccaa	aatgttccct	ttaaggacag	ttctggaatc	176760
agcaatagcc	ttccgagtgc	tcagggatgt	ataaatactt	agaaaaacttc	ccctggagaa	176820
atgagacca	gggtacactg	ctctcagagc	tgcccagaaa	gttgtttatc	ctggattcat	176880
ttcagccttc	ctaactgttc	aggcattcag	aggtcacttc	tgtagttagc	aatgtctaaa	176940
aaggctaaac	tactgctcag	catggctgtg	gtacttgcca	ttatcatttt	gtgactgggt	177000
ttgtagttat	gcagaattca	agagtatatg	catcatgaaa	gtttccacca	agttcctgat	177060
ccagtcacct	cttaaaagtt	ggatgcacca	agtgcccttg	gggtgataaa	ttatatccaa	177120
ataatggtat	tccaccctaa	tccccaaga	cttctggcca	tctcataatg	taaaatgctg	177180
agccatcgca	ccagcccatg	gccttgaact	cttgatggto	ctgtctcagc	ctgtgttttg	177240
attataaatc	tgttggtgag	gtattccctt	gctgataata	caagcaaat	cttcaagctt	177300
ccatcctaga	ctgaagacca	gcagctctcc	aggagtcctc	aatgcagact	ggcccagctg	177360
ggacattgag	cctcatggac	tcagccgcta	ctagattcgc	aacctattca	gacaagccac	177420
tgttggacta	ccagacaaat	actatgtaag	ccaatcccat	tttaatacac	atattcatct	177480
gggtgtgtg	cacacacctc	tactcccagc	acgcaagagg	cagaggcagg	cagatctctg	177540
atttcgaggc	ctgggtctata	gagtgaattc	caggccagcc	agggctacac	agagaaaaac	177600
tgtttcaaca	aaacccaaac	cgtaaattca	ttctatcagc	tctatttcct	tagagaattc	177660
taatacatgt	gggtaccagg	ggttgaactc	aaagtcttca	tgtttacgta	gcaagtttcc	177720
ttctgctagc	ctagtgaagc	tgaggcaggt	acagccgggt	ctttactgct	ccttgcaaat	177780
ggctcctctg	agctttcctt	tgagagccta	caaagaactc	tttttttctt	taggtctcca	177840
ggtttttggtc	ttaagaggtt	ctggacttgg	atctgtagct	gtcatatcac	agacattcaa	177900
catctggcaa	atgtccttgac	aaaggagatc	acttgtgttt	gctgcagtgt	cccttctggc	177960
tgtgagattt	tgtcctcac	cactgcagga	ctgcagatct	attctgcctt	tttagttgac	178020
ttttcattcc	tgagaactgg	ggaaaactga	ctttgtattt	gggctttgaa	tttgtccatt	178080
tgtcaatcca	tcacaccaga	cctaaccac	tgccaagagt	tctgctgact	tttgttttct	178140
ctaggggtggt	cacttttgctg	ggctcatcct	catccttggc	ctgcagttta	tccccaggaa	178200
agaaaatggc	taacgactgc	taagaagcag	tctttccttc	cagaaatttt	agtctatcta	178260
gaccttgctg	cagtctgaag	tctttaaaat	gtgtttgtta	tggtagaata	ttttgagttg	178320
ccttaggagt	attgcttgct	gtcacctatc	atattctatc	aggaagcaga	cgtcccattt	178380
accaaattgtg	aagaaatatg	gcatcaatac	ccactgcaaa	aagtgtaaat	aaataataaa	178440
aaaatagatt	tattacagag	tgcaagggaa	aagaaaaaaa	tcagccagtt	tcagaattgt	178500
aactggacaa	atgttggtag	agttcatgaa	gaggttctac	aaaatggctg	ggggtgggaa	178560
cataatgagt	tagtttgctt	tttttttctt	ctttccttcc	ctttccttcc	cttacaaggt	178620
ctcatgtagt	ctatgggtctc	aaactcacca	ctgtaaatca	ccttgaactt	ctgatccttc	178680
tgcacgctgg	caatgtaagc	atgtgccacc	aggcctggct	cacacatttg	gttttttcaat	178740
acagaatagc	tctgtgatga	ttaacttcaa	tcatcaactt	gacataacca	agaatcgtct	178800
gaggaagagt	ctcagtgact	gggtgggcta	agggcatgct	cataagggat	tatcctgatt	178860
gttaattgac	atggaaagat	caagtccatt	gtgagcagca	acacgcctcg	aacagaagtc	178920

ttctgaagta	taagaggaga	aagcttgatg	agagcaagca	ggcaagcaag	ccaggatcca	178980
cgtgttttatt	ctctgtctgt	tcttgaccgt	agatgtgatg	gctgtcttgg	cttcctggga	179040
aacatgaact	gcacctgga	attgcaaggc	aaacaaacct	ttctctcttc	caagtgctt	179100
tatgctaaga	tattttatcg	cagcaataga	aatgaaactt	agaacaggcc	cataactgcc	179160
agctttgga	ctgaacctaa	ggctgttata	attcactagg	atagggacca	ctggaagtga	179220
atctgatttt	gatggtaaaa	tcatgtgttt	gtttctggat	atgatagatt	tatcaatttg	179280
agactcagaa	aagaagttag	gacttgaatt	ccgtttttaga	gacattccag	agaaaaactga	179340
tgtcattgtt	ctgaatgtaa	gtgcctcagc	tgaataatca	aagagtacag	ggaagaaagc	179400
ccaggctaga	atctgaagga	actcctctat	ttttgttttg	cttgtttgtt	tggttgggtt	179460
tttgagacag	ggtttctctg	tgtagccctg	actgtcctgg	aactcacttt	gtagaccagg	179520
ctggcctcga	actaagaaat	ctgcctgcct	ctgcttccca	agtgtctggga	ttaaaggcgt	179580
gtgccaccac	accaggctag	gaactcgtct	attacacatt	aacacccctc	tttaattaac	179640
tgttcctgcc	aatgtaccaa	atagtcaatt	gattcctgtt	tatttaccac	atgtttctgt	179700
tagtaaacca	gaataactta	tctagccaaa	gtctgcctat	tagccatatt	ttcatcagtt	179760
cccaaccatt	tttggaaatt	tgtgagggga	atccacagat	gctgtagacc	gctttagaca	179820
tttttcagct	tttttcaagt	tgcaggtcat	gattcagttg	gtcatgaaat	taatttagtg	179880
ggttctgatt	agcatttcaa	aatgaggcaa	gcagagggca	tattgtcaca	gcacagcaca	179940
tgcggtaagc	agccacacac	tcttgcttgg	aggcttagtc	agtttctggc	tctaaacgcc	180000
ccagggtttgt	ttctctatcc	taggcctctc	tcttaaattc	caaacatagt	tagacattac	180060
cattggggca	cgtgcaactc	aaacacggag	tgtgactcct	ttcccctctc	gcgggtccca	180120
gaattggcaa	tgtcacccctc	ctcccttctc	cctagggtca	gttttacctc	tcacactcca	180180
caacacaaac	ctctcatctc	caagaattgc	cattagggct	ggtgagatgg	ctcagagggt	180240
aagagcaccg	actgctcttc	tgaaggttct	gagttcaaat	cccagcaacc	acatgggtggc	180300
tcacaacctat	ctgtaatggg	atctgattac	ctcttctggg	gtgtctgaag	acagctacag	180360
tgtactcaca	tatattaaat	aaataaatct	aaaaaaaata	aaaaaaagaa	ttgccattaa	180420
atgtacctca	gagtcctaat	gcttcttctc	cccctgacta	cactcacgct	ggcctgagtc	180480
cattttctta	ttgaggttac	tgtctctctg	cttctacctc	ggctccttct	gctgcctatc	180540
cttgacacag	cagacaagca	gttctttaa	gcagggtcct	ggaccagtga	gactgatcgg	180600
ctctgggtgc	acttctctgc	atctaaggtt	aatcaaggtt	aagcctagaa	cccacgaggt	180660
agaagcaaac	gacctactct	ccaaagccgt	cctctgacca	ccatgtgtaa	actgcacatg	180720
tacatgcatg	cacatggtac	acacacatac	acagaagtaa	aaagagattt	aaattgaaaa	180780
tcatttaaaaa	gaaaaatcag	ggctcagcaa	actttccgtg	tagaaaaacta	gagtacttag	180840
gctttgaaag	ccaagaagtg	gatatttaatt	atagtatttc	attatagcag	agatttctaa	180900
aaacttttga	aaaaactaaa	aaatataaca	gagtgatttt	tttttgtaat	gtaagtttac	180960
taatggcagc	agtgaggatta	gtttcttttt	tagattattg	ttattatttt	tatttaattat	181020
tagtgttttt	gtgtttatcc	atattccaca	gcattgtgtg	ggaattggat	ttctgcttcc	181080
acctttgtgt	gggtctctaga	gattgaactc	aagtcaccaa	gcttgacacg	taggtgggtca	181140
ggctttacaca	gtagggtggc	aggcttggtat	ctttgggaag	caagcatttt	acttctctgt	181200
ccagctcact	ggccttcttt	gtttaaaaaa	agaaaaaaa	agtccttttt	tgtttaatta	181260
gggttctatgg	ccagtgtctc	ttatcttaaa	atcaactgca	aaacttttatc	tggtaaaaaa	181320
ccatccctag	ctgtgtctct	aggagaaaaa	catacagttg	gatggcttta	tcctgcaggc	181380
ttagtttgat	catctctctt	tgaagatata	atcagctcac	atcacactca	agcctctgcc	181440
aacgagtttt	ctacttctgt	tcaacaaact	acccaagctg	agcagctcca	aacaacagcc	181500
agttatgatc	ctcacagtcc	gggtgggtcag	aagcctaagc	gggcgtggct	acctcgctgc	181560
tattgcctga	ccctgctcgg	tgcattccaca	ttcacatcct	ttcctggtga	gtgtggttct	181620
ttgactggtt	ttgttccaat	tttttagtata	tgtgctgctg	aaacaaactt	tttgctctctg	181680
cctccagact	gcagggatta	atgttcttga	ctgccacaga	gcactaatat	ttactgaaca	181740
tgtgatcatg	tgggtgctcag	cactcttgca	cccaaggctc	ggggaacatg	gagggaagagg	181800
gggtggaaag	attccaagaa	ccagagggaag	aagaaagtca	gaggtgagac	tgcattctct	181860
agaaatgtca	gggacatttc	tagacctctg	aagtctcaag	aaacaaggcct	gaaagtctta	181920
tttatatagg	ttaacctgaa	aggggaaaaa	attcttacag	gggtccaaag	ttagacaaaag	181980
aactctaaag	aactaaggaa	tgttgggggg	ggggtagtct	tccccaggga	acactcctct	182040
acccttcaag	ccccacccaa	gctgtgtatc	caaaaacaa	tggtcagtcc	tgaagccata	182100
tacgcacaag	taacatcata	tggatggggc	gattgcattt	aggaatacac	acatacacac	182160
acacaactta	aaaagagagg	ccatgaattt	aagagagagc	aaagcaaaagt	gggaagggggt	182220
acatgggaag	gttgagggca	gnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	182280
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	182340
nagagagaga	gcctattatg	tctgtgtgtg	cttctaatac	ttagaaaacc	actctcttag	182400
gctgagtcag	aactagccta	gctgagacac	tgtccaccca	ctgtcccaga	gcaaggccat	182460
cgctgtccca	gatttgcctt	tgggggcccc	tgaatatgaa	gtcaccagca	ggctccggaa	182520
gctgcctcat	gctaatacag	tgagggttgt	aaaatagccc	tcagtggtt	cctgggcctg	182580
cagctggggc	agagccacta	aggggagttc	ggtccttttg	agcagagtta	acagtcacat	182640
gtgctttttt	ttttttttta	aatgttccct	gcttttaggt	cagtgtctgt	cgctacttct	182700
aatccttgca	ataggctgca	agacaggcaa	gaatatcatc	cctgttttgc	cctcaggcaa	182760

