



REPÚBLICA FEDERATIVA DO BRASIL



Ministério do Desenvolvimento, Indústria e Comércio Exterior
Instituto Nacional da Propriedade Industrial

CARTA PATENTE N.º PI 9915969-4

Patente de Invenção

O INSTITUTO NACIONAL DA PROPRIEDADE INDUSTRIAL concede a presente PATENTE, que outorga ao seu titular a propriedade da invenção caracterizada neste título, em todo o território nacional, garantindo os direitos dela decorrentes, previstos na legislação em vigor.

(21) Número do Depósito : PI 9915969-4

(22) Data do Depósito : 06/05/1999

(43) Data da Publicação do Pedido : 11/11/1999

(51) Classificação Internacional : C12N 15/32; C07K 14/325; A01N 63/00

(30) Prioridade Unionista : 06/05/1998 US 09/073,898

(54) Título : "HOSPEDEIRO RECOMBINANTE E MÉTODO PARA CONTROLAR UMA PESTE LEPIDOPTERANA".

(73) Titular : MYCOGEN CORPORATION, Sociedade Norte Americana. Endereço: 5501 Oberlin Drive - San Diego, California 92121, Estados Unidos (US).

(72) Inventor : JERALD S. FEITELSON, Técnico(a). Endereço: 4387 Mistral Place, San Diego, CA 92130, Estados Unidos. Cidadania: Norte Americana.; H. ERNEST SCHNEPF, Técnico(a). Endereço: 7954 Handel Court, San Diego, CA 92126, Estados Unidos. Cidadania: Norte Americana.; KENNETH E. NARVA, Técnico(a). Endereço: 12123 Caminito Mira Del Mar, San Diego, CA 92130, Estados Unidos. Cidadania: Norte Americana.; BRIAN A. STOCKHOFF, Técnico(a). Endereço: 11771 Ramsdell Court, San Diego, CA 92131, Estados Unidos. Cidadania: Norte Americana.; JAMES SCHMEITS, Técnico(a). Endereço: Apartment 133, 8215 Jade Coast Road, San Diego, CA 92126, Estados Unidos. Cidadania: Norte Americana.; DAVID LOEWER, Técnico(a). Endereço: 3937 Nobel Drive 210, San Diego, CA 92126, Estados Unidos. Cidadania: Norte Americana.; CHARLES JOSEPH DULLUM, Técnico(a). Endereço: 4147 Loma Alta Drive, San Diego, CA 92115, Estados Unidos. Cidadania: Norte Americana.; JUDY MULLER-COHN, Técnico(a). Endereço: 12744 Via Donada, Del Mar, CA 92014, Estados Unidos. Cidadania: Norte Americana.; LISA STAMP, Técnico(a). Endereço: 675 South Sierra Avenue 7, Solana Beach, CA 92075, Estados Unidos. Cidadania: Norte Americana.; GEORGE MORRILL, Técnico(a). Endereço: 2141 Crest Drive, El Cajon, CA 92021, Estados Unidos. Cidadania: Norte Americana.; STACEY FINSTAD-LEE, Técnico(a). Endereço: 13831 Paseo Cardiel, San Diego, CA 92129, Estados Unidos. Cidadania: Norte Americana.

Prazo de Validade : 10 (dez) anos contados a partir de 14/10/2014, observadas as condições legais.

Expedida em : 14 de Outubro de 2014.

Assinado digitalmente por
Júlio César Castelo Branco Reis Moreira
Diretor de Patentes

Relatório Descritivo da Patente de Invenção para **"HOSPEDEIRO RECOMBINANTE E MÉTODO PARA CONTROLAR UMA PESTE LEPIDOPTERANA"**.

Fundamentos da Invenção

5 Insetos e outras pragas custam aos agricultores bilhões de dólares anualmente em perdas de colheita e na despesa de se manter estas pragas sob controle. As perdas causadas por pragas de inseto nos ambientes de produção de agricultura incluem diminuição no rendimento de colheita, qualidade de colheita reduzida e custos de colheita aumentados.

10 Métodos de cultivo, tais como rotação de colheita e a aplicação de altos níveis de nitrogênio para estimular o crescimento de um sistema de raiz adventícia, parcialmente leva em conta problemas causados por pragas de agricultura. Demandas econômicas na utilização de gleba de terra cultivada restringem o uso de rotação de colheita. Além disso, trilhas de alguns insetos que ultrapassam o inverno estão rompendo rotações de colheita em algumas áreas. Assim, conta-se com a maior intensidade com inseticidas químicos para garantir o nível desejado de controle. Inseticidas são ou associados sobre ou incorporados no solo.

20 O uso de inseticidas químicos tem várias desvantagens. O uso contínuo de inseticidas permitiu que insetos resistentes se desenvolvam. Situações tais como populações extremamente altas de larvas, chuvas pesadas e calibração imprópria de equipamento de aplicação de inseticida podem resultar em controle pobre. O uso de inseticidas frequentemente aumenta as preocupações ambientais tais como contaminação de solo e tanto de suprimentos de água de subsolo quanto de superfície. O público também se tornou preocupado com a quantidade de produtos químicos sintéticos, residuais que pode ser encontrado no alimento. O trabalho com inseticidas pode também apresentar riscos as pessoas que os aplicam. Portanto, pesticidas químicos sintéticos estão sendo crescentemente minuciosamente examinados e corretamente assim, quanto as suas consequências ambientais tóxicas potenciais. Exemplos de pesticidas químicos sintéticos amplamente usados incluem os organocloro, por exemplo, DDT, mirex, kepona,

30

lindano, aldrina, clordano, aldicarb e dieldrin; os organofosfatos, por exemplo, clorpirifos, paration, malation e diazinon; e carbamatos. Novas restrições rigorosas sobre o uso de pesticidas e da eliminação de alguns pesticidas eficazes oriundos do mercado poderia limitar opções econômicas e eficazes para o controle de pestes prejudiciais e dispendiosas.

Por causa dos problemas associados ao uso de pesticidas químicos sintéticos orgânicos, existe uma clara necessidade de limitar o uso destes agentes e uma necessidade de indentificar agentes de controle alternativos. A substituição de pesticidas químicos sintéticos ou combinações destes agentes com pesticidas biológicos, poderiam reduzir os níveis de produtos químicos tóxicos no ambiente.

Um agente pesticida biológico que está desfrutando de popularidade crescente é o micróbio de solo *Bacillus thuringiensis* (B.t.). O micróbio de solo *Bacillus thuringiensis* (B.t.) é uma bactéria formadora de esporo, gram-positiva. A maior parte das cepas de B.t. não exibem atividade pesticida. Algumas cepas de B.t. produzem, e podem ser caracterizadas por, inclusões de proteína cristalina paraesporal. Estas inclusões freqüentemente aparecem microscopicamente como cristais de forma distinta. Algumas proteínas de B.t. são altamente tóxicas a pestes, tais como insetos, e são específicas em sua atividade tóxica. Certas proteínas de B.t. inseticidas são associadas às inclusões. Estas "δ-endotoxinas", são diferentes de exotoxinas, que têm uma faixa de hospedeiro não específica.

Certos genes de toxina *Bacillus* foram isolados e foram seqüenciados, e produtos à base de DNA recombinante foram produzidos e foram aprovados para o uso. Além disso, com o uso de técnicas de engenhieramento genético, novas abordagens para o fornecimento destas toxinas aos ambientes de agricultura estão sob desenvolvimento. Estas incluem o uso de plantas geneticamente engenheiradas com genes de toxina para resistência a inseto e o uso de células microbianas intactas estabilizadas como veículos de fornecimento de toxina. Assim, genes de toxina de *Bacillus* isolados estão se tornando comercialmente de alto valor.

Até os últimos quinze anos, o uso comercial de pesticidas de

B.t. foi grandemente restrito ao direcionamento de uma faixa estreita de pestes de lepidopterano (caterpillar). Preparações dos esporos e cristais de subespécie de *B. thuringiensis kurstaki* foram usadas por muitos anos como inseticidas comerciais para pestes de lepidopterano. Por exemplo, *B. thuringiensis* var. *kurstaki* HD-1 produz uma δ -endotoxina cristalina que é tóxica às larvas de numerosos insetos de lepidopterano.

Em anos recentes, no entanto, investigadores revelaram pesticidas de B.t. com especificidades para uma faixa muito mais ampla de pestes. Por exemplo, outras espécies de B.t., a saber *israelensis* e *morrisoni* (a.k.a. *tenebrionis*, a.k.a. B.t. M-7, a.k.a B.t. *san diego*), foram usadas comercialmente para controlar insetos das ordens Diptera e Coleoptera, respectivamente, *Bacillus thuringiensis* var. *tenebrionis* foi relatado ser ativo contra dois besouros na ordem Coleoptera (besouros de batata Colorado, *Leptinotarsa decemlineata* e *Agelastica alni*).

Mais recentemente novas subespécies de B.t. foram identificadas, e genes responsáveis por proteínas de δ -endotoxina ativas foram isoladas. Höfte e Whiteley classificaram genes de proteína de cristal de B.t. em quatro classes principais (Höfte, H., H.R. Whiteley [1989] *Microbiological Reviews* 52(2): 242 – 255). As classes eram CryI (específica de Lepidoptera), CryII (específica de Lepidoptera e Diptera), CryIII (específica de Coleoptera), e CryIV (específica de Diptera). A revelação de cepas especificamente tóxicas a outras pestes foi relatada. Por exemplo, CryV e CryVI foram propostas designar uma classe de genes de toxina que são específica de nematóide.

O esquema de classificação e nomenclatura de 1989 de Höfte e Whiteley para proteínas de cristal foi à base da seqüência de aminoácidos e a faixa de hospedeiro da toxina. Aquele sistema foi adaptado para cobrir 14 diferentes tipos de genes de toxina que foram divididos em cinco classes principais. O número de genes de proteína de cristal de *Bacillus thuringiensis* seqüenciados fica em mais do que 50. Um esquema de nomenclatura revisada foi proposto que é baseado exclusivamente na identidade de aminoácido (Crickmore e outros [1996] *Society for Invertebrate Pathology*, 29^a An-

nual Meeting, III rd International Colloquium on *Bacillus thuringiensis*, University of Cordoba, Cordoba, Espanha, em 1 a 6 de setembro de 1996, resumo). A "cry" mnemônica foi conservada para todos os genes de toxina exceto cytA e cytB, que permanecem uma classe separada. Números romanos foram trocados para números arábicos na classe primária, e os parênteses na classe terciária foram removidos. Muitos dos nomes originais foram conservados, com as exceções observadas, embora um número tenha sido reclassificado.

Muitos outros genes de B.t. foram agora identificados. A WO 94/21795, WO 96/10083, WO 98/44137 e Estruch, J.J. e outros (1996) PNAS 93: 5389 – 5394 descrevem toxinas de Vip1A(a), de Vip1A(b), de Vip2A(a), de Vip2A(b), de Vip3A(a), e de Vip3A(b) obtidas a partir de micróbios de *Bacillus*. Aquelas toxinas são relatadas serem produzidas durante o crescimento celular vegetativo e foram assim chamadas de proteínas inseticidas vegetativas (VIP). A atividade destas toxinas contra certas pestes de lepidopterano e certas pestes de coleopterano foi relatada. A WO 98/18932 revela novas classes de toxinas pesticidas.

Obstáculos ao uso em agricultura bem sucedida de toxinas de *Bacillus* incluem o desenvolvimento de resistência às toxinas de B.t. por insetos. Além disso, certos insetos podem ser refratários aos efeitos de toxinas de *Bacillus*. O último inclui insetos tais como gorgulho de "boll" e gramíola negra bem como insetos adultos da maioria das espécies que até agora demonstraram não aparentar sensibilidade significativa a δ -endotoxinas de B.t. Embora estratégias de gerenciamento em tecnologia de planta transgene de B.t. tenha se tornado de grande interesse, permanece uma grande necessidade de desenvolver genes adicionais que podem ser expressos em plantas a fim de eficazmente controlar vários insetos.

O presente pedido provê novas classes de toxinas e genes, além daquelas descritas na WO 98/18932, e que são distintas daquelas reveladas na WO 94/21795, na WO 96/10083, na WO 98/44137, e Estruch e outros.

Breve Sumário da Invenção

A presente invenção refere-se a materiais e a métodos úteis no controle de pestes de não mamífero e, particularmente, pestes de planta. Em uma modalidade, a presente invenção provê novos isolados de *Bacillus* tendo atividade vantajosa contra pestes de não mamífero. Em uma modalidade ulterior, a presente invenção provê novas toxinas úteis para o controle de pestes de não mamífero. Em uma modalidade preferida, estas pestes são lepidopteranos e/ou coleopteranos. As toxinas da presente invenção incluem, δ -endotoxinas bem como tóxicas solúveis que podem ser obtidas a partir do sobrenadante das culturas de *Bacillus*.

A presente invenção ulteriormente provê seqüências de nucleotídeos que codificam as toxinas da presente invenção. A presente invenção ulteriormente provê métodos e seqüências de nucleotídeos úteis na identificação e na caracterização de genes que codificam toxinas pesticidas.

Em uma modalidade, a presente invenção refere-se a seqüências de nucleotídeos únicas que são úteis como sondas de hibridização e/ou iniciadores em técnicas de PCR. Os iniciadores produzem fragmentos de gene característicos que podem ser usados na identificação, caracterização, e/ou isolamento de genes de toxina específicos. As seqüências de nucleotídeos da presente invenção codificam toxinas que são distintas das toxinas descritas anteriormente.

Em uma modalidade específica, a presente invenção provê novas classes de toxinas tendo atividades pesticidas vantajosas. Estas classes de toxinas podem ser codificadas por seqüências de polinucleotídeos que são caracterizados por sua capacidade de hibridizar com certas seqüências exemplificadas e/ou por sua capacidade de ser ampliada por PCR usando-se certos iniciadores exemplificados.

Um aspecto da presente invenção pertence à identificação e à caracterização de novas famílias inteiras de toxinas de *Bacillus* tendo propriedades pesticidas vantajosas. A presente invenção inclui novas classes de genes e toxinas chamadas aqui de MIS-7 e MIS-8. Genes e toxinas da nova classe WAR e SUP são também revelados. Certos genes e toxinas

MIS-1 e MIS-2 são também ulteriormente caracterizados aqui.

Estas famílias de toxinas, e os genes que os codificam, podem ser caracterizados em termos, por exemplo, do tamanho do gene ou da toxina, da seqüência de aminoácidos ou de DNA, da atividade pesticida, e/ou da reatividade de anticorpo. Com relação aos genes que codificam as novas famílias de toxina da presente invenção, a revelação atual provê sondas de hibridização únicas e iniciadores de PCR que podem ser usados para identificar e caracterizar DNA dentro das famílias exemplificadas.

Em uma modalidade da presente invenção, isolados de *Bacillus* podem ser cultivados sob condições que resultam em alta multiplicação do micróbio. Depois do tratamento do micróbio para prover ácido nucléico genômico de fio único, o DNA pode ser posto em contato com os iniciadores da invenção e submetê-lo a ampliação por PCR. Fragmentos característicos de genes de codificação de toxina serão ampliados pelo procedimento, assim a identificação da presença do(s) gene(s) de codificação de toxina.

Um outro aspecto da presente invenção é o uso das seqüências de nucleotídeos reveladas como sondas para detectar genes que codificam toxinas de *Bacillus* que são ativos contra pestes.

Outros aspectos da presente invenção incluem os genes e isolados identificados usando-se os métodos e as seqüências de nucleotídeos revelados aqui. Os genes assim identificados codificam toxinas ativas contra pestes. Similarmente, os isolados terão atividade contra estas pestes. Em uma modalidade preferida, estas pestes são pestes de lepidopterano ou de coleopterano.

Em uma modalidade preferida, a presente invenção refere-se a células de plantas transformadas com pelo menos uma seqüência de polinucleotídeos da presente invenção de modo que as células de planta transformadas expressam toxinas pesticidas em tecidos consumidos por pestes alvo. Como descrito aqui, as toxinas úteis de acordo com a presente invenção podem ser toxinas quiméricas por combinação de porções de toxinas múltiplas. Além disso, misturas e/ou combinações de toxinas podem ser usadas de acordo com a presente invenção.

Transformação de plantas com os construtos genéticos revelados aqui pode ser realizada usando-se técnicas bem conhecidas por aquele versado na técnica e tipicamente envolveriam modificação do gene para otimizar expressão da toxina em plantas.

- 5 Alternativamente, os isolados de *Bacillus* da presente invenção, ou micróbios recombinantes que expressam toxinas descritas aqui, podem ser usados para controlar pestes. A este respeito, a invenção inclui o tratamento de células de *Bacillus* substancialmente intactas, e/ou células recombinantes contendo as toxinas expressas da invenção, tratadas para prolongar a atividade pesticida quando as células substancialmente intactas são aplicadas ao ambiente de uma peste alvo. A célula tratada age como um revestimento protetor para a toxina pesticida. A toxina se torna ativa após a ingestão por um inseto alvo.

Breve Descrição das Seqüências

- 15 SEQ ID NO. 1 é uma seqüência de nucleotídeos que codificam um toxina oriunda de cepa B.t. Javelin 1990.

SEQ ID NO. 2 é uma seqüência de aminoácidos para a toxina Javelin 1990.

- 20 SEQ ID NO. 3 é um iniciador de avanço usado de acordo com a presente invenção.

SEQ ID NO. 4 é um iniciador reverso usado de acordo com a presente invenção.

SEQ ID NO. 5 é uma seqüência de nucleotídeos de um gene de toxina oriundo de cepa B.t. PS66D3

- 25 SEQ ID NO. 6 é uma seqüência de aminoácidos oriunda da toxina 66D3.

SEQ ID NO. 7 é uma seqüência de nucleotídeos de um gene de toxina MIS oriundo de cepa de B.t. PS177C8.

- 30 SEQ ID NO. 8 é uma seqüência de aminoácidos oriunda de toxina 177C8-MIS.

SEQ ID NO. 9 é uma seqüência de nucleotídeos de um gene de toxina oriundo de cepa de B.t. PS177I8

SEQ ID NO. 10 é uma seqüência de aminoácidos oriunda da toxina 177I8.

SEQ ID NO. 11 é uma seqüência de nucleotídeos que codifica um gene de toxina 177C8-WAR oriunda de cepa de B.t. PS177C8.

5 SEQ ID NO. 12 é uma seqüência de aminoácidos de uma toxina 177C8-WAR oriundo de cepa de B.t. PS177C8.

SEQ ID NO. 13-21 são iniciadores usados de acordo com a invenção.

10 SEQ ID NO. 22 é o complemento reverso do iniciador de SEQ ID NO. 14.

SEQ ID NO. 23 é o complemento reverso do iniciador de SEQ ID NO. 15.

SEQ ID NO. 24 é o complemento reverso do iniciador de SEQ ID NO. 17.

15 SEQ ID NO. 25 é o complemento reverso do iniciador de SEQ ID NO. 18.

SEQ ID NO. 26 é o complemento reverso do iniciador de SEQ ID NO. 19.

20 SEQ ID NO. 27 é o complemento reverso do iniciador de SEQ ID NO. 20.

SEQ ID NO. 28 é o complemento reverso do iniciador de SEQ ID NO. 21.

SEQ ID NO. 29 é um iniciador de avanço MIS-7.

SEQ ID NO. 30 é um iniciador de reverso MIS-7.

25 SEQ ID NO. 31 é um iniciador de avanço MIS-8.

SEQ ID NO. 32 é um iniciador de reverso MIS-8.

SEQ ID NO. 33 é uma seqüência de nucleotídeos de um gene de toxina MIS-7 designado 157C1-A oriundo da cepa de B.t. PS157C1.

30 SEQ ID NO. 34 é uma seqüência de aminoácidos de uma toxina MIS-7 designado 157C1-A oriunda da cepa de B.t. PS157C1.

SEQ ID NO. 35 é uma seqüência de nucleotídeos de um gene de toxina MIS-7 oriundo da cepa de B.t. PS201Z.

SEQ ID NO. 36 é uma seqüência de nucleotídeos de um gene de toxina MIS-8 oriundo da cepa de B.t. PS31F2.

SEQ ID NO. 37 é uma seqüência de nucleotídeos de um gene de toxina MIS-8 oriundo da cepa de B.t. PS185Y2.

5 SEQ ID NO. 38 é uma seqüência de nucleotídeos de um gene de toxina MIS-1 oriundo da cepa de B.t. PS33F1.

SEQ ID NO. 39 é um iniciador MIS para o uso de acordo com a presente invenção.

10 SEQ ID NO. 40 é um iniciador MIS para o uso de acordo com a presente invenção.

SEQ ID NO. 41 é um iniciador WAR para o uso de acordo com a presente invenção.

SEQ ID NO. 42 é um iniciador WAR para o uso de acordo com a presente invenção.

15 SEQ ID NO. 43 é uma seqüência de nucleotídeos parcial para um gene MIS-7 oriundo de PS205C.

SEQ ID NO. 44 é uma seqüência de aminoácidos parcial para uma toxina MIS-7 oriunda de PS205C.

20 SEQ ID NO. 45 é uma seqüência de aminoácidos parcial para um gene WAR oriundo de PS205C.

SEQ ID NO. 46 é uma seqüência de aminoácidos parcial para uma toxina WAR oriunda de PS205C.

SEQ ID NO. 47 é uma seqüência de nucleotídeos para um gene MIS-8 oriundo de PS31F2.

25 SEQ ID NO. 48 é uma seqüência de aminoácidos para uma toxina MIS-8 oriunda de PS31F2.

SEQ ID NO. 49 é uma seqüência de nucleotídeos para um gene WAR oriundo de PS31F2.

30 SEQ ID NO. 50 é uma seqüência de aminoácidos para uma toxina WAR oriunda de PS31F2.

SEQ ID NO. 51 é um iniciador SUP para o uso de acordo com a presente invenção.

SEQ ID NO. 52 é um iniciador SUP para o uso de acordo com a presente invenção.

SEQ ID NO. 53 é uma seqüência de nucleotídeos para um gene SUP oriundo de KB59A4-6.

5 SEQ ID NO. 54 é uma seqüência de aminoácidos para uma toxina SUP oriunda de KB59A4-6.

Descrição Detalhada da Invenção

A presente invenção refere-se a materiais e a métodos para o controle de pestes de não mamífero. Em modalidades específicas, a presente invenção pertence a novos isolados de *Bacillus thuringiensis* e toxinas que têm atividade contra lepidópteros e/ou coleópteros. A presente invenção ulteriormente refere-se a novos genes que codificam toxinas pesticidas e a novos métodos para a identificação e caracterização de genes de *Bacillus* que codificam toxinas com propriedades úteis. A presente invenção refere-se não apenas às seqüências de polinucleotídeos que codificam estas toxinas, mas também ao uso destas seqüências de polinucleotídeos para produzir hospedeiros recombinantes que expressam as toxinas. As proteínas da presente invenção são distintas das toxinas de proteína que foram anteriormente isoladas de *Bacillus thuringiensis*.

20 Isolados de B.t. úteis de acordo com a presente invenção foram depositados na coleção permanente da Agricultural Research Service Patent Culture Collection (NRRL), Northern Regional Research Center, 1815 North University Street, Peoria, Illinois 61604, USA. Os números repositórios de cultura das cepas de B.t. são como se segue:

Tabela 1

Cultura	Nº Repositório	Dados de Depósito	Patente nº
<i>B.t.</i> PS157C1 (MT104)	NRRL B-18240	17 de Julho de 1987	5.262.159
<i>B.t.</i> PS31F2	NRRL B-21876	24 de Outubro de 1997	
<i>B.t.</i> PS66D3	NRRL B-21858	24 de Outubro de 1997	
<i>B.t.</i> PS177C8a	NRRL B-21867	24 de Outubro de 1997	

Continuação

Tabela 1			
Cultura	Nº Repositório	Dados de Depósito	Patente nº
<i>B.t.</i> PS17718	NRRL B-21868	24 de Outubro de 1997	
KB53A49-4	NRRL B-21879	24 de Outubro de 1997	
KB68B46-2	NRRL B-21877	24 de Outubro de 1997	
KB68B51-2	NRRL B-21880	24 de Outubro de 1997	
KB68B55-2	NRRL B-21878	24 de Outubro de 1997	
PS33F1	NRRL B-21977	24 de Abril de 1998	
PS71G4	NRRL B-21978	24 de Abril de 1998	
PS86D1	NRRL B-21979	24 de Abril de 1998	
PS185V2	NRRL B-21980	24 de Abril de 1998	
PS191A21	NRRL B-21981	24 de Abril de 1998	
PS201Z	NRRL B-21982	24 de Abril de 1998	
PS205A3	NRRL B-21983	24 de Abril de 1998	
PS205C	NRRL B-21984	24 de Abril de 1998	
PS234E1	NRRL B-21985	24 de Abril de 1998	
PS248N10	NRRL B-21986	24 de Abril de 1998	
KB63B19-13	NRRL B-21990	29 de Abril de 1998	
KB63B19-7	NRRL B-21989	29 de Abril de 1998	
KB68B62-7	NRRL B-21991	29 de Abril de 1998	
KB68B63-2	NRRL B-21992	29 de Abril de 1998	
KB69A125-1	NRRL B-21993	29 de Abril de 1998	
KB69A125-3	NRRL B-21994	29 de Abril de 1998	
KB69A125-5	NRRL B-21995	29 de Abril de 1998	
KB69A127-7	NRRL B-21996	29 de Abril de 1998	
KB69A132-1	NRRL B-21997	29 de Abril de 1998	
KB69B2-1	NRRL B-21998	29 de Abril de 1998	
KB70B5-3	NRRL B-21999	29 de Abril de 1998	
KB71A125-15	NRRL B-30001	29 de Abril de 1998	
KB71A35-6	NRRL B-30000	29 de Abril de 1998	

Continuação

Tabela 1			
Cultura	Nº Repositório	Dados de Depósito	Patente nº
KB71A72-1	NRRL B-21987	29 de Abril de 1998	
KB71A134-2	NRRL B-21988	29 de Abril de 1998	
PS185Y2	NRRL B-30121	4 de Maio de 1999	
KB59A4-6	NRRL B-		
MR992	NRRL B-30124	4 de Maio de 1999	
MR983	NRRL B-30123	4 de Maio de 1999	
MR993	NRRL B-30125	4 de Maio de 1999	
MR951	NRRL B-30122	4 de Maio de 1999	

Culturas que foram depositadas para as finalidades deste pedido de patente foram depositadas sob condições de que o acesso às culturas está disponível durante a pendência deste pedido de patente por alguém determinado pelo Comissário de Patentes e Marcas Registradas a ser entitulado ao mesmo sob 37 CFR 1.14 e 35 U.S.C. 122. Os depósitos estarão disponíveis quando necessários pelas leis de Patente estrangeiras em países em que cópias da presente invenção, ou sua progênia, são arquivadas. No entanto, deve ser entendido que a disponibilidade de um depósito não constitui uma licença para praticar a presente invenção em detrimento dos direitos de patente concedidos pela ação governamental.