attctgaagt	ctggcacaatg	aaagatgtgg	gatttgaaca	cagacttgtt	tggccaaaga	182820
attctcactc	tacttctgcc	tgtgccacct	tcctctcatg	cacggggagg	ggaggggagc	182880
ccacctccca	tgctcagggg	ctaggaagtg	gggagaagat	ggatgtccctc	aaagcaggggt	182940
gagaatgaag	tagaagccag	cttcaaatct	aaactaagca	atgtttttatt	tccattttccc	183000
tgacaataaa	gttcagttac	atttggttta	aaaaaaaaaa	tccttacaca	actggttctt	183060
gagaaatgtc	aagtgttaca	attcagtgga	tgtggatgaa	acaatcaaaa	tgttgaacac	183120
ccccaacag	atacaaacct	tcatacaagt	ctcttccaaa	ggctgggtct	gaaaagagcg	183180
actcatgttc	cagcccagtt	ggctccttct	catgtgagct	ccgacttcca	aagactgctt	183240
gcaccaggag	gaaatataat	agatgtcctt	tttaaggggg	gtgggggtctg	tctgacaacc	183300
tcccacagtg	actgtggata	cagcccagtt	agtagagttc	tcgcctagca	agcgtgcggc	183360
cctgggcttg	agatctagca	ccctaaggca	tagtggtaca	tgcccatgac	cacagcactt	183420
gggacgtaga	ggcagaagga	tcagttcaag	gtcagatgag	gggtggggag	gcatttcctt	183480
aatgccagca	tttgagaggc	agaaacagat	gaatttgtga	gttcaaggcc	agcctgggtct	183540
acagactgag	ttccaagaca	gccaaaggcta	cacagagaaa	ccctgtcctt	tcaggaaaaa	183600
agatagtggtg	agagaattca	aggttatctt	ggactgcata	agactttgat	tccaaaataa	183660
acaaaaatgg	agcatgaatg	cttgcaactg	tggaacaatat	tggtttcata	catattctgt	183720
tttgtcacct	acataccaat	tatacaaatc	agattcagct	gggcccactg	gtgcatgttt	183780
gttcccagca	tctggggaggt	agagatgggc	agagctctat	gactttaagg	ctagcctggt	183840
gtacaaaagta	agttctagga	cagccaaccc	tacagagaga	tacactactt	ctaaatcaat	183900
caatcaatcc	atcagtcaat	catgggctgg	agagatagct	cagtgatcaa	tagtgccatt	183960
cctccagag	acctgggttt	gattcccagc	acccacatgg	cagctcacag	atgtctgtaa	184020
ctccaatccc	aaggggacatg	acaacttcta	ctggtctctt	tggtcaacag	gcattgcacgc	184080
agcatacaat	atatatatgg	gtaaaatgct	atatatataa	aaatcagatt	cacaaatcaa	184140
gtacagaaga	agatcaacta	taatgaaaca	accataacac	atactgtttt	aaaagctacc	184200
tttctgtct	gacatctgac	ttcttgcgtg	acccagtgcc	tccacaatga	gacaagcaga	184260
tgtgttagtc	cctgtgaccc	aggattaggt	aaccactgcc	cctgggaagc	cattcctagt	184320
aacaggatca	aatcccacag	tgctccactg	taccccgcca	gaggacgtga	agcctccctt	184380
ccccggctgt	ctgtgcagca	ccgccagat	gcttgggtgc	actgagtgac	catcggagcc	184440
caactgagga	gtgccctcagt	gctcctcagg	gcagtggtgca	attgaaaact	gcattgtgatt	184500
tctggaattt	gtcatttgat	atttcagac	tccagctgac	ccttgataac	agcacagagg	184560
gcagaactgc	agaggaaaag	gggcactact	gtgactgtta	tcccctgcat	gattatagaa	184620
gggtctgtgc	tttctcttac	aaatcattgc	ggcctctctt	cacttccctc	ctttgaaagtc	184680
gaaaaaaaata	gatgcatcct	cacagtacag	gatggcaggt	acgagggcggg	ttccgggaact	184740
gaggcgagga	tcattcatgca	gcctctttcc	ctcaactaca	ccgcgcgccc	tgcaaggttc	184800
cctctgatca	aatcatttcc	aggcctggaa	agaaaacggcc	actcaggctg	gggatgtggt	184860
tccattagag	gcctgcatgc	acgaagctct	ggatttgatc	ttcagcacgg	gcataagccc	184920
agtatgttag	tatctgttta	gcattgtgtg	gaggtatacc	tatctctgca	cagggagacc	184980
agaaagttcag	agttatcctt	gcacttacag	taagttcaaa	gctagcttgg	gcaacatgaa	185040
gttttgtctt	aacaaaaacga	aacaagaggg	gccggggaga	tgcttcgctg	tgctgagtca	185100
tttgctgccca	agttttagta	cctgagtttg	gtcccttgag	cctatgggtg	aaggagagaa	185160
atgacgcttg	aactccacgc	acatgccaca	gccacgcgat	gaatgtgtat	acacacacac	185220
actaagtggga	taaatgttaa	aaataaaata	atcaaaaggaa	tgccactca	aaatctacca	185280
tcgttgggaa	gggaggggaa	aaggcaggcg	agggagatag	ataaccctga	tatgaacacg	185340
gaaagagcca	gtgtgcccac	aaagctgccc	agtgtgccac	caaagctgcc	cagtgtgcca	185400
ccaaagctgc	ccagtgtgcc	accaaagctg	cccagacttg	attacagatt	tgccagggga	185460
cacaggaggc	cagcaggagc	agccagggtt	cacctcagag	gtggagccac	aaacctggaa	185520
atgaaaacgtc	tttccctttt	ttcagaccac	agcagtgaca	gctgtcctgc	agagtctgga	185580
gggctgagca	ggctcatcca	ctctagtgtg	cctgtggcca	gaacaggcct	cagtccacag	185640
tgctttttcca	aggtcttagt	gtctaattaa	ggtttagcagc	caaattggag	agagaagggt	185700
gctggacttt	actctgtgtg	aaggactttg	ggcattgttc	cattccgtga	tcaaatacca	185760
ctgggctctg	caaccacat	gtcagtgggt	cttcagaggt	agaagaactc	atcctttttt	185820
gagagggttg	gtctgttctt	tgtctaattg	aaaatgcctg	gggcaccagg	ttaatgtcaa	185880
ctcaaaggca	agtgtgtgtc	cagcatgtgt	ggaatcctaa	gttcaatacc	catcagagcc	185940
ccaagcccta	gaggacagac	atgctttaa	aaaaagtcac	gctttaaata	aattctgttg	186000
agggggctgg	agaaatagct	cagcagttaa	gagcactagc	tgctctttca	gaggaccag	186060
attccgttcc	taacattttc	atgggtggctc	aaagctgtct	ataattcaag	tcctagagg	186120
gaatctgttg	ccctctctgg	ttttctcagg	caccaagaac	acatgtggtg	caaacatata	186180
cgagggcaaa	acactcatat	atatgaaaca	ttttttataa	acctgctcgg	atgtgtgggc	186240
tcattgccagc	gatgctctca	gcactcagat	ggcagaggca	gggggatttt	gtgttgagcc	186300
cagcctgagc	tatagaatga	gatgtgtct	caaaaagaaa	aaaaaaacaa	aaaaaaacaa	186360
aaaaaataaa	caaatctgtg	ggcttaatac	ttcctagcca	gaaggagctg	gctccagcaa	186420
caatggatcc	ctagagctct	gctttgcccc	ctggtctgga	gagcttgctc	tagaaaggaa	186480
ttctccatag	accacatttc	tattttgggg	accactctgg	gcctgcacat	ggaaaaatgg	186540
agagttgag	tttaacatggt	catttttttt	gtgtgtacag	taatgagagc	tttgaaatga	186600

tcacaaaagg	aaaataggaa	aatatgcctc	ctaaaaagag	ccgaggcaaa	ttattatagc	186660
aaccgaatcc	taaaagggaat	gttcaagtaa	aaaaaaaaaa	atatgggtca	tgcgaaagttg	186720
ttatggcaac	tgacatttaa	aagtaacagg	atttggggcaa	gtttctttct	ccaccctctc	186780
tgcaaaagtct	tgacaggaact	tcatgtctaa	attatgcttt	aattttaatt	agccaaatag	186840
gtcataaata	tagctttattc	ccaaatcctt	agaatttcta	ccctgcaagg	agctgaacta	186900
ctaagtaggga	aatattttcca	caaaaaccca	cttaccatta	agaaccccca	tccgatattt	186960
ttctaataata	attttatcaat	ttaaacacat	tgcataaatgt	gccactctgt	agcattccat	187020
taaaatgata	aatagcaata	gtggtgaggg	gtgggggggca	aaaaccaggga	gattaaacat	187080
atccaaagca	gtgtagctat	atttaaatac	ctcaatccat	tgtagaggaa	aacacactgt	187140
tctcatccgc	agatacagtc	tagactcaga	gcagcatatc	cttgactgta	agggtattaa	187200
taggacagac	gaaggggggc	aataagaaat	gacaggaaac	ttcagaagaa	ataaaatttc	187260
tattagcctt	tgttataaga	ttacatcaaa	gcagttcata	tgttttaatc	tggggaggaa	187320
aaaaagcaac	tacttgggggt	ttgcgcctgg	gggctgcctc	tggtactga	accagacagt	187380
ttgcataatg	aacaattttc	attcaatcag	gatctcagca	gagatagctc	ctactcaaag	187440
gaacccggca	caggctcata	gtttttatct	cccagctcca	cctgctggag	aaaccttgta	187500
ttgcaggag	agaaagcagt	cgggaggcat	tgtcctagt	gctgtgtacc	taaagttaca	187560
gacctgactt	taaacagttt	ctctctggag	gttgaaaggg	gctctgtaag	ataccagagt	187620
ggattgtctc	caaagactct	cggactcctg	ttacagycaa	gtaaggtcct	agcagatggg	187680
agcatggatc	tccggccctt	ctcactgctt	tcttgaatca	gggatttaga	aattgctatt	187740
tgcataccag	gaggactgaa	gtttggctcc	cggtgaccag	aggacaaggt	cattgtttaa	187800
aaccacccaa	actcatttcc	gacttggttg	gtcaaatttt	caagtttccc	agcagctcaa	187860
ggatttcata	aataaggcag	aggcagagaa	acggagggtg	tgtgtgtgtg	tgtgtgtgtg	187920
tacccaaaat	ggaactgcat	tttcatgcac	aatagaaaaa	ttaaagactg	aaccaatcat	187980
tttggaaaac	tggcacagct	gacattggct	agaggaagga	acggccaggg	cgagccagct	188040
gcaccaagac	ccagggctga	ggcctaattc	gcctttatcc	gagggtttag	tgaggctccc	188100
gcccgctcac	atcccgctc	ggagcccgag	aagagctctc	ttcacttggc	tcagttccag	188160
cacagtcgca	ctatgctctc	ccgtggggag	gccgctccgg	gagggggagc	gacatcaagc	188220
tttgtgaaac	tgttttcgaa	aacctgggat	gatcatttaa	atgtttaaaa	tatgcacatg	188280
gttaattcaaa	actaattacc	ctgagcacat	ttgaaacatt	tatgccatca	tcttggatcc	188340
tgctactga	ttgtgcgctg	cagctcactc	tggtgtttct	ataaactgct	tcagcgattt	188400
taacttccag	gctaaatcag	gcagccacag	gcgctgcctc	cagccctggg	ttgggtggag	188460
gaccccccac	cctgacttcc	aggcgaggag	gcggcccggt	tctccagaga	gccgtttgtc	188520
agggtcttgt	agttctggct	gccgaattat	tgctcttata	cgtgttcata	attctcatct	188580
gattatttta	atttaggcta	gaatgacctc	tttccctccc	gagttctcct	ccctcattcc	188640
catttctctc	tcttcaattc	gtggccccc	ttttctgatt	ggtccaaata	tatagacaaa	188700
tatccttgat	cgtcccaccc	cacttggcta	catcttcata	tgggagccaa	tgtgtgtgag	188760
tttctgggtt	tgcaggtggg	tcaggtccac	cagtcacctc	aaggtgtgtg	agagaggtag	188820
accaacatga	gcggcgacac	gccgccatca	ctgagagagc	acgtgcctcg	cagctcaggc	188880
acaggctcag	acacaccggc	agacatgtgc	acatgcgctt	tccccagcaa	accctgcttg	188940
cagagtaatt	aggcctaggc	agttcctgaa	gcaaatcat	ttcccctttt	tccagaataa	189000
aatgagttct	cttcccttgg	gggtgctaaa	ccagcatgcc	agtggctaga	agcctgagat	189060
gggtgagtgt	gctgaaacca	tttctgcagc	caagcctgtg	ggcagaagct	aaccttgggc	189120
tggggagctg	cagtcggaag	aggcacaatt	ctgggatcaa	gaaatgagca	ctgggtttata	189180
ggtacactcc	cagaaaataga	cagatgaggg	ctgcctcctt	attagcgctt	tgaagatgcc	189240
catggcggggt	ttttagacat	ttagggaatat	aaaagttaggt	tggattccca	cagtcagctg	189300
aagtttgaca	gagtgatatt	accgggttta	actagagcca	ttaagagact	cttcattatc	189360
ccacaccacc	gccacccaag	ttatcacatg	agccataatg	caagagaatt	ttcattccat	189420
caacaagaga	gggagccggg	ctatctttgt	ccaaaggaaa	tgagcagccc	agcgtgaagc	189480
ttgtgaggaa	ttgagtgtac	aacactccaa	taacatcccc	tgaggatttg	cctctgcgat	189540
ttagtccgtg	aagcaggggt	aactgcgctc	gagcagtcct	cctgtgtacc	tggttgcaa	189600
gaacaccagc	tcgagggaaca	ccaaaagggc	cgattaatga	caaaggacac	tcatagaggc	189660
ccgaattcca	cagggtctaa	gtattaaagg	ccaaagaaat	caaggtctag	gccattctcc	189720
tggcgctcag	caatctcatt	tattattttc	ctacaaagat	ccaacactca	atttcccagg	189780
tatccctctg	atctgactca	cattctcctg	ctcagtaagc	catcctgggt	tgaaaacggg	189840
ctccctcctc	cctgcctatg	catgctttgc	gtcttcacaa	cgacagctgg	taatttgcaa	189900
gacccctctc	actggactct	ctcacccac	atacttgtaa	ctactccttg	gaactacttg	189960
tttatcaagt	gttctgttgg	tgagccttct	cttgcattaa	agctgtgaga	aggaaccaca	190020
gtttctattt	cctttacatt	tcttgtagcg	tctcacatgg	gagacaccca	ggttagatat	190080
actgagggtc	ctggtagtgt	tagagttgga	gttagatgac	ccagcaacat	gccttcccc	190140
accacgcacc	aagcaaaaat	tgacccacc	cttccctcag	atgttctctg	catcttataa	190200
ctgcgccaaa	gccagattta	ttgctcctgc	tgtaaaagtgt	atcttctcta	agcctcactt	190260
aaaagctacc	acttggcaga	agatcaagtc	agaagtgcag	gctagcaggt	gacggtgagg	190320
acagggcggg	atggggcggg	taggggtggg	cgaataattg	aagctccaag	agttaccagc	190380
tcaatatatta	acctaactgg	taatttgctg	tgacaattac	gccatgaagg	gaacgctgcg	190440

actatgcaag	aatgttgctc	tctaattaag	agggctctgc	atttcctagt	cacccgcact	190500
ttaataacac	acagaatgag	ccttggtccc	gggagctaaa	ggttccatta	ggagcacggg	190560
cagcatatgg	ctgtgcacat	agcccgtag	tgatgcagcc	cagttaagcc	cgctaaccac	190620
ttcaattcgt	cctcagatag	agcccagaga	gcgcggctca	ggccctcacg	ccacgagccc	190680
catttgactg	acaggcatct	tcccggaaag	cctgcgcgtg	cctacactgc	aaatggacct	190740
gcttcccaca	gcccggcttt	caaccaggaa	ggcttgcggt	gggtctgac	cttcaagagt	190800
aactttaata	aggattttct	cacagaaaga	aaagtccatg	ggaacaaatc	ctcctcttaa	190860
gagcgtgaga	caggaatggg	gacacaagcc	aacaccccaa	ttgctaggct	aactctgata	190920
tgagacaaaa	gaatattaat	atcttggtca	tgaaggagga	tggtgccatc	ttctgaattg	190980
atgggagttt	tgaggcatgg	ctaagctggg	caaaccattt	tctttttttt	ctcttcttaa	191040
ttagtgttc	atttatggag	ggcttgctgc	ccggagagcc	catcagaaga	gagctcgctt	191100
tatggagatg	tagcttataa	aactactcag	attttaaaca	aacagtgcag	gaggccagag	191160
gtagaagtgg	tgggggtggg	gtggggcaag	agaacaattg	catctgcaga	aggctagccc	191220
tgaccccaa	gcctatgttt	agggttgatc	agcttcccga	ggcaagccca	gaagcctcta	191280
aaattttagg	ccaatagaaa	tgacctctgc	accacggctg	actgaagcta	taaataagcc	191340
tcgagttgag	cagtggtgtc	aacggagaga	gcagaggaaa	gtccaatcag	agcttcattt	191400
ttttttttta	aaagtcactt	gcttgggact	cacctgaagg	cagggcattg	agtagagcct	191460
tggtccctcg	cagcgagagg	ctccagtttt	cccaggcacc	agcccatcgg	ttggttaact	191520
aaccacgaa	agggaactgc	acagcacaca	agttaaatat	aggctgggtt	atctgcattt	191580
tacaagctct	gagcaagcta	tctgaagaag	ctgtcatttt	taatgacggc	acaaacttcc	191640
aattaccgac	tgggtaatcc	actagggagc	aggtagtttt	ggaagaacag	ttcaccatta	191700
ttaaaagtgt	acacaatcac	ttttgagtgg	actataagta	tttcacacga	ggcagggtgg	191760
attagggact	ttttgggtgg	tttactcgag	gctgcaacca	acaatgagtg	ttttctcaag	191820
aattatacat	tgagatttgt	caactgctgg	ggagtagtgg	agggtcctgg	taatgcagaa	191880
aggttatgaa	atggccaggt	aaggttggtg	gcttccaagt	ctcaaatata	ctcctaaggg	191940
cagctccaag	tcataagctc	aaacaagtct	tcaagggggc	tggagagtta	agacaaataa	192000
ggatcactta	ggctaccac	ggacaagcac	ttctcatata	aggaccggct	acctccaaca	192060
ccatcttccc	aaatggctt	ctatgttgct	tcaacaacca	gggcagggtg	aattagggtt	192120
gggtctctcc	aatgtggact	caaatcatga	ctacagcntg	gggttttttt	tttttttttt	192180
tttttttttt	tnnttggttt	ttcgagacag	ggtttctcca	tatagccctg	gctgtcctgg	192240
aactcacttt	gtagaccagg	ctggcctcgg	actcagaaac	ccgcctgcct	ctgcctctgc	192300
ctctgagtg	ctggaaatga	aggtgtatgc	taccacgccc	ggccgagtc	gtcttgataa	192360
tgaagtcccc	agtgcactgg	atgtcaactg	aagttggatt	ttactgtgat	gactactgag	192420
tccggctcag	aattttgggg	ggacaaggta	ccttgattta	actgggcact	acacgactgt	192480
aacccccaca	tggggagagg	cagaggcaga	ggcagaggca	gaaagtgggt	tggaggctag	192540
gcaaggctac	acagcaagaa	gctgtctcaa	aaccaaagac	atctttcttg	atccaaatcc	192600
tgtcggaggg	tgtgaggcct	tgggggccag	aacaaggtag	tcaaggaa	ccactgactc	192660
tgctccttgg	tccattactt	aatcagaatc	gccatcacag	atatagctag	gagattttta	192720
gccttggtgg	ctgcaatctg	catttaagag	ctaagtggga	taaactcagg	ggtgggcccc	192780
atgcctccct	ccccaccctc	cctgcacccc	tccatttacc	tggttccagg	gatctgctta	192840
atttacctgc	cagcctttgg	tgggacacag	gcttagtggc	ttagcgtgct	tccggggcac	192900
agagaccctc	acagaagcac	ctgaatgtac	tttcagcgct	gcagagcacg	cacggctcag	192960
gcccatacaga	agaacccagg	cttatgctaa	ggagccagaa	agtagaagca	gctggcaaga	193020
gtgattcagc	cccataaatt	tacacatccg	tacagccaaa	cccacttgaa	gtgatccaga	193080
gccactttta	ttgaaataga	aaagatgcct	attctggagt	gctaagtggg	acaggagggt	193140
gggtatataa	gagataatcc	catgttgtct	ttgatgtggg	gctaggggaga	taaccacagg	193200
cctcacgcct	gcctgcaagg	tagccaccaa	gccacaccca	caacctctat	ttatacacac	193260
actaagtgtg	gaggtatgga	taaaaaaa	tgtcccaaga	cctcacgaat	ctgcaaacat	193320
ggtgcctggt	tgggtggcacc	gtttggggag	gcagtggacc	atttggctct	gcaggaggaa	193380
ggttatgtca	tgggtatggg	ctttgagagt	ttgtagcttt	gctccccttc	cagttaaact	193440
tgctctcgta	aggttcctgc	cacctgtttt	cctctgccat	tatggacacc	tggtcctcta	193500
gaactgtaag	ccacttactc	tcaggtctct	ttcagtcctg	gagtcttata	atagcaatga	193560
aaagtaactt	gtgtggcagc	cagctaagca	agggctgtgg	ccgactgctt	gggattatgg	193620
ttgtgtctgt	ctgtctgtct	gtcattccat	ttatatagtc	ctgagaattg	aaccacttta	193680
ccactgacat	gtctcagtc	tcttggtatc	atatattcac	ttaagacaag	atctcattaa	193740
gtcattcaga	ctggttttga	gcttgcaatc	ctcctgcctc	tgccctcaag	cgataggatc	193800
cctagggtac	tgcaccagac	tgggagtagc	aggttctgtt	ctcttagctt	tctacagtga	193860
ttgtggatta	tttgtgtata	aagatctgat	ggcccagacc	actcccttcc	ctttaagtga	193920
acatcaacag	tatttagcat	caacttaata	aactcatttg	gtaaagccat	ctccccacct	193980
cttgaaacaaa	tgaanaatcaa	acagcagtac	ctgttctcct	agagcagcgg	ctctcagcct	194040
tccggccttt	taatacagtt	cctcgtgttg	cggtgacccc	ccccccccc	agccgtagaa	194100
ttatttccatt	gcttaaccag	agttaactgg	aagggttaat	aataaaacca	gtctgggaga	194160
ctaaggttac	caaaccacgc	taggaaggag	aggaaagggc	cactcgaca	aacctgtctt	194220
tgagatgaag	aacaatcaac	ataacaggga	cagagcagtc	cttctaacaa	gtgcaaaagg	194280

gagagagagg	ctgagtttct	acttctataa	ataaacccctt	ggcaggcgga	tcactaaagg	194340
aacacaagtc	aatataaaac	tttagacatg	gggctgccaa	acttcacttt	tcgacagtat	194400
attaatttatg	tagtcaatag	ccatgggttt	cattagcgta	ttaaataacca	cgatcaatat	194460
tattttatact	tttcgaagac	aagccactca	gggaaaaaat	ggtgggggga	ggaggaggaa	194520
caatttgacc	ctgtagtcca	aaaaaagtca	gaacagcaca	ctagagatta	gcaagggttt	194580
aatggaaggc	ataaaacact	ggaaatatgg	acagaaatca	gatccctgcc	ttcatttttc	194640
tgcccttttac	aaagagactg	gagggaaattc	agaaactatt	taaaataaag	gcaaaatgat	194700
tagagccctc	ccctcccttc	agctgcttaa	cactgggggtt	gtggtggacg	caaaataagc	194760
attgagctct	aagtgataga	tgagaatcag	aacaggaaca	gtgtttttga	ggcaaaatat	194820
gtccaagaga	attcaaaaga	ctgtggggcca	gaatctactt	aggcagtcct	ctgggacccg	194880
aatccctcac	aggcgtaaac	agtggaaacca	atltccaagg	cagccctgct	ggtgatctga	194940
tttttgagta	gggaaatctg	ttaaacatcg	tcccacgagg	gagccagctt	ctttcactcc	195000
ccacgggttt	ctacatgcag	ctgtgctaga	tctgtctgaag	tggccgggtga	ggaggtgtgg	195060
ggattgggtt	agcgacctca	gaggacattc	ttgttcacta	gccctcgtgc	actggggcgca	195120
tgaccgaatg	ctgtgagcag	gagatatcaa	aggccggcta	ctggactgaa	aactagatca	195180
ccatctctaa	cctgcaattt	gtcaatctca	gacagcaatg	aagactgtga	ttttctagtc	195240
aacgctttgt	aagcaaggte	agatagaggc	tccataaaaa	ttgttcaggg	ttcaggcaga	195300
gaatcaagt	taactcaatc	cctatctcct	gagattaggg	aagggaagga	aggctgtgtc	195360
tactaaacca	gtgagcctca	agcaaagcct	gtctgttctc	agcaagggtga	gccacccacc	195420
aaagatgcc	acagctaagg	gccagggatg	tagtgagagg	tgctgtgata	tcaacagctg	195480
ggagacagaa	acaggaggat	caggacttca	aggtagtttg	ggctataaaa	tataagcttg	195540
aagctaccca	cttgaagact	gtccccaaca	aaacaaacaa	gctgggtatg	gtggtcgatg	195600
cttgatattc	tagtgtatga	gacgaaggaa	gaaagctcaa	gcccagagacc	tgcttgggtt	195660
acatagggaa	gatgtgctc	caaaaaacag	acagccgagg	agcagacaga	cagggcagac	195720
aggggtgcatt	gactctaaatc	cacatacctg	gatttaaaagt	aatatctggg	agactgggtc	195780
gtgaggggcg	ttccagagat	ttataaaaga	cccacccctga	ctgagtatgg	gcaacaccca	195840
tgggtggccc	aggggtccag	actgaataaa	gggaaaaatgg	gagggaagttc	agctggtagt	195900
gtttccaagt	attaggacca	cagcctggcc	cctggcatgt	gctggccagc	tagtctagct	195960
ccagtatcaa	gcttcaggcc	agcggcaggg	cactggacag	ttccacacac	cgacacacac	196020
acacacagag	cactagcatt	cacctcctgg	tctcttcttg	acaacagata	aaatgtaact	196080
ggctgccaca	gtgagagtcc	cctaccttcc	tcaccgttaa	ggatggtaca	aactgtgagc	196140
cagcagcagc	catttctcca	gttaacttgc	ttccacagat	actgttatag	cactaagaaa	196200
agcaactgaa	acatgggggtg	ctgtgacccc	ttggcaccac	aaagccatgg	caagctgaag	196260
tgcacacatc	acaggccagg	cctgaagatg	ctgggggact	gcaatgctgc	ctggattctg	196320
gcagagatgt	gcagagatg	ccaagagggt	ggctgcagca	accagagata	attaatatga	196380
ttagggaacac	actgagcagg	catgctcttg	ccgaatgaaa	agcctcgag	tgtaatgact	196440
gttttcttcc	tcgatcacgg	tctccacggt	tcagagttgg	cttggtgtta	ggctgccggc	196500
taaacatcaa	tcacaacccc	aggggccaga	tcatcggtgt	tcctgggctc	aatcgcttt	196560
ccttttgtgt	tttcattcat	ttaaagatgc	attccagggt	tgcaaacatt	agtgagaatc	196620
atctccaggc	ctcagttctaa	tctctgagtc	tgtaatgagt	taacatcttt	ccctagtgaa	196680
tatttattat	gaaggctaat	taattgtctt	ccagttacaa	gaatccttta	cagtcaaaga	196740
aagttaggatc	caaaaagata	tactgtttat	tcaaaacaa	caaaggaaac	aaagcttctt	196800
tcttaaatcc	tatttaacat	agctttaata	aaggtacaca	ggtcggcctg	gcaaccgaac	196860
ggtaactgat	gcaaaactgaa	gccatgctct	gtagcagcct	ggatgtccca	gtgccacctc	196920
tgtctgcagg	ctttgtcgga	tttactaaga	ttctgttatc	ttcaaacagg	gattgtgtct	196980
caagtaactg	acccactat	gtggataatg	aagtaaatga	tgcaatttgg	gggtttgctt	197040
ttccccaagg	ggacagcaag	ccagtgcctt	tcagccgtcc	tcagaggaga	caattctgat	197100
taatatcaga	gtcatctgac	tcagtctatt	aaacctatca	aacctgaa	gaaggatatt	197160
cagatattaa	cgataggcct	ttgattaata	attctacctt	gttgccattc	taagcattaa	197220
caaccatgca	gtaactctgc	aaaacagacc	ccttgattcc	aggcagacgc	accctctgaa	197280
cacctgggtt	ctccctact	ccttccccc	caggaggaac	tcaagacaaa	aagtgccac	197340
cactggaaaa	gcacactcca	ggttacataa	tttgccctcat	tatccagagt	gggtttaatg	197400
acttgtgaca	taatttctgt	ttgaagataa	caaaatttca	tgaaatccga	caaagccgga	197460
aggcaggagg	aggggactgc	tgccacacta	ccggtggctg	agaactggag	cgggaagggtc	197520
acacacagcc	ctctgagctc	actgtctttg	cttatcagtg	agtcceaaga	ggggcccaga	197580
tgggttgcca	gcctccocta	gaggatcttc	attgtggagc	tgtcccatgg	ggcgggaagg	197640
aagccattct	atttctgttc	ttctctcttc	cgttctggcc	accagtggta	cttgctccca	197700
tcacatgttc	ttcctgatgt	tcgcgatcag	ccgtctgcca	tagtctctga	agtcacaggg	197760
cttcacgtcc	atcacagtgg	ccttaattcg	agattcatcc	tttagaaaag	agagaagctg	197820
tttgtgagtc	gcagagccctg	gcgtgcagcg	gaagagagaa	ctttctttgc	ttcagtggtc	197880
tcaatgagtc	tcacaggaag	aaggaaagtt	tacaagtctc	agagagaaag	tgctgtgact	197940
tcctggagtt	gggccagatc	ctcttccaca	gaccttttcc	ccatcctagg	tgccctgtgc	198000
tcagacctag	catcctccc	gagaagcctc	tgtctttcta	tgggtgcagt	gggggcccag	198060
agcagacagg	taactcacc	taaagcatca	ctttcatcta	gaggagctct	gtggtagtag	198120