Além disso, os presentes depósitos de cultura serão armazenados e tornados disponíveis ao público de acordo com as provisões do Tratado de Budapeste para o Depósito de Microorganismo, isto é, eles serão armazenados com todo o cuidado necessário para mantê-los viáveis e não contaminados por um período de pelo menos cinco anos depois da exigência mais recente para o fornecimento de uma amostra do depósito, e em alguns casos, por um período de pelo menos trinta (30) anos depois da data do depósito ou pela vida obrigatória de qualquer patente que pode expedir a revelação da(s) cultura(s). O depositante reconhece a obrigação de substituir o(s) depósito(s) caso o depositário seja incapaz de fornecer uma amostra quando exigido, devido a condição de um depósito. Todas as restrições

sobre a disponibilidade ao público dos presentes depósitos de cultura serão irrevogavelmente removidos após a concessão de uma patente que as revelem.

5 Muitas das cepas úteis de acordo com a presente invenção estarão prontamente disponíveis em virtude da expedição de patentes que revelam estas cepas ou por seu depósito em coleções públicas ou por sua inclusão em produtos comerciais. Por exemplo, a cepa de B.t. usada no produto comercial, Javelin, e os isolados de HD são todos disponíveis publicamente.

10 Mutantes dos isolados mencionados aqui podem ser produzidos por procedimentos bem conhecidos na técnica. Por exemplo, um mutante asporogenoso pode ser obtido através de mutagênese de sulfonato de etilmetano (EMS) de um isolado. Os mutantes podem ser produzidos usando-se luz ultravioleta e guanidina nitrosa por procedimentos bem conhecidos na
15 técnica.

Em uma modalidade, a presente invenção refere-se a materiais e a métodos incluindo iniciadores de nucleotídeo e sondas para o isolamento, caracterização e identificação de genes de *Bacillus* que codificam toxinas de proteína que são ativas contra pestes de não mamífero. As se-
20 quências de nucleotídeos descritos aqui podem também ser usadas para identificar novos isolados de *Bacillus* pesticidas. A invenção ulteriormente refere-se a genes, a isolados, a toxinas identificadas usando-se os métodos e os materiais revelados aqui.

As novas toxinas e as seqüências de nucleotídeos providas aqui
25 são definidas de acordo com vários parâmetros. Uma característica das toxinas descrita aqui é a atividade pesticida. Em uma modalidade específica, estas toxinas têm atividade contra as pestes de coleopterano e/ou lepidopterano. As toxinas e os genes da presente invenção podem ser ulteriormente definidos por suas seqüências de nucleotídeos e de aminoácidos. As
30 seqüências das moléculas podem ser definidas em termos de homologia a certas seqüências exemplificadas bem como em termos da capacidade de hibridizar com, ou ser ampliada por, certos iniciadores e sondas exemplifi-

cados. As toxinas providas aqui podem também ser identificadas com base em sua imunoreatividade com certos anticorpos.

Um aspecto importante da presente invenção é a identificação e caracterização de novas famílias de toxinas de *Bacillus*, e genes que codificam estas toxinas. Estas famílias foram designadas MIS-7 e MIS-8. Novas famílias de toxina de tipo WAR E SUP são também reveladas aqui. Toxinas dentro destas famílias, bem como genes que codificam toxinas dentro destas famílias, podem prontamente ser identificados como descritos aqui por, por exemplo, tamanho, seqüência de aminoácidos ou de DNA, e reatividade de anticorpo. Características de seqüência de aminoácidos e de DNA incluem homologia com seqüências exemplificadas, capacidade de hibridizar com sondas de DNA, e capacidade de ser ampliada com iniciadores específicos.

Um gene e uma toxina (que são obteníveis a partir de PS33F1) da família de MIS-1 e um gene e uma toxina (que são obteníveis a partir de PS66D3) da família MIS-2 são também ulteriormente caracterizados aqui.

Uma nova família de toxinas identificadas aqui é a família de MIS-7. Esta família inclui toxinas que podem ser obtidas a partir de isolados de B.t. PS157C1, PS205C e PS201Z. A presente invenção ulteriormente provê sondas e iniciadores para a identificação das toxinas e genes MIS-7.

Uma outra nova família de toxinas identificadas aqui é a família de MIS-8. Esta família inclui toxinas que podem ser obtidas a partir de isolados de B.t. PS31F2 e PS185Y2. A presente invenção ulteriormente provê sondas e iniciadores para a identificação das toxinas e genes MIS-8.

Em uma modalidade preferida, os genes da família de MIS codificam toxinas tendo um peso molecular de cerca de 70 a cerca de 100 kDa e, mais de preferência, as toxinas têm um tamanho de cerca de 80 kDa. Tipicamente, estas toxinas são solúveis e podem ser obtidas a partir do sobrenadante de culturas de *Bacillus* como descritas aqui. Estas toxinas têm toxicidade contra pestes de não mamífero. Em uma modalidade preferida, estas toxinas têm atividade contra pestes de coleopterano. As proteínas MIS são ulteriormente úteis devido a sua capacidade de formar poros em células.

las. Estas proteínas podem ser usadas com segundas entidades incluindo, por exemplo, outras proteínas. Quando usada com uma segunda entidade, a proteína MIS facilitará entrada do segundo agente em uma célula alvo. Em uma modalidade preferida, a proteína MIS interage com receptores MIS em
5 uma célula alvo e causa formação de poro na célula alvo. A segunda entidade pode ser uma toxina ou uma outra molécula cuja entrada na célula alvo é desejada.

A presente invenção ulteriormente refere-se a uma família de toxinas designadas toxinas de tipo WAR. As toxinas WAR tipicamente têm
10 um tamanho de cerca de 30 – 50 kDa e, mais tipicamente, têm um tamanho de cerca de 40 kDa. Tipicamente, estas toxinas são solúveis e podem ser obtidas a partir do sobrenadante de culturas de *Bacillus* como descritas aqui. As toxinas WAR podem ser identificadas com iniciadores descritos aqui bem como anticorpos.

15 Uma família adicional de toxinas provida de acordo com a presente invenção são as toxinas designadas toxinas de tipo SUP. Tipicamente, estas toxinas são solúveis e podem ser obtidas a partir do sobrenadante de culturas de *Bacillus* como descritas aqui. Em uma modalidade preferida, as toxinas SUP são ativas contra pestes de lepidopterano. As toxinas SUP
20 tipicamente têm um tamanho de cerca de 70-100 kDa e, de preferência cerca de 80 kDa. A família SUP é exemplificada aqui por toxinas oriundas de isolado KB59A4-6. A presente invenção provê sondas e iniciadores úteis para a identificação de toxinas e genes na família SUP.

A presente invenção também provê genes e toxinas de *Bacillus*
25 adicionais, incluindo genes e toxinas MIS, WAR e SUP adicionais.

Toxinas nas famílias MIS, WAR e SUP são todas solúveis e podem ser obtidas como descritas aqui a partir das culturas de *Bacillus*. Estas toxinas podem ser usadas sozinhas ou em combinação com outras toxinas para controlar pestes. Por exemplo, toxinas oriundas das famílias MIS po-
30 dem ser usadas em combinação com toxinas de tipo WAR para conseguir o controle de pestes, particularmente pestes de coleopterano. Estas toxinas podem ser usadas, por exemplo, com δ -endotoxinas que são obtidas a partir

de isolados de *Bacillus*.

A tabela 2 provê um sumário das novas famílias de toxinas e genes da presente invenção. Certas famílias MIS são especificamente exemplificadas aqui por toxinas que podem ser obtidas a partir de isolados de B.t. como mostrado na tabela 2. Genes que codificam toxinas em cada uma destas famílias podem ser identificados por uma variedade de parâmetros altamente específicos, incluindo a capacidade de hibridizar com as sondas particulares indicadas na tabela 2. A identidade de seqüência em excesso de cerca de 80% com as sondas indicadas na tabela 2 pode também ser usada para identificar os genes das várias famílias. Também exemplificados são pares de iniciador particulares que podem ser usados para ampliar os genes da presente invenção. Uma porção de um gene dentro das famílias indicadas tipicamente seria amplificável com pelo menos um dos pares de iniciador enumerados. Em uma modalidade, a porção ampliada seria de cerca do tamanho de fragmento indicado. Iniciadores mostrados na tabela 2 consiste em seqüências de polinucleotídeos que codificam peptídeos como mostrados na listagem de seqüência ligada até aqui. Sondas e iniciadores adicionais podem prontamente ser construídos por aqueles versados na técnica de modo que alternam seqüências de polinucleotídeos que codificam as mesmas seqüências de aminoácidos que podem ser usadas para identificar e/ou caracterizar genes adicionais que codificam toxinas pesticidas.

Em uma modalidade preferida, estas toxinas adicionais, e seus genes, poderiam ser obtidos a partir de isolados de *Bacillus*.

Tabela 2				
Família	Isolados	Sondas (SEQ ID NO.)	Pares de Iniciador (SEQ ID NO.)	Tamanho de Fragmento (nt)
MIS-1	PS33F1	37	13 e 22	69
			13 e 23	506
			14 e 23	458

Continuação

Tabela 2				
Família	Isolados	Sondas (SEQ ID NO.)	Pares de Iniciador (SEQ ID NO.)	Tamanho de Fragmento (nt)
MIS-2	PS66D3	5	16 e 24	160
			16 e 25	239
			16 e 26	400
			16 e 27	509
			16 e 28	703
			17 e 25	102
			17 e 26	263
			17 e 27	372
			17 e 28	566
			18 e 26	191
			18 e 27	300
			18 e 28	494
			19 e 27	131
			19 e 28	325
			20 e 28	213
MIS-7	PS205C, PS157C1 (157C1-A), PS201Z	33,35	20 e 30	598
MIS-8	PS31F2, PS185Y2	36,37	31 e 32	585
SUP	KB59A4-6	1	51 e 52	

Além do mais, toxinas quiméricas podem ser usadas de acordo com a presente invenção. Métodos foram desenvolvidos para a produção de toxinas quiméricas úteis por combinação de porções de proteínas de B.t. As

5 porções que são combinadas não necessitam, elas próprias, serem pesticidas contanto que a combinação de porções cria uma proteína quimérica que seja pesticida. Isto pode ser feito usando-se enzimas de restrição, como descrito na, por exemplo, Patente Européia 0 228 838; Ge, A.Z., N.L. Shivarova, D. H. Dean (1989) *Proc. Natl. Acad. Sci. USA* 86: 4037 – 4041; Ge, A.

10 Z., D. Rivers, R. Milne, D. H. Dean (1991) *J. Biol. Chem.* 266: 17954-17958;

Schnepf, H. E., K. Tomczak, J. P. Ortega, H.R. Whiteley (1990) *J. Biol. Chem.* 265: 20923-20930; Honee, G., D. Convents, J. Van Rie, S. Jansens, M. Peferoen, B. Visser (1991) *Mol. Microbiol.* 5:2799-2806. Alternativamente, recombinação usando-se mecanismos de recombinação celular pode ser usada para se conseguir resultados similares. Ver, por exemplo, Caramori, T., A. M. Albertini, A. Galizzi (1991) *Gene* 98:37 – 44; Widner, W. R., H.R. Whiteley (1990) *J. Bacteriol.* 172:2826-2832; Bosch, D., B. Schipper, H. van der Kliei, R.A. de Maagd, W. J. Stickema (1994) *Biotechnology* 12:915-918. Numerosos outros métodos são conhecidos na técnica pelos quais tais DNAs quiméricos podem ser produzidos. A presente invenção destina-se a incluir proteínas quiméricas que utilizam novas seqüências identificadas no presente pedido.

Com os ensinamentos providos aqui, aquele versado na técnica poderia prontamente produzir e usar as várias toxinas e seqüências de polinucleotídeos descritas aqui.

Genes e Toxinas. Os genes e as toxinas úteis de acordo com a presente invenção não incluem apenas as seqüências de comprimento completo mas também fragmentos das seqüências, variantes, mutantes e proteínas de fusão que retêm a atividade pesticida característica das toxinas especificamente exemplificadas aqui. Toxinas e genes quiméricos, produzidos por combinação de porções oriundos de mais do que um gene ou toxina de *Bacillus*, podem também ser utilizados de acordo com os ensinamentos da presente invenção. Como usado aqui, os termos "variantes" ou "variações" de genes referem-se a seqüência de nucleotídeos que codificam as mesmas toxinas ou que codificam toxinas equivalentes tendo atividade pesticida. Como usado aqui, o termo "toxinas equivalentes" refere-se a toxinas tendo a mesma ou essencialmente a mesma atividade biológica contra as pestes alvo que as toxinas exemplificadas. Por exemplo, a patente U.S. no. 5.605.793 descreve métodos para a geração de diversidade molecular adicional por uso de reagrupamento de DNA depois da fragmentação aleatória.

É evidente para uma pessoa versada na técnica que genes que codificam toxinas ativas podem ser identificados e obtidos através de vários

meios. Os genes específicos exemplificados aqui podem ser obtidos a partir de isolados depositados em um depósito de cultura como descrito acima.

Estes genes ou porções ou suas variantes, podem também ser contruídos sinteticamente, por exemplo, por uso de um sintetizador de gene.

5 Variações de genes podem ser prontamente construídas usando-se técnicas padrão para a produção de mutações de ponto. Também, fragmentos destes genes podem ser produzidos usando-se exonucleases disponíveis comercialmente ou endonucleases de acordo com os procedimentos padrão. Por exemplo, enzimas tal como mutagênese de sítio-dirigida ou *Ba/31* podem
10 ser usadas para sinteticamente cortar nucleotídeos oriundos das extremidades destes genes. Também, genes que codificam fragmentos ativos podem ser obtidos usando-se uma variedade de enzimas de restrição. Proteases podem ser usadas para diretamente se obter fragmentos ativos destas toxinas.

15 Genes e/ou toxinas equivalentes que codificam estas toxinas equivalentes podem ser derivadas de isolados de *Bacillus* e/ou bibliotecas de DNA usando-se os ensinamentos providos aqui. Há numerosos métodos para a obtenção das toxinas pesticidas da presente invenção. Por exemplo, anticorpos para as toxinas pesticidas reveladas e reivindicadas aqui podem
20 ser usados para identificar e isolar toxinas de uma mistura de proteínas. Especificamente, anticorpos podem ser criados para as porções das toxinas que são mais constantes e mais distintas de outras toxinas de *Bacillus*. Estes anticorpos podem então ser usados para especificamente identificar toxinas equivalentes com a atividade característica por imunoprecipitação, ensaio imunossorvente ligado por enzima (ELISA), ou Western blotting. Anticorpos para as toxinas reveladas aqui, ou para as toxinas equivalentes, ou fragmentos destas toxinas, podem prontamente ser preparados usando-se procedimentos padrão nesta técnica. Os genes que codificam estas toxinas podem então ser obtidos a partir do microorganismo.

30 Fragmentos e equivalentes que retêm a atividade pesticida das toxinas exemplificadas estão dentro do escopo da presente invenção. Também, por causa da redundância do código genético, uma variedade de dife-

rentes seqüências de DNA podem codificar as seqüência de aminoácidos reveladas aqui. Está bem dentro da habilidade de uma pessoa versada na técnica criar estas seqüências de DNA alternativas que codificam as mesmas, ou essencialmente as mesmas toxinas. Estas seqüências de DNA variantes estão dentro do escopo da presente invenção. Como usado aqui, referência a seqüência "essencialmente a mesma" refere-se a seqüências que têm inserções, adições, deleções ou substituições de aminoácido que não materialmente afetam a atividade pesticida. Fragmentos que retêm atividade pesticida estão também incluídos nesta definição.

Um outro método para a identificação das toxinas e genes da presente invenção é através do uso de sondas de oligonucleotídeos. Estas sondas são seqüências detectáveis. Sondas provêem um método rápido para a identificação de genes que codificam toxina da presente invenção. Os segmentos de nucleotídeo que são usados como sonda de acordo com a invenção podem ser sintetizados usando-se um sintetizador de DNA e procedimentos padrão.

Certas toxinas da presente invenção foram especificamente exemplificadas aqui. Uma vez que estas toxinas são meramente exemplares das toxinas da presente invenção, será prontamente evidente que a presente invenção compreende toxinas variantes ou equivalentes (e seqüência de nucleotídeos que codificam toxinas equivalentes) tendo a mesma atividade pesticida ou similar da toxinas exemplificada. Toxinas equivalentes terão homologia de aminoácido com uma toxina exemplificada. Esta identidade de aminoácido tipicamente será maior do que 60%, de preferência será maior do que 75%, mais de preferência maior do que 80%, de preferência será maior do que 90%, e pode ser maior do que 95%. Estas identidades são como determinadas usando-se técnicas de alinhamento padrão. A homologia de aminoácido será mais alta em regiões críticas da toxina que levam em conta a atividade biológica ou estão envolvidas na determinação de configuração tridimensional que afinal de contas é responsável pela atividade biológica. A este respeito, certas substituições de aminoácido são aceitáveis e podem ser esperadas se estas substituições são em regiões

que não são críticas a atividade ou são substituições de aminoácido conservadores que não afetam a configuração tridimensional da molécula. Por exemplo, aminoácidos podem ser colocados nas seguintes classes: não polar, polar não carregado, básico e ácido. Substituições conservadoras pelo que um aminoácido de uma classe é substituído com um outro aminoácido do mesmo tipo cai dentro de escopo da presente invenção, contanto que a substituição não materialmente afeta a atividade biológica do composto. A tabela 3 provê uma listagem de exemplos de aminoácidos que pertencem a cada classe.

Tabela 3

Classe de aminoácido	Exemplos de aminoácidos
Não polar	Ala, Val, Leu, Ile, Pro, Met, Phe, Trp
Polar não carregado	Gly, Ser, Thr, Cys, Tyr, Asn, Gln
Ácido	Asp, Glu
Básico	Lys, Arg, His

Em alguns casos, substituições não conservadoras podem também ser feitas. O fator crítico é que estas substituições não devem significativamente depreciar a atividade biológica da toxina.

As δ -endotoxinas da presente invenção podem também ser caracterizadas em termos da forma e da localização de inclusões de toxina, que são descritas acima.

Como usado aqui, referência a polinucleotídeos "isolados" e/ou toxinas "purificadas" refere-se a estas moléculas quando elas não estão associadas às outras moléculas com as quais elas seriam encontradas na natureza. Assim, referência a "isolado e purificado" significa o envolvimento da "mão do homem" como descrito aqui. Genes e toxinas quiméricos também envolvem a "mão do homem".

Hospedeiros Recombinantes. Os genes de codificação de toxina da presente invenção podem ser introduzidos para dentro de uma ampla variedade de hospedeiros de planta e microbianos. A expressão do gene de toxina resulta, diretamente ou indiretamente, na produção e na manutenção

do pesticida. Com hospedeiros microbianos adequados, por exemplo, *Pseudomonas*, os micróbios podem ser aplicados aos sítios da peste, onde eles proliferarão e serão ingeridos. O resultado é um controle da peste. Alternativamente, o micróbio que hospeda o gene de toxina pode ser morto ou tratado sob condições que prolongam a atividade da toxina e estabilizam a célula. A célula tratada, que retém a atividade tóxica, então pode ser aplicada ao ambiente da peste alvo.

Onde o gene de toxina de *Bacillus* via um vetor adequado para dentro de um hospedeiro microbiano, e o dito hospedeiro é aplicado ao ambiente em um estado vivo, é essencial que certos micróbios hospedeiros sejam usados. Hospedeiros de microorganismos selecionados que são conhecidos como ocupando a "fitoesfera" (filoplano, filoesfera, rizoosfera e/ou rizoplano) de uma ou mais colheitas de interesse. Estes microorganismos são selecionados de modo a ser capaz de competir com sucesso no ambiente particular (colheita ou outros habitats de insetos) com o microorganismo de tipo selvagem, provêem manutenção estável e expressão do gene que expressa o pesticida de polipeptídeo, e, desejavelmente, provêem proteção aperfeiçoada do pesticida contra a inativação e degradação ambiental.

Um grande número de microorganismos são conhecidos como habitando o filoplano (a superfície das folhas da planta) e/ou rizoosfera (o solo que envolve raízes de planta) de uma ampla variedade de colheitas importantes. Estes microorganismos incluem bactérias, algas e fungos. De particular interesse são microorganismos, tais como bactérias, por exemplo, gêneros *Pseudomonas*, *Erwinia*, *Serratia*, *Klebsiella*, *Xanthomonas*, *Streptomyces*, *Rhizobium*, *Rhodopseudomonas*, *Methylophilus*, *Agrobacterium*, *Acetobacter*, *Lactobacillus*, *Arthrobacter*, *Azotobacter*, *Leuconostoc*, e *Alcaligenes*; fungos, particularmente levedura, por exemplo, gêneros *Saccharomyces*, *Cryptococcus*, *Kluyveromyces*, *Sporobolomyces*, *Rhodotorula*, e *Aureobasidium*. De particular interesse são tais espécies bacterianas da fitoesfera como *Pseudomonas syringae*, *Pseudomonas fluorescens*, *Serratia marcescens*, *Acetobacter xylinum*, *Agrobacterium tumefaciens*, *Rhodopseu-*

domonas spheroides, *Xanthomonas campestris*, *Rhizobium melioli*, *Alcaligenes entrophus*, e *Azotobacter vinlandii*; e espécies de levedura da fitoesfera tais como *Rhodotorula rubra*, *R. glutinis*, *R. marina*, *R. aurantiaca*, *Cryptococcus albidus*, *C. diffluens*, *C. laurentii*, *Saccharomyces rosei*, *S. pretorien-*
 5 *sis*, *S. cerevisiae*, *Sporobolomyces roseus*, *S. odorus*, *Kluyveromyces veronae*, e *Aureobasidium pollulans*. De particular interesse são os microorganismos pigmentados.

Uma ampla variedade de meios estão disponíveis para a introdução de um gene de *Bacillus* que codifica uma toxina para dentro um microorganismo hospedeiro sob condições que permitem manutenção estável
 10 e expressão do gene. Estes métodos são bem conhecidos por aqueles versados na técnica e são descritos, por exemplo, na patente U.S. no. 5.135.867, que é incorporada aqui por referência.

Genes sintéticos que são funcionalmente equivalentes às toxinas da presente invenção podem também ser usados para transformar hospedeiros. Métodos para a produção de genes sintéticos podem ser encontrados na, por exemplo, patente U.S. no. 5.380.831.

Tratamento de Células. Como mencionado acima, células de *Bacillus* ou recombinantes que expressam uma toxina de *Bacillus* podem ser
 20 tratadas para prolongar a atividade de toxina e estabilizar a célula. A microcápsula pesticida que é formada compreende a toxina de *Bacillus* dentro de uma estrutura celular que foi estabilizada e protegerá a toxina quando a microcápsula é aplicada ao ambiente da peste alvo. Células de hospedeiro adequadas podem incluir ou procariotas ou eucariotas. Como hospedeiros,
 25 de particular interesse serão as procariotas e as eucariotas inferiores, tais como fungos. A célula usualmente estará intacta e estará substancialmente na forma proliferativa quando tratada, ao invés de em uma forma de esporo.

Tratamento da célula microbiana, por exemplo, um micróbio contendo o gene de toxina de *Bacillus*, pode ser por meio químico ou físico,
 30 ou por uma combinação de meios químico e/ou físico, contanto que a técnica não nocivamente afete as propriedades da toxina, nem diminua a capacidade celular de proteção da toxina. Métodos para o tratamento de células

microbianas são revelados nas patentes U.S. nos. 4.695.455 e 4.695.462, que são incorporadas aqui por referência.

Métodos e Formulações para o Controle de Pestes. Controle de pestes usando-se os isolados, toxinas e genes da presente invenção pode
5 ser realizado por uma variedade de métodos conhecidos por aqueles versados na técnica. Estes métodos incluem, por exemplo, a aplicação de isolados de *Bacillus* às pestes (ou sua localização), a aplicação de micróbios recombinantes às pestes (ou suas localizações), e a transformação de plantas com genes que codificam as toxinas pesticidas da presente inven-
10 ção. Transformações podem ser feitas por aqueles versados na técnica usando-se técnicas padrão. Materiais necessários para estas transformações são revelados aqui ou estão de outra maneira prontamente disponíveis pelo especializado versado.

Grânulos de isca formulados contendo um atrativo e as toxinas
15 dos isolados de *Bacillus*, ou micróbios recombinantes compreendendo os genes obteníveis a partir de isolados de *Bacillus* revelados aqui, podem ser aplicados ao solo. Produtos formulados podem também ser aplicados como um revestimento de planta ou tratamento de raiz ou tratamento de planta total em estágios posteriores do ciclo de colheita. Tratamentos de planta e
20 de solo de células de *Bacillus* podem ser empregados como poeiras, grânulos ou pós molháveis, por misturação com vários materiais inertes, tais como minerais inorgânicos (filossilicatos, carbonatos, sulfatos, fosfatos, e semelhantes) ou materiais botânicos (espigas de milho em pó, cascas de arroz, cascas de nozes, e semelhantes). As formulações podem incluir auxi-
25 liares espalhadores-adesivos, agentes de estabilização, outros aditivos pesticidas ou tensoativos. Formulações líquidas podem ser de base aquosa ou não aquosa e empregadas como espumas, géis, suspensões, concentrados emulsificáveis, ou semelhantes. Os ingredientes podem incluir agentes reológicos, tensoativos, emulsificadores, dispersantes ou polímeros.