ggactgaggc	ttctgctcca	gctctgggca	aggttacttc	tctgctcttc	accattcctg	198180
tccatcccag	gaagacagaa	aatccctaca	ctctcccttg	atctacccga	ctttctgaca	198240
ccagccctacc	tatgttcatt	taatacaaca	actaaaatat	ctattcacag	gcactaagct	198300
ggtgataacg	cagaatgcac	aaactctcg	gctgcagggg	agacggcaga	gttcctcctc	198360
cacttgctctc	cttgaactaa	acagtgctct	tgaggcagaa	caggggtgaca	cctagggaca	198420
cacaagtcta	gctggggggc	ttcatgcttc	catgtgctta	gtaattaatt	actacatgca	198480
ccgtgtgtta	caagtatggt	taggagcccg	actgctcggt	ttggcctctc	gcctctgcca	198540
ctcatggct	ttaggttcag	agtcattctc	tgcatgcctc	tgctgtctc	tcggttggtg	198600
aagcttgcaa	caacagctcc	aacacagaaa	gtgctgtgag	ggctgacagt	ggatagatgg	198660
ctagatagat	ggggcaggac	ggactgtcca	gtaagcaggg	ttcatcatgg	ctatgcagct	198720
ctggacatca	ggattagttt	aaacacttgt	caggtggggc	acttttacc	gcacgtgcta	198780
tttgtttaat	attctgagtt	ttagaacctc	aactgtggga	aacaagagtc	cacacataac	198840
annnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	198900
nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	nnnnnnnnnn	ngccttgaa	198960
ctctgctgg	ggttacaggg	ttgagccacc	atgacagctt	tagcaatagc	tttgtaaatc	199020
cacagtgtca	agctggatat	gatggcacat	gcttgcaata	ctaacctcca	aagattccct	199080
gaactcaggt	tggtgaaata	gtccaccagg	taaaagagct	tgctgcccac	gcctgaatct	199140
gattccctgg	tcacatgctt	taaggaaaga	acttgccgaa	gatgtcctct	gagtgccacg	199200
tgtaccgatg	catgcttgca	gccaccacac	caccacacac	agtgcactgt	ctcacacagt	199260
gagAACagca	agtgaacaaa	caaacagcc	gggggggggg	ggattgtgac	cagaataatt	199320
gagggggggg	gtaaagctct	tggcaggtgg	ctggctcctg	gtaacactcc	ataagtgggg	199380
aagtccacac	tgtaagggtca	tgtgatcgag	tacatctggg	cctccaacag	tccttgngga	199440
agaAACagat	gcagtctgtc	atatctaaac	catgtgtgtc	gtatctctgg	gtagtctctc	199500
ttttctcctt	cctttctttt	ttctctccct	ttctctttta	aaaaattatt	tattttattat	199560
tatatctaa	tacactgtag	ctgtcttcaa	acacaccaga	agaggggtgc	tgtcagatct	199620
cattatggat	ggttgtgagc	catcatgtgg	ttgctgggat	ttgaactcag	gaccttcaga	199680
agaacagtca	gtgtctttaa	ccgctgagcc	atctctccat	cccccaaccc	ctttctcttt	199740
tgagttaggt	tttgtgtagc	cctgggtggc	gttaacctaa	ctacactggc	tttgaacttg	199800
caatgatgact	ctgctgatac	tgtcttaact	attttgagat	agggactcac	tacatagcct	199860
ttgtgtggtc	ggaactaaca	gagatctgcc	tgtttctgcc	ttgcaaatgc	tggaataaaa	199920
ggtatgtacc	accacacctg	gagtttaagg	gtttttgttt	tggttggttt	tcgagacagg	199980
gtttctctgg	ttgtctctgg	ctgtctctga	actcgtctcg	tagaccaggc	tggtcctgaa	200040
ctcagaaatc	cgctgcctc	tgctctccaa	gtgtctggat	taaaggcgtg	tgccaccacg	200100
cccagtttaa	gggttttttt	ttgtttgttt	tttcttgaga	caggggtttct	ctgtgtagct	200160
ctggatgttc	tggaactcac	tctgtagagc	aggttagcct	tgaaactcaga	aatctgactg	200220
cctctgcctc	ccaagtgtcg	ggattaaagg	cgtgtgccat	cactgcccag	tgattttttt	200280
ttttttaatg	tgtgtatttg	tatgggtgtg	tggtgtcttg	tggaagccag	gtgtcagatc	200340
cccagagcgg	aagtgttctt	aaccgctgaa	ccatctctct	tccctcttcc	ctaaactctga	200400
ttttaaaagg	accaaaactct	taggttaggag	actatacaca	cacacacaca	cacacacaca	200460
cacaccctga	cacaccctga	cacaccacat	gaccatgcct	gagcacacaa	gtgggtttat	200520
tgctgggtcg	gcctgtgtat	gagctggaac	caaaaccttt	gtcgggagat	cgcagctctg	200580
cagtttgagc	acaggctctc	tggtttctgt	tctctgtcct	gtgtcgcac	ttgactagag	200640
gcagagaagc	atctgcaagg	ctgtgaccac	gctggctggt	gctctgccat	ctacattttc	200700
aacaggaaat	ctcaggagag	tatttccttt	taagaacgcc	agacttttgt	gcctgggcca	200760
cttctctact	tcccagaaca	ttgtgtgcca	agtggaaggt	tattaaacaa	gtgcttttga	200820
aaattaaact	ccttggtttg	cagagtagca	tgggagcatt	gagaggggtg	atgcctaaag	200880
gcctgggtct	gctgctggca	gagctgacac	ttggctaaag	ggctggcatt	tctgagatga	200940
gcctcactag	atccgcgtct	cagagctctc	aggagaaatc	agagagggga	gaagggtccag	201000
tggcctgttc	aggatgatct	tcctctgcat	ttaaaggcgg	ctgggttgcc	cacgtagccc	201060
cagaaccaaa	cgagcctcgg	acgaagcccc	ctaaaggcag	taggagagac	tgagccttgg	201120
ctcttcagca	gggttgggga	caagagcaag	aggcgggac	tcgcccggcc	ctttagagac	201180
acgtgcggtt	gtttccgtgt	ctgggagatc	acatgaccgc	catcagctga	cccgctacgg	201240
tggagctcag	cgctgggtgt	tcgcgctccc	cgccctgtcg	cgccccggag	cgaggagccc	201300
tgcggagggg	taagaaaacc	cccaggtttt	ctttcctttg	tcgctgggtc	gcgcagctac	201360
ctgcacccta	cccccgctc	ctcgttccat	ccagttcttc	cggtcctggc	ccccggaagc	201420
cactgcgagg	agggccgtgg	ccaggctcag	ccttgccgtg	ccccaggcgg	gccaggacca	201480
aatggcccag	gggagcagaa	ggcggaaggt	ggttcttaca	gcagggtccg	agggctggtc	201540
cccttcctca	ggacctgaca	tggaggagct	gctccggagc	gtggagagag	atctgaacat	201600
tgatgcccg	cagctggccc	tggcgccggg	gggcactcat	gtagtggccc	tagtgctcac	201660
gcgttggtct	gctagtctcc	gggagcgccg	actgggaccc	tgtccccggg	ctgagggcct	201720
gggtgaagca	gaagtcagga	ctttactgca	acgttcggtc	cagaggctgc	ccccaggctg	201780
gactcgagtg	gaggtgcatg	ggctgcccga	acggagactg	tcctaccgcg	tggttgaggg	201840
cgtgcccttt	gaggaggggt	cctgtagccc	tgaactctc	actcggttca	tgaggagggt	201900
ggctgcccag	aattaccgga	acctgtggcg	ccatgcatac	cacacttatg	gacagcctta	201960

cagccacagc	actgccccct	cagctctacc	tgccctagac	tctatacgac	aagctctcca	202020
gaggggtgat	ggatgcacct	tcttgccagt	gggtgaatcc	atcccatgtc	tatcaaatgt	202080
cagggatggg	ccctgccccct	ctcggggcag	ccctgacctg	cccagccttt	tgcgagctga	202140
ggctttgtctg	gagtcgccccg	agatgctcta	tgtggtacac	ccttatgtgc	aattctccct	202200
gcatgatgta	gttaccttca	gccctgccaa	gctgaccaac	agccaagcca	aggtgctctt	202260
tcttctcttc	ctgtttctga	gggcoatgga	tgctgtcac	cgccaggggc	tgccctgtgg	202320
ggctctgtct	ttgcaccaca	ttgctgtaga	cgagaagcta	tgcaagtgagc	tccggctgga	202380
cctgagcgct	tacgagatgc	cttccgagga	tgaaaaccag	gagggctctg	aagagaaaaa	202440
tgggacaggg	attaagtctg	aaaaagagg	ggaagggaga	actgagtgct	ccacctgcca	202500
gaaagaactt	cggggccttg	tgctagactg	gggtccatgg	cgaatcagca	acttccacta	202560
cctatgcag	ctgaatcggt	tgccaggtcg	acggcagggg	gatcccaact	atcacccagt	202620
gctgccctgg	gtggtggact	ttaccacacc	ttatgggctg	ttccgagacc	ttcgtaaatc	202680
caagtctcga	ctcaacaagg	gagataagca	attggacttc	acctatgaga	tgacccggca	202740
ggcattttgtt	gcaggtgggtg	caggaagtgg	ggagccaccc	catgttcttc	accacatctc	202800
tgacgtgttc	tctgacatca	cgtactatgt	atacaaggcc	cgctgcacac	cgcgctcggg	202860
gctctgtgga	catgtccgag	cgcaagtggga	acccacagag	tatcctgcca	ccatggagcg	202920
gatgcagacc	tgacacccgg	atgagtgcac	acccgagttc	tacacggacc	cctctatctt	202980
ttgctctatc	caccctgaca	tgcccgacct	ggatgtgcgg	gcctgggtga	gttctaaaca	203040
ggaattttgtg	gctgcccata	gagccctcct	ggagagctgg	gaggtgtccc	aagacctgca	203100
tcactggatt	gatcttacct	ttggtacaaa	actccagggc	aaagaagctg	tgaaggagaa	203160
gaatgtgtgt	ctgcacctgg	tggaagctca	cacctatctg	accagctatg	gcgtgggtaca	203220
gctatttgat	cagccacacc	cccaacgect	ggctggatct	cctgccctgg	cccctgaacc	203280
tccactcacc	ccccggctgt	tggtccagcc	tattcgggag	gccacaggcc	aggaggacat	203340
ttcaggacaa	ctataaaatg	gtgcgggcag	gcttgtcgta	gagggcactc	catgtgagac	203400
tggtctgact	agagataggc	ctgggacagg	agaagatgat	ttagaacagg	ctacagagac	203460
tctggattcc	atctccctcc	ccgggaaagc	aggtgaccag	ccaggctctt	cctccagtca	203520
agacatcact	ggcctgttgt	ctttttctgc	accctcgggg	tctcgaccag	gccgtaggag	203580
caaaagctgc	gggttggaac	ctggggaggg	tgaaagaggc	aagattgtcc	ttccagaggg	203640
cttcagctcc	atacaggcct	tggaagagct	ggagaaagtg	ggtaacttcc	tgcccaaagg	203700
cctagggagc	cagttggagg	agcctgaaaa	gcctcacgcc	cagccacctg	tgccacctga	203760
gagcctcttc	catcgagaca	tgcaagtcct	gggtgtcctg	ttggctgaga	tgtgtgttgc	203820
caccagggtc	cggatactgc	agcctgatgc	acctttgtgg	gtacgctttg	aggctgttcg	203880
gggtctctgc	atacgccact	ccaaaggacat	ccccgtgtct	ctgcagcctg	tgctagacac	203940
actcctacag	ctgagcggac	ccaaaagtcc	catggtgtcg	aagaaggcca	agctagaccc	204000
actgttttgg	tataggccgg	tttcccaggg	attaccccca	cccagccacg	cccagctcct	204060
cagccctctc	agctccgtgg	tccctctccc	tccatacttc	ccagcactgc	acaagttcat	204120
tcttttatat	caggcccggc	gtgtggagga	tgaggtccag	ggtcgggagc	tggcgtttgc	204180
tctgtggcag	cagctgggtg	cggtgttaaa	tgacatcact	cccaggggct	tagagatcct	204240
cctgcctttc	gtgctgtcgc	tcatgtctga	ggagcacacg	gctgtgtaca	cagcctggtga	204300
cctatttgaa	cccgttgcca	aggccctggg	ccccaaaaat	gccaaacaag	acctcctgaa	204360
gcctctctac	ggtgcctatg	agagccctcg	ccgctgcac	ggcgccttct	acctgtacac	204420
cgactgtttt	gtggcccggt	tggtggtgcg	gctgggcttg	caggccttcc	tcacccacct	204480
gctgccccat	gtcctccagg	tactggctgg	ggtggaggct	tcccaggagg	agggcaagg	204540
cctggtcggg	accactgagg	atgaggaaa	tgagctcccg	gtgtccgggc	ctggctcctg	204600
tgccctttgg	gaagagattc	agatggatgg	gcagccggct	gcttcctcag	gactggggct	204660
cccagactac	aggtcggggc	tcagcttcca	tgaccaggcc	gacctgccgg	acacggagga	204720
cttccaagct	ggactctacg	tggtgaatc	tccacagccc	caggaggctg	aggccgtgag	204780
cctggggccag	ctgagtgata	agagcagtag	cagcgaagcc	tcccaggggc	aggagagggg	204840
tggggatgat	ggcggtgccc	ctgcgacaaa	gaacagcgct	aagtcagggg	acagcagcca	204900
ggacttgaa	cagagcgaa	gctctgagga	agaggaggag	gaggaaggct	gtgtggtgtt	204960
ggaggaggac	caggaggatg	aagtcacggg	aacatccgag	ctcactctgt	ctgacacgat	205020
gctgtccatg	gagacgggtg	tggtcctcgg	tgatgggaga	gacagagaa	aggaagagga	205080
gccgtgaca	gagcagacag	aaggcaaa	acaaaagatc	ctccttgggt	agcccgtggg	205140
ctgagggggc	atgggtcagg	tgcttttctc	tcaggctctc	atatgctggg	tgtgggtcca	205200
accagatcca	ctgtagcacg	cacagccaca	gtcagacaca	gtgcatggaa	tgtggaagtg	205260
ctgtgtgtga	gtggaaagtg	gggcttagat	ttagctttca	ggagacagaa	agctccttta	205320
aaagccatac	cttgggctga	ggctgggagt	ggagttagat	ggtagagcac	ttgtctggtg	205380
tactcgaggg	ccctgggtgt	cttatctcta	gccccagaag	aagtattaa	aaataaaagc	205440
aagtgggtgt	tgagatgtga	atggagccag	aactggccgg	aacagtcggg	tggaagtggg	205500
aagagtgttc	cagacaggga	acagtgtgtg	tgtacctctg	aggctctcat	ggttccatca	205560
gagaggcagg	gaaaggctaa	aatgggtttc	ttaagagagt	ccagaagggc	tggtcttggt	205620
ggtgcaggcc	tttaatccca	gcactcagga	ggcagaggca	ggcggatttt	tgagttcgag	205680
gccagcctgg	tctacaaagt	gagttccagg	acagccaggg	ctatacagag	aaacctgtc	205740
tcgaaaaaaa	aaaaaaaaaa	aaagagtgtg	gaagggtgga	agccaggggc	aagtctgtac	205800

aagaaggaac	ttgggagcat	tgcggaag	atgacctctc	tgaggtcct	gcccaggag	205860
ccagtttctg	gggaccttga	ccatggctag	gtgaatggac	ccaggatggg	atggtcaggc	205920
ttgctagcag	agccacagcc	gagttggctg	ggtgggggtg	ggtgggggtg	gaagggtgag	205980
ttatctgatg	agctcaggac	cttttcctgc	cctgcagata	cagcctgcaa	gatggtcceg	206040
tggtgtctcg	ccaagcttgg	ccccacagta	gcctctcgcc	atgtggcccc	gaacctgctg	206100
cgctgtctga	catcttgtta	tggtggtaag	gtctgtgggt	agtgtgtggg	accaggttcc	206160
ccagccaggc	ttctgcccc	ccttagccct	ctctaggcga	ctccttccct	aacttcccag	206220
cactccctga	gcagggcctg	ggcttcaccc	attaagctgg	gttttcttgg	gtaagtgggg	206280
aagagcccg	tattgaatga	atagaagcca	ccccacagtc	tcagaaggcc	ggcttccctc	206340
ctgccctcca	ctggcttctc	aacgctgctg	cccttccctg	gtagggccca	ctcgacagca	206400
gttcaccgtc	agcagtgatg	acaccctctc	actgaatgcc	ggcaacatct	accagaagag	206460
gccagtccca	gggtgacatg	tgctggggcc	tggtctcagc	tgcttccctc	acattgccta	206520
cctgtatgga	gaaccctgtc	tcacctacca	gtacctgccc	tacatcagct	acctggtcag	206580
tccttggttc	gtcaaacccc	ggcttggggg	tgggggcaag	gatccaagga	ccagcccgag	206640
gtcttggggg	ttccaggagg	tctgtggggg	gacctgtccc	tcctcatctc	attctgtggt	206700
tctaggtagc	cccagggagc	aactcaaacc	ccagccgact	gaacagccgc	aaggaggccg	206760
ggctgtctgg	agcgttgaca	ctgacgcaga	aaatcatcgt	atacctctct	gacacgacct	206820
tcattggacat	tctgccccgc	attagccacg	aggtcttgct	gcctgtgctt	ggcttccctc	206880
cctcctctgt	cacagggtag	gccccgtgct	cttgggagag	ccacctggct	gaggggggccc	206940
ccagggaagg	ctagggaagc	cagggagaa	cagataccgg	cctgagtcct	ggttctgatg	207000
ttgggggtag	tggtcacagg	ctttcattcc	agcaccagca	ggaggggca	tttctgtgag	207060
tctgagacta	gcctgtgtct	cagagagagc	tcacaggctat	ctaaggctcc	atagtaagac	207120
tctgacttaa	gaaaagagtc	gtggttcatt	ctgggttgtg	ggtgtggctt	ggtgatggga	207180
cactttccca	gcattgcagga	ggagctatgc	ttgagttcca	gccccttcaga	aaaacaaaaa	207240
tgggggctgg	aaagaaatag	tcagggttta	agagcactgg	ttgctcttcc	agaggatcca	207300
ggttagattc	ccagctgcca	catggttagc	cataaccatc	cggcagttct	atggaaacct	207360
ccacctctct	tcggtctctg	tgggcactgc	aaacatgtgc	acagacatac	atgcaggcag	207420
aaaaaacacc	catacacata	aaattagacc	aaaaaagttc	atgttctctc	ctacctgtag	207480
ctctgactaa	gctacactgc	ttccctgtgc	ctcagtttcc	tcctctggtc	tggaactgac	207540
agccttacat	gcagctcctg	ttattttga	ttcctggtaa	attggtcaag	tccttcaggg	207600
aagggtctgg	aactcttgca	ctttgattct	aggttcccca	gtggggccca	ggcccgactc	207660
gtcctatgcy	tgaaaaccat	cagctctcat	gccctcatct	gcttgccgat	cgggcaggag	207720
atggtccagc	agcactgtag	tgagccagtg	gccaccttct	tcacagtctt	ctctcatctg	207780
catgagcttc	ggcagcaggt	aggcaggcag	cttctgggct	gggtggggca	ggccaggcca	207840
ggccaggcca	ggcagtgga	cccactgaat	ctgtggtctt	cctaccgcga	ggatctgcca	207900
ctggatccca	agggctgtac	tgagggccag	ctgccagagg	cgaccttctc	tgatgggcag	207960
cgacgaccag	tggaacccac	cctgctggaa	gagctgcaga	aggtgttcac	cctggaaatg	208020
gcgtacacaa	tctacgtaac	tttctcctgc	ctgttgggta	ttgcccatca	cgttcccttg	208080
cacagagttg	gtgactacat	ctcttccctg	gggtggggcc	cgatgcttcc	acctccagag	208140
tcagcaatgg	aatcttttta	tttttatttt	gacatggggt	ctcatttagc	ccagggtgac	208200
cttttaactc	agcttccctc	agcttccacc	gtctcctgtt	ggcattgtag	tcattgtgga	208260
ttgctcaggc	ttcttncatg	ttcttatttt	taaatgacct	gtgtgtgtgt	gtgatatctg	208320
tgtgagtgtg	gaggtgacag	aataacagtt	gggggggtcag	cagatgcctt	gcctgatgag	208380
catctctcta	gatccagttt	ttggttttgt	gggcttttat	gtgtgtgttt	gtttgtctgt	208440
ttttgtagac	agggctctct	tgtgtagcct	ggccatccct	gaactcattc	agtagaccaa	208500
gctggccttg	agctcacaga	gattcacctg	cctctgcctc	ccagtgtctg	gattaaaggc	208560
gtgtaccact	cctgcctggc	ttgttttttg	ttaccaccca	tcctcctgcc	tcagcatctg	208620
cctccctctg	gctgggatta	cagggtgtgt	ctatcacacc	cagctaacag	tggttttaaa	208680
cgtaggaatt	ttaggatcag	agtgaaccaga	tttgggtcta	gggcccattt	tcacagtgga	208740
ttatctatct	tagttaggat	ctctgttgaa	aatcatgtgt	gaacatcatt	accaaatgca	208800
acttggggag	gaaaagggtt	attttgtctg	acaactctca	ggtcaccaag	ggaagtccag	208860
gcagggaactc	gaggcagaag	ctgaagcaaa	agccatggaa	gaactctggc	ttgttctcca	208920
tggtctgtgc	agtctgtgtg	acccctctcc	caccacacct	cccacaatgg	tttctctgct	208980
tatgcctggc	tgctctagaa	ctcactctgt	agactagget	ggcctcaaac	tcaagagatc	209040
cccttgcctc	tgcatctcaa	gtgctaagat	taaaggcggg	tgccatcacc	cctgccccag	209100
gggtggcact	acccactgta	aattgggtccc	ttcccatatc	agttgttaaa	taagaaaact	209160
cctccatagg	ccaatctggt	gggggaattt	tctcagttga	gggtttctct	tctcaaatga	209220
ctgtagctga	tgccaaattg	ataaaaacaa	tctcaaacca	ccaccaccaa	caacaataaa	209280
accaaacaac	ccaaacaact	aaccaagaca	gtgacttata	aagagaatct	gaacattttc	209340
cagcaggaaa	ggctcaggag	ctggccattc	aagtctgggg	aacagaatgt	aggggaatat	209400
gatggtctcc	agaagctacc	tgcaaggaa	tgcaacagct	gctgggtttt	gtggcttccc	209460
ttatgggatg	ggcgtgttac	tgggttctct	tctgagttag	atggggccac	ctgtagtgtg	209520
gaatattttg	ctcctacaga	attgtgaagt	cccagaggca	ggacacatct	gtcttattct	209580
tcattgtgtg	tctgatgcta	gaatgggtgcc	tggcatacac	gtgtgtgtct	ctatagagac	209640