30 Como seria apreciado por uma pessoa versada na técnica, a concentração pesticida variará dependendo da natureza da formulação particular, particularmente se ela é um concentrado ou será usada diretamente.

O pesticida estará presente em pelo menos 1% em peso e pode ser de 100% em peso. As formulações secas terão de cerca de 1-95% em peso do pesticida enquanto as formulações líquidas em geral será de cerca de 1-60% em peso dos sólidos na fase líquida. As formulações que contêm células em geral terão de cerca de 10^2 a cerca de 10^4 células/mg. Estas formulações serão administradas a cerca de 50 mg (líquido ou seco) a 1 kg ou mais por hectare.

As formulações podem ser aplicadas ao ambiente da peste, por exemplo, sole e folhagem, por borrifamento, polvilhamento, regamento, ou semelhantes.

Sondas de Polinucleotídeos. É bem conhecido que DNA possui uma propriedade fundamental chamada complementaridade de base. Na natureza, DNA comumente existe na forma de pares de padrões antiparalelos, as bases sobre cada fio que se projeta daquele fio em relação ao fio oposto. A base adenina (A) sobre um fio sempre estará oposta à base timina (T) sobre outro fio, e a base guanina (G) estará oposta à base citosina (C). As bases são mantidas em aposição por sua capacidade de ligar hidrogênio deste modo específico. Embora cada ligação individual seja relativamente fraca, o efeito global de muitas bases de ligação de hidrogênio adjacentes, juntamente com efeitos de empilhamento de base, é uma ligação estável dos dois fios complementares. Estas ligações podem ser rompidas por tratamento tal como pH alto ou alta temperatura, e estas condições resultam na dissociação, ou "desnaturação" dos dois fios. Se o DNA é então colocado em condições que produzem ligação de hidrogênio das bases termodinamicamente favoráveis, os fios de DNA se anelarão ou "hibridizarão", e reformarão o DNA de fio duplo original. Se realizada sob condições apropriadas, esta hibridização pode ser altamente específica. Isto é, apenas fios com um alto grau de complementaridade de base será capaz de formar estruturas de fios duplos estáveis. A relação da especificidade de hibridização para dar condições de reação é bem conhecida. Assim, hibridização pode ser usada para testar se dois pedaços de DNA são complementares em suas seqüências de base. É este mecanismo de hibridização que facilita o uso

de sondas da presente invenção para prontamente detectar e caracterizar seqüências de DNA de interesse.

As sondas podem ser RNA, DNA ou PNA (ácido nucléico de peptídeo). A sonda normalmente terá pelo menos cerca de 10 bases, mais
5 usualmente pelo menos cerca de 17 bases, e pode ter até cerca de 100 bases ou mais. Sondas mais longas podem prontamente ser utilizadas, e tais sondas podem ter, por exemplo, vários quilobases de comprimento. A seqüência de sondas é projetada ser pelo menos substancialmente complementar a uma porção de um gene que codifica uma toxina de interesse. A
10 sonda não necessita ter complementaridade perfeita à seqüência a qual ela hibridiza. As sondas podem ser marcadas utilizando-se técnicas que são bem conhecidas por aqueles versados na técnica.

Uma abordagem para o uso da presente invenção como sondas vincula primeiramente a identificação por análise de Southern blot de um
15 banco de gene do isolado de *Bacillus* todos os segmentos de DNA homólogos com as seqüências de nucleotídeos reveladas. Assim, é possível, sem o auxílio de análise biológica, conhecer antecipadamente a atividade provável de muitos novos isolados de *Bacillus*, e dos produtos de gene individuais expressos por um dado isolado de *Bacillus*. Uma tal análise de sonda provê
20 um método rápido para a identificação de genes de toxina de alto valor potencialmente comerciais dentro das subespécies variadas de B.t.

Um procedimento de hibridização útil de acordo com a presente invenção tipicamente inclui as etapas iniciais de isolamento da amostra de DNA de interesse e purificação dela quimicamente. Ou bactérias lisadas ou
25 ácido nucléico fracionado total isolados de bactérias podem ser usados. Células podem ser tratadas usando-se técnicas conhecidas para liberar seu DNA (e/ou RNA). A amostra de DNA pode ser cortada em pedaços com uma enzima de restrição apropriada. Os pedaços podem ser separados pelo tamanho através de eletroforese em um gel, usualmente agarose ou acrilamida. Os pedaços de interesse podem ser transferidos para uma membrana de imobilização.
30

A técnica de hibridização particular não é essencial a presente

invenção. Quando aperfeiçoamentos são feitos nas técnicas de hibridização, eles podem ser prontamente aplicados.

A sonda e a amostra podem então ser combinadas em uma solução de tampão de hibridização e mantidas em uma temperatura apropriada até que o anelamento ocorra. Depois disso, a membrana é lavada livre de materiais estranhos, deixando a amostra e as moléculas de sonda ligadas tipicamente detectadas e quantificadas por auto-radiografia e/ou contagem de cintilação líquida. Como é bem conhecido na técnica, se a molécula de sonda e a amostra de ácido nucléico se hibridizam por formação de uma ligação não covalente forte entre as duas moléculas, pode ser razoavelmente admitido que a sonda e a amostra são essencialmente idênticas. O rótulo detectável da sonda provê um meio para a determinação de um modo conhecido se hibridização ocorreu.

No uso dos segmentos de nucleotídeo como sondas, a sonda particular é marcada com qualquer rótulo adequado conhecido por aqueles versados na técnica, incluindo rótulos radioativos ou não radioativos. Rótulos radioativos típicos incluem ^{32}P , ^{35}S , ou semelhantes. Rótulos não radioativos incluem, por exemplo, ligando tal como biotina ou tiroxina, bem com enzimas tais como hidrolases e peroxidases, ou os vários quimiluminescentes tais como luciferina ou compostos fluorescentes como fluoresceína e seus derivados. As sondas podem ser produzidas inerentemente como descritas no Pedido Internacional no. WO 93/16094.

Vários graus de severidade de hibridização podem ser empregados. Quanto mais severa é a condição, maior é a complementaridade que é necessária para formação duplex. Severidade pode ser controlada por temperatura, concentração de sonda, comprimento de sonda, concentração iônica, tempo, e semelhantes. De preferência, hibridização é conduzida sob condições de severidade moderada a alta por técnicas bem conhecidas na técnica, como descrito, por exemplo, em Keller, G. H., M.M Manak (1987) DNA Probes, Stockton Press, Nova York, NY, págs. 169-170.

Como usado aqui, condições de "severidade de moderada a alta" para hibridização refere-se a condições que conseguem o mesmo, ou

cerca do mesmo grau de especificidade de hibridização que as condições empregadas pelas requerentes atuais. São providos exemplos de condições de severidade de moderada e alta aqui. Especificamente, hibridização de DNA imobilizado sobre Southern blots com sondas específicas de gene marcadas por ^{32}P foi realizada por métodos padrão (Maniatis e outros). Em geral, hibridização e lavagens subseqüentes são realizadas sob condições de severidade de moderada a alta que permitiram a detecção de seqüências alvo com homologia aos genes de toxina exemplificados. Para sondas de gene de DNA de fio duplo, hibridização foi realizada de um dia para o outro a 20 – 25°C abaixo da temperatura de fusão (T_m) do híbrido de DNA em 6X SSPE, solução de 5X Denhardt, SDS a 0,1%, 0,1 mg/ml de DNA desnaturado. A temperatura de fusão é descrita pela seguinte fórmula (Beltz, G.A., K.A. Jacobs, T.H. Eickbush, P.T. Cherbas e F.C. Kafatos [1983] *Methods of Enzymology*, R. Wu, L. Grossman e K. Moldave [eds.] Academic Press, Nova York 100: 266-285).

$$T_m = 81,5^\circ\text{C} + 16,6 \text{ Log}[\text{Na}^+] + 0,41(\% \text{G+C}) - 0,61(\% \text{ de formamida}) - 600/\text{comprimento de duplex em pares de base}.$$

Lavagens são tipicamente realizadas como se segue:

(1) duas vezes a temperatura ambiente por 15 minutos em 1X SSPE, SDS a 0,1% (lavagem de baixa severidade).

(2) Uma vez a $T_m - 20^\circ\text{C}$ por 15 minutos em 0,2X SSPE, SDS a 0,1% (lavagem de severidade moderada).

Para sondas de oligonucleotídeos, hibridização foi realizada de um dia para o outro a 10-20°C abaixo da temperatura de fusão (T_m) do híbrido em 6X SSPE, solução de 5X Denhardt, SDS a 0,1%, 0,1 mg/ml de DNA desnaturado. T_m para sondas de oligonucleotídeos foi determinada pela seguinte fórmula:

$$T_m (^\circ\text{C}) = 2 (\text{número de pares de base T/A}) + 4 (\text{número de pares de base G/C})$$

(Suggs, S.V., T. Miyake, E. H. Kawashime, M.J. Johnson, k. Itakura, e R. B. Wallace [1981] *ICN-UCLA Symp. Dev. Bio. Using Purified Genes*, D.D. Brown [ed], Academic Press, Nova York, 23:683-693).

Lavagens foram tipicamente realizadas como se segue:

(1) duas vezes a temperatura ambiente por 15 minutos em 1X SSPE, SDS a 0,1% (lavagem de baixa severidade).

(2) Uma vez na temperatura de hibridização por 15 minutos em 1X SSPE, SDS a 0,1% (lavagem de severidade moderada).

5 Em geral, sal e/ou temperatura podem ser alterados para mudar severidade. Com o fragmento de DNA marcado > 70 bases mais ou menos de comprimento, as seguintes condições podem ser usadas:

baixa: 1 ou 2X SSPE, temperatura ambiente

baixa: 1 ou 2X SSPE, 42°C

10 moderada: 0,2X ou 1X SSPE, 65°C

alta: 0,1X SSPE, 65°C

Formação duplex e estabilidade dependem da complementaridade substancial entre os dois fios de um híbrido, e, como observado acima, um certo grau de mal emparelhamento pode ser tolerado. Portanto, as se-
 15 quências de sondas da presente invenção incluem mutações (tanto simples quanto múltiplas), deleções, inserções das seqüências descritas e suas combinações, em que as ditas mutações, inserções e deleções permitem a formação de híbridos estáveis com o polinucleotídeo alvo de interesse. Mu-
 tações, inserções e deleções podem ser produzidas em uma dada seqüên-
 20 cia de polinucleotídeos de muitos modos, e estes métodos são conhecidos por um artesão normalmente versado. Outros métodos podem se tornar conhecidos no futuro.

Assim, variantes mutacionais, insercionais e delecionais das seqüências de nucleotídeos reveladas podem ser prontamente preparadas
 25 que são bem cohecidas por aquele versado na técnica. Estas variantes podem ser usadas do mesmo modo que as seqüências de iniciadores exemplificadas contanto que as variantes têm homologia de seqüência substancial com a seqüência original. Como usado aqui, homologia de seqüência substancial refere-se a homologia que é suficiente para possibilitar a sonda
 30 de variante funcionar na mesma capacidade que a sonda original. De preferência, esta homologia é maior do que 50%; mais de preferência, esta homologia é maior do que 75%; e mais de preferência, esta homologia é maior

do que 90%. O grau de homologia necessário para a variante funcionar em sua capacidade pretendida dependerá do uso pretendido da sequência. Está bem dentro da habilidade de um pessoa treinada nesta técnica produzir mutações mutacionais, insercionais e delecionais que são projetadas para aperfeiçoar a função da sequência ou de outra maneira prover uma vantagem metodológica.

Tecnologia por PCR. Reação de cadeia de polimerase (PCR) é uma síntese inicializada, enzimática, repetitiva de uma sequência de ácidos nucleicos. Este procedimento é bem conhecido e é comumente usado por aqueles versados na técnica (ver Mullis, nas patentes U.S. nos. 4.683.195, 4.683.202 e 4.800.159; Saiki, Randall K., Stephen Scharf, Fred Faloona, Kary B. Mullis, Glenn T. Horn, Henry A. Erlich, Norman Arnheim [1985] "Enzymatic Amplification of beta-Globin Genomic Sequences and Restriction Site Analysis for Diagnosis of Sickle Cell Anemia" *Science* 230:1350-1354).

PCR é a base da ampliação enzimática de um fragmento de DNA de interesse que é flanqueado por dois iniciadores de oligonucleotídeo que hibridizam para dar fios da sequência alvo. Os iniciadores são orientados com as extremidades 3' apontando em direção a outra. Ciclos repetidos de desnaturação a quente do modelo, anelamento dos iniciadores com suas sequências complementares, e extensão dos iniciadores anelados com uma polimerase de DNA resultam na ampliação do segmento definido pelas extremidades 5' dos iniciadores de PCR. Uma vez que o produto de extensão de cada iniciador pode servir como um modelo para o outro iniciador, cada ciclo essencialmente dobra a quantidade de fragmento de DNA produzido no ciclo anterior. Isto resulta no acúmulo exponencial do fragmento alvo específico, até várias milhões de vezes em umas poucas horas. Por uso de uma polimerase de DNA termoestável tal como *Taq* polimerase, que é isolada da bactéria termofílica *Thermus aquaticus*, o processo de ampliação pode ser totalmente automatizado. Outras enzimas que podem ser usadas são conhecidas por aqueles versados na técnica.

As sequências de DNA da presente invenção podem ser usadas como iniciadores para ampliação de PCR. Na realização de ampliação de

PCR, um certo grau de desparelhamento pode ser tolerado entre o iniciador e o modelo. Portanto, mutações, deleções e inserções (especialmente adições de nucleotídeos a extremidade 5') dos iniciadores exemplificados caem dentro do escopo da presente invenção. Mutações, inserções e deleções
 5 podem ser produzidas em um dado iniciador por métodos conhecidos por um artesão normalmente versado.

Todas as referências citadas aqui são aqui incorporadas por referência.

O que se segue são exemplos que ilustram os procedimentos
 10 para a prática da invenção. Estes exemplos não devem ser interpretados como limitantes. Todas as percentagens são em peso e todas as proporções de mistura de solvente são em volume a não ser que de outra maneira observado.

Exemplo 1 – Cultivo de Isolados de *Bacillus* Úteis de Acordo com a Presente Invenção 15

O hospedeiro celular contendo o gene inseticida de *Bacillus* pode ser desenvolvido em qualquer meio nutriente conveniente. Estas células podem então ser coletadas de acordo com os meios convencionais. Alternativamente, as células podem ser tratadas antes da coleta.

20 As células de *Bacillus* da invenção podem ser cultivadas usando-se meios padrão e técnicas de fermentação. Durante o ciclo de fermentação, as bactérias podem ser coletadas por primeiramente separação dos fragmentos celulares lisados, cristais, esporos, e células vegetativas de *Bacillus* do caldo de fermentação por meio bem conhecido na técnica. Quais-
 25 quer δ -endotoxinas de cristal ou esporos de *Bacillus* formados podem ser recuperados empregando-se técnicas bem conhecidas e usadas como uma preparação de B.t. de δ -endotoxina convencional. O sobrenadante oriundo do processo de fermentação contém toxinas da presente invenção. As toxinas são isoladas e purificadas empregando-se técnicas bem conhecidas.

30 Uma subcultura de isolados de *Bacillus*, ou seus mutantes, podem ser usados para inocular o seguinte meio, conhecido como caldo de TB:

	Triptona	12 g/l
	Extrato de levedura	24 g/l
	Glicerol	4 g/l
	KH_2PO_4	2,1 g/l
5	K_2HPO_4	14,7 g/l
	pH	7,4

O fosfato de potássio foi adicionado ao caldo tratado em autoclave depois do resfriamento. Frascos foram incubados a 30°C sobre um agitador rotativo a 250 rpm por 24-36 horas.

10 O procedimento acima pode ser prontamente aumentado de escala para fermentadores grandes por procedimentos conhecidos na técnica.

O *Bacillus* obtidos na fermentação acima, pode ser isolado por procedimentos bem conhecidos na técnica. Um procedimento frequentemente usado é submeter o caldo de fermentação coletado a técnicas de separação, por exemplo, centrifugação. Em uma modalidade específica, proteínas de *Bacillus* úteis de acordo com a presente invenção podem ser obtidas a partir do sobrenadante. O sobrenadante de cultura contendo a(s) proteína(s) ativa(s) pode ser usado em bioensaios.

Alternativamente, uma subcultura de isolados de *Bacillus*, ou seus mutantes, pode ser usado para inocular a seguinte peptona, glicose, meio de sais:

	bacto peptona	7,5 g/l
	glicose	1,0 g/l
	KH_2PO_4	3,4 g/l
25	K_2HPO_4	4,35 g/l
	solução de sal	5,0 ml/l
	solução de CaCl_2	5,0 ml/l
	pH	7,2
	Solução de sais (100 ml)	
30	$\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$	2,46 g
	$\text{MnSO}_4 \cdot \text{H}_2\text{O}$	0,04 g
	$\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$	0,28 g

$\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$	0,40 g
Solução de CaCl_2 (100 ml)	
$\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$	3,66 g

A solução de sais e a solução de CaCl_2 são esterilizadas por
 5 filtragem e são adicionadas ao caldo cozido e tratado em autoclave na hora
 de inoculação. Frascos são incubados a 30°C sobre um agitador rotativo a
 200 rpm por 64 hs.

O procedimento acima pode ser prontamente aumentado de es-
 cala para fermentadores grandes por procedimentos bem conhecidos na
 10 técnica.

Esporos e/ou cristais de *Bacillus*, obtidos na fermentação acima,
 podem ser isolados por procedimentos bem conhecidos na técnica. Um pro-
 cedimento freqüentemente usado é para submeter o caldo de fermentação
 coletado a técnicas de separação, por exemplo, centrifugação.

15 Exemplo 2 – Isolamento e Preparação de DNA Celular para PCR

DNA pode ser preparado a partir de células desenvolvidas sobre
 ágar de Spizizen, ou outro ágar enriquecido ou mínimo conhecido por
 aqueles versados na técnica, por cerca de 16 horas. Ágar de ácido casami-
 no de Spizizen compreende 23,2 g/l de sais mínimos de Spizizen
 20 $[(\text{NH}_4)_2\text{SO}_4, 120 \text{ g}; \text{K}_2\text{HPO}_4, 840 \text{ g}; \text{KH}_2\text{PO}_4, 360 \text{ g}; \text{citrato de sódio}, 60 \text{ g};$
 $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}, 12 \text{ g}.$ Total: 1392 g], 1,0 g/l de ácidos casaminos livres de vi-
 tamina; 15,0 g/l de ágar Difco. Na preparação do ágar, a mistura foi tratada
 em autoclave por 30 minutos então uma solução de glicose a 50%, estéril
 pode ser adicionada para dar uma concentração final de 0,05% (1/100 vol).
 25 Uma vez que as células são desenvolvidas por cerca de 16 horas, um em-
 plastro de cerca de 1 cm^2 de células pode ser raspado a partir da ágar para
 dentro de 300 μl de Tris-HCl a 10 mM (pH 8,0)-EDTA a 1 mM. Proteinase K
 foi adicionada a 50 $\mu\text{g/ml}$ e foi incubada a 55°C por 15 minutos. Outras pro-
 teases adequadas que carecem de atividade de nuclease podem ser usa-
 30 das. As amostras foram então colocadas em um banho de água fervente por
 15 minutos para inativar a proteinase e desnaturam o DNA. Isto também
 precipita componentes indesejados. As amostras são então centrifugadas a

14.000 x g em uma micrófuga de Eppendorf a temperatura ambiente por 5 minutos para se remover fragmento celular. Os sobrenadantes contendo DNA bruto foram transferidos para tubos frescos e foram congelados a – 20°C até que fossem usados em reações de PCR.

- 5 Alternativamente, DNA celular total pode ser preparado a partir de células desenvolvidas em placas usando-se o kit de Tecido QIAmp da Qiagen (Santa Clarita, CA) seguindo as instruções do fabricante.

Exemplo 3 – Iniciadores Úteis para a Caracterização e/ou Identificação de Genes de Toxina

- 10 O seguinte conjunto de iniciadores de PCR pode ser usado para identificar e/ou caracterizar genes da presente invenção, que codificam toxinas pesticidas:

GGRTTAMTTGGRTAYTATTT (SEQ ID NO. 3)

ATATCKWAYATTKGCATTTA (SEQ ID NO. 4)

- 15 Códigos de nucleotídeo redundantes usados através de toda a presente revelação estão de acordo com a convenção de IUPAC e incluem:

R = A ou G

M = A ou C

Y = C ou T

- 20 K = G ou T

W = A ou T

Exemplo 4 – Identificação e Seqüênciação de Genes que Codificam Novas Toxinas de Proteína Solúveis Oriundas de Cepas de *Bacillus*

- 25 PCR usando-se iniciadores de SEQ ID NO. 3 e SEQ ID NO. 4 foi realizado sobre DNA genômico celular total isolado de uma ampla faixa de cepas de B.t. Aquelas amostras que fornecem uma tira de cerca de 1 kb para a caracterização por seqüênciação de DNA. Fragmentos de DNA ampliados foram primeiramente clonados para dar vetor de plasmídeo de clonagem de TA de DNA por PCR, pCR2.1, como descrito pelo fornecedor (Invitrogen, San Diego, CA). Plasmídios foram isolados dos clones recombinantes e foram testados quanto a presença de um inserto de cerca de 1 kpb por
- 30 PCR usando-se os iniciadores de vetor de plasmídeo, T3 e T7.

As seguintes cepas forneceram a tira esperada de cerca de 1000 pb, assim indicando a presença de um gene de toxina de tipo MIS: PS66D3, PS177C8, PS177I8, PS33F1, PS157C1 (157C1-A), PS201Z, PS31F2 e PS185Y2.

5 Plasmídios foram então isolados para o uso como a seqüênciação de modelos usando-se kits miniprep QIAGEN (Santa Clarita, CA) como descritos pelo fornecedor. Reações de seqüênciação foram realizadas usando-se um Kit de Reação Pronta de Seqüênciação de Ciclo Terminador de Corante da PE Applied Biosystems. As reações de seqüênciação foram
10 realizadas sobre um Sequenciador Automatizado ABI PRISM 377. Dados de seqüência foram coletados, foram editados e foram montados usando-se a Coleção ABI PRISM 377, software Factura e AutoAssembler da PE ABI.

Seqüências de DNA foram determinadas para porções de novas toxinas de gene oriundas dos seguintes isolados: PS66D3, PS177C8,
15 PS177I8, PS33F1, PS157C1 (157C1-A), PS201Z, PS31F2 e PS185Y2. Estas seqüências de nucleotídeos são mostradas nas SEQ ID NOs. 5, 7, 9, 38, 33, 35, 36 e 37, respectivamente. Seqüências de polipeptídeos foram deduzidas para porções das novas toxinas solúveis, codificadas oriundas dos seguintes isolados: PS66D3, PS177C8, PS177I8, PS157C1 (toxina 157C1-A).
20 Estas seqüências de nucleotídeos são mostradas nas SEQ ID NOs. 6, 8, 10 e 34, respectivamente.

Exemplo 5 – Polimorfismo de Comprimento de Fragmento de Restrição (RFLP) de Toxinas Oriundas de Cepas de *Bacillus thuringiensis*

O DNA celular total foi preparado a partir de várias cepas de
25 *Bacillus thuringiensis* (B.t.) foi desenvolvido para uma densidade ótica de 0,5 – 0,8 em luz visível de 600 nm. DNA foi extraído usando-se o kit Qiagen Genomic-tip 500/G e Conjunto de Tampão de DNA Genômico de acordo com o protocolo para bactérias gram-positivas (Qiagen Inc.; Valencia, CA).

Hibridizações Southern padrão usando-se sondas ³²P-marcadas
30 foram usadas para identificar e caracterizar novos genes de toxina das preparações de DNA genômico totais. DNA genômico total preparado foi digerido com várias enzimas de restrição, foi eletroforado sobre um gel de aga-

rose a 1%, e foi imobilizado sobre uma membrana de náilon de suporte usando-se métodos padrão (Maniatis e outros).

Fragmentos DNA ampliados por PCR de 1,0-1,1 kb de comprimento foram purificados por gel para o uso como sondas. Cerca de 25 ng de cada fragmento de DNA foram usados como um modelo para a inicialização de síntese de DNA nascente usando-se fragmento de Klenow de polimerase I de DNA (New England Biolabs), iniciadores de hexanucleotídeo aleatórios (Boehringer Mannheim) e ^{32}P dCTP.

Cada fragmento ^{32}P -marcado serviu como uma sonda específica a sua mancha de DNA genômico correspondente. Hibridizações de DNA imobilizado com sondas ^{32}P -marcadas foram realizadas em tampão aquoso padrão que consiste em 5X SSPE, solução de 5X Denhardt, SDS a 0,5%, 0,1 mg/ml a 65°C de um dia para o outro. Manchas ("blots") foram lavadas sob severidade moderada em 0,2X SSC, 0,1% de SDS a 65°C e foram expostas a película. Dados de RFLP que mostram tiras de hibridização específicas contendo todas as partes do novo gene de interesse foram obtidos para cada cepa.

Tabela 3

(cepa)/nome do gene	Sonda SEQ I.D. número	Dados de RFLP (tamanhos de tira aproximados)
(PS)66D3	24	BamHI: 4,5 kbp, HindIII: >23 kbp, KpnI: 23 kbp, PstI: 15 kbp, XbaI: >23 kbp
(PS)177I8	33	BamHI: >23 kbp, EcoRI: 10 kbp, HindIII: 2 kbp, SalI: >23 kbp, XbaI: 3,5 kbp

Em experiências separadas, sondas alternativas de genes MIS e WAR foram usadas para detectar novos genes de toxina em Southern blots de DNA genômico por auto-radiografia de ^{32}P ou por métodos não radioativos usando-se o sistema de detecção e marcação de ácido nucléico por DIG (Boehringer Mannheim; Indianapolis, IN). Fragmentos de DNA cerca de 2,6 kbp (gene de toxina PS177C8 MIS; SEQ ID NO.7) e 1,3 kbp (gene de toxina

PS177C8 WAR); SEQ ID NO. 11) de comprimento foram ampliados por PCR a partir de plasmídio pMYC2450 usando-se iniciadores homólogos às extremidades 5'e 3' de cada gene respectivo. pMYC2450 é um plasmídio recombinante contendo PS177C8 MIS e WAR sobre um fragmento de ClaI de
 5 cerca de 14 kpb em pHTBlueII (um vetor de transporte nos dois sentido de *E. coli*/*B. thuringiensis* constituído de pBluescript S/K (Stratagene, La Jolla, CA) e a origem de replicação de um plasmídio de B.t. residente [D. Lereclus e outros, 1989; FEMS Microbiology Letters 60:211-218]). Estes fragmentos de DNA foram usados como sondas para MIS RFLP classes de A até N e
 10 WAR RFLP classes de A até L. Dados de RFLP na tabela 4 para classe O foram gerados usando-se fragmentos MIS de cerca de 1636 pb ampliados com iniciadores S1-633F (CACTCAAAAATGAAAGGGAAA; SEQ ID NO. 39) e S1-2269R (CCGGTTTTATTGATGCTAC; SEQ ID NO. 40). Dados de RFLP na tabela 5 para classe M foram gerados usando-se fragmentos WAR
 15 de cerca de 495 pb ampliados com iniciadores S2-501F (AGAACAAATTTTATGATAGGG; SEQ ID NO 41) e S2-995R (TCCCTAAAGCATCAGAAATA; SEQ ID NO. 42).

Fragmentos foram purificados por gel e cerca de 25 ng de cada fragmento de DNA foram aleatoriamente marcados com ^{32}P para detecção
 20 radioativa ou cerca de 30 ng de cada fragmento foram aleatoriamente marcados com kit de DIG High Prime para detecção não radioativa. Hibridização de DNA imobilizado com sondas ^{32}P marcadas aleatoriamente foi realizada em condições de formamida padrão: formamida a 50%, 5X SSPE, solução de 5X Denhardt, SDS a 2%, 0,1 mg/ml de DNA de esperma sonificado a 42°C
 25 de um dia para o outro. Manchas foram lavadas sob baixa severidade em 2X SSC, SDS a 0,1% a 42°C e foram expostas a película. Dados de RFLP que mostram tiras de DNA contendo todo ou parte do novo gene de interesse foram obtidos para cada cepa.

Dados de RFLP usando-se sondas MIS como discutido acima
 30 eram como se segue:

Tabela 4

RFLP Classe	Nome da cepa	Dados de RFLP (tamanho de tira aproximado em pares de base)
A	177C8, 74H3, 66D3	HindIII: 2.454; 1.645 XbaI: 14.820; 9.612; 8.138; 5.642; 1.440
B	177I8	HindIII: 2.454 XbaI: 3.500 (muito fraco 7.000)
C	66D3	HindIII: 2.454 (fraco 20.000) XbaI: 3.500 (fraco 7.000)
D	28M, 31F2, 71G5, 71G7, 71I1, 71N1, 146F, 185Y2, 201JJ7, KB73, KB68B46-2, KB71A35- 4, KB71A116-1	HindIII: 11.738; 7.614 XbaI: 10.622; 6.030
D ₁	70B2, 71C2	HindIII: 11.738; 8.698; 7.614 XbaI: 11.354; 10.622; 6.030
E	KB68B51-2, KB68B55-2	HindIII: 6.975; 2.527 XbaI: 10.000; 6.144
F	KB53A49-4	HindIII: 5.766 XbaI: 6.757
G	86D1	HindIII: 4.920 XbaI: 11.961
H	HD573B, 33F1, 67B3	HindIII: 6.558; 1.978 XbaI: 7.815; 6.558
I	205C, 40C1	HindIII: 6.752 XbaI: 4.618
J	130A3, 143A2, 157C1	HindIII: 9.639; 3.943; 1.954; 1.210 XbaI: 7.005; 6.165; 4.480; 3.699
K	201Z	HindIII: 9.639; 3.943; 1.954; 1.210 XbaI: 7.232; 6.365
L	71G4	HindIII: 7.005 XbaI: 9.639
M	KB42A33-8, KB71A72-1, KB71A133-11	HindIII: 3.721 XbaI: 3.274
N	KB71A134-2	HindIII: 7.523 XbaI: 10.360; 3.490
O	KB69A125-3, KB69A127-7, KB69A136,2, KB71A20-4	HindIII: 6.360; 3.726; 1.874; 1.098 XbaI: 6.360; 5.893; 5.058; 3.726

Dados de RFLP usando-se sondas WAR como discutido acima eram como se segue:

Tabela 5		
RFLP Classe	Nome da cepa	Dados de RFLP (tamanho de tira aproximado em pares de base)
A	177C8, 74H3	HindIII: 3.659, 2.454, 606 XbaI: 5.457, 4.469, 1.440, 966
B	177I8, 66D3	Dados indisponíveis
C	28M, 31F2, 71G5, 71G7, 71I1, 71N1, 146F, 185Y2, 201JJ7, KB73, KB68B46-2, KB71A35-4, KB71A116-1	HindIII: 7.614 XbaI: 10.985, 6.235
C ₁	70B2, 71C2	HindIII: 8.698, 7.614 XbaI: 11.354, 6.235
D	KB68B51-2, KB68B55-2	HindIII: 7.200 XbaI: 6.342 (e 11.225 para 51-2) (e 9.888 para 55-2)
E	KB53A49-4	HindIII: 5.766 XbaI: 6.757
F	HD573B, 33F1, 67B3	HindIII: 3.348, 2.037 (e 6.558 para HD573B apenas) XbaI: 6.953 (e 7.815, 6.185 para HD573B apenas)
G	205C, 40C1	HindIII: 3.158 XbaI: 6.558, 2.809
H	130A3, 143A2, 157C1	HindIII: 4.339, 3.361, 1.954, 660, 349 XbaI: 9.043, 4.203, 3.583, 2.958, 581, 454
I	201Z	HindIII: 4.480, 3.819, 703 XbaI: 9.336, 3.256, 495
J	71G4	HindIII: 7.005 XbaI: 9.639
K	KB42A33-8, KB71A72-1, KB71A133-11	Nenhum sinal de hibridização
L	KB71A134-2	HindIII: 7.523 XbaI: 10.360
M	KB69A125-3, KB69A127-7, KB69A136,2, KB71A20-4	HindIII: 5.058, 3.726, 3.198, 2.745, 257 XbaI: 5.255, 4.341, 3.452, 1.490, 474

Exemplo 6 – Caracterização e/ou Identificação de Toxinas WAR

Em uma modalidade ulterior da presente invenção, toxinas pesticidas podem ser caracterizadas e/ou identificadas por seu nível de reatividade com anticorpos para toxinas pesticidas exemplificadas aqui. Em uma

5 modalidade específica, anticorpos podem ser criados para toxinas WAR tal como a toxina obtível a partir de PS177C8a. Outras toxinas WAR podem então ser identificadas e/ou caracterizadas por sua reatividade com os anticorpos. Em uma modalidade preferida, os anticorpos são anticorpos policlonais. Neste exemplo, toxinas com similaridade maior a toxina 177C8a-WAR

10 teria a reatividade maior com os anticorpos policlonais. Toxinas WAR com diversidade maior reagem com os anticorpos policlonais 177C8a, mas em uma extensão menor. Toxinas que imunoreagem com anticorpos policlonais criados para a toxina 177C8a WAR podem ser obtidos a partir dos, por exemplo, isolados designados PS177C8a, PS177I8, PS66D3, KB68B55-2,

15 PS185Y2, KB53A49-4, KB68B51-2, PS31F2, PS74H3, PS28M, PS71G6, PS71G7, PS71I1, PS71N1, PS201JJ7, KB73, KB68B46-2, KB71A35-4, KB71A116-1, PS70B2, PS71C2, PS86D1, HD573B, PS33F1, PS67B3, PS205C, PS40C1, PS130A3, PS143A2, PS157C1, PS201Z, PS71G4, KB42A33-8, KB71A72-1, KB71A133-11, KB71A134-2, KB69A125-3,

20 KB69A127-7, KB69A136-2, e KB71A20-4. Isolados PS31F2 e KB68B46-2 mostram reatividade de anticorpo muito fraca, sugerindo diversidade vantajosa.

Exemplo 7 – Clonagem Molecular e Análise de Sequência de DNA de Genes de Proteína Inseticida Solúvel (MIS e WAR) Oridundos de Cepa de *Bacillus thuringiensis* PS205C

DNA celular total foi preparado a partir de cepa de *Bacillus thuringiensis* PS205C densenvolvida para uma densidade ótica de 0,5-0,8 a 600 nm de luz visível em caldo de Luria Bertani (LB). DNA foi extraído usando-se o kit Qiagen Genomic-tip 500/G e Conjunto de Tampão de DNA

30 Genômico de acordo com o protocolo para bactérias gram-positivas (Qiagen Inc.; Valencia, CA). Uma biblioteca de cosmídio PS205C foi construída no vetor SuperCos (Stratagene) usando-se insertos de DNA celular total

PS205C parcialmente digerido com NdeI. Células XL1-Blue (Stratagene) foram transfectadas com cosmídio embalados para se obter clones resistentes a carbenicilina e canamicina. 576 Colônias de cosmídios foram desenvolvidas em blocos de 96 cavidades em 1 ml de LB + carbenicilina (100
5 µg/ml) + canamicina (50 µg/ml) a 37°C por 18 horas e foram laminadas em réplica sobre filtros de náilon para a triagem por hibridização.

Um "amplicon" de PCR contendo cerca de 1000 pb do gene PS205 MIS foi ampliado a partir de DNA genômico PS205 usando-se iniciadores SEQ ID NO. 3 e SEQ ID NO. 4 como descrito no exemplo 4. O frag-
10 mento de DNA foi purificado por gel usando-se extração QiaexII (Qiagen). A sonda foi radiomarcada com ³²P-dCTP usando-se o kit Prime-It II (Stratagene) e foi usada em solução de hibridização aquosa (6X SSPE, solução de 5X Denhardt, SDS a 0,1%, 0,1 mg/ml de DNA desnaturado) com os filtros de levantamento de colônia a 65°C por 16 horas. Os filtros de levantamento de
15 colônia foram brevemente lavados 1X em 2X SSC/SDS a 0,1% a temperatura ambiente seguido por duas lavagens adicionais por 10 minutos em 0,5X SSC/SDS a 0,1%. Os filtros foram então expostos a película de raios X por 5,5 horas. Um clone de cosmídio que hibridizou fortemente para dar a sonda foi selecionado para outra análise. Este clone de cosmídio foi confirmado
20 conter o gene MIS por ampliação de PCR com iniciadores SEQ ID NO. 3 e SEQ ID NO. 4. Este clone de cosmídio foi designado como pMYC3105; células XL-1 Blue MR de *E. coli* contendo pMYC3105 são designadas MR992.

Uma subcultura de MR992 foi depositada na coleção permanente da Patent Culture Collection (NRRL), Regional Research Center,
25 1815 North University Street, Peoria, Illinois 61604 USA em 4 de maio de 1999. O número de acesso é NRRL B-30124. Um clone de plasmídio truncado para PS205C foi também depositado em 4 de maio de 1999. O número de acesso é NRRL B-30122.

Para seqüenciar os genes PS205C MIS e WAR, inserções de
30 transposon aleatórias em pMYC3105 foram geradas usando-se os protocolos e Sistema de Inicialização de Genoma GPS-1 (New England Biolabs). O vetor de transposição de GPS2 que codifica resistência a clorofenicol foi

escolhido para a seleção de cosmídios contendo inserções. Cosmídios pMYC3105 que adquiriram transposons foram identificados por transformação e seleção de XL1-Blue MR de *E. coli* em meios contendo ampicilina, canamicina e cloranfenicol. Modelos de cosmídios foram preparados a partir

5 de colônias individuais para o uso como modelos de seqüenciação usando-se plasmídio prep Multiscreen de 96 cavidades (Millipore). Os genes de toxina MIS e WAR codificados por pMYC3105 foram seqüenciados com iniciadores GPS2 usando-se o sistema de seqüenciação automatizado ABI377 e o programa associado. Os genes MIS e WAR demonstraram estar localiza-

10 dos perto um do outro em um óperon transcricional aparente. As seqüências de polipeptídeos deduzidas e nucleotídeos são designadas como novos SEQ ID NOs. 43-46.

Exemplo 8 – Clonagem Molecular e Análise de Seqüência de DNA de Genes de Proteína Inseticida Solúvel (MIS e WAR) Oridundos de Cepa de *Bacillus thuringiensis* PS31F2

15

a. Preparação e Clonagem de DNA Genômico

DNA celular foi preparado a partir da cepa de *Bacillus thuringiensis* PS31F2 desenvolvidas para uma densidade ótica de 0,5-0,8 a 600 nm de luz visível em caldo de Luria Bertani (LB). DNA foi extraído usando-se o

20 kit Qiagen Genomic-tip 500/G ou Genomic-Tip 20/G e Conjunto de Tampão de DNA Genômico (Qiagen Inc.; Valencia, CA) de acordo com o protocolo para bactérias gram positivas.

Biblioteca lambda contendo DNA genômico total oriundo da cepa de *Bacillus thuringiensis* PS31F2 foram preparadas a partir de DNA

25 parcialmente digerido com NdeII. Digestos de restrição de NdeII parciais foram eletroforados sobre uma gel de agarose a 0,7% e a região do gel contendo fragmentos de DNA dentro da faixa de tamanho de 9-20 kpb foram extirpados do gel. DNA foi eletroeluído do fragmento de gel em tampão de 0,1X TAE a cerca de 30 V por uma hora e foi purificado usando-se colunas

30 de Elutip-d (Schleicher e Schuell; Keene, NH).

DNA fracionado, purificado estava ligado em braços de Lambda-GEM-11 digerido por BamHI (Promega Corp., Madison, WI). DNA ligado foi

então embalado em fago lambda usando-se extrato de embalagem Giga-pack III Gold (Stratagene Corp., La Jolla, CA). Cepa de *E. coli* KW251 foi infectada com fago recombinante e foi laminada sobre placas de LB em agarose de topo de LB. Placas foram levantadas sobre filtros de nitrocelulose e

5 foram preparadas por hibridização usando-se métodos padrão (Maniatis, e outros). Fragmentos de DNA de cerca de 1,1 kb (PS177C8 MIS) ou 700 pb (PS177C8 WAR) de comprimento foram ampliados por PCR a partir de plasmídeo pMYC2450 e foram usados como as sondas. Fragmentos foram

10 purificados por gel e cerca de 25 ng de cada fragmento de DNA foi aleatoriamente marcado com ^{32}P -dCTP. Hibridização de DNA imobilizado com sondas PS177C8 ^{32}P -marcadas aleatoriamente foi realizada em condições de formamida padrão: formamida a 50%, 5X SSPE, solução de 5X Denhardt, SDS a 2%, 0,1 mg/ ml a 42°C de um dia para o outro. Manchas foram lava-

15 das sob baixa severidade em 2X SSC, SDS a 0,1% a 42°C e foram expostas a película. Placas de hibridização foram isoladas das placas e foram suspensas em tampão de SM. Fago de DNA foi preparado usando-se fago de LambdaSorb adsorvente (Promega, Madison, WI). PCR usando-se os iniciadores de oligonucleotídeo SEQ ID NO. 3 e SEQ ID NO. 4 foi realizado usando-se modelos de DNA de fago para verificar a presença do gene alvo.

20 As reações de PCR forneceram a tira de 1 kb esperada em ambas as amostras de DNA que confirmam que aqueles clones de fago contêm o gene de interesse. Para a subclonagem, DNA de fago foi digerido com várias enzimas, foi fracionado sobre um gel de agarose a 1% e foi manchado para análise de Southern. Análise de Southern foi realizada como descrita acima.

25 Um fragmento de HindIII de cerca de 8 kb de tamanho foi identificado que continha os genes de toxina PS31F2. Este fragmento foi purificado por gel e foi clonado para dar o sítio HindIII de pBluescript (SK+); este clone de plasmídeo é designado pMYC2610. A cepa XL10 Gold [pMYC2610] de *E. coli* foi designada MR983.

30 Uma subcultura de MR983 foi depositada na coleção de sobrenadante da Patent Culture Collection (NRRL), Regional Research Center, 1815 North University Street, Peoria, Illinois 61604 USA em 4 de maio de

1999. O número de acesso é NRRL B-30123.

b. Sequênciação de DNA

O fragmento HindIII pMYC2610 contendo os genes de toxina PS31F2 foi isolado por digestão de restrição, fracionamento sobre um gel de agarose a 0,7% e purificação da matriz de gel usando-se um kit QiaexII (Qiagen Inc.; Valencia, CA). DNA de inserto purificado por gel foi então digerido separadamente com enzimas de restrição AluI, MseI ou RsaI e foi fracionado sobre um gel de agarose a 1%. Fragmentos de DNA entre 0,5 e 1,5 kb foram extirpados do gel e foram purificados usando-se o kit QioexII. Fragmentos recuperados estavam ligados em pBluescriptII digeridos por EcoRV e foram transformados em células XL10 Gold de *E. coli*. DNA de plasmídio foi preparado a partir de transformantes escolhidos aleatórios, foi digerido com NotI e ApaI para verificar o tamanho do inserto e foi usado como modelos de seqüênciação com iniciadores homólogos às seqüências de vetor de plasmídio. Iniciador ambulante foi usado para completar a seqüência.

Reações de seqüênciação foram realizadas usando-se kit de seqüênciação dRhodamine ou BigDye (ABI Prism/Perkin Elmer Applied Biosystems) e foram realizadas sobre seqüenciadores automatizados ABI 373 e 377. Dados foram analisados usando-se programas Factura, Autoassembler (ABI Prism) e Genetics Computer Group (Madison, WI). Os genes MIS e WAR demonstraram estar localizados pertos um do outro em um óperon transcripcional aparente. O gene WAR é 5' em relação ao gene MIS, e os dois genes são separados por 4 bases de nucleotídeo.

As seqüências de nucleotídeos e seqüências de peptídeos deduzidos para os novos genes MIS e WAR oriundos de PS31F2 são relatadas como novos SEQ ID NO. 47-50.

c. Subclonagem e Transformação de *B. thuringiensis*

Os genes de toxina PS31F2 foram subclonados sobre o fragmento de HindIII de 8 kpb oriundos de pMYC2610 para dar o vetor de transporte de ida e volta *E. coli*/B.t., pHT370 (O. Arantes e D. Lereclus. 1991. Gene 108: 115-119), para a expressão oriunda do promotor de *Bacillus* nativo. O construto de plasmídio resultante foi designado pMYC2615. DNA de

plasmídeo pMYC2415 foi preparado a partir de XL10Gold de *E. coli* recombinante para a transformação no hospedeiro de B.t. acristalífero, CryB (A. Aronson, Purdue University, West Lafayette, IN), por eletroporação. A cepa CryB recombinante [pMYC2615] foi designada MR558.

5 Exemplo 9 – Clonagem Molecular e Análise de Seqüência de DNA de um Novo Gene de Toxina SUP Oriundo de Cepa de *Bacillus thuringiensis* KB59A4-6

DNA celular total foi preparado a partir da cepa de *Bacillus thuringiensis* KB59A4-6 desenvolvido para uma densidade ótica de 0,5-0,8 a 600
 10 nm de luz visível em caldo de Luria Bertani (LB). DNA foi extraído usando-se kit Qiagen Genomic-tip 500/G e Conjunto de Tampão de DNA Genômico de acordo com o protocolo para bactérias gram-positivas (Qiagen Inc.; Valencia, CA). DNA foi digerido com HindIII foi realizado sobre géis de agarose a 0,7% para análise de Southern blot por métodos padrão (Maniatis e
 15 outros). Um "amplicon" de PCR contendo o gene de tipo SUP (SEQ ID NO. 1) oriundo de DNA genômico Javelin-90 foi obtido por uso dos oligos "3A-atg
 (GCTCTAGAAGGAGGTAAGTTATGAACAAGAATAATACTAAATTAAGC)
 (SEQ ID NO. 51) E "3A-Taa" (GGGGTACCTTACTTAATAGAGACATCG)
 20 (SEQ ID NO. 52). Este fragmento de DNA foi purificado por gel e foi marcado com ³²P-dCTP radioativo usando-se kit de Marcação de Iniciador Aleatório Prime-It (Stratagene) para o uso como uma sonda. Hibridização de filtros de Southern blot foi realizada em uma solução de 6X SSPE, solução de 5X Denhardt, SDS a 0,1%, 0,1 mg/ml de DNA desnaturado a 42°C de um dia
 25 para o outro em um banho de água de agitação. Os filtros foram subsequentemente lavados em 1X SSPE e SDS a 0,1% uma vez a 25°C seguido por duas lavagens adicionais a 37°C. Filtros hibridizados foram então expostos a película de raios X a -80°C. Um fragmento de HindIII de cerca de 1 kpb de DNA genômico KBS59A4-6 foi identificado que hibridizou para dar a
 30 sonda SUP Javelin 90.

Uma biblioteca lambda de DNA genômico KB59A4-6 foi construída como se segue. DNA foi parcialmente digerido com Sau3A e foi fraci-

onado pelo tamanho sobre géis de agarose. A região de gel contendo fragmentos entre 9,0 e 23 kpb foi extirpada e DNA foi isolado por eletroeluição em tampão 0,1X TAE seguido por purificação sobre colunas Elutip-d (Sheleicher e Schuell, Keene, NH). Insetos de DNA fracionados pelo tamanho
5 estavam ligados em Lamda-Gem 11 digerido por BamHI (Promega) e fago recombinante foram embalados usando-se Extrato de Embalagem Giga-packIII (Stratagene). Fago foram laminados sobre células VCS257 de *E. coli* para a triagem por hibridização. Placas foram transferidas par filtros de náilon e foram secas sob vácuo a 80°C. Hibridização foi então realizada com a
10 sonda de gene Sup Javelin 90 como descrita acima. Uma placa que deu um sinal positivo foi selecionada usando-se uma pipeta Pateur para se obter um tampão("plug"). O tampão foi encharcado durante a noite a temperatura ambiente em 1 ml de tampão de SM + 10 ul de CHI₃. Preparações de DNA de fago de escala grande (Maniatis e outros) foram obtidas a partir de lisados
15 líquidos de KW251 de *E. coli* foram infectadas com este fago.

O gene de toxina KB59A4-6 foi subclonado para dar o vetor de transporte de ida e volta de *B.thuringiensis* de *E. coli*, pHT370 (O. Arantes e D. Lereclus. 1991. Gene 108: 115-119), sobre um fragmento de cerca de 5,5 kpb de SacI/XbaI identificado por hibridização Southern. Este subclone de
20 plasmídeo foi designado pMYC2473. Células XL10-Gold de *E. coli* recombinante (Stratagene) contendo este constructo são desingadas MR993. O gene de toxina inseticida foi seqüenciado por iniciador ambulante usando-se o plasmídeo pMYC2473 e amplicons de PCR como modelos de DNA. Reações de seqüenciação foram realizadas usando-se Kit de Reação de Pronta de
25 Seqüenciação de Ciclo Terminador de Corante da PE Applied Biosystems e foram realizadas sobre um Sequenciador Automatizado ABI PRISM 377. Dados de seqüência foram analisados usando-se o programa PE ABI PRISM 377, Collection, Factura e Auto Assembler. A seqüência de DNA e seqüência de peptídeos deduzida da toxina KB59A4-6 são relatadas como
30 novos SEQ ID NOs. 53 e 54, respectivaente.

Uma subcultura de MR993 foi depositada na coleção permanente da Patente Culture Collection (NRRL), Regional Research Center,

1815 North University Street, Peoria, Illinois 61604 USA em 4 de maio de 1999. O número de acesso é NRRL B-30125.

Exemplo 10 – Bioensaios para Atividade Contra Lepidopteros e Coleopteros

5 Atividade biológica das toxinas e dos isolados da presente invenção pode ser confirmada usando-se procedimentos de bioensaio padrão. Um tal ensaio é o ensaio de gorgulho de botão de flor-"bollworm" (*Heliothis virescens* [Fabricius] e *Helioverpa zea* [Boddie]). Bioensaios de Lepidoptera foram conduzidos ou com aplicação de superfície a dieta de inseto artificial
10 ou incorporação de dieta das amostras. Todos os insetos Lepidoptero foram testados a partir do estágio neonato até o segundo ínstar. Todos os ensaios foram conduzidos ou com dieta artificial de farinha de soja tostada ou dieta artificial de gramiola negra (BioServ, Frenchtown, NJ).

 A incorporação de dieta pode ser conduzida por misturação das
15 amostras com dieta artificial em uma taxa de 6 ml de suspensão mais dieta de 45 ml. Depois do turbilhonamento, esta mistura é despejada para dentro de bandejas de plástico com cavidades de 3 ml compartimentalizadas (Nutrend Container Corporation, Jacksonville, FL). Um material em branco aquoso não contendo nenhum B.t. serve como o controle. Larvas de primeiro
20 ínstar (USDA-ARS, Stoneville, MS) são colocadas sobre a mistura de dieta. Cavidades são então seladas com cobertura de Mylar (ClearLam Packaging, IL) usando-se um ferro de bordejamento, e vários furos foram feitos em cada cavidade para prover troca de gás. Larvas foram mantidas a 25°C por 6 dias em uma câmara de retenção 14:10 (claro:escuro). A mortalidade e retarda-
25 mento de crescimento são registrados depois de 6 dias.

 O bioensaio pelo método de carregamento pelo topo utiliza a mesma amostra e preparações de dieta que listadas acima. As amostras são aplicadas à superfície da dieta de inseto. Em uma modalidade específica, área de superfície variou de 0,3 a cerca de 0,8 cm² dependendo do tamanho de bandeja, placas de cultura de tecido de 96 cavidades foram usa-
30 das além do formato listado acima. Após a aplicação, amostras são deixadas secar ao ar antes da infestação de inseto. Um material em branco

aquoso não contendo nenhum B.t. pode servir como o controle. Ovos são aplicados a cada cavidade tratada e foram então selados com cobertura de Mylar (ClearLam Packaging, IL) usando-se um ferro de bordejamento, e furos foram feitos em cada cavidade para prover troca de gás. Bioensaios são mantidos a 25°C por 7 dias em um câmara de retenção de 14:10 (claro:escuro) ou 28°C por 4 dias em uma câmara de retenção de 14:10 (claro:escuro). A mortalidade e retardamento de crescimento são registrados no fim de cada bioensaio.