agcactcatg	tctacgtatc	gataaaggaa	gctgttttgg	ggggaggaaa	caggcttaca	209700
gacgagaact	taataaccca	gagtagccca	gtcagtagct	tgccttggct	tctgttggtt	209760
ctaagctctg	ggtagatagc	taccttgcca	tcttccctga	tcttagaact	ttccccactc	209820
ccctgtaggt	gacatcatcc	ggaaaatcat	ccccaacccat	gagttggtcg	gggagctggc	209880
agggtcttat	ctggaaagca	tgagcccagc	ctctcgaaac	ccagccagca	tggaaaccac	209940
catggctagt	gcccggccctg	aatgggaccc	tcagagtggg	agctgtctcc	aggacgatgg	210000
ccactcaggg	acctttggga	gtgtccctgt	tggaaatcgc	atccagatcc	ctgactctca	210060
gccccagagt	cctggggccac	tgggtccctc	ctctggagtg	ggtagtagcg	gaggccctcag	210120
caacaggaa	gaagacaacg	ccctgaagcg	ggagctgcct	cggagtgcgc	atgggctgag	210180
cggaagctgg	ctggcgctact	ggcagtagca	gatcggtgtg	agccagcagg	atgcccactt	210240
ccacttccac	cagatccgcc	tgcagagctt	cccagggcac	acggggggccg	tcaaatgcgt	210300
ggccgccctg	agcagtgaag	acttctttct	gagtggcagc	aaggaccgga	ctgtgcgcct	210360
ctggccgctg	tacaactatg	gggacgggac	caatgagacg	gcttcccgcc	tcatctatgc	210420
ccagcaccgc	aaaagcgtct	tctacgtggg	ccagcttgag	gccccgcagt	atgtgggtgag	210480
ctgtgatggg	gcagtgccag	tctgggaccc	cttcacaggt	gagcggggccc	aggtgagggc	210540
tggttcgacgg	ctgtcttact	gtgccttagc	caggccctctg	ggaacgggac	ctagtgcgaa	210600
acgtacaatg	gcgtattttg	acgggggaaga	ttcagtgagg	caggaagaga	agaagagtca	210660
ggacttagaa	ctctgtggac	ccaagtttga	atccactccc	ccaacttacc	agcaatcggc	210720
tcagttgctg	caggcgtctg	ccttctacct	gtaagaacca	aaaattttaga	agattccacg	210780
agtagggctt	tggcttcttg	tacgacgtca	cctgtcgtcg	ttgtaaagag	aatgatcgag	210840
tggaggaggg	tcaggggcaga	cggaggtcgc	agctagttag	agcatgctat	gtgaagagag	210900
cagactgttc	tggggctgga	cccttgactt	cactgtggaa	gcagcaagat	gagaaagccc	210960
tgagattgtg	ttttctgagg	gtcactgggg	aatgggatgc	aggtgtgggg	tgagttggag	211020
tttgaagtag	ccagggtctct	ttgtagacca	ctaagtcccc	agatgtgtcc	tttttcagga	211080
aagacccttc	gcacagtgga	tcttctcagc	agccgggtgc	ccctgcaggg	tgtggctgtc	211140
atgcctgccc	cacacaccag	catcaccatg	gccagctccg	actccactct	gcgttttgtg	211200
gactgcagga	agccaggctt	gcaggtcagg	aggggtgcag	ttcctgggct	actgggggtc	211260
tctaggtacc	agtcaggaaa	gacactcagg	ggactccacc	aggaacgctg	cagtgcacag	211320
cagccctgtg	tgggtggggc	gctggcacgg	atggggcttt	tctcttccgg	ggatggagtg	211380
ggagggtcag	gcctactggt	ttcgtggggc	tgaatgggtg	gagctgcagt	aggggtgggtg	211440
gcagtgtatg	atggcgacgg	gcacttgaac	acaatctcct	cctatagcat	gagttccgac	211500
tgggtggagg	gctgaacctt	gggcttgttc	gctcgttggc	cgtcagcccc	agtggccgga	211560
gtgttgtggc	tggcttctcc	tccggcttca	tgggtctcct	agataccgcg	acgggcccgg	211620
ttctacgagg	ctggccagcc	catgaagggg	acattctaca	gatcaagggtg	actgactgcc	211680
tgaggtcccta	tcttttcatt	tctacttagg	gcttggtctg	ggagaggaca	ggttttatgct	211740
gggtgtccctt	ataactactc	ggggacattc	agtggggtgg	gaaaatggcc	ctcgtaggcc	211800
agctcaggaa	ccagctgcac	aggaggcagg	ctaggggcag	gaatcagggc	tagaactgac	211860
cctgatgctc	cacagcgatg	ttctaatgag	taacccttgt	ccataattgt	cttgcttggg	211920
ggatcagggg	tcacgcctctg	tccgtgaccc	agttcagggt	aaataaagcc	aggaggctgt	211980
ttactgctctg	gagaccactg	agcagagtcc	atgccccctg	ctgggctgtc	ctgatggggg	212040
gcagggaacag	gcgcaggcct	gcgcctcgtg	ttcctgcctc	ctatatccaa	tcatagacct	212100
cagagctcag	caggttctgg	gaggggagaa	atagggctgc	ttgtgggagg	atttctccct	212160
gcagtgggaa	ctctcctccc	cgcgcgtcca	gatggagggtg	aaagacaggg	actgttgctt	212220
acagggaagg	caggctgccc	ccagctctat	ccaggacccc	aggggacccct	gggtctcagt	212280
gtctctaaat	cccaacatc	taagaaagtg	tcaggatggc	tctgggggtca	tcttgggtgt	212340
tagtcccagc	tctcggagtc	ttctctgagc	accagttctt	ttcctatggg	aagtaaggac	212400
atgccaggtg	ttctttgaga	ggacctgagt	ttggttctca	gcactgtcta	gctctggctc	212460
cagggggttc	aacacccctt	atggcttctg	tggacacata	ttcttatgtg	gcctgcacgc	212520
acacacacga	acacaaataa	aaataaatgt	taaaagaaga	cagcggcacc	ttgtacctca	212580
catgttagta	cgattggatg	tggcagtgcc	tcacagaatc	tgtgggactt	tattttattta	212640
tttatttttt	ggtctttaaaa	ttttaaaaga	ttggttttagg	agtgggtgta	cacaccatta	212700
atcgcaacac	tcaggagcag	aggcagggtg	atctctatgg	gtttgaggcc	agcctggtct	212760
acagagcaag	tttcagggca	gccaagggtta	cacagagaaa	ctctttctca	aaaaataaaa	212820
acaaaatatt	taaaagattt	acttactttc	tatttgagct	aggatctccc	tattagccct	212880
ggctgtcctg	gaactcactg	tatagaccag	gctggcacct	taaaactcacg	aagatcctcc	212940
tgtctctgcc	tcctaagtgc	tgagattaaa	gtagtgttat	accatgcccc	actattttct	213000
ttatagattt	ggtgttttgc	ctgcttgtgt	atatatgcac	taccttcacg	cagtgctaat	213060
aggggtcaga	ggcgagtatc	agctcttctc	ggaactagag	ttatgggaagg	ttgggagtca	213120
ccatgctggg	actgggtcat	ttgcaagagt	tacaagtact	tctgagccat	ctccagcccc	213180
ctagagtttt	tttccccctt	ggctgtcctg	aagtagaatc	tgttcttgtt	ttgttttgtt	213240
tttcagagaca	gggttttctc	acataggcct	ggctgtcctg	gaactcactc	tgtagaccag	213300
gctggcctcg	aactcagaaa	tccgcctgcc	tctggctctc	agaatactgg	aattaaaggt	213360
gtgcgccacc	acgcctggct	cagaatctgt	ttttaaatga	gagtaaatgt	tacaggtttt	213420
ttgttttgtt	tttttctttt	ctttttttgt	ttttgttttt	tggctcattt	gtttttatttg	213480

```

ttttgagaca ggtctcactc tgaaccccta ggtggcctgg agcttgctat gtagaacaca 213540
ctgacttttaa acttggttttc tgagtgtctgg atttatgggc ttgtgctatt ttgcccagcc 213600
tctgatgggt gttaataaca atattattta gcttttcttt tggagatagg ctctcactgt 213660
atatcaccaca gacagtggct ggtctggaaa tcactgtgta ggccaggctg accttgaatt 213720
cacagagatc tgcttccga gttcagaaat taaaagcact ctgggatggg tttggagttt 213780
ggtgagtacc caagcctcca ttgatgctat ctgtccctcc cgctctctgc aggtgtaga 213840
gggcagcgtg ctcatcagct cctcttccga ccattccttg actgtttgga aggagctgga 213900
acagaagccc acgcaccact acaagtcaag gtccgaccca atccacacct ttgacctgta 213960
cggcagcgag gtggtcaccg gcaactgtagc caacaagatt ggtgtctgtt cctgtcttga 214020
gccaccctct caggccacca caaagctcag ttccgagaac ttccgtggca cgctcactag 214080
tctggctttg ctgcccacga aacgccacct cctgtctggc tcggacaatg gcatcatccg 214140
cctcctggca tagggccagc caggagttgg ctgagggcag ggcgagatga catctctcag 214200
ggcccgctcc tcattcttga tctcgaagcc gattcttcta ggcaagcccc aggtctctggc 214260
taccacatg gcctgctgtc tgggattgca cagctcctga atctccaaag ccttgaagtg 214320
gcttcatgaa actcgggaga tactgttcct aaccagcaag aattggggca aggaaagcac 214380
tgtgatcccc attgtcccc agttctgcct tctggattca catggggaca gggcagctcc 214440
aggaaatgaa aggagtggg ccttgctca gccagcttcc tctagccacg ctctccttag 214500
ctctgtttct cccttgggtg ggaactgct cctgtctagg gttctgatgg tactgggact 214560
ccaggctcag gagggtggc caggacctac gactttcagg gcttggtctg gggtttttagc 214620
attcattcag ccaggtcttc agtatgggac cagaaaaaag gggatgtgag aacagggcta 214680
gggaaggggt tatatgggcc cagctggtcc aggaatgaat ccatgccttg ccttggtagc 214740
cctaaccaca gcgtttgtgc cttcagccgg ggaggcagcc cttgggacca gcatccctag 214800
ggacaggagg cagcgggaat catctctgta tctcgggttc tgcccagggg atgggcagac 214860
tctgccaact cttagtggtt cgtttggaga agcctgagat gtggcccctg ctgccttctc 214920
actagttgca gtctatgtaa ataagggtcaa taaattcttt ggaagagcca cggagctgag 214980
tgaggctgtg ttgtgttttg ctttgcctag gctgggtcga ggcagctctg cctcagcctc 215040
ccaaggagct ggggaactgg tataatgtac tgtatatgtc actgtgcctg gcttatggct 215100
tggcttggct ttttttcaga tgggtctcaag tgcctcaggt tggccttgat cttgggatga 215160
ccttcctgct tgaacacag tagtgggctt ataggcatga cccaccaggt ccaattttta 215220
ttttttaaag gcattgattt ttatacgtgt atggttgttt tgcccacttg tacatatgca 215280
caccatactt gtgtctggtc cctgcggagg tcagaagagg gcatcgggat cacctgggaa 215340
cgaagttaat gaatggttat gagccacatc tcatgctgta agattgaacc tggatccttt 215400
tcaagagcag ccagtggtct taccactga gccatctcta agccccacac ccagcttctt 215460
ttgatacaag gtctgtagc tcaaaactga tatgcagccg aggaggttga cctggtatct 215520
cctacctacc ctcttctctc taccttccaa gtgctgatat tatacatagg catggatagt 215580
catgccacc agtttgcctt gatggcacca gagtccaggaa agtccaaacc tggtagttgc 215640
aaacacagca agagggtaga ggcagccatt gtcctctggc tgccttggat acagagcttc 215700
tgggttgggt ggccctgggt cagttttccg aatggttcac ccttggggaa agggaaacct 215760
gctgaagagg tgggaccctg ggagggccgg cctccagctg ggtctctcca gccctgcct 215820
tggaacctag gctggaggga gccaaaccag atcctggact tgctacagtt aggtgaacag 215880
gctcctgcag cctccccttc ccttgggtag ctgtggtggg ggtggtggtg gtggtggtg 215940
tggtggtggt ggtggtggtg gtgggggggg gggngngngt

```

<210> 17

<211> 473

<212> PRT

<213> *Mus sp.*

<400> 17

```

Met Lys Arg Ala Ser Ser Gly Gly Ser Arg Leu Leu Ala Trp Val Leu
 1                      5                      10                      15

Trp Leu Gln Ala Trp Arg Val Ala Thr Pro Cys Pro Gly Ala Cys Val
                20                      25                      30

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

Gln Ala Val Pro Thr Gly Ile Pro Ala Ser Ser Gln Arg Ile Phe Leu
          50                      55                      60

His Gly Asn Arg Ile Ser His Val Pro Ala Ala Ser Phe Gln Ser Cys
          65                      70                      75                      80

```

Arg Asn Leu Thr Ile Leu Trp Leu His Ser Asn Ala Leu Ala Arg Ile
 85 90 95
 Asp Ala Ala Ala Phe Thr Gly Leu Thr Leu Leu Glu Gln Leu Asp Leu
 100 105 110
 Ser Asp Asn Ala Gln Leu His Val Val Asp Pro Thr Thr Phe His Gly
 115 120 125
 Leu Gly His Leu His Thr Leu His Leu Asp Arg Cys Gly Leu Arg Glu
 130 135 140
 Leu Gly Pro Gly Leu Phe Arg Gly Leu Ala Ala Leu Gln Tyr Leu Tyr
 145 150 155 160
 Leu Gln Asp Asn Asn Leu Gln Ala Leu Pro Asp Asn Thr Phe Arg Asp
 165 170 175
 Leu Gly Asn Leu Thr His Leu Phe Leu His Gly Asn Arg Ile Pro Ser
 180 185 190
 Val Pro Glu His Ala Phe Arg Gly Leu His Ser Leu Asp Arg Leu Leu
 195 200 205
 Leu His Gln Asn His Val Ala Arg Val His Pro His Ala Phe Arg Asp
 210 215 220
 Leu Gly Arg Leu Met Thr Leu Tyr Leu Phe Ala Asn Asn Leu Ser Met
 225 230 235 240
 Leu Pro Ala Glu Val Leu Met Pro Leu Arg Ser Leu Gln Tyr Leu Arg
 245 250 255
 Leu Asn Asp Asn Pro Trp Val Cys Asp Cys Arg Ala Arg Pro Leu Trp
 260 265 270
 Ala Trp Leu Gln Lys Phe Arg Gly Ser Ser Ser Glu Val Pro Cys Asn
 275 280 285
 Leu Pro Gln Arg Leu Ala Asp Arg Asp Leu Lys Arg Leu Ala Ala Ser
 290 295 300
 Asp Leu Glu Gly Cys Ala Val Ala Ser Gly Pro Phe Arg Pro Ile Gln
 305 310 315 320
 Thr Ser Gln Leu Thr Asp Glu Glu Leu Leu Ser Leu Pro Lys Cys Cys
 325 330 335
 Gln Pro Asp Ala Ala Asp Lys Ala Ser Val Leu Glu Pro Gly Arg Pro
 340 345 350
 Ala Ser Ala Gly Asn Ala Leu Lys Gly Arg Val Pro Pro Gly Asp Thr
 355 360 365
 Pro Pro Gly Asn Gly Ser Gly Pro Arg His Ile Asn Asp Ser Pro Phe
 370 375 380
 Gly Thr Leu Pro Ser Ser Ala Glu Pro Pro Leu Thr Ala Leu Arg Pro
 385 390 395 400
 Gly Gly Ser Glu Pro Pro Gly Leu Pro Thr Thr Gly Pro Arg Arg Arg
 405 410 415
 Pro Gly Cys Ser Arg Lys Asn Arg Thr Arg Ser His Cys Arg Leu Gly