Um outro ensaio útil de acordo com a presente invenção é o ensaio de "rootworm de milho do oeste. Amostras podem ser bioanalisadas contra larvas de "rootworm" milho do oeste neonatas (*Diabrotica virgifera virgifera*) via carregamento pelo topo de amostra sobre uma dieta artificial à base de ágar em uma taxa de 160 ml/cm². Dieta artificial pode ser distribuída em cavidades de 0,78 cm² em cultura de tecido de 48 cavidades ou placas similares e deixada endurecer. Depois que a dieta solidifica, amostras são distribuídas por pipeta sobre a superfície de dieta. Líquido em excesso é então evaporado a partir da superfície antes da transferência de aproximadamente três larvas neonatas por cavidade sobre a superfície de dieta por escova de cabelo de camelo. Para prevenir escape de inseto enquanto permitindo troca de gás, cavidades são seladas a quente com película de poliéster de punção de 2 mil (5,8 micra) com o adesivo 27HT (Oliver Products Company, Grand Rapids, Michigan). Bioensaios são mantidos no escuro a 25°C, e a mortalidade foi classificada depois de 4 dias.

Bioensaios análogos podem ser realizados por aqueles versados na técnica para avaliar atividade contra outras pestes, tal como gramíola negra (*Agrotis ipsilon*).

Resultados são mostrados na tabela 6.

Tabela 6. Genética e função de sobrenadantes de B.t. concentrados triados quanto a atividade de lepidopterano e coleopterano								
Cepa	Fragmento de PCR de cerca de 339 pb	Proteína total (µg/cm²)	Proteína de cerca de 80-100 kDa	H. virescens		H. zen		Diabrotica % de mortalidade
				% de mortalidade	retardamento de crescimento	% de mortalidade	retardamento de crescimento	
PS157C1 (#1)	-	24	2	43	sim	13	sim	-
PS157C1 (#2)	-	93	8	-	-	-	-	40
PS157C1 (#3)	-	35	3	-	-	-	-	18
Javelin 1990	++	43,2	3,6	100	sim	96	sim	NT
Água				0-8	-	0-4	-	12

Exemplo 11 – Resultados de Bioensaios de "Rootworm" de Milho do Oeste e Caracterização Ulterior das Toxinas

Soluções líquidas concentradas, obtidas de acordo com a presente invenção, foram testadas quanto a atividade contra "rootworm" de milho do oeste (WCRW). Sobrenadantes oriundos dos seguintes isolados demonstraram causar mortalidade contra WCRW: PS31F2, PS66D3, PS177I8, KB53A49-4, KB68B46-2, KB68B51-2, KB68B55-2, e PS177C8.

Sobrenadantes oriundos dos seguintes isolados foram demonstraram causar mortalidade contra WCRW: PS205A3, PS185V2, PS234E1, PS71G4, PS248N10, PS191A21, KB63B19-13, KB63B19-7, KB68B62-7, KB68B63-2, KB69A125-1, KB69A125-3, KB69A125-5, KB69A127-7, KB69A132-1, KB69B2-1, KB70B5-3, KB71A125-15, e KB71A35-6; foi confirmado que esta atividade era instável quente. Além do mais, foi determinado que os sobrenadantes dos seguintes isolados não reagiram (fornecer resultados de teste negativos) com o anticorpo WAR (ver, o exemplo 12), e não reagiram com as sondas MIS (SEQ ID NO. 31) e WAR (SEQ ID NO. 51): PS205A3, PS185V2, PS234E1, PS71G4, PS248N10, PS191A21, KB63B19-13, KB63B19-7, KB68B62-7, KB68B63-2, KB69A125-1, KB69A125-5, KB69A132-1, KB69B2-1, KB70B5-3, KB71A125-15, e KB71A35-6; os sobrenadantes de isolados KB69A125-3 e KB69A127-7 forneceram resultados positivos.

Exemplo 12 – Cultivo de Clones 31F2 e Bioensaio de Toxinas 31F2 sobre "Rootworm" de Milho do Oeste (wCRW).

MR983 de *E. coli* e a cepa de controle negativo MR948 (controle de vetor XL1-Blue de *E. coli* [pSupercros]); foram desenvolvidos em frascos providos de defletor fundo de 250 ml contendo 50 ml de meio de Caldo Ter-rific DIFCO. Culturas foram incubadas em agitador New Brunswick que agitam a 250 RPM, 30°C por ~23 horas. Depois de 23 horas de amostras de incubação foram assepticamente tomadas para examinar as culturas sob o microscópio para checar quanto a presença de contaminantes. 30 ml de cultura foram distribuídas para dentro de um tubo de centrífuga de 50 ml e foram centrifugados em uma centrífuga Sorvall a 15.000 rpm por 20 minutos.

O sobrenadante 1X foi protegido e foi submetido para bioensaio contra wCRW. O pélete foi ressuspensão 5X com tampão de Tris a 10 mM, e foi sonificada antes da submissão para bioensaio contra wCRW.

5 Cepa de B.t. MR558 e o controle negativo MR539 (controle de vetor cry de B.t. B[pHT BlueII]); foram desenvolvidos do mesmo modo exceto quanto a omissão de glicerol do meio de Caldo Terrific. Péletes de célula de B.t. foram resuspenso em água ao invés de tampão antes da sonificação.

10 Ensaios para o clone de *E. coli* MR983 e clone de *B. thuringiensis* MR558 contendo os genes de toxina 31F2 foram conduzidos usando-se o mesmo projeto experimental que no exemplo 10 para "rootworm" de milho do oeste com as seguintes exceções: Amostras de sobrenadantes foram carregadas pelo topo sobre a dieta em uma dose de $\sim 160 \text{ ul/cm}^2$. Amostras de péletes de B.t. celular a uma concentração de 5X foram carregadas pelo
15 topo sobre a dieta a uma dose de $\sim 150 \text{ ul/cm}^2$ para ambos os clones, e a ~ 75 , e em doses de $\sim 35 \text{ ul/cm}^2$ para cada clone de *B. thuringiensis* MR558 (quantidade de toxina ativa desconhecida para qualquer clone). Cerca de 6-8 larvas foram transferidas sobre a dieta imediatamente depois que a amostra evaporou. A placa de bioensaio foi selada com cobertura de Mylar usando-se um ferro de bordejamento e furos foram feitos acima de cada cavidade para provê troca de gás. Tanto os clones MR983 quanto MR558 demonstraram graus de bioatividade (mortalidade maior) contra "rootworm" de milho do oeste quando comparadas com os clones negativos à toxina MR948 e MR539.

25 A tabela 7 apresenta os resultados que mostram a bioatividade de toxinas PS31F2 clonadas contra "rootworm" de milho do oeste.

Tabela 7

			Percentagem de mortalidade de wCRW			
			Sobrena- dante	Pélete 5X	Pélete 5X	Pélete 5X
Cepa	Genes de toxina	Taxa ⇒	160 ul/cm ²	150 ul/cm ²	75 ul/cm ²	35 ul/cm ²
MR983	31F2		7% (4/56)	19% (5/27)	--	--
MR948	nenhum		4% (1/24)	26% (6/23)	--	--
MR983	31F2		3% (5/147)	--	20% (49/245)	--
MR948	nenhum		27% (19/70)	--	51% (79/154)	--
MR983	31F2		13% (32/243)	--	33% (85/259)	--
MR948	nenhum		9% (14/155)	--	20% (52/273)	--
MR558	31F2		35% (41/118)	88% (43/49)	9% (9/100)	13% (13/97)
MR539	nenhum		10% (14/134)	14% (3/21)	15% (17/111)	17% (19/111)
MR558	31F2		3% (1/29)	35% (17/48)	29% (15/52)	13% (7/55)
MR539	nenhum		19% (5/27)	20% (9/46)	31% (18/57)	18% (9/49)
MR558	31F2		13% (9/69)	38% (19/50)	18% (15/85)	15% (10/65)
MR539	nenhum		29% (16/55)	24% (13/58)	14% (13/91)	28% (18/64)
MR558	31F2		7% (5/74)	14% (9/66)	17% (14/83)	11% (6/57)
MR539	nenhum		11% (9/79)	32% (19/59)	9% (7/78)	15% (10/67)

Exemplo 13 – Pestes Alvo

As toxinas da presente invenção podem ser usadas, sozinhas ou em combinação com outras toxinas, para controlar uma ou mais pestes de não mamífero. Estas pestes podem ser, por exemplo, aquelas listadas na

tabela 8. A atividade pode prontamente ser confirmada usando-se bioensaios providos aqui, adaptações destes bioensaios, e/ou outros bioensaios bem conhecidos por aqueles versados na técnica.

Tabela 8. Espécies de Peste Alvo	
ORDEM/Nome Comum	Nome em Latim
LEPIDOPTERA	
Broca de milho europeu	<i>Ostrinia nubilalis</i>
Broca de milho europeu resistente a classe Cry1A de toxinas	<i>Ostrinia nubilalis</i>
Gramiola negra	<i>Agrotis ipsilon</i>
lagarta de cereais do outono	<i>Spodoptera frugiperda</i>
broca de milho do sudoeste	<i>Diatraea grandiosella</i>
"earworm/bollworm" de milho	<i>Helicoverpa zea</i>
gorgulho de botão de flor de tabaco	<i>Heliothis virescens</i>
gorgulho de botão de flor de tabaco resistente a classe Cry1-A de toxinas	<i>Heliothis virescens</i>
traça de "head" de girassol	<i>Homeosoma electellum</i>
traça de girassol listrada	<i>Cochylis hospes</i>
"looper" argentino	<i>Rachiplusia nu</i>
"spilosoma"	<i>Spilosoma virginica</i>
lagarta de cereais "bertha"	<i>Mamestra configurata</i>
traça "diamondback"	<i>Plutella xylostells</i>
traça diamondback" resistente a classe de Cry1A de toxinas	<i>Plutella xylostells</i>
COLEOPTERA	
gorgulho de semente de girassol verde	<i>Smicronyx fulvus</i>
gorgulho de tronco de girassol	<i>Cylindrocopturus adspersus</i>
besouro de girassol	<i>Zygogramma exclamationis</i>
"flea beetle" de canola	<i>Phyllotreta cruciferae</i>
"rootworm" de milho do oeste	<i>Diabrotica virgifera virgifera</i>
DIPTERA	
mosca de Hesse	<i>Mayetiola destructor</i>
HOMOPTERA	
percevejo verde	<i>Schizaphis graminum</i>
HEMIPTERA	
percevejo "lygus"	<i>Lygus lineolaris</i>
NEMATODA	
	<i>Heterodera glycines</i>

Exemplo 14 – Inserção de Genes de Toxina em Plantas

Um aspecto da presente invenção é a transformação de plantas com genes que codificam a toxina inseticida da presente invenção. As plantas transformadas são resistentes ao ataque pela peste alvo.

5 Genes que codificam toxinas pesticidas, como revelados aqui, podem ser inseridos em células de planta usando-se uma variedade de técnicas que são bem conhecidas na técnica. Por exemplo, um grande número de vetores de clonagem compreendendo um sistema de replicação em *E. coli* e um marcador que permite que seleção para as células transformadas
10 esteja disponível para a preparação para a inserção de genes estranhos em plantas superiores. Os vetores compreendem, por exemplo, pBR322, série pUC, espécie M13mp, pACYC184, etc. Correspondentemente, a sequência que codifica a toxina de *Bacillus* pode ser inserida no vetor em um sítio de reação adequado. O plasmídeo resultante é usado para a transformação em
15 *E. coli*. As células de *E. coli* são cultivadas em um meio nutriente adequado, então são coletadas e são lisadas. O plasmídeo é recuperado. Análise de sequência, análise de restrição, eletroforese, e outros métodos bioquímicos-biológicos moleculares são em geral realizados como métodos de análise. Depois de cada manipulação, a sequência de DNA usada pode ser clivada e
20 ligada à próxima sequência de DNA. Cada sequência de plasmídeo pode ser clonada na mesma ou em outros plasmídios. Dependendo do método de inserção do genes desejados na planta, outras sequências de DNA podem ser necessárias. Se, por exemplo, o plasmídeo Ti ou Ri é usado para a transformação da célula de planta, então pelo menos a borda à direita, mas
25 freqüentemente a borda à direita e à esquerda do T-DNA de plasmídeo Ti ou Ri, tem de estar ligada como a região de flaqueamento dos genes a serem inseridos.

O uso de T-DNA para a transformação de células de planta foi intensivamente pesquisado e suficientemente descrito na EP 120 516; Hoekema (1985) Em: *The Binary Plant Vector System*, Offset-durkkerij Kanters
30 B. V., Alblasterdam, capítulo 5, Fraley e outros, Crit. Ver. Plant Sci. 4:1-46; e An e outros (1985) *EMBO J.* 4:277-287.

Uma vez que o DNA inserido foi integrado no genoma, ele é relativamente estável lá e, por via regra, não sai novamente. Ele normalmente contém um marcador de seleção que confere sobre as células de planta transformadas resistência a um biocida ou a um antibiótico, tal como canamicina, G 418, bleomicina, higromicina ou cloranfenicol, entre outros. Mar-
5 marcador individualmente empregado deve correspondentemente permitir a seleção de células transformadas ao invés de células que não contêm o DNA inserido.

Um grande número de técnicas estão disponíveis para a inserção
10 de DNA em uma célula de hospedeiro de planta. Aquelas técnicas incluem transformação com T-DNA usando-se *Agrobacterium tumefaciens* ou *Agrobacterium rhizogenes* como agente de transformação, fusão, injeção, biolístico (bombardeamento de micropartículas) ou eletroporação bem como outros métodos possíveis. Se *Agrobacteria* são usadas para a transformação,
15 o DNA a ser inserido tem de ser clonado para dar plasmídios especiais, a saber ou para dar um vetor intermediário ou para dar um vetor binário. Os vetores intermediários podem ser integrados no plasmídio Ti ou Ri por recombinação homóloga devido a seqüências que são homólogas a seqüências no T-DNA. O plasmídio Ti ou Ri também compreende a região vir ne-
20 cessária para a transferência do T-DNA. Vetores intermediários não podem replicar eles próprios nas *Agrobacteria*. O vetor intermediário pode ser transferido para dentro de *Agrobacterium tumefaciens* por meio de um plasmídio auxiliar (conjugação). Vetores binários podem replicar eles próprios tanto em *E. coli* quanto em *Agrobacteria*. Eles compreendem um gene de
25 marcador de seleção e um ligador ou poliligador que são modelados pelas regiões de borda de T-DNA a direita e a esquerda. Eles podem ser transformados diretamente para dentro de *Agrobacteria* (Holsters e outros [1978] *Mol. Gen. Genet.* 163: 181-187). A *Agrobacterium* usada como célula de hospedeiro é para compreender um plasmídio que porta uma região vir. A
30 região vir é necessária para a transferência do T-DNA para dentro da célula de planta. T-DNA adicional pode ser contaminado. A bactéria assim transformada é usada para a transformação de células de planta. Explantes de

planta pode vantajosamente por cultivo com *Agrobacterium tumefaciens* e *Agrobacterium rhizogenes* para a transferência do DNA para dentro da célula de planta. Plantas inteiras podem então ser regeneradas a partir de material de planta infectado (por exemplo, pedaços de folha, segmentos de talo, raízes, mas também protoplastos ou células de suspensão cultivadas) em um meio adequado, que pode conter antibióticos ou biocidas para seleção. As plantas assim obtidas podem então ser testadas quanto a presença do DNA inserido. Nenhuma das demandas especiais são feitas dos plasmídios no caso de injeção e eletroporação. É possível usar plasmídios comuns, tais como, por exemplo, derivados de pUC. Em transformação bioestática, DNA de plasmídio ou DNA linear pode ser empregado.

As células transformadas são regeneradas a partir de plantas morfológicamente normais do modo usual. Se um evento de transformação envolve uma célula de linha de germe, então o DNA inserido e característica(s) fenotípica(s) correspondente(s) serão transmitidas para plantas de pro- gênia. Tais plantas podem ser desenvolvidas no modo normal e cruzadas com plantas que têm os mesmos fatores hereditários transformados ou outros fatores hereditários. Os híbridos resultantes individuais têm as propriedades fenotípicas correspondentes.

Em uma modalidade preferida da presente invenção, plantas serão transformadas com genes em que o uso de códon foi otimizado para plantas. Ver, por exemplo, a patente U.S. no. 5.380.831. Também, vantajosamente, plantas que codificam uma toxina truncada serão usadas. A toxina truncada tipicamente codificará cerca de 55% a cerca de 80% da toxina de comprimento total. Métodos para a criação de genes de *Bacillus* sintéticos para o uso em plantas são conhecidos na técnica.

Deve ser entendido que os exemplos e as concretizações descritas aqui são para as finalidades ilustrativas apenas e que várias modificações ou mudanças à luz das mesmas serão sugeridas para pessoas versadas na técnica e devem estar incluídas dentro do espírito e campo de ação deste pedido e das reivindicações anexas.

LISTAGEM DE SEQUÊNCIA

(1) INFORMAÇÃO GERAL:

(i) REQUERENTE:

Nome(s) da Requerente:	MYCOGEN CORPORATION
Rua:	5501 Oberlin Drive
Cidade:	San Diego
Estado/Província:	California
País:	Estados Unidos
Código Postal/CEP:	92121
Número de telefone:	(800) 745-7475
Número do fax:	(619) 453-0142

(ii) TÍTULO DA INVENÇÃO: NOVAS TOXINAS PESTICIDAS E SE-
QÜÊNCIAS DE NUCLEOTÍDEOS QUE CODIFICAM ESTAS TOXINAS

(iii) NÚMERO DE SEQÜÊNCIAS: 54

15 (iv) ENDEREÇO DE CORRESPONDÊNCIA:

(A) DESTINATÁRIO: SALIWANCHIK, LLOYD 7 SALIWANCHIK
(B) RUA: 2441 N.W. 41st STREET, SUITE A-1
(C) CIDADE: GAINESVILLE
(D) ESTADO: FL
(E) PAÍS: ESTADOS UNIDOS DA AMÉRICA
(F) CEP: 32606-6669

(v) FORMA LEGÍVEL POR COMPUTADOR:

25

(A) TIPO DE MEIO: DISQUETE

(B) COMPUTADOR: IBM PC compatível

(C) SISTEMA DE OPERAÇÃO: PC-DOS/MS-DOS

(D) PROGRAMA: PatentIn Release #1.0, Versão # 1.30

(vi) DADOS DO PEDIDO ATUAL:

30 (A) NÚMERO DO PEDIDO:
(B) DATA DO ARQUIVAMENTO:
(C) CLASSIFICAÇÃO:

(vii) DADOS DO PEDIDO ANTERIOR:

(A) NÚMERO DO PEDIDO: US. 09/073.898

(B) DATA DO ARQUIVAMENTO: 5 de maio de 1998

(viii) INFORMAÇÃO DO PROCURADOR/AGENTE

(A) NOME: SANDERS, JAY M.

(B) NÚMERO DE REGISTRO: 39.355

5 (C) REFERÊNCIA/NÚMERO DO DOCUMENTO: MA-708C2

(ix) INFORMAÇÃO DE TELECOMUNICAÇÃO:

(A) TELEFONE: 352-375-8100

(B) TELEFAX: 352-372-5800

(2) INFORMAÇÃO PARA A SEQ ID NO: 1:

10 (i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 2375 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

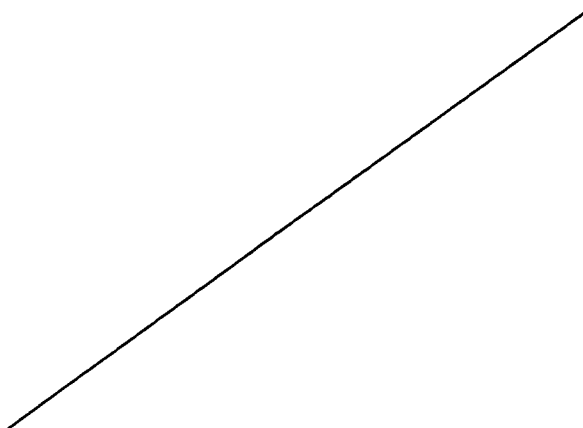
(D) TOPOLOGIA: LINEAR

15 (ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: Jav90

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



ATGAACAAGA ATAATACTAA ATTAAGCACA AGAGCCTTAC CAAGTTTTAT TGATTATTTT	60
AATGGCATTT ATGGATTGTC CACTGGTATC AAAGACATTA TGAACATGAT TTTTAAAACG	120
GATACAGGTG GTGATCTAAC CCTAGACGAA ATTTTAAAGA ATCAGCAGTT ACTAAATGAT	180
ATTTCTGGTA AATTGGATGG GGTGAATGGA AGCTTAAATG ATCTTATCGC ACAGGGAAAC	240
TTAAATACAG AATTATCTAA GGAAATATTA AAAATTGCAA ATGAACAAAA TCAAGTTTTA	300
AATGATGTTA ATAACAAACT CGATGCGATA AATACGATGC TTCGGGTATA TCTACCTAAA	360
ATTACCTCTA TGTTGAGTGA TGTAATGAAA CAAAATTATG CGCTAAGTCT GCAAATAGAA	420
TACTTAAGTA AACAATTGCA AGAGATTCTT GATAAGTTGG ATATTATTAA TGTAATGTGA	480
CTTATTAACT CTACACTTAC TGAAATTACA CCTGCGTATC AAAGGATTAA ATATGTGAAC	540
GAAAAATTG AGGAATTAAC TTTTGCTACA GAACTAGTT CAAAAGTAAA AAAGGATGGC	600
TCTCCTGCAG ATATTCTTGA TGAGTTAACT GAGTTAACTG AACTAGCGAA AAGTGTAACA	660
AAAAATGATG TGGATGGTTT TGAATTTTAC CTTAATACAT TCCACGATGT AATGGTAGGA	720
AATAATTTAT TCGGGCGTTC AGCTTTAAAA ACTGCATCGG AATTAATTAC TAAAGAAAAT	780
GTGAAAACAA GTGGCAGTGA GGTCGGAAAT GTTTATAACT TCTTAATTGT ATTAACAGCT	840
CTGCAAGCAA AAGCTTTTCT TACTTTAACA ACATGCCGAA AATTATTAGG CTTAGCAGAT	900
ATTGATTATA CTTCTATTAT GAATGAACAT TTAAATAAGG AAAAAGAGGA ATTTAGAGTA	960
AACATCCTCC CTACACTTTC TAATACTTTT TCTAATCCTA ATTATGCAAA AGTTAAAGGA	1020
AGTGATGAAG ATGCAAAGAT GATTGTGGAA GCTAAACCAG GACATGCATT GATTGGGTTT	1080
GAAATTAGTA ATGATTCAAT TACAGTATTA AAAGTATATG AGGCTAAGCT AAAACAAAAT	1140
TATCAAGTCG ATAAGGATTC CTTATCGGAA GTTATTTATG GTGATATGGA TAAATTATTG	1200
TGCCCAGATC AATCTGAACA AATCTATTAT ACAAATAACA TAGTATTTCC AAATGAATAT	1260
GTAATTACTA AAATTGATTT CACTAAAAAA ATGAAAACCT TAAGATATGA GGTAACAGCG	1320
AATTTTTATG ATTCTTCTAC AGGAGAAATT GACTTAAATA AGAAAAAGT AGAATCAAGT	1380
GAAGCGGAGT ATAGAACGTT AAGTGCTAAT GATGATGGGG TGTATATGCC GTTAGGTGTC	1440
ATCAGTGAAA CATTTTTGAC TCCGATTAAT GGGTTTGGCC TCCAAGCTGA TGAAAATTCA	1500

AGATTAAATTA CTTTAACATG TAAATCATAT TTAAGAGAAC TACTGCTAGC AACAGACTTA 1560
 AGCAATAAAG AACTAAATT GATYGTCCCG CCAAGTGGTT TTATTAGCAA TATTGTAGAG 1620
 AACGGGTCCA TAGAAGAGGA CAATTTAGAG CCGTGGAAG CAAATAATAA GAATGCGTAT 1680
 GTAGATCATA CAGGCGGAGT GAATGGAAGT AAAGCTTTAT ATGTTTCATAA GGACGGAGGA 1740
 ATTTACAAAT TTATTGGAGA TAAGTTAAAA CCGAAAAGT AGTATGTAAT CCAATATACT 1800
 GTTAAAGGAA AACCTTCTAT TCATTTAAAA GATGAAAATA CTGGATATAT TCATTATGAA 1860
 GATACAAATA ATAATTTAGA AGATTATCAA ACTATTAATA AACGTTTTAC TACAGGAACT 1920
 GATTTAAAGG GAGTGTATTT AATTTTAAAA AGTCAAAATG GAGATGAAGC TTGGGGAGAT 1980
 AACTTTATTA TTTTGGAAAT TAGTCCTTCT GAAAAGTTAT TAAGTCCAGA ATTAATTAAT 2040
 ACAAATAATT GGACGAGTAC GGGATCAACT AATATTAGCG GTAATACACT CACTCTTTAT 2100
 CAGGGAGGAC GAGGGATTCT AAAACAAAAC CTTCAATTAG ATAGTTTTTC AACTTATAGA 2160
 GTGTATTTT CTGTGTCCCG AGATGCTAAT GTAAGGATTA GAAATTCTAG GGAAGTGTTA 2220
 TTTGAAAAAA GATATATGAG CGGTGCTAAA GATGTTTCTG AAATGTTTAC TACAAAATTT 2280
 GAGAAAGATA ACTTTTATAT AGAGCTTTCT CAAGGGAATA ATTTATATGG TGGTCCTATT 2340
 GTACATTTTT ACGATGTCTC TATTAAGTAA CCCAA 2375

(2) INFORMAÇÃO PARA A SEQ ID NO: 2:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 790 aminoácidos

(B) TIPO: AMINOÁCIDO

5

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: PROTEÍNA

(vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: Jav90

10

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 2:

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

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

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

Ser Thr Asn Ile Ser Gly Asn Thr Leu Thr Leu Tyr Gln Gly Gly Arg
 690 695 700
 Gly Ile Leu Lys Gln Asn Leu Gln Leu Asp Ser Phe Ser Thr Tyr Arg
 705 710 715 720
 Val Tyr Phe Ser Val Ser Gly Asp Ala Asn Val Arg Ile Arg Asn Ser
 725 730 735
 Arg Glu Val Leu Phe Glu Lys Arg Tyr Met Ser Gly Ala Lys Asp Val
 740 745 750
 Ser Glu Met Phe Thr Thr Lys Phe Glu Lys Asp Asn Phe Tyr Ile Glu
 755 760 765
 Leu Ser Gln Gly Asn Asn Leu Tyr Gly Gly Pro Ile Val His Phe Tyr
 770 775 780
 Asp Val Ser Ile Lys Pro
 785 790

(2) INFORMAÇÃO PARA A SEQ ID NO: 3:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- 5 (A) COMPRIMENTO: 20 pares de base
 (B) TIPO: ÁCIDO NUCLÉICO
 (C) TIPO DE FIO: ÚNICO
 (D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 3:

10 GGRTTAMTTG GRTAYTATTT

20

(2) INFORMAÇÃO PARA A SEQ ID NO: 4:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- 15 (A) COMPRIMENTO: 20 pares de base
 (B) TIPO: ÁCIDO NUCLÉICO
 (C) TIPO DE FIO: ÚNICO
 (D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

ATATCKWAYA TTKGCATTTA

20

20 (2) INFORMAÇÃO PARA A SEQ ID NO: 5:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- (A) COMPRIMENTO: 1042 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

5 (vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: 66D3

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

TTAATTGGGT ACTATTTTAA AGGAAAAGAT TTAAATAATC TTACTATATT TGCTCCAACA	60
CGTGAGAATA CTCTTATTTA TGATTTAGAA ACAGCGAATT CTTTATTAGA TAAGCAACAA	120
CAACCTATC AATCTATTCG TTGGATCGGT TTAATAAAAA GCAAAAAGC TGGAGATTTT	180
ACCTTTCAAT TATCGGATGA TGAGCATGCT ATTATAGAAA TCGATGGGAA AGTTATTTTCG	240
CAAAAAGGCC AAAAGAAACA AGTTGTTTAT TTAGAAAAAG ATAAATTAGT TCCCATCAAA	300
ATTGAATATC AATCTGATAA AGCGTTAAAC CCAGATAGTC AAATGTTTAA AGAATTGAAA	360
TTATTTAAAA TAAATAGTCA AAAACAATCT CAGCAAGTGC AACAAAGACGA ATTGAGAAAT	420
CCTGAATTG GTAAAGAAAA AACTCAAACA TATTTAAAGA AAGCATCGAA AAGCAGCCTG	480
TTAGCAATA AAAGTAAACG AGATATAGAT GAAGATATAG ATGAGGATAC AGATACAGAT	540
GGAGATGCCA TTCCTGATGT ATGGGAAGAA AATGGGTATA CCATCAAAGG AAGAGTAGCT	600
GTTAAATGGG ACGAAGGATT AGCTGATAAG GGATATAAAA AGTTTGTTTC CAATCCTTTT	660
AGACAGCACA CTGCTGGTGA CCCCTATAGT GACTATGAAA AGGCATCAAA AGATTGGAT	720
TTATCTAATG CAAAAGAAAC ATTTAATCCA TTGGTGGCTG CTTTCCAAG TGTCAATGTT	780
AGCTTGAAAA ATGTCACCAT ATCAAAAGAT GAAAATAAAA CTGCTGAAAT TGCCTCTACT	840
TCATCGAATA ATTGGTCCTA TACAAATACA GAGGGGGCAT CTATTGAAGC TGGGAATTGGA	900
CCAGAAGGTT TGTGTCTTTT TGGAGTAAGT GCCAATTATC AACATTCTGA AACAGTGGCC	960
AAAGAGTGGG GTACAATAA GGGAGACGCA ACACAATATA ATACAGCTTC AGCAGGATAT	1020
CTAAATGCCA ATGTACGATA TA	1042

(2) INFORMAÇÃO PARA A SEQ ID NO: 6:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

10 (A) COMPRIMENTO: 347 AMINOÁCIDO

(B) TIPO: AMINOÁCIDO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: PEPTÍDEO

15 (vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: 66D3

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

```

Leu Ile Gly Tyr Tyr Phe Lys Gly Lys Asp Phe Asn Asn Leu Thr Ile
1           5           10           15

Phe Ala Pro Thr Arg Glu Asn Thr Leu Ile Tyr Asp Leu Glu Thr Ala
20           25           30

Asn Ser Leu Leu Asp Lys Gln Gln Gln Thr Tyr Gln Ser Ile Arg Trp
35           40           45

Ile Gly Leu Ile Lys Ser Lys Lys Ala Gly Asp Phe Thr Phe Gln Leu
50           55           60

Ser Asp Asp Glu His Ala Ile Ile Glu Ile Asp Gly Lys Val Ile Ser
65           70           75           80

Gln Lys Gly Gln Lys Lys Gln Val Val His Leu Glu Lys Asp Lys Leu
85           90           95

Val Pro Ile Lys Ile Glu Tyr Gln Ser Asp Lys Ala Leu Asn Pro Asp
100          105          110

Ser Gln Met Phe Lys Glu Leu Lys Leu Phe Lys Ile Asn Ser Gln Lys
115          120          125

Gln Ser Gln Gln Val Gln Gln Asp Glu Leu Arg Asn Pro Glu Phe Gly
130          135          140

Lys Glu Lys Thr Gln Thr Tyr Leu Lys Lys Ala Ser Lys Ser Ser Leu
145          150          155          160

Phe Ser Asn Lys Ser Lys Arg Asp Ile Asp Glu Asp Ile Asp Glu Asp
165          170          175

Thr Asp Thr Asp Gly Asp Ala Ile Pro Asp Val Trp Glu Glu Asn Gly
180          185          190

Tyr Thr Ile Lys Gly Arg Val Ala Val Lys Trp Asp Glu Gly Leu Ala
195          200          205

Asp Lys Gly Tyr Lys Lys Phe Val Ser Asn Pro Phe Arg Gln His Thr
210          215          220

Ala Gly Asp Pro Tyr Ser Asp Tyr Glu Lys Ala Ser Lys Asp Leu Asp
225          230          235          240

Leu Ser Asn Ala Lys Glu Thr Phe Asn Pro Leu Val Ala Ala Phe Pro
245          250          255

Ser Val Asn Val Ser Leu Glu Asn Val Thr Ile Ser Lys Asp Glu Asn
260          265          270

Lys Thr Ala Glu Ile Ala Ser Thr Ser Ser Asn Asn Trp Ser Tyr Thr
275          280          285

Asn Thr Glu Gly Ala Ser Ile Glu Ala Gly Ile Gly Pro Glu Gly Leu
290          295          300

Leu Ser Phe Gly Val Ser Ala Asn Tyr Gln His Ser Glu Thr Val Ala

```

305		310		315		320									
Lys	Glu	Trp	Gly	Thr	Thr	Lys	Gly	Asp	Ala	Thr	Gln	Tyr	Asn	Thr	Ala
			325						330				335		
Ser	Ala	Gly	Tyr	Leu	Asn	Ala	Asn	Val	Arg	Tyr					
			340					345							

(2) INFORMAÇÃO PARA A SEQ ID NO: 7:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 2645 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

5

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: PS177C8A

10

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

ATGAAGAAGA	AGTTAGCAAG	TGTTGTAACG	TGTACGTTAT	TAGCTCCTAT	GTTTTTGAAT	60
GGAAATGTGA	ATGCTGTTTA	CGCAGACAGC	AAAACAAATC	AAATTTCTAC	AACACAGAAA	120
AATCAACAGA	AAGAGATGGA	CCGAAAAGGA	TTACTTGGGT	ATTATTTCAC	AGGAAAAGAT	180
TTTAGTAATC	TTACTATGTT	TGCACCGACA	CGTGATAGTA	CTCTTATTTA	TGATCAACAA	240
ACAGCAAATA	AACTATTAGA	TAAAAACAA	CAAGAATATC	AGTCTATTCG	TTGGATTGGT	300
TTGATTTCAGA	GTAAAGAAAC	GGGAGATTTC	ACATTTAACT	TATCTGAGGA	TGAACAGGCA	360
ATTATAGAAA	TCAATGGGAA	AATTATTTCT	AATAAAGGGA	AAGAAAAGCA	AGTTGTCCAT	420
TTAGAAAAAG	GAAAATTAGT	TCCAATCAAA	ATAGAGTATC	AATCAGATAC	AAAATTTAAT	480
ATTGACAGTA	AAACATTTAA	AGAACTTAAA	TTATTTAAAA	TAGATAGTCA	AAACCAACCC	540
CAGCAAGTCC	AGCAAGATGA	ACTGAGAAAT	CCTGAATTTA	ACAAGAAAGA	ATCACAGGAA	600
TTCTTAGCGA	AACCATCGAA	AATAAATCTT	TTCACTCAAA	AAATGAAAAG	GGAAATTGAT	660
GAAGACACGG	ATACGGATGG	GGACTCTATT	CCTGACCTTT	GGGAAGAAAA	TGGGTATACG	720
ATTCAAAATA	GAATCGCTGT	AAAGTGGGAC	GATTCTYTAG	CAAGTAAAGG	GTATACGAAA	780
TTTGTTTCAA	ATCCGCTAGA	AAGTCACACA	GTTGGTGATC	CTTATACAGA	TTATGAAAAG	840
GCAGCAAGAG	ACCTAGATTT	GTCAAATGCA	AAGGAAACGT	TTAACCATT	GGTAGCTGCT	900
TTTCCAAGTG	TGAATGTTAG	TATGGAAAAG	GTGATATTAT	CACCAAATGA	AAATTTATCC	960
AATAGTGTAG	AGTCTCATT	ATCCACGAAT	TGGTCTTATA	CAAATACAGA	AGGTGCTTCT	1020
GTTGAAGCGG	GGATTGGACC	AAAAGGTATT	TCGTTCCGAG	TTAGCGTAAA	CTATCAACAC	1080

TCTGAAACAG TTGCACAAGA ATGGGGAACA TCTACAGGAA ATACTTCGCA ATTCAATACG	1140
GCTTCAGCGG GATATTTAAA TGCAAAATGTT CGATATAACA ATGTAGGAAC TGGTGCCATC	1200
TACGATGTAA AACCTACAAC AAGTTTTGTA TTAAATAACG ATACTATCGC AACTATTACG	1260
GCGAAATCTA ATTCTACAGC CTTAAATATA TCTCCTGGAG AAAGTTACCC GAAAAAAGGA	1320
CAAAATGGAA TCGCAATAAC ATCAATGGAT GATTTTAAAT CCCATCCGAT TACATTAAAT	1380
AAAAACAAG TAGATAATCT GCTAAATAAT AAACCTATGA TGTGGGAAAC AAACCAAACA	1440
GATGGTGTTT ATAAGATAAA AGATACACAT GGAAATATAG TAACTGGCGG AGAATGGAAT	1500
GGTGTCTATC AACAAATCAA GGCTAAAACA GCGTCTATTA TTGTGGATGA TGGGGAACGT	1560
GTAGCAGAAA AACGTGTAGC GGCAAAAGAT TATGAAAATC CAGAAGATAA AACACCGTCT	1620
TTAACTTTAA AAGATGCCCT GAAGCTTTCA TATCCAGATG AAATAAAAGA AATAGAGGGA	1680
TTATTATATT ATAAAAACAA ACCGATATAC GAATCGAGCG TTATGACTTA CTTAGATGAA	1740
AATACAGCAA AAGAAGTGAC CAAACAATTA AATGATACCA CTGGGAAATT TAAAGATGTA	1800
AGTCATTTAT ATGATGTAAA ACTGACTCCA AAAATGAATG TTACAATCAA ATTGTCTATA	1860
CTTTATGATA ATGCTGAGTC TAATGATAAC TCAATTGGTA AATGGACAAA CACAAATATT	1920
GTTTCAGGTG GAAATAACGG AAAAAACAA TATTCTTCTA ATAATCCGGA TGCTAATTTG	1980
ACATTAAATA CAGATGCTCA AGAAAAATTA AATAAAAATC GTACTATTAT ATAAGTTTAT	2040
ATATGAAGTC AGAAAAAAC ACACAATGTG AGATTACTAT AGATGGGGAG ATTTATCCGA	2100
TCACTACAAA AACAGTGAAT GTGAATAAAG ACAATTACAA AAGATTAGAT ATTATAGCTC	2160
ATAATATAAA AAGTAATCCA ATTTCTTCAA TTCATATTAA AACGAATGAT GAAATAACTT	2220
TATTTTGGGA TGATATTTCT ATAACAGATG TAGCATCAAT AAAACCGGAA AATTTAACAG	2280
ATTCAGAAAT TAAACAGATT TATAGTAGGT ATGGTATTAA GTTAGAAGAT GGAATCCTTA	2340
TTGATAAAAA AGGTGGGATT CATTATGGTG AATTTATTAA TGAAGCTAGT TTAAATATTG	2400
AACCATTGCA AAATTATGTG ACAAATATA AAGTTACTTA TAGTAGTGAG TTAGGACAAA	2460
ACGTGAGTGA CACACTTGAA AGTGATAAAA TTTACAAGGA TGGGACAATT AAATTTGATT	2520
TTACAAAATA TAGTRAAAAT GAACAAGGAT TATTTTATGA CAGTGGATTA AATTGGGACT	2580
TTAAAATTAA TGCTATTACT TATGATGGTA AAGAGATGAA TGTTTTTCAT AGATATAATA	2640
AATAG	2645

(2) INFORMAÇÃO PARA A SEQ ID NO: 8:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 881 AMINOÁCIDOS

(B) TIPO: AMINOÁCIDO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: PEPTÍDEO

(vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: PS177C8A

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

```

Met Lys Lys Lys Leu Ala Ser Val Val Thr Cys Thr Leu Leu Ala Pro
1           5           10           15

Met Phe Leu Asn Gly Asn Val Asn Ala Val Tyr Ala Asp Ser Lys Thr
          20           25           30

Asn Gln Ile Ser Thr Thr Gln Lys Asn Gln Gln Lys Glu Met Asp Arg
          35           40           45

Lys Gly Leu Leu Gly Tyr Tyr Phe Lys Gly Lys Asp Phe Ser Asn Leu
          50           55           60

Thr Met Phe Ala Pro Thr Arg Asp Ser Thr Leu Ile Tyr Asp Gln Gln
65           70           75           80

Thr Ala Asn Lys Leu Leu Asp Lys Lys Gln Gln Glu Tyr Gln Ser Ile
          85           90           95

Arg Trp Ile Gly Leu Ile Gln Ser Lys Glu Thr Gly Asp Phe Thr Phe
          100          105          110

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

Ile Ser Asn Lys Gly Lys Glu Lys Gln Val Val His Leu Glu Lys Gly
          130          135          140

Lys Leu Val Pro Ile Lys Ile Glu Tyr Gln Ser Asp Thr Lys Phe Asn
145          150          155          160

Ile Asp Ser Lys Thr Phe Lys Glu Leu Lys Leu Phe Lys Ile Asp Ser
          165          170          175

Gln Asn Gln Pro Gln Gln Val Gln Gln Asp Glu Leu Arg Asn Pro Glu
          180          185          190

Phe Asn Lys Lys Glu Ser Gln Glu Phe Leu Ala Lys Pro Ser Lys Ile
          195          200          205

Asn Leu Phe Thr Gln Lys Met Lys Arg Glu Ile Asp Glu Asp Thr Asp
          210          215          220

Thr Asp Gly Asp Ser Ile Pro Asp Leu Trp Glu Glu Asn Gly Tyr Thr
225          230          235          240

Ile Gln Asn Arg Ile Ala Val Lys Trp Asp Asp Ser Leu Ala Ser Lys
          245          250          255

```

Gly Tyr Thr Lys Phe Val Ser Asn Pro Leu Glu Ser His Thr Val Gly
 260 265 270
 Asp Pro Tyr Thr Asp Tyr Glu Lys Ala Ala Arg Asp Leu Asp Leu Ser
 275 280 285
 Asn Ala Lys Glu Thr Phe Asn Pro Leu Val Ala Ala Phe Pro Ser Val
 290 295 300
 Asn Val Ser Met Glu Lys Val Ile Leu Ser Pro Asn Glu Asn Leu Ser
 305 310 315 320
 Asn Ser Val Glu Ser His Ser Ser Thr Asn Trp Ser Tyr Thr Asn Thr
 325 330 335
 Glu Gly Ala Ser Val Glu Ala Gly Ile Gly Pro Lys Gly Ile Ser Phe
 340 345 350
 Gly Val Ser Val Asn Tyr Gln His Ser Glu Thr Val Ala Gln Glu Trp
 355 360 365
 Gly Thr Ser Thr Gly Asn Thr Ser Gln Phe Asn Thr Ala Ser Ala Gly
 370 375 380
 Tyr Leu Asn Ala Asn Val Arg Tyr Asn Asn Val Gly Thr Gly Ala Ile
 385 390 395 400
 Tyr Asp Val Lys Pro Thr Thr Ser Phe Val Leu Asn Asn Asp Thr Ile
 405 410 415
 Ala Thr Ile Thr Ala Lys Ser Asn Ser Thr Ala Leu Asn Ile Ser Pro
 420 425 430
 Gly Glu Ser Tyr Pro Lys Lys Gly Gln Asn Gly Ile Ala Ile Thr Ser
 435 440 445
 Met Asp Asp Phe Asn Ser His Pro Ile Thr Leu Asn Lys Lys Gln Val
 450 455 460
 Asp Asn Leu Leu Asn Asn Lys Pro Met Met Leu Glu Thr Asn Gln Thr
 465 470 475 480
 Asp Gly Val Tyr Lys Ile Lys Asp Thr His Gly Asn Ile Val Thr Gly
 485 490 495
 Gly Glu Trp Asn Gly Val Ile Gln Gln Ile Lys Ala Lys Thr Ala Ser
 500 505 510
 Ile Ile Val Asp Asp Gly Glu Arg Val Ala Glu Lys Arg Val Ala Ala
 515 520 525
 Lys Asp Tyr Glu Asn Pro Glu Asp Lys Thr Pro Ser Leu Thr Leu Lys
 530 535 540
 Asp Ala Leu Lys Leu Ser Tyr Pro Asp Glu Ile Lys Glu Ile Glu Gly
 545 550 555 560
 Leu Leu Tyr Tyr Lys Asn Lys Pro Ile Tyr Glu Ser Ser Val Met Thr

565	570	575
Tyr Leu Asp Glu Asn Thr Ala Lys	Glu Val Thr Lys Gln Leu Asn Asp	
580	585	590
Thr Thr Gly Lys Phe Lys Asp Val Ser His Leu Tyr Asp Val Lys Leu		
595	600	605
Thr Pro Lys Met Asn Val Thr Ile Lys Leu Ser Ile Leu Tyr Asp Asn		
610	615	620
Ala Glu Ser Asn Asp Asn Ser Ile Gly Lys Trp Thr Asn Thr Asn Ile		
625	630	635 640
Val Ser Gly Gly Asn Asn Gly Lys Lys Gln Tyr Ser Ser Asn Asn Pro		
645	650	655
Asp Ala Asn Leu Thr Leu Asn Thr Asp Ala Gln Glu Lys Leu Asn Lys		
660	665	670
Asn Arg Asp Tyr Tyr Ile Ser Leu Tyr Met Lys Ser Glu Lys Asn Thr		
675	680	685
Gln Cys Glu Ile Thr Ile Asp Gly Glu Ile Tyr Pro Ile Thr Thr Lys		
690	695	700
Thr Val Asn Val Asn Lys Asp Asn Tyr Lys Arg Leu Asp Ile Ile Ala		
705	710	715 720
His Asn Ile Lys Ser Asn Pro Ile Ser Ser Ile His Ile Lys Thr Asn		
725	730	735
Asp Glu Ile Thr Leu Phe Trp Asp Asp Ile Ser Ile Thr Asp Val Ala		
740	745	750
Ser Ile Lys Pro Glu Asn Leu Thr Asp Ser Glu Ile Lys Gln Ile Tyr		
755	760	765
Ser Arg Tyr Gly Ile Lys Leu Glu Asp Gly Ile Leu Ile Asp Lys Lys		
770	775	780
Gly Gly Ile His Tyr Gly Glu Phe Ile Asn Glu Ala Ser Phe Asn Ile		
785	790	795 800
Glu Pro Leu Gln Asn Tyr Val Thr Lys Tyr Lys Val Thr Tyr Ser Ser		
805	810	815
Glu Leu Gly Gln Asn Val Ser Asp Thr Leu Glu Ser Asp Lys Ile Tyr		
820	825	830
Lys Asp Gly Thr Ile Lys Phe Asp Phe Thr Lys Tyr Ser Xaa Asn Glu		
835	840	845
Gln Gly Leu Phe Tyr Asp Ser Gly Leu Asn Trp Asp Phe Lys Ile Asn		
850	855	860
Ala Ile Thr Tyr Asp Gly Lys Glu Met Asn Val Phe His Arg Tyr Asn		
865	870	875 880

(2) INFORMAÇÃO PARA A SEQ ID NO: 9:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 1022 pares de base

(B) TIPO: ÁCIDO NUCLEÍCO

5 (C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: 17718

10 (xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 9:

TGGATTAATT GGGTATTATT TCAAAGGAAA AGATTTTAAT AATCTTACTA TGTTTGCACC	60
GACACGTGAT AATACCCCTTA TGTATGACCA ACAAACAGCG AATGCATTAT TAGATAAAAA	120
ACAACAAGAA TATCAGTCCA TTCGTTGGAT TGGTTTGATT CAGAGTAAAG AAACGGGCGA	180
TTTCACATTT AACTTATCAA AGGATGAACA GGCAATTATA GAAATCGATG GGAAAATCAT	240
TTCTAATAAA GGGAAAGAAA AGCAAGTTGT CCATTTAGAA AAAGAAAAAT TAGTTCCAAT	300
CAAAATAGAG TATCAATCAG ATACGAAATT TAATATTGAT AGTAAACAT TTAAGAACT	360
TAAATTATTT AAAATAGATA GTCAAAACCA ATCTCAACAA GTTCAACTGA GAAACCCTGA	420
ATTTAACAAA AAAGAATCAC AGGAATTTT AGCAAAAGCA TCAAAAACAA ACCTTTTTAA	480
GCAAAAAATG AAAAGAGATA TTGATGAAGA TACGGATACA GATGGAGACT CCATTCCCTGA	540
TCTTTGGGAA GAAAATGGGT ACACGATTCA AAATAAAGTT GCTGTCAAAT GGGATGATTC	600
GCTAGCAAGT AAGGGATATA CAAAATTTGT TTCGAATCCA TTAGACAGCC ACACAGTTGG	660
CGATCCCTAT ACTGATTATG AAAAGGCCGC AAGGGATTTA GATTTATCAA ATGCAAAGGA	720
AACGTTCAAC CCATTGGTAG CTGCTTTYCC AAGTGTGAAT GTTAGTATGG AAAAGGTGAT	780
ATTATCACCA AATGAAAATT TATCCAATAG TGTAGAGTCT CATTCAATCCA CGAATTGGTC	840
TTATACGAAT ACAGAAGGAG CTTCATTGA AGCTGGTGGC GGTCCATTAG GCCTTTCTTT	900
TGGAGTGAGT GTTAATTATC AACACTCTGA AACAGTTGCA CAAGAATGGG GAACATCTAC	960
AGGAAATACT TCACAATTCA ATACGGCTTC AGCGGGATAT TTAATGCCA ATATACGATA	1020
TA	1022

(2) INFORMAÇÃO PARA A SEQ ID NO: 10:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 340 aminoácidos

(B) TIPO: AMINOÁCIDO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: PEPTÍDEO

5

(vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: 17718

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

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

Thr Phe Asn Pro Leu Val Ala Ala Xaa Pro Ser Val Asn Val Ser Met
 245 250 255
 Glu Lys Val Ile Leu Ser Pro Asn Glu Asn Leu Ser Asn Ser Val Glu
 260 265 270
 Ser His Ser Ser Thr Asn Trp Ser Tyr Thr Asn Thr Glu Gly Ala Ser
 275 280 285
 Ile Glu Ala Gly Gly Gly Pro Leu Gly Leu Ser Phe Gly Val Ser Val
 290 295 300
 Asn Tyr Gln His Ser Glu Thr Val Ala Gln Glu Trp Gly Thr Ser Thr
 305 310 315 320
 Gly Asn Thr Ser Gln Phe Asn Thr Ala Ser Ala Gly Tyr Leu Asn Ala
 325 330 335
 Asn Ile Arg Tyr
 340

(2) INFORMAÇÃO PARA A SEQ ID NO: 11:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 1341 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

5

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: PS177C8a

10

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

ATGTTTATGG TTTCTAAAAA ATTACAAGTA GTTACTAAAA CTGTATTGCT TAGTACAGTT 60
 TTCTCTATAT CTTTATTAAA TAATGAAGTG ATAAAAGCTG AACCAATTAAA TATAAATTCT 120
 CAAAGTAAAT ATACTAAGTT GCAAAATCTA AAAATCACTG ACAAGGTAGA GGATTTTAAA 180
 GAAGATAAGG AAAAAGCGAA AGAATGGGGG AAAGAAAAAG AAAAAGAGTG GAACTAACT 240
 GCTACTGAAA AAGGAAAAAT GAATAATTTT TTAGATAATA AAAATGATAT AAAGACAAAT 300
 TATAAAGAAA TTACTTTTTT TATGGCAGGC TCATTGTAAG ATGAAATAAA AGATTAAAA 360
 GAAATTGATA AGATGTTTGA TAAACCAAT CTATCAAATT CTATTATCAC CTATAAAAT 420
 GTGGAACCGA CAACAATTGG ATTTAATAAA TCTTTAACAG AAGGTAATAC GATTAATTCT 480
 GATGCAATGG CACAGTTTAA AGAACAATTT TTAGATAGGG ATATTAAGTT TGATAGTTAT 540
 CTAGATACGC ATTTAACTGC TCAACAAGTT TCCAGTAAAG AAAGAGTTAT TTTGAAGGTT 600
 ACGGTTCCGA GTGGGAAAGG TTCTACTACT CCAACAAAAG CAGGTGTCAT TTTAAATAAT 660

```

AGTGAATACA AAATGCTCAT TGATAATGGG TATATGGTCC ATGTAGATAA GGTATCAAAA      720
GTGGTGAAAA AAGGGGTGGA GTGCTTACAA ATTGAAGGGA CTTTAAAAAA GAGTCTTGAC      780
TTTAAAAATG ATATAAATGC TGAAGCGCAT AGCTGGGGTA TGAAGAATTA TGAAGAGTGG      840
GCTAAAGATT TAACCGATTG GCAAAGGGAA GCTTTAGATG GGTATGCTAG GCAAGATTAT      900
AAAGAAATCA ATAATTATTT AAGAAATCAA GGCGGAAGTG GAAATGAAAA ACTAGATGCT      960
CAAATAAAAA ATATTTCTGA TGCTTTAGGG AAGAAACCAA TACCGGAAAA TATTACTGTG     1020
TATAGATGGT GTGGCATGCC GGAATTGGT TATCAAATTA GTGATCCGTT ACCTTCTTTA     1080
AAAGATTTTG AAGAACAATT TTTAAATACA ATCAAAGAAG ACAAAGGATA TATGAGTACA     1140
AGCTTATCGA GTGAACGTCT TGCAGCTTTT GGATCTAGAA AAATTATATT ACGATTACAA     1200
GTTCCGAAAG GAAGTACGGG TCGGTATTTA AGTGCCATTG GTGGATTTGC AAGTGAAAAA     1260
GAGATCCTAC TTGATAAAGA TAGTAAATAT CATATTGATA AAGTAACAGA GGTAATTATT     1320
AAGGTGTTAA GCGATATGTA G                                     1341