420					425					430						
Gln	Ala	Gly	Ser	Gly	Ala	Ser	Gly	Thr	Gly	Asp	Ala	Glu	Gly	Ser	Gly	
435					440					445						
Ala	Leu	Pro	Ala	Leu	Ala	Cys	Ser	Leu	Ala	Pro	Leu	Gly	Leu	Ala	Leu	
450					455					460						
Val	Leu	Trp	Thr	Val	Leu	Gly	Pro	Cys								
465					470											

Lisboa, 2010-05-24

REIVINDICAÇÕES

1. Polipéptido compreendendo uma sequência de aminoácidos seleccionada de entre o grupo que consiste em:

- a) aminoácidos 31 a 310 de SEQ ID NO:2;
- b) aminoácidos 31 a 395 de SEQ ID NO:2;
- c) aminoácidos 1 a 395 de SEQ ID NO:2;
- d) aminoácidos 31 a 310 de SEQ ID NO:2 excepto 20 ou menos substituições de aminoácidos conservativas por 100 resíduos de aminoácidos;
- e) aminoácidos 31 a 310 de SEQ ID NO:2 e até cinco aminoácidos no terminal N e/ou no terminal C do referido polipéptido;
- f) aminoácidos 31 a 395 de SEQ 10 NO:2 excepto 20 ou menos substituições de aminoácidos conservativas por 100 resíduos de aminoácidos;
- g) aminoácidos 31 a 395 de SEQ ID NO:2 e até cinco aminoácidos no terminal N e/ou terminal C do referido polipéptido; e
- h) aminoácidos 1 a 395 de SEQ ID NO:2 excepto 20 ou menos substituições de aminoácidos conservativas por 100 resíduos de aminoácidos;

em que o referido polipéptido diminui a inibição do crescimento axonal em neurónios do SNC humano.

2. Polipéptido compreendendo uma sequência de aminoácidos seleccionada de entre o grupo que consiste em:

- a) uma sequência de aminoácidos pelo menos 70% idêntica aos aminoácidos 1 a 420 de SEQ ID NO:2;
- b) aminoácidos 1 a 420 de SEQ ID NO:2;
- c) uma sequência de aminoácidos pelo menos 70% idêntica aos aminoácidos 31 a 420 de SEQ ID NO:2; e
- d) aminoácidos 31 a 420 de SEQ ID NO:2;

em que o referido polipéptido diminui a inibição do crescimento axonal em neurónios do SNC humano.

3. Polipéptido de qualquer uma das reivindicações 1-2 compreendendo ainda um polipéptido heterólogo.

4. Polipéptido da reivindicação 3, em que o referido polipéptido heterólogo é seleccionado de entre o grupo que consiste em Fc, Glutathione-S-transferase (GST), e uma etiqueta de Histidina (etiqueta His).

5. Polinucleótido compreendendo um ácido nucleico codificando o polipéptido de qualquer uma das reivindicações 1-4.

6. Polinucleótido da reivindicação 5, compreendendo ainda um ou mais elementos de controlo da expressão operativamente ligados ao referido ácido nucleico.

7. Vector compreendendo o polinucleótido de qualquer uma das reivindicações 5-6.

8. Célula hospedeira compreendendo o polinucleótido de qualquer uma das reivindicações 5-6 ou o vector da reivindicação 7, em que a célula hospedeira exclui uma célula capaz de formar um ser humano.

9. Célula hospedeira da reivindicação 8, em que a referida célula hospedeira é uma célula eucariótica.

10. Polipéptido receptor de NOGO de acordo com qualquer uma das reivindicações 1-4 expresso pela célula hospedeira da reivindicação 8 ou 9.

11. Anticorpo ou seu fragmento de ligação ao antigénio que se liga especificamente a um polipéptido de qualquer uma das reivindicações 1-4 ou 10.

12. Anticorpo ou seu fragmento da reivindicação 11, em que o referido anticorpo ou seu fragmento de ligação ao antigénio é seleccionado de entre o grupo que consiste em anticorpo policlonal, anticorpo monoclonal, um anticorpo humanizado e um anticorpo humano.

13. Composição compreendendo o polinucleótido de qualquer uma das reivindicações 5-6, o vector da reivindicação 7, o polipéptido de qualquer uma das reivindicações 1-4 ou 10, ou o anticorpo da reivindicação 11 ou reivindicação 12, e um transportador.

14. Método *in vitro* para diminuição da inibição do crescimento axonal de um neurónio do SNC compreendendo o contacto de um neurónio com o polipéptido de qualquer uma das reivindicações 1-4 ou 10 ou o anticorpo da reivindicação 11 ou reivindicação 12.

15. Utilização do polipéptido de qualquer uma das reivindicações 1-4 ou 10 ou o anticorpo da reivindicação 11 ou da reivindicação 12 para a preparação de uma composição farmacêutica para modulação da inibição do crescimento axonal.

16. Utilização do polipéptido de qualquer uma das reivindicações 1-4 ou 10 ou do anticorpo da reivindicação 11 ou da reivindicação 12 para a preparação de uma composição farmacêutica para tratamento de uma doença, distúrbio ou lesão do sistema nervoso central (SNC), num mamífero, em que a referida doença, distúrbio ou lesão do SNC são seleccionados de entre o grupo que consiste em lesão cerebral, lesão da medula espinal, AVC ou uma doença desmielinizante.

17. Utilização da reivindicação 15 ou da reivindicação 16, em que a referida composição farmacêutica compreende ainda um transportador.

18. Utilização de qualquer uma das reivindicações 15 a 17, em que a referida composição farmacêutica é formulada para administração através de uma via seleccionada de entre o grupo que consiste em administração parentérica, administração subcutânea, administração intravenosa, administração transmucosa, administração intradérmica e administração transdérmica.

19. Utilização da reivindicação 16, em que a referida doença desmielinizante é seleccionada de entre o grupo que

consiste em esclerose múltipla, desmielinização monofásica, encefalomielite, leucoencefalopatia multifocal, panencefalite, doença de Marchiafava-Bignami, degenerescência Esponjosa, doença de Alexander, doença de Canavan, leucodistrofia metacromática e doença de Krabbe.

Lisboa, 2010-05-24

RESUMO

"Homólogos do receptor de Nogo"

	1		50
NOGO-R1 Humana	-----	MKEASA GGSRLLANVL WLQANVAAP	CPGACVCYNE
NOGO-R1 Murídea	-----	NKRASS GGSRLLANVL WLQANRVATP	CPGACVCYNE
NOGO-R3 Murídea	MSWQSGTTVT	QSPVQAAQVS GCCVELLLLL	LAGELFLOGG CPRDCVCYPA
NOGO-R3 Humana	-----	-----	-----
NOGO-R2 Humana	-----	NLPGLRRLIQ APASACLLLM	LLALPLAAPS CPMICTCYSS
Consenso	-----	-----	CP--C-CY--
	51	LRR NT	LRR 1
NOGO-R1 Humana	FKVTTSCFQQ	GLQAVPTGIP	AASQRIFLHG NRISHVPAAS FRACHNLTL
NOGO-R1 Murídea	FKVTTSCFQQ	GLQAVPTGIP	ASSQRIFLHG NRISNVPAAS FQSCNHLTL
NOGO-R3 Murídea	P.MTVSCQAH	NFAAIEGIP	EDSERIFLQN NRITFLQGGH FSP..AMVTL
NOGO-R3 Humana	-----	-----EGIP	VDSESVFLQN NRIGLLQGGH FSP..AMVTL
NOGO-R2 Humana	P.PTVSCQAH	NFSSVPLSLP	PSTQRLFLQN NLITLRPGT FGS..NLTL
Consenso	P--T-SC--	----P---P	---R-FL-- N-I-----F-----L
	101	LRR 2	LRR 3
NOGO-R1 Humana	WLHSNVLARI	DAAFTGLA	LEQLDLSDNA QLESVDPATF HGLGHLHTLH
NOGO-R1 Murídea	WLHSNALARI	DAAFTGLT	LEQLDLSDNA QLVVDPTTF HGLGHLHTLH
NOGO-R3 Murídea	WIYSNNITFI	APNTFEGFVH	LEELDLDGDNR QLRTAPETP QGLVKLHALY
NOGO-R3 Humana	WIYSNNITYI	HPSTFEGFVH	LEELDLDGDNR QLRTAPETP QGLVKLHALY
NOGO-R2 Humana	WLFSNNLSTI	YPOTFRHLQA	LEELDLDGDNR NLESLEPTP QGLERLQSLH
Consenso	W--SN----I	----F----	LE-LDL-DN- -L----P-TF -GL--L--L-
	151	LRR 4	LRR 5
NOGO-R1 Humana	LDRCGLQELG	PGLFRGLAAL	QVLYLQDNAL QALPDPTFED LGNLTHLFLH
NOGO-R1 Murídea	LDRCGLRELG	PGLFRGLAAL	QVLYLQDNAL QALPDNTFRD LGNLTHLFLH
NOGO-R3 Murídea	LYKCGLSALP	AGIFGGLHEL	QVLYLQDNHI EYLQDDIFVD LYNLSHLFLH
NOGO-R3 Humana	LYKCGLSALP	AGVFGGLHEL	QVLYLQDNHI EYLQDDIFVD LYNLSHLFLH
NOGO-R2 Humana	LYRCQLSSLP	GNIFRGLVEL	QVLYLQENSL LRLQDDLPAQ LANLSHLFLH
Consenso	L--C-L--L-	---F-GL--L	QVLYLQ-N- -L-D--F-D L-NL-NLFLH

O presente invento refere-se genericamente a genes que codificam proteínas que inibem o crescimento axonal. O presente invento refere-se especificamente a genes codificando homólogos da proteína NgR em humanos e ratinhos. O presente invento inclui também composições e métodos para modulação da expressão e actividade de Nogo e das proteínas NgR. Especificamente, o presente invento inclui péptidos, proteínas e anticorpos que bloqueiem a inibição mediada por Nogo do alongamento axonal. As composições e os métodos do presente invento são úteis no tratamento de traumatismo craniano ou cerebral, lesão da medula espinal, AVC ou uma doença desmielinizante.

FIG. 1A

FIG. 1A

	1	50
NOGO-R2	~~~~~MLPG, LRRLQAPAS AC...LLLML LA..LPLAAP SCPMLCTCYS	
NOGO-R3	MSWQSGTTVT QSPVQAAQVS GCCVELLLL LAGELPLGG. GCPRDCVCYP	
NOGO-R1	~~~~~MKRAS AGGSRLAWV LWLQAWQVAA PCPGACVCYN	
Consenso	~~~~~S ----LL--- L----- -CP--C-CY-	
	51	100
NOGO-R2	SP.PTVSCQA NNFSSVPLSL PPSTQRLFLQ>NNLIRTLRPG TFCS..NLLT	
NOGO-R3	AP.MTVSCQA HNFAAIPEGI PEDSERIFLQ>NNRITFLQQG HFSP..AMVT	
NOGO-R1	EPKVTTS CPQ QGLQAVPVGI PAASQRIFLH GNRISHVPAA SFRACRNLT	
Consenso	-P--T-SC-- -----P--- P----R-FL- -N-I----- -F-----	
	101	150
NOGO-R2	LWLFSNNLST IYPCTFRHLQ ALEELDLDGN RHLRSLEPDT FQGLERLQSL	
NOGO-R3	LWIYSNNITF IAPNTFEGFV HLEELDLDGN RQLRTLAPET FQGLVKLHAL	
NOGO-R1	LWLHSNVLAR IDAAFTGLA LLEQLDSDN AQLRSVDPAT FHGLGRHLTL	
Consenso	LW--SN---- I----F---- -LE-LDL-DN --LR---P-T F-GL--L--L	
	151	200
NOGO-R2	HLRYCQLSSL PGNIFRGLVS LQYLYQENS LLHLQDDLFA DLANLSHLFL	
NOGO-R3	YLYKCGLSAL PAGIFGGLHS LQYLYQDNH IEYLQDDIFV DLVNLSHLFL	
NOGO-R1	HLDRCGLOEL GPGLFRGLAA LQYLYQDNA LQALPDDTFR DLGNLTHLFL	
Consenso	-L--C-L--L ----F-GL-- LQYLYQ-N- ---L-DD-F- DL-NL-HLFL	
	201	250
NOGO-R2	HGNRLRLLTE HVFRGLGSLD RLLHGNRLQ GVHRAAFRGL SRLTILYLFN	
NOGO-R3	HGNKLWSLQ GIFRGLVND RLLHENQLQ WVHKAHFDL HRLTTLFLFN	
NOGO-R1	HGNRISSVPE RAFRGLHSLD RLLHQNRVA HVHPHAFRDL GRMLTYLFLA	
Consenso	HGN----- --FRGL--LD RLLH-N--- -VH--AF--L -RL--L-LF-	