```

(2) INFORMAÇÃO PARA A SEQ ID NO: 12:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 446 aminoácidos

(B) TIPO: AMINOÁCIDO

5

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: PEPTÍDEO

(vi) FONTE ORIGINAL:

(C) ISOLADO INDIVIDUAL: PS177C8a

10

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

```

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

```

100					105					110					
Glu	Asp	Glu	Ile	Lys	Asp	Leu	Lys	Glu	Ile	Asp	Lys	Met	Phe	Asp	Lys
		115					120					125			
Thr	Asn	Leu	Ser	Asn	Ser	Ile	Ile	Thr	Tyr	Lys	Asn	Val	Glu	Pro	Thr
		130					135					140			
Thr	Ile	Gly	Phe	Asn	Lys	Ser	Leu	Thr	Glu	Gly	Asn	Thr	Ile	Asn	Ser
		145					150					155			160
Asp	Ala	Met	Ala	Gln	Phe	Lys	Glu	Gln	Phe	Leu	Asp	Arg	Asp	Ile	Lys
				165					170					175	
Phe	Asp	Ser	Tyr	Leu	Asp	Thr	His	Leu	Thr	Ala	Gln	Gln	Val	Ser	Ser
			180					185					190		
Lys	Glu	Arg	Val	Ile	Leu	Lys	Val	Thr	Val	Pro	Ser	Gly	Lys	Gly	Ser
		195					200					205			
Thr	Thr	Pro	Thr	Lys	Ala	Gly	Val	Ile	Leu	Asn	Asn	Ser	Glu	Tyr	Lys
		210					215					220			
Met	Leu	Ile	Asp	Asn	Gly	Tyr	Met	Val	His	Val	Asp	Lys	Val	Ser	Lys
		225					230					235			240
Val	Val	Lys	Lys	Gly	Val	Glu	Cys	Leu	Gln	Ile	Glu	Gly	Thr	Leu	Lys
				245					250					255	
Lys	Ser	Leu	Asp	Phe	Lys	Asn	Asp	Ile	Asn	Ala	Glu	Ala	His	Ser	Trp
			260					265					270		
Gly	Met	Lys	Asn	Tyr	Glu	Glu	Trp	Ala	Lys	Asp	Leu	Thr	Asp	Ser	Gln
		275					280					285			
Arg	Glu	Ala	Leu	Asp	Gly	Tyr	Ala	Arg	Gln	Asp	Tyr	Lys	Glu	Ile	Asn
		290					295					300			
Asn	Tyr	Leu	Arg	Asn	Gln	Gly	Gly	Ser	Gly	Asn	Glu	Lys	Leu	Asp	Ala
		305					310					315			320
Gln	Ile	Lys	Asn	Ile	Ser	Asp	Ala	Leu	Gly	Lys	Lys	Pro	Ile	Pro	Glu
			325						330				335		
Asn	Ile	Thr	Val	Tyr	Arg	Trp	Cys	Gly	Met	Pro	Glu	Phe	Gly	Tyr	Gln
			340					345					350		
Ile	Ser	Asp	Pro	Leu	Pro	Ser	Leu	Lys	Asp	Phe	Glu	Glu	Gln	Phe	Leu
		355					360					365			
Asn	Thr	Ile	Lys	Glu	Asp	Lys	Gly	Tyr	Met	Ser	Thr	Ser	Leu	Ser	Ser
		370					375					380			
Glu	Arg	Leu	Ala	Ala	Phe	Gly	Ser	Arg	Lys	Ile	Ile	Leu	Arg	Leu	Gln
		385					390					395			400
Val	Pro	Lys	Gly	Ser	Thr	Gly	Ala	Tyr	Leu	Ser	Ala	Ile	Gly	Gly	Phe
			405					410					415		

Ala Ser Glu Lys Glu Ile Leu Leu Asp Lys Asp Ser Lys Tyr His Ile
 420 425 430

Asp Lys Val Thr Glu Val Ile Ile Lys Val Leu Ser Asp Met
 435 440 445

(2) INFORMAÇÃO PARA A SEQ ID NO: 13:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 21 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

5 (C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

GCTGATGAAC CATTTAATGC C

21

10 (2) INFORMAÇÃO PARA A SEQ ID NO: 14:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 22 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

15 (D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

CTCTTTAAAG TAGATACTAA GC

22

(2) INFORMAÇÃO PARA A SEQ ID NO: 15:

20 (i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 24 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

25 (ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

GATGAGAACT TATCAATAG TATC

24

(2) INFORMAÇÃO PARA A SEQ ID NO: 16:

(i) CARACTERÍSTICAS DE SEQÜÊNCIA:

(A) COMPRIMENTO: 33 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

5 (D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

CGAATTCTTT ATTAGATAAG CAACAACAAA CCT

33

(2) INFORMAÇÃO PARA A SEQ ID NO: 17:

10 (i) CARACTERÍSTICAS DE SEQÜÊNCIA:

(A) COMPRIMENTO: 24 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

15 (ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

GTTATTTTCGC AAAAAGGCCA AAAG

24

(2) INFORMAÇÃO PARA A SEQ ID NO: 18:

(i) CARACTERÍSTICAS DE SEQÜÊNCIA:

20 (A) COMPRIMENTO: 31 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

25 (xi) DESCRIÇÃO DA SEQÜÊNCIA: SEQ ID NO: 18:

GAATATCAAT CTGATAAAGC GTTAAACCCA G

31

(2) INFORMAÇÃO PARA A SEQ ID NO: 19:

(i) CARACTERÍSTICAS DE SEQÜÊNCIA:

(A) COMPRIMENTO: 23 pares de base

30 (B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

GCAGCYTGTT TAGCAATAAA AGT

23

(2) INFORMAÇÃO PARA A SEQ ID NO: 20:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- 5 (A) COMPRIMENTO: 20 pares de base
(B) TIPO: ÁCIDO NUCLÉICO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

CAAAGGAAGA GTAGCTGTTA

20

(2) INFORMAÇÃO PARA A SEQ ID NO: 21:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- 15 (A) COMPRIMENTO: 25 pares de base
(B) TIPO: ÁCIDO NUCLÉICO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

20 CAATGTTAGC TTGGAAAATG TCACC

25

(2) INFORMAÇÃO PARA A SEQ ID NO: 22:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- 25 (A) COMPRIMENTO: 22 pares de base
(B) TIPO: ÁCIDO NUCLÉICO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

GCTTAGTATC TACTTTAAAG AG

22

30 (2) INFORMAÇÃO PARA A SEQ ID NO: 23:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- (A) COMPRIMENTO: 24 pares de base
(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

5 GATACTATTT GATAAGTTCT CATC

24

(2) INFORMAÇÃO PARA A SEQ ID NO: 24:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 24 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

10 (C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

CTTTTGGCCT TTTTGC AAA TAAC

24

15 (2) INFORMAÇÃO PARA A SEQ ID NO: 25:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 31 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

20 (D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

CTGGGTTTAA CGCTTTATCA GATTGATATTC

31

(2) INFORMAÇÃO PARA A SEQ ID NO: 26:

25 (i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 23 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

30 (ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 26:

ACTTTTATTG CTAAACARGC TGC

23

(2) INFORMAÇÃO PARA A SEQ ID NO: 27:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 20 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

5 (D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

TAACAGCTAC TCTTCCTTTG

20

(2) INFORMAÇÃO PARA A SEQ ID NO: 28:

10 (i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 25 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

15 (ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

GGTGACATTT TCCAAGCTAA CATTG

25

(2) INFORMAÇÃO PARA A SEQ ID NO: 29:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

20 (A) COMPRIMENTO: 21 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

CCAGTCCAAT GAACCTCTTA C

21

(2) INFORMAÇÃO PARA A SEQ ID NO: 30:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 21 pares de base

30 (B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

AGGGAACAAA CCTTCCCAAC C

21

(2) INFORMAÇÃO PARA A SEQ ID NO: 31:

5 (i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 20 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

10 (ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

CARMTAKTAA MTAGGGATAG

20

(2) INFORMAÇÃO PARA A SEQ ID NO: 32:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

15 (A) COMPRIMENTO: 22 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

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

AGYTTCTATC GAAGCTGGGR ST

22

(2) INFORMAÇÃO PARA A SEQ ID NO: 33:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 1035 pares de base

25 (B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 33:

GGGTTAATTG GGTATTATTT TAAAGGGAAA GATTTTAATA ATCTGACTAT GTTTCACCA	60
ACCATAAATA ATACGCTTAT TTATGATCGG CAAACAGCAG ATACACTATT AAATAAGCAG	120
CAACAAGAGT TCAATTCTAT TCGATGGATT GGTTTAATAC AAAGTAAAGA AACAGGTGAC	180
TTTACATTCC AATTATCAGA TGATAAAAT GCCATCATTG AAATAGATGG AAAAGTTGTT	240
TCTCGTAGAG GAGAAGATAA ACAAATATC CATTTAGAAA AAGGAAAGAT GGTTCCAATC	300
AAAATTGAGT ACCAGTCCAA TGAACCTCTT ACTGTAGATA GTAAAGTATT TAACGATCTT	360
AAACTATTTA AAATAGATGG TCATAATCAA TCGCATCAA TACAGCAAGA TGATTGAAA	420
ATCCTGAATT TAATAAAAAG GAAACGAAAG AGCTTTTATC AAAACAGCA AAAAGAACCT	480
TTTCTCTTCA AAACGGGGTT GAGAAGCGAT GAGGATGATG ATCTAGGATA CAGATGGTGA	540
TAGCATTCTT GGATAATTGG GAAATGAATG GATATACCAT TCAAACGAAA AATGGCAGTC	600
AAATGGGATG ATTCATTTGC AGAAAAGGA TATACAAAAT TTGTTTCGAA TCCATATGAA	660
GCCCATACAG CAGGAGATCC TTATACCGAT TATGAAAAAG CAGCAAAAGA TATTCCTTTA	720
TCGAACGCAA AAGAAGCCTT TAATCCTCTT GTAGCTGCTT TTCCATCTGT CAATGTAGGA	780
TTAGAAAAAG TAGTAATTTT TAAAAATGAG GATATGAGTC AGGGTGTATC ATCCAGCACT	840
TCGAATAGTG CCTCTAATAC AAATTCAATT GGTGTTACCG TAGATGCTGG TTGGGAAGGT	900
TTGTTCCCTA AATTTGGTAT TTCAACTAAT TATCAAAACA CATGGACCAC TGCACAAGAA	960
TGGGGCTCTT CTAAAGAAGA TTCTACCCAT ATAAATGGAG CACAATCAGC CTTTTTAAAT	1020
GCAAATGTAC GATAT	1035

(2) INFORMAÇÃO PARA A SEQ ID NO: 34:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 345 aminoácidos

(B) TIPO: AMINOÁCIDO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: PROTEÍNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 34:

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

Ser Asn Ala Lys Glu Ala Phe Asn Pro Leu Val Ala Ala Phe Pro Ser
 245 250 255
 Val Asn Val Gly Leu Glu Lys Val Val Ile Ser Lys Asn Glu Asp Met
 260 265 270
 Ser Gln Gly Val Ser Ser Ser Thr Ser Asn Ser Ala Ser Asn Thr Asn
 275 280 285
 Ser Ile Gly Val Thr Val Asp Ala Gly Trp Glu Gly Leu Phe Pro Lys
 290 295 300
 Phe Gly Ile Ser Thr Asn Tyr Gln Asn Thr Trp Thr Thr Ala Gln Glu
 305 310 315 320
 Trp Gly Ser Ser Lys Glu Asp Ser Thr His Ile Asn Gly Ala Gln Ser
 325 330 335
 Ala Phe Leu Asn Ala Asn Val Arg Tyr
 340 345

(2) INFORMAÇÃO PARA A SEQ ID NO: 35:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 1037 pares de base

(B) TIPO: ÁCIDO NUCLEÍCO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 35:

GGGTTAATTG GGTATTATTT TAAAGGGAAA GATTTTAATA ATCTGACTAT GTTTGCACCA 60
 ACCATAAATA ATACGCTTAT TTATGATCGG CAAACAGCAG ATACACTATT AAATAAGCAG 120
 CAACAAGAGT TCAATTCTAT TCGATGGATT GGTTTAATAC AAAGTAAAGA AACAGGTGAC 180
 TTTACATTCC AATTATCAGA TGATAAAAAT GCCATCATTG AAATAGATGG AAAAGTTGTT 240
 TCTCGTAGAG GAGAAGATAA ACAAATATC CATTTAGAAA AAGGAAAGAT GGTTCCAATC 300
 AAAATTGAGT ACCAGTCCAA TGAACCTCTT ACTGTAGATA GTAAAGTATT TAACGATCTT 360
 AAACTATTTA AAATAGATGG TCATAATCAA TCGCATCAAA TACAGCAAGA TGATTTGAAA 420
 AATCCTGAAT TTAATAAAAA AGAAACGAAA GAGCTTTTAT CAAAAACAGC AAAAAGRAAC 480
 CTTTTCTCTT CAAACGRRGT KGAGAAGCGA TGAGGATGAT RATCYTAGAT ACAGGTGGKG 540
 ATAGCATTCG YKGATAATTG GGGAAATGAA WGGRTATACC ATTCAACSGA AAAATGGSAG 600
 TCAAATGGGA TGATTCATTT GCGGAAAAAG GATATACAAA ATTTGTTTCG AATCCATATG 660
 AAGCCCATAC AGCAGGAGAT CCTTATACCG ATTATGAAAA AGCAGCAAAA GATATTCCTT 720

TATCGAACGC AAAAGAAGCC TTTAATCCTC TTGTAGCTGC TTTTCCATCT GTCAATGTAG	780
GATTAGAAAA AGTAGTAATT TCTAAAAATG AGGATATGAG TCAGGGTGTA TCATCCAGCA	840
CTTCGAATAG TGCCTCTAAT ACAAATTCAA TTGGTGTAC CGTAGATGCT GGTGGGAAG	900
GTTTGTTCCTC TAAATTTGGT ATTTCAACTA ATTATCAAAA CACATGGACC ACTGCACAAG	960
AATGGGGCTC TTCTAAAGAA GATTCTACCC ATATAAATGG AGCACAATCA GCCTTTTTAA	1020
ATGCAAATGT ACGATAT	1037

(2) INFORMAÇÃO PARA A SEQ ID NO: 36:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 1048 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 36:

TGGGTTAATT GGGTATTATT TTAAAGGGCA AGAGTTTAAT CATCTTACTT TGTTGCGACC	60
AACACGTGAT AATACCCTTA TTTATGATCA ACAAACAGCG AATTCCTTAT TAGATACCAA	120
GCAACAAGAA TATCAATCTA TTCGCTGGAT TGGTTTAATT CAAAGTAAAG AAACGGGTGA	180
TTTCACATTT AACTTATCAG ATGATCAACA TGCAATTATA GAAATCGATG GCAAAATCAT	240
TTGCGATAAA GGACAGAATA AACAGTTGT TCACTTAGAA AAAGGAAAGT TAGTCCCGAT	300
AAAAATTGAG TATCAATCAG ATCAACTATT AAATAGGGAT AGTAACATCT TTAAAGAGTT	360
TAAATTATTC AAAGTAGATA GTCAGCAACA CGCTCACCAA GTTCAACTAG ACGAATTAAG	420
AAACCCTGCG TTTAATAAAA AGGAAACACA ACAATCTTAA GAAAAAGCAT CCAAAAACAA	480
TCTTTTTACA CCAGGGACAT TAAAGGAAG ATACTGATGA TGATGATAAG GATAACAGGA	540
TGGGAGATTG TATTCCTGGA CCTTTTGGGG GAAGAAAATG GGTATACCAA TCCAAAATA	600
AAATAGCTGG TCCAAGTGGG ATGTTTATTC GCCGCGAAAG GGTATACAAA TTTGTTTCTT	660
AATCCACTTG ATAGTCATAC AGTTGGAGAT CCCTATACGG ATTATGAAAA AGCAGCAAGA	720
GATTTAGACT TGGCCCAATG CAAAGAAAC ATTTAACCCA TTAGTAGCTG CTTTTCCAAG	780
TGTGAATGTG AATTGGAAG AAGTCATTTT ATCTAAAGAT GAAAATCTAT CCAATAGTGT	840
AGAGTCACAT TCCTCCACCA ACTGGTCTTA TACGAATACA GAAGGAGCTT CTATCGAAGC	900
TGGGGCTAAA CCAGAGGGTC CTACTTTTGG AGTGAGTGCT ACTTATCAAC ACTCTGAAAC	960
AGTTGCAAAA GAATGGGGAA CATCTACAGG AAATACCTCG CAATTTAATA CAGCTTCAGC	1020

AGGATATTTA AATGCAAATG TACGATAT

1048

(2) INFORMAÇÃO PARA A SEQ ID NO: 37:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 1175 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 37:

ACCTCTAGAT GCANGCTCGA GCGGCCGCCA GTGTGATGGA TATCTGCAGA ATTCGGATTA 60
 CTTGGGTATT ATTTTAAAGG GAAAGAGTTT AATCATCTTA CTTTGTTTCGC ACCAACACGT 120
 GATAATACCC TTATTTATGA TCAACAAACA GCGAATTCCT TATTAGATAC CAAACAACAA 180
 GAATATCAAT CTATTCGCTG GATTGGTTTG ATTCAAAGTA AAGAAACAGG TGATTTTCACG 240
 TTAACTTAT CTGATGATCA AAATGCAATT ATAGAAATAG ATGGCAAAAT CATTTTCGCAT 300
 AAAGGACAGA ATAAACAAGT TGTTCACTTA GAAAAAGGAA AGTTAGTCCC GATAAAAATT 360
 GAGTATCAAT CAGATCAGAT ATTAAGTAGG GATAGTAACA TCTTTAAAGA GTTCAATTAT 420
 TCAAAGTAGA TAGTCAAGCA AACTCTCAC CAAAGTTCAA CTTAGGNCNG AATTAAGNAA 480
 CCTNGGATT TTAANTTNA AAAAAAGGAAC CCNCANCATT CTTTAGGAAA AAGCAGCAAN 540
 AACCAAATCC TTTTTFACCA CAGGATATTG AAAAGGAGAT ACGGGNTNGA TGATGGATTG 600
 ATACCGGGAT ACCAGTTGGG GNTTCTANTC CCTGACCTTT GGGGAAAGAA AATNGGTATA 660
 CCNATCCCAA AANTTAAGCC AGCTGTCCAG GTGGGATGAT TCAATTCGCC CGCGAAAGGG 720
 TATACCAAAA TTTGTTTCTT AATCCACTTG AGAGTCATAC AGTTGGAGAT CCCTATACGG 780
 ATTATGAAAA AGCAGCAAGA GATTAGACT TGGCCAATGC AAAAGAAACA TTAAACCCAT 840
 TAGTAGCTGC TTTTCCAAGT GTGAATGTGA ATTTGGAAAA AGTAATATTA TCCCCAGATG 900
 AGAATTTATC TAACAGTGTA GAATCTCATT CGTCTACAAA TTGGTCTTAT ACGAATACTG 960
 AAGGAGCTTC TATCGAAGCT GGGGGTGGTC CATTAGGTAT TTCATTTGGA GTGAGTGCTA 1020
 ATTATCAACA CTCTGAAACA GTTGCAAAAG AATGGGGAAC ATCTACAGGA AATACCTCGC 1080
 AATTTAATAC AGCTTCAGCA GGATATTTAA ATGCCAATGG TCGATNTAAG CCGAATNCCA 1140
 NCACACTGNC GGCCGTTAGT AGTGGCACCG AGCCC 1175

10 (2) INFORMAÇÃO PARA A SEQ ID NO: 38:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 1030 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

5 (ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 38:

GGRTTAMTTG GGTATTATTT TAAAGGGAAA GATTTTAAATG ATCTTACTGT ATTTGCACCA	60
ACGCGTGGGA ATACTCTTGT ATATGATCAA CAAACAGCAA ATACATTACT AAATCAAAAA	120
CAACAAGACT TTCAGTCTAT TCGTTGGGTT GGTTTAATTC AAAGTAAAGA AGCAGGCGAT	180
TTTACATTTA ACTTATCAGA TGATGAACAT ACGATGATAG AAATCGATGG GAAAGTTAAT	240
TCTAATAAAG GGAAAGAAAA ACAAGTTGTC CATTTAGAAA AAGGACAGTT CGTTTCTATC	300
AAAATAGAAT ATCAAGCTGA TGAACCATTT AATGCGGATA GTCAAACCTT TAAAAATTTG	360
AAACTCYTTA AAGTAGATAC TAAGCAACAG TCCCAGCAA TTCAACTAGA TGAATTAAGA	420
AACCCTGRAA TTTAATAAAA AAGAAACACA AGAATTTCTA ACAAAGCAA CAAAAACAA	480
CCTTATTACT CAAAAGTGA AGAGTACTAG GGATGAAGAC ACGGATACAG ATGGAGATTC	540
TATTCCAGAC ATTTGGGAAG AAAATGGGTA TACCATCCAA AATAAGATTG CCGTCAAATG	600
GGATGATTCA TTAGCAAGTA AAGGATATAC GAAATTTGTT TCAAACCCAC TAGATACTCA	660
CACGGTTGGA GATCCTTATA CAGATTATGA AAAAGCAGCA AGGGATTTAG ATTTGTCAAA	720
TGCAAAAGAA ACATTTAACC CATTAGTTGC GGCTTTTCCA AGTGTGAATG TGAGTATGGA	780
AAAAGTGATA TTGTCTCCAG ATGAGAACTT ATCAAATAGT ATCGAGTCTC ATTCATCTAC	840
GAATTGGTCG TATACGAATA CAGAAGGGGC TTCTATTGAA GCTGGTGGGG GAGCATTAGG	900
CCTATCTTTT GGTGTAAGTG CAACTATCA ACATTCTGAA ACAGTTGGGT ATGAATGGGG	960
AACATCTACG GGAAATACTT CGCAATTTAA TACAGCTTCA GCGGGGTATT TAAATGCCAA	1020
TRTAMGATAT	1030

(2) INFORMAÇÃO PARA A SEQ ID NO: 39:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 23 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 39:

CACTCAAAAA ATGAAAAGGG AAA

23

(2) INFORMAÇÃO PARA A SEQ ID NO: 40:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- 5 (A) COMPRIMENTO: 19 pares de base
(B) TIPO: ÁCIDO NUCLÉICO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

10 (xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 40:

CCGGTTTTAT TGATGCTAC

19

(2) INFORMAÇÃO PARA A SEQ ID NO: 41:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- 15 (A) COMPRIMENTO: 20 pares de base
(B) TIPO: ÁCIDO NUCLÉICO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 41:

20 AGAACAATTT TTAGATAGGG

20

(2) INFORMAÇÃO PARA A SEQ ID NO: 42:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- 25 (A) COMPRIMENTO: 20 pares de base
(B) TIPO: ÁCIDO NUCLÉICO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 42:

TCCCTAAAGC ATCAGAAATA

20

30 (2) INFORMAÇÃO PARA A SEQ ID NO: 43:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

- (A) COMPRIMENTO: 1170 pares de base
(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 43:

ATGAAGAAAC AAATAGCAAG CGTTGTAAC TGTACGCTAT TAGCCCCTAT GCTTTTAAAT	60
GGAGATATGA ACGCTGCTTA CGCAGCTAGT CAAACAAAAC AAACACCTGC AGCTCAGGTA	120
AACCAAGAGA AAGAAGTAGA TCGAAAAGGA TTACTTGGCT ATTACTTTAA AGGGAAAGAT	180
TTTAATGATC TTACTGTATT TGCACCAACG CGTGGGAATA CTCTTGATA TGATCAACAA	240
ACAGCAAATA CATTACTAAA TCAAAAACAA CAAGACTTTC AGTCTATTCG TTGGGTGGT	300
TTAATTCAAA GTAAAGAAGC AGGCGATTTT ACATTTAACT TATCAGATGA TGAACATACG	360
ATGATAGAAA TCGATGGGAA AGTTATTTCT AATAAAGGGA AAGAAAAACA AGTTGTCCAT	420
TTAGAAAAAG GACAGTTCGT TTCTATAAAA TGATTCAGCT GATGAACCAT TTAATGCGGT	480
AGTAAACCTT TAAAAATTTG AAACCTCTTA AAGTAGATAC TAAGCAACAG TCCCAGCAAA	540
TTCAACTAGA TGAATTAAGA AACCCTGAAT TTAATAAAAA AGAAACACAA GAATTTCTAA	600
CAAAAGCAAC AAAACAAAAC CTTATTACTC AAAAAGTGAA GAGTACTAGG GATGAAGACA	660
CGGATACAGA TGGAGATTCT ATTCCAGACA TTTGGGAAGA AAATGGGTAT ACCATCCAAA	720
ATAAATTGCC GTCAAATGGG ATGATTCATT AGCAAGTAAA GGATATACGA AATTTGTTTC	780
AAACCCACTA GATACTCACA CGGTGGGAGA TCCTTATACA GATTATGAAA AAGCAGCAAG	840
GGATTTAGAT TTGTCAAATG CAAAAGAAAC ATTTAACCCA TTAGTTGCGG CTTTCCAAG	900
TGTAATTGAG TATGGAAAAA GGATTTGTTC CAGATGAGAA CTTATCAAAT AGTATCGAGT	960
TCATTCATTC CTACAAATGG TCGATACGAA TACAGAAGGG GCTTCTATTG AAGCTGGTGG	1020
GGGAGCATTA GGCCTATCTT TTGGTGTAAG TGCAAACTAT CAACATTCTG AAACAGTTGG	1080
GTATGAATGG GGAACATCTA CGGGAAATAC TTCGCAATTT AATACAGCTT CAGCGGGGTA	1140
TTTAAATGCG AATGTTGCTA CAATAACGTG	1170

5 (2) INFORMAÇÃO PARA A SEQ ID NO: 44:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 348 aminoácidos

(B) TIPO: AMINOÁCIDO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: PROTEÍNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 44:

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

(i) CARACTERÍSTICAS DE SEQÜÊNCIA:

- (A) COMPRIMENTO: 3 pares de base
(B) TIPO: ÁCIDO NUCLÉICO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(xi) DESCRIÇÃO DA SEQÜÊNCIA: SEQ ID NO: 45:

3

(i) CARACTERÍSTICAS DE SEQÜÊNCIA:

- (A) COMPRIMENTO: 1 aminoácido
(B) TIPO: AMINOÁCIDO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(xi) DESCRIÇÃO DA SEQÜÊNCIA: SEQ ID NO: 46:

1

(i) CARACTERÍSTICAS DE SEQÜÊNCIA:

- (A) COMPRIMENTO: 2583 pares de base

(B) TIPO: ÁCIDO NUCLEÍCO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

5 (xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 47:

ATGACATATA TGAAAAAAAA GTTAGTTAGT GTTGTAAGTT GCACGTTATT GGCTCCGATA	60
TTTTTGACTG GAAATGTACA TCCTGTTAAT GCAGACAGTA AAAAAAGTCA GCCTTCTACA	120
GCGCAGGAAA AACAAGAAAA GCCGGTTGAT CGAAAAGGGT TACTCGGCTA TTTTTTTAAA	180
GGGAAAGAGT TTAATCATCT TACTTTGTTC GCACCAACAC GTGATAATAC CCTTATTTAT	240
GATCAACAAA CAGCGAATTC CTTATTAGAT ACCAAACAAC AAGAATATCA ATCTATTCGC	300
TGGATTGGTT TGATTCAAAG TAAAGAAACA GGTGATTTC ACGTTTAACTT ATCTGATGAT	360
CAAAATGCAA TTATAGAAAT AGATGGCAAA ATCATTTCGC ATAAAGGACA GAATAACAA	420
GTTGTTCACT TAGAAAAAGG AAAGTTAGTC CCGATAAAAA TTGAGTATCA ATCAGATCAG	480
ATATTAACATA GGGATAGTAA CATCTTTAAA GAGTTTCAAT TATTCAAAGT AGATAGTCAG	540
CAACACTCTC ACCAAGTTCA ACTAGACGAA TTAAGAAACC CTGATTTTAA TAAAAAGAA	600
ACACAACAAT TCCTAGAAAA AGCAGCAAAA ACAAATCTTT TTACACAGAA TATGAAAAGA	660
GATACGGATG ATGATGATGA TACGGATACA GATGGAGATT CTATTCCTGA CCTTTGGGAA	720
GAAAATGGGT ATACCATCCA AAATAAAGTA GCTGTCAAGT GGGATGATTC ATTCGCCGCG	780
AAAGGGTATA CAAAATTTGT TTCTAATCCA CTTGAGAGTC ATACAGTTGG AGATCCCTAT	840
ACGGATTATG AAAAAGCAGC AAGAGATTTA GACTTGGCCA ATGCAAAAGA AACATTTAAC	900
CCATTAGTAG CTGCTTTTCC AAGTGTGAAT GTGAATTTGG AAAAAGTAAT ATTATCCCCA	960
GATGAGAATT TATCTAACAG TGTAGAATCT CATTTCGTCTA CAAATTTGGTC TTATACGAAT	1020
ACTGAAGGAG CTTCTATCGA AGCTGGGGGT GGTCCATTAG GTATTTTATT TGGAGTGAGT	1080
GCTAATTATC AACACTCTGA AACAGTTGCA AAAGAATGGG GAACATCTAC AGGAAATACC	1140
TCGCAATTTA ATACAGCTTC AGCAGGATAT TTGAATGCGA ATGTTCGATA CAATAATGTG	1200
GGAACAGGTG CGATTTATGA GGTGAAACCT ACAACAAGTT TTGTATTAGA TAAAGATACT	1260
GTAGCAACAA TTACCGCAAA ATCGAATTCG ACAGCTTTAA GTATATCTCC AGGAGAAAGT	1320
TATCCCAAAA AAGGACAAAA TGGAATTGCA ATTAATACAA TGGATGATTT TAATTCCCAT	1380
CCGATTACAT TAAATAAACA ACAATTAGAT CAACTATTAA ATAATAAACC TCTTATGTTA	1440
GAAACAAATC AGGCAGATGG TGTTTATAAA ATAAAGGATA CAAGCGGTAA TATTGTGACT	1500

```

GGTGGAGAAT GGAACGGTGT TATCCAACAA ATTCAAGCAA AAACAGCCTC TATTATCGTT 1560
GATACGGGAG AAAGTGTTC AGAAAAGCGT GTCGCAGCAA AAGATTATGA TAATCCTGAG 1620
GATAAAACAC CTTCTTTATC TTTAAAAGAG GCACTTAAAC TTGGATATCC AGAAGAAATT 1680
AAAGAAAAAG ATGGATTGTT GTACTATAAG GACAAGCCAA TTTACGAATC TAGTGTTATG 1740
ACTTATCTAG ATGAGAATAC AGCCAAGGAA GTGGAAAAAC AATTACAGGA TACAACCGGA 1800
ATATATAAAG ATATCAATCA TTTATATGAT GTGAAATTAA CACCTACAAT GAATTTTACG 1860
ATTAAATTAG CTTCTTTATA TGATGGAGCT GAAAATAATG ATGTGAAGAA TGGTCCTATA 1920
GGACATTGGT ATTATACCTA TAATACAGGG GGAGGAAATA CTGGAAAACA CCAATATAGG 1980
TCTGCTAATC CCAGTGCAA TGTAGTTTTA TCTTCTGAAG CGAAAAGTAA GTTAGATAAA 2040
AATACAAATT ACTACCTTAG TATGTATATG AAAGCTGAGT CTGATACAGA GCCTACRATA 2100
GAAGTAAGTG GTGAGAATTC TACGATAACG AGTAAAAAGG TAAACTAAA CAGTGAGGGC 2160
TATCAAAGAG TAGATATTTT AGTGCCGAAT TCTGAAAGAA ATCCAATAAA TCAAATATAT 2220
GTAAGAGGAA ATAATACAAC AAATGTATAC TGGGATGATG TTTCAATTAC AAATATTTCA 2280
GCTATAAACC CAAAACTTT AACAGATGAA GAAATTAAAG AAATATATAA AGATTTTAGT 2340
GAGTCTAAAG ACTGGCCTTG GTTCAATGAT GTTACGTTTA AAAATATTAA ACCATTAGAG 2400
AATTATGTAA AACAAATATAG AGTTGATTTT TGGAATACTA ATAGTGATAG ATCATTTAAT 2460
AGGATTAAGG ACAGTTACCC AGTTAATGAA GATGGAAAGTG TTAAAGTCAA CATGACAGAA 2520
TATAATGAAG GATATCCACT TAGAATTGAA TCCGCCTACC ATTTAAATAT TTCAGATCTA 2580
TAA 2583

```

(2) INFORMAÇÃO PARA A SEQ ID NO: 48:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 860 aminoácidos

(B) TIPO: AMINOÁCIDO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: PROTEÍNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 48:

```

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

```

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

340	345	350
Leu Gly Ile Ser Phe Gly Val Ser Ala Asn Tyr Gln His Ser Glu Thr 355 360 365		
Val Ala Lys Glu Trp Gly Thr Ser Thr Gly Asn Thr Ser Gln Phe Asn 370 375 380		
Thr Ala Ser Ala Gly Tyr Leu Asn Ala Asn Val Arg Tyr Asn Asn Val 385 390 395 400		
Gly Thr Gly Ala Ile Tyr Glu Val Lys Pro Thr Thr Ser Phe Val Leu 405 410 415		
Asp Lys Asp Thr Val Ala Thr Ile Thr Ala Lys Ser Asn Ser Thr Ala 420 425 430		
Leu Ser Ile Ser Pro Gly Glu Ser Tyr Pro Lys Lys Gly Gln Asn Gly 435 440 445		
Ile Ala Ile Asn Thr Met Asp Asp Phe Asn Ser His Pro Ile Thr Leu 450 455 460		
Asn Lys Gln Gln Leu Asp Gln Leu Leu Asn Asn Lys Pro Leu Met Leu 465 470 475 480		
Glu Thr Asn Gln Ala Asp Gly Val Tyr Lys Ile Lys Asp Thr Ser Gly 485 490 495		
Asn Ile Val Thr Gly Gly Glu Trp Asn Gly Val Ile Gln Gln Ile Gln 500 505 510		
Ala Lys Thr Ala Ser Ile Ile Val Asp Thr Gly Glu Ser Val Ser Glu 515 520 525		
Lys Arg Val Ala Ala Lys Asp Tyr Asp Asn Pro Glu Asp Lys Thr Pro 530 535 540		
Ser Leu Ser Leu Lys Glu Ala Leu Lys Leu Gly Tyr Pro Glu Glu Ile 545 550 555 560		
Lys Glu Lys Asp Gly Leu Leu Tyr Tyr Lys Asp Lys Pro Ile Tyr Glu 565 570 575		
Ser Ser Val Met Thr Tyr Leu Asp Glu Asn Thr Ala Lys Glu Val Glu 580 585 590		
Lys Gln Leu Gln Asp Thr Thr Gly Ile Tyr Lys Asp Ile Asn His Leu 595 600 605		
Tyr Asp Val Lys Leu Thr Pro Thr Met Asn Phe Thr Ile Lys Leu Ala 610 615 620		
Ser Leu Tyr Asp Gly Ala Glu Asn Asn Asp Val Lys Asn Gly Pro Ile 625 630 635 640		
Gly His Trp Tyr Tyr Thr Tyr Asn Thr Gly Gly Gly Asn Thr Gly Lys 645 650 655		

His Gln Tyr Arg Ser Ala Asn Pro Ser Ala Asn Val Val Leu Ser Ser
 660 665 670
 Glu Ala Lys Ser Lys Leu Asp Lys Asn Thr Asn Tyr Tyr Leu Ser Met
 675 680 685
 Tyr Met Lys Ala Glu Ser Asp Thr Glu Pro Thr Ile Glu Val Ser Gly
 690 695 700
 Glu Asn Ser Thr Ile Thr Ser Lys Lys Val Lys Leu Asn Ser Glu Gly
 705 710 715 720
 Tyr Gln Arg Val Asp Ile Leu Val Pro Asn Ser Glu Arg Asn Pro Ile
 725 730 735
 Asn Gln Ile Tyr Val Arg Gly Asn Asn Thr Thr Asn Val Tyr Trp Asp
 740 745 750
 Asp Val Ser Ile Thr Asn Ile Ser Ala Ile Asn Pro Lys Thr Leu Thr
 755 760 765
 Asp Glu Glu Ile Lys Glu Ile Tyr Lys Asp Phe Ser Glu Ser Lys Asp
 770 775 780
 Trp Pro Trp Phe Asn Asp Val Thr Phe Lys Asn Ile Lys Pro Leu Glu
 785 790 795 800
 Asn Tyr Val Lys Gln Tyr Arg Val Asp Phe Trp Asn Thr Asn Ser Asp
 805 810 815
 Arg Ser Phe Asn Arg Ile Lys Asp Ser Tyr Pro Val Asn Glu Asp Gly
 820 825 830
 Ser Val Lys Val Asn Met Thr Glu Tyr Asn Glu Gly Tyr Pro Leu Arg
 835 840 845
 Ile Glu Ser Ala Tyr His Leu Asn Ile Ser Asp Leu
 850 855 860

(2) INFORMAÇÃO PARA A SEQ ID NO: 49:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 1356 pares de base

(B) TIPO: ÁCIDO NUCLÉICO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 49:

ATGGTATCCA AAAAGTTACA ATTAGTCACA AAAACTTTAG TGTTTAGTAC AGTTTGTGCA 60
 ATACCGTTAT TAAATAATAG TGAGATAAAA GCGGAACAAT TAAATATGAA TTCTCAAATT 120

AAATATCCTA ACTTCCAAAA TATAAATATC GCTGATAAGC CAGTAGATTT TAAAGAGGAT	180
AAAGAAAAAG CACGAGAATG GGGAAAAGAA AAAGAAAAAG AGTGGAAACT AACTGCTACT	240
GAAAAAGGGA AAATTAATGA TTTTITAGAT GATAAAGATG GATTAAAAAC AAAATACAAA	300
GAAATTAATT TTTCTAAGAA TTTTGAATAT GAAACAGAGT TAAAACAGCT TGAAAAAATT	360
AATAGCATGC TAGATAAAGC AAATCTAACA AATTCAATTG TCACGTATAA AAACGTTGAG	420
CCTACAACAA TAGGATTCAA TCACTCTTTG ACTGATGGGA ATCAAATTAA TTCCGAAGCT	480
CAACAGAAGT TCAAGGAACA GTTTTITAGGA AATGATATTA AATTTGATAG TTATTTGGAT	540
ATGCACTTAA CTGAACAAAA TGTTCCTGGT AAAGAAAGGG TTATTTTAAA AGTTACAGTA	600
CTTAGTGGGA AAGGTTCTAC TCCAACAAAA GCAGGTGTTG TTTTAAATAA TAAAGAATAC	660
AAAATGTTGA TTGATAATGG ATATATACTA CATGTAGAAA ACATAACGAA AGTTGTAAAA	720
AAAGGACAGG AATGTTTACA AGTTGAAGGA ACGTTAAAAA AGAGCTTGGG CTTTAAAAAT	780
GATAGTGACG GTAAGGGAGA TTCCTGGGGA AAGAAAAATT ACAAGGAATG GTCTGATTCT	840
TTAACAAATG ATCAGAGAAA AGACTTAAAT GATTATGGTG CGCGAGGTTA TACCGAAATA	900
AATAAATATT TACGTGAAGG GGGTACCGGA AATACAGAGT TGGAGGAAAA AATTAAAAAT	960
ATTTCTGACG CACTAGAAAA GAATCCTATC CCTGAAAACA TTACTGTTTA TAGATATTGC	1020
GGAATGGCGG AATTTGGTTA TCCAATTCAA CCCGAGGCTC CCTCCGTACA AGATTTTGAA	1080
GAGAAATTTT TGGATAAAAT TAAGGAAGAA AAAGGATATA TGAGTACGAG CTTATCAAGT	1140
GATGCGACTT CTTTGGCGC AAGAAAAATT ATCTTAAGAT TGCAGATACC AAAAGGAAGT	1200
TCAGGAGCAT ATGTAGCTGG TTTAGATGGA TTAAACCAG CAGAGAAGGA GATTCTTATT	1260
GATAAGGGAA GCAAGTATCA TATTGATAAA GTAACAGAAG TAGTTGTGAA AGGTATTAGA	1320
AAACTCGTAG TAGATGCGAC ATTATTATTA AAATAA	1356

(2) INFORMAÇÃO PARA A SEQ ID NO: 50:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 451 aminoácidos

(B) TIPO: AMINOÁCIDO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 50:

Met	Val	Ser	Lys	Lys	Leu	Gln	Leu	Val	Thr	Lys	Thr	Leu	Val	Phe	Ser
1				5					10					15	

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

(i) CARACTERÍSTICAS DE SEQÜÊNCIA:

- (A) COMPRIMENTO: 47 pares de base
(B) TIPO: ÁCIDO NUCLÉICO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 51:

GCTCTAGAAG GAGGTAAGTT ATGAACAAGA ATAATACTAA ATTAAGC

47

(2) INFORMAÇÃO PARA A SEQ ID NO: 52:

- (i) CARACTERÍSTICAS DE SEQUÊNCIA:

- (A) COMPRIMENTO: 27 pares de base
(B) TIPO: ÁCIDO NUCLÉICO
(C) TIPO DE FIO: ÚNICO
(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 52:

GGGGTACCTT ACTTAATAGA GACATCG

(2) INFORMAÇÃO PARA A SEQ ID NO: 53:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 2364 pares de base

(B) TIPO: ÁCIDO NUCLEÍCO

5

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: DNA (GENÔMICO)

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 53:

ATGAATATGA ATAATACTAA ATTAAACGCA AGGGCCCTAC CGAGTTTTAT TGATTATTTT	60
AATGGCATT T ATGGATTTC CACTGGTATC AAAGACATTA TGAATATGAT TTTTAAAACG	120
GATACAGGTG GTAATCTAAC CTTAGACGAA ATCCTAAAGA ATCAGCAGTT ACTAAATGAG	180
ATTTCTGGTA AATTGGATGG GGTAAATGGG AGCTTAAATG ATCTTATCGC ACAGGGAAAC	240
TTAAATACAG AATTATCTAA GGAAATCTTA AAAATTGCAA ATGAACAGAA TCAAGTCTTA	300
AATGATGTTA ATAACAACT CGATGCGATA AATACGATGC TTCATATATA TCTACCTAAA	360
ATCACATCTA TGTTAAGTGA TGTAATGAAG CAAAATTATG CGCTAAGTCT GCAAGTAGAA	420
TACTTAAGTA AACAATTGAA AGAAATTTCT GATAAATTAG ATGTTATTAA CGTAAATGTT	480
CTTATTAACT CTACACTTAC TGAAATTACA CCTGCATATC AACGGATTAA ATATGTAAAT	540
GAAAAATTTG AAGAATTAAC TTTTGCTACA GAAACCACTT TAAAAGTAAA AAAGGATAGC	600
TCGCCTGCTG ATATTCTTGA CGAGTTAACT GAATTAACTG AACTAGCGAA AAGTGTTACA	660
AAAAATGACG TGGATGGTTT TGAATTTTAC CTTAATACAT TCCACGATGT AATGGTAGGA	720
AATAATTTAT TCGGGCGTTC AGCTTTAAAA ACTGCTTCAG AATTAAATTGC TAAAGAAAAT	780
GTGAAAACAA GTGGCAGTGA AGTAGGAAAT GTTTATAATT TCTTAATTGT ATTAACAGCT	840
CTACAAGCAA AAGCTTTTCT TACTTTAACA ACATGCCGAA AATTATTAGG CTTAGCAGAT	900
ATTGATTATA CATCTATTAT GAATGAACAT TTAAATAAGG AAAAGAGGA ATTTAGAGTA	960
AACATCCTTC CTACACTTTC TAATACTTTT TCTAATCCTA ATTATGCAA AGTTAAAGGA	1020
AGTGATGAAG ATGCAAAGAT GATTGTGGAA GCTAAACCAG GACATGCATT GGTGGGTTT	1080
GAAATTAGTA ATGATTCAAT GACAGTATTA AAAGTATATG AAGCTAAGCT AAACAAAAT	1140
TACCAAGTTG ATAAGGATTC CTTATCGGAA GTCATTTATA GTGATATGGA TAAATTATTG	1200
TGCCCAGATC AATCTGAACA AATTTATTAT ACAAATAATA TAGTATTTCC AAATGAATAT	1260
GTAATTACTA AAATTGATTT TACTAAGAAA ATGAAAACCT TAAGATATGA GGTAACAGCT	1320

```

AATTCTTACG ATTCTTCTAC AGGAGAAATT GACTTAAATA AGAAGAAAGT AGAATCAAGT 1380
GAAGCGGAGT ATAGGACGTT AAGTGCTAAT AATGATGGAG TATATATGCC GTTAGGTGTC 1440
ATCAGTGAAA CATTTTTGAC TCCAATTAAT GGATTTGGCC TCCAAGCTGA TGAAAATTCA 1500
AGATTAAATTA CTTTAAACATG TAAATCATAT TTAAGGGAAC TACTACTAGC GACAGACTTA 1560
AGCAATAAAG AAATAAAATT GATTGTCCCG CCTATTAGTT TTATTAGTAA TATTGTAGAA 1620
AATGGGAACT TAGAGGGAGA AAACCTAGAG CCGTGGATAG CAAATAACAA AAATGCGTAT 1680
GTAGATCATA CAGGTGGTAT AAATGGAACCT AAAGTTTTAT ATGTTTCATAA GGATGGTGAG 1740
TTTTCACAAT TTGTTGGAGG TAAGTTAAAA TCGAAAACAG AATATGTAAT TCAATATATT 1800
GTAAAGGGAA AAGCTTCTAT TTATTTAAAA GATAAAAAAA ATGAGAATTC CATTTATGAA 1860
GAAATAAATA ATGATTTAGA AGGTTTTCAA ACTGTTACTA AACGTTTTAT TACAGGAACG 1920
GATTCTTCAG GGATTCATTT AATTTTTACC AGTCAAAATG GCGAGGGAGC ATTTGGAGGA 1980
AACITTTATTA TCTCAGAAAT TAGGACATCC GAAGAGTTAT TAAGTCCAGA ATTGATTATG 2040
TCGGATGCTT GGGTTGGATC CCAGGGAACCT TGGATCTCAG GAAATTCTCT CACTATTAAT 2100
AGTAATGTAA ATGGAACCTT TCGACAAAAT CTTCCGTTAG AAAGTTATTC AACCTATAGT 2160
ATGAACTTTA CTGTGAATGG ATTTGGCAAG GTGACAGTAA GAAATTCTCG TGAAGTATTA 2220
TTTGAAAAAA GTTATCCGCA GCTTTCACCT AAAGATATTT CTGAAAAATT TACAACTGCA 2280
GCCAATAATA CCGGATTATA TGTAGAGCTT TCTCGCTCAA CGTCGGGTGG TGCAATAAAT 2340
TTCCGAGATT TTTCAATTAA GTAA 2364

```

(2) INFORMAÇÃO PARA A SEQ ID NO: 54:

(i) CARACTERÍSTICAS DE SEQUÊNCIA:

(A) COMPRIMENTO: 787 aminoácidos

(B) TIPO: AMINOÁCIDO

(C) TIPO DE FIO: ÚNICO

(D) TOPOLOGIA: LINEAR

(ii) TIPO DE MOLÉCULA: PROTEÍNA

(xi) DESCRIÇÃO DA SEQUÊNCIA: SEQ ID NO: 54:

```

Met Asn Met Asn Asn Thr Lys Leu Asn Ala Arg Ala Leu Pro Ser Phe
1           5           10           15

```

```

Ile Asp Tyr Phe Asn Gly Ile Tyr Gly Phe Ala Thr Gly Ile Lys Asp
20           25           30

```

```

Ile Met Asn Met Ile Phe Lys Thr Asp Thr Gly Gly Asn Leu Thr Leu

```

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

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

660	665	670
Leu Leu Ser Pro Glu Leu Ile Met Ser Asp Ala Trp Val Gly Ser Gln		
675	680	685
Gly Thr Trp Ile Ser Gly Asn Ser Leu Thr Ile Asn Ser Asn Val Asn		
690	695	700
Gly Thr Phe Arg Gln Asn Leu Pro Leu Glu Ser Tyr Ser Thr Tyr Ser		
705	710	715
Met Asn Phe Thr Val Asn Gly Phe Gly Lys Val Thr Val Arg Asn Ser		
725	730	735
Arg Glu Val Leu Phe Glu Lys Ser Tyr Pro Gln Leu Ser Pro Lys Asp		
740	745	750
Ile Ser Glu Lys Phe Thr Thr Ala Ala Asn Asn Thr Gly Leu Tyr Val		
755	760	765
Glu Leu Ser Arg Ser Thr Ser Gly Gly Ala Ile Asn Phe Arg Asp Phe		
770	775	780
Ser Ile Lys		
785		

REIVINDICAÇÕES

1. Hospedeiro recombinante, caracterizado pelo fato de que compreende o polinucleotídeo compreendendo a sequência de nucleotídeos da SEQ ID NO: 53 e codificando uma proteína pesticidamente ativa, em que
- 5 a referida proteína compreende a sequência de aminoácidos da SEQ ID NO: 54, em que o referido polinucleotídeo está em um constructo genético para a transformação de plantas e não está associado com uma molécula natural ou não é natural ou encontrado na natureza ou ainda que dela isolado, em que o referido hospedeiro consiste de uma célula microbiana.
- 10 2. Hospedeiro de acordo com a reivindicação 1, caracterizado pelo fato de que a célula microbiana é uma célula bacteriana.
3. Método para controlar uma peste lepidopterana, caracterizado pelo fato de que compreende por em contato a referida peste com uma proteína, em que a referida proteína compreende a sequência de aminoácidos
- 15 da SEQ ID NO: 54.
4. Método de acordo com a reivindicação 3, caracterizado pelo fato de que a referida proteína é produzida por e está presente em uma planta, a referida peste lepidopterana ingere uma porção da referida planta, dessa forma entrando em contato com a referida proteína.

RESUMO

Patente de Invenção: **"HOSPEDEIRO RECOMBINANTE E MÉTODO PARA CONTROLAR UMA PESTE LEPIDOPTERANA"**.

Revelados e reivindicados são novos isolados de *Bacillus thuringiensis*, toxinas pesticidas, genes e iniciadores e sondas de nucleotídeo para a identificação de genes que codificam toxinas ativas contra pestes. Os iniciadores são úteis em técnicas de PCR para produzir fragmentos de gene que são característicos de genes que codificam estas toxinas. A presente invenção provê totalmente novas famílias de toxinas que podem ser obtidas a partir de sobrenadantes de culturas de *Bacillus*.