FIG. 1B

Fig. 1B

	251	300
NOGO-R2	NSLASLPGEA LADLPSLEFL RLNANPWACD CRARPLWAWF QRARVSSSDV	
NOGO-R3	NSLTELOGDC LAPLVALEFL RLNGNAWDCC CRARSLWEWL RRFRCSSSAV	
NOGO-R1	NNLSALPTEA LAPLRALQYL RLNDNPWVCD CRARPLWAWL QKFRCSSEV	
Consenso	N-L--L---- LA-L--L--L RLN-N-W-C- CRAR-LW-W- ---R-SSS-V	
	301	350
NOGO-R2	TCATPPERQG RDLRALREAD FQAC...P.P AAPTRPGSRA	
NOGO-R3	PCATPELRQG QDLKLLRVED FRNC...TGP VSPHQIKSHT	
NOGO-R1	PCSLPQRLAG RDLKRLAAND LQCAVATGP YHPIWTGRAT DEEPLGLPKC	
Consenso	-C--P----G -DL--L---D ---C-----P --P-----	
	351	400
NOGO-R2	RGN ..SSSNH.LY G.VAE..... AGAPPADPS. . TLYRDLPA	
NOGO-R3	. .LTTSDDAA ..RKEHHPH G.ASRDKGHP HGHPPGSRSG YKKGKNCNTS	
NOGO-R1	CQPDAAKAS VLEPCRPASA GNALKGRVPP GDSPPCNCSC PRHI .NDSPF	
Consenso	----- G----- ---PP---S- -----	
	401	450
NOGO-R2	EDSRGR... QGGDAPTE.D DYWGGY.... ..GGEQ QRGEQMCPGA	
NOGO-R3	HRNRNQISKV SSGKELTELQ DYAPDYQHKF SFDIMPTARP KRKGKCARRT	
NOGO-R1	GTLPGSAEPP LTAVRPEGSE P..PGFPTSG PRRRPCCSRK NRTRSHCRLC	
Consenso	----- -R-----	
	451	491
NOGO-R2	ACQAPPDSRG PALSAGLPSP LLCLLLLVPH HL~~~~~	
NOGO-R3	PIRAPSGVQQ ASSCTALGAP LLAWILGLAV TLR~~~~~	
NOGO-R1	QAGSGGGGTG DSESGALPS LTCSLTPLGL ALVLTWTVLGP C	
Consenso	----- L----- -L~~~~~	

FIG. 2

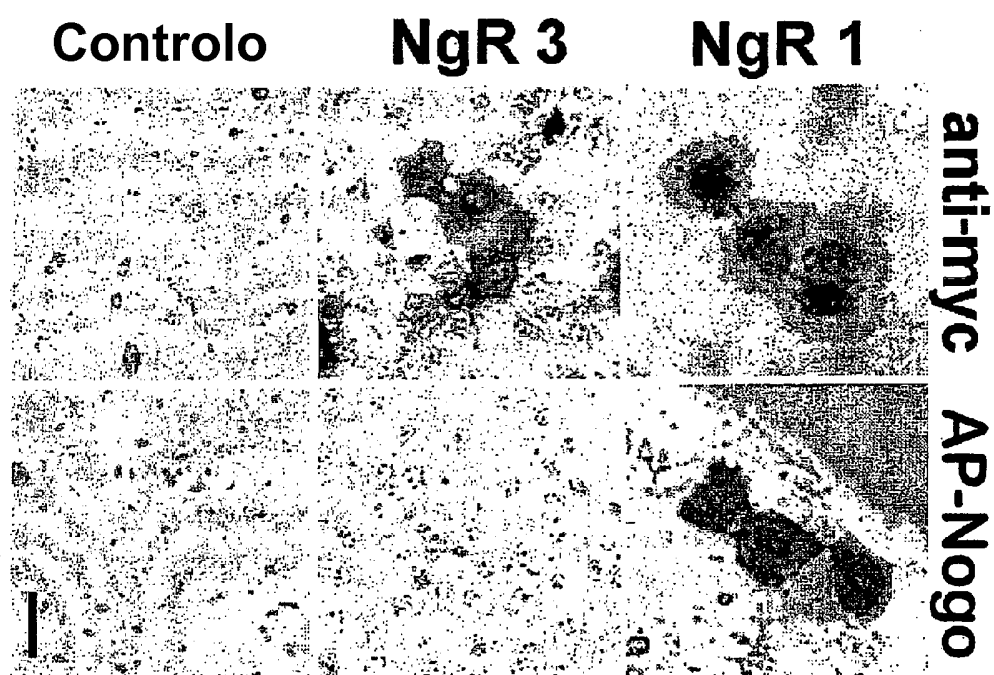


FIG. 3

	1		50
NOGO-R1 Humana	~~~~~	~~~~MKRASA GGSRLLAWL WLQAWQAAP	CPGACVCYNE
NOGO-R1 Murídea	~~~~~	~~~~MKRASS GGSRLLAWL WLQAWRVATP	CPGACVCYNE
NOGO-R3 Murídea	MSWQSGTTVT	QSPVQAAQVS GCCVELLLLL LAGELPLGGG	CPRDCVCYPA
NOGO-R3 Humana	~~~~~	~~~~~	~~~~~
NOGO-R2 Humana	~~~~~	MLPGLRRLQ APASACLLM LLALPLAAPS	CPMLCTCYSS
Consenso	~~~~~	-----	CP--C-CY--
	51	LRR NT	LRR 1
			100
NOGO-R1 Humana	PKVTTSCPQQ	GLQAVPVGIP	AASQRIFLHG NRISHVPAAS FRACRNLTL
NOGO-R1 Murídea	PKVTTSCPQQ	GLQAVPTGIP	ASSQRIFLHG NRISHVPAAS FQSCRNLTL
NOGO-R3 Murídea	P.MTVSCQAH	NFAAIPEGIP	EDSERIFLQN NRITFLQOGH FSP. AMVTL
NOGO-R3 Humana	~~~~~	~~~~~EGIP	VDSERVFLQN NRIGLQPGH FSP. AMVTL
NOGO-R2 Humana	P.PTVSCQAN	NFSSVPLSLP	PSTQRLFLQN NLRTLRPGT FGS. NLTL
Consenso	P--T-SC--	-----P--P	---R-FL-- N-I----- F-----L
	101	LRR 2	LRR 3
			150
NOGO-R1 Humana	WLHSNVLARI	DAAFTGLAL	LEQLDLSDNA QLRSDPATF HGLGRHLTLH
NOGO-R1 Murídea	WLHSNALARI	DAAFTGLTL	LEQLDLSDNA QLHVDPPTF HGLGHLTLH
NOGO-R3 Murídea	WIYSNNITFI	APNTFEGFVH	LEELDLGDNR QLRTLAPETF QGLVKLHALY
NOGO-R3 Humana	WIYSNNITYI	HPSTFEGFVH	LEELDLGDNR QLRTLAPETF QGLVKLHALY
NOGO-R2 Humana	WLFSNNLSTI	YPGTFRHLQA	LEELDLGDNR HLRSLBPTF QGLERLQSLH
Consenso	W--SN----I	----F-----	LE-LDL-DN- -L----P-TF -GL--L--L-
	151	LRR 4	LRR 5
			200
NOGO-R1 Humana	LDRCGLQELG	PGLFRGLAAL	QYLYQDNAL QALPDDTFRD LGNLTHLFLH
NOGO-R1 Murídea	LDRCGLRELG	PGLFRGLAAL	QYLYQDNNL QALPDNTFRD LGNLTHLFLH
NOGO-R3 Murídea	LYKCGLSALP	AGIFGGLHSL	QYLYQDNHI EYLQDDIFVD LWNLSHLFLH
NOGO-R3 Humana	LYKCGLSALP	AGVFGGLHSL	QYLYQDNHI EYLQDDIFVD LWNLSHLFLH
NOGO-R2 Humana	LYRCQLSSLP	GNIFRGLVSL	QYLYQENSL LHLQDDLFAD LANLSHLFLH
Consenso	L--C-L--L-	---F-GL--L	QYLYQ-N-- --L-D--F-D L-NL-HLFLH

FIG. 3, cont.

	201	LRR 6	LRR 7	250
NOGO-R1 Humana	GNRISSVPER	AFRGLHS LDR	LLLHQNRVAH VHPHAFRDLG	RLMTLYLFAN
NOGO-R1 Muridea	GNRIPSVPEH	AFRGLHS LDR	LLLHQNHVAR VHPHAFRDLG	RLMTLYLFAN
NOGO-R3 Muridea	GNKLWSLGQG	IFRGLVNLDR	LLLHENQLQW VHHKAFHDLH	RLTTLELFNN
NOGO-R3 Humana	GNKLWSLGPG	TFRGLVNLDR	LLLHENQLQW VHHKAFHDLR	RLTTLELFNN
NOGO-R2 Humana	GNRLRLLTEH	VFRGLGSLDR	LLLHGNRLQG VHRAAFRGLS	RLTILYLFNN
Consenso	GN-----	-FRGL--LDR	LLLH-N---- VH--AF--L-	RL--L-LF-N
	251	LRR 8	LRR CT	300
NOGO-R1 Humana	NLSALPTEAL	APLRALQYLR	LNDNPWVCDC RARPLWAWLQ	KFRGSSSEVP
NOGO-R1 Muridea	NLSMLPAEVL	MPLRSLOLYR	LNDNPWVCDC RARPLWAWLQ	KFRGSSSEVP
NOGO-R3 Muridea	SLTELQGDCL	APLVALEFLR	LNGNWDGCG RARSLWEWLR	RFRGSSSAVP
NOGO-R3 Humana	SLSELQGECL	APLGALEFLR	LNGNPWDGCG RARSLWEWLQ	RFRGSSSAVP
NOGO-R2 Humana	SLASLPGEAL	ADLPSLEFLR	LNANPWACDC RARPLWAWFQ	RARVSSSDVT
Consenso	-L--L---L-	--L--L--LR	LN-N-W-C-C RAR-LW-W--	--R-SSS-V-
	301			350
NOGO-R1 Humana	CSLPQRLAGR	DLKRLAANDL	QGCATVATGPY HPIWTGRATD	EEPLGLPKCC
NOGO-R1 Muridea	CNLPQRLADR	DLKRLAASDL	EGCAVASGPF RPIQTSQ L TD	EELSLPKCC
NOGO-R3 Muridea	CATPELRQGG	DLKLLRVEDF	RNCTGVPVSP. HQIKSHTLTT	SDRAARKEHH
NOGO-R3 Humana	CVSPGLRHGQ	DLKLLRAEDF	RNCTGPASP. HQIKSHTLTT	TDRAARKEHH
NOGO-R2 Humana	CATPPERQGR	DLRALREADF	QACP.PAAP. TRPGSRA...	..RGNSSSNH
Consenso	C--P-----	DL--L---D-	--C-----P-	-----
	351			400
NOGO-R1 Humana	QPDAAKASV	LEPGRPASAG	NALKGRVPPG DSPPGNGSGP	RHINDSPFGT
NOGO-R1 Muridea	QPDAAKASV	LEPGRPASAG	NALKGRVPPG DTPPGNGSGP	RHINDSPFGT
NOGO-R3 Muridea	PSHGASRDKG	HPHGHPPGSR	SGYK..... .KAGKNTSH	RNRNQISKVS
NOGO-R3 Humana	SPHGPTRSKG	HPH...GPR	PGHR..... .KPGKNTNP	RNRNQISKAG
NOGO-R2 Humana	.LYGVA.EAG	AP...PADPS	TLYR..... .DLPA.....	EDSRGR
Consenso	-----	-----	-----	-----

FIG. 3, cont.

	401		450
NOGO-R1 Humana	LPGSAEPPLT	AVRPEGSEPP GF...PTSGP	RRRPGCSRKN RTRSHCRLGQ
NOGO-R1 Murídea	LPSSAEPPLT	ALRPGGSEPP GL...PTTGP	RRRPGCSRKN RTRSHCRLGQ
NOGO-R3 Murídea	.SGKELTELQ	DYAPDYQHKF SFDIMPTARP	KRKGKCARRT PIRAPSGVQQ
NOGO-R3 Humana	.AGKQAPELP	DYAPDYQHKF SFDIMPTARP	KRKGKCARRT PIRAPSGVQQ
NOGO-R2 Humana	.QGGDAPTED	DYWGGY....	GGED QRGEQMCPGA ACQAPPD...
Consenso	-----	-----	-R-----
	451	Sinais de GPI Putativos	490
NOGO-R1 Humana	AGSGGGGTGD	SEGSALPSL	TCSLTPLGLA LVLWTVLGPC
NOGO-R1 Murídea	AGSGASGTGD	AEGSGALPAL	ACSLAPLGLA LVLWTVLGPC
NOGO-R3 Murídea		..ASSGTALG	APLLAWILGL AVTLR~~~~
NOGO-R3 Humana		..ASSASSLG	ASLLAWTLGL AVTLR~~~~
NOGO-R2 Humana		..SRGPALSA	GLPS PLLCL LLLVPHHL~~
Consenso	-----	-----	-----