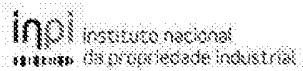


---

(11) Número de Publicação: **PT 2001900 E**



(51) Classificação Internacional:  
**C07K 14/18 (2015.01)**

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

---

(22) Data de pedido: **2007.03.15**

(30) Prioridade(s): **2006.03.15 CA 2538898  
2006.04.04 CA 2545597**

(43) Data de publicação do pedido: **2008.12.17**

(45) Data e BPI da concessão: **2015.04.29  
180/2015**

(73) Titular(es):

**INSTITUT PASTEUR  
28, RUE DU DOCTEUR ROUX 75015 PARIS FR**

(72) Inventor(es):

**PIERRE CHARNEAU FR  
PHILIPPE DESPRES FR  
ANNE-CLAIRE BREHIN FR  
VALÉRIE MARECHAL FR  
PHILIPPE SOUQUE FR**

(74) Mandatário:

**ALBERTO HERMÍNIO MANIQUE CANELAS  
RUA VÍCTOR CORDON, 14 1249-103 LISBOA PT**

(54) Epígrafe: **NOVAS ESTIRPES ISOLADAS E PURIFICADAS DO VÍRUS CHIKUNGUNYA E SEQUÊNCIAS POLINUCLEOTÍDICAS E POLIPEPTÍDICAS, SUAS UTILIZAÇÕES DIAGNÓSTICAS E IMUNOGÉNICAS**

(57) Resumo:

A PRESENTE INVENÇÃO REFERE-SE A ESTIRPES SELVAGENS DO VÍRUS CHIKUNGUNYA ISOLADAS A PARTIR DE DOENTES QUE APRESENTAM FORMAS GRAVES DE INFECÇÃO E QUE PROVÊM DE UMA EPIDEMIA DE ARBOVIROSE HUMANA. A PRESENTE INVENÇÃO REFERE-SE TAMBÉM A SEQUÊNCIAS POLIPEPTÍDICAS E FRAGMENTOS DAS MESMAS DERIVADOS DO SEU GENOMA, AO POLINUCLEÓTIDO QUE CODIFICA AS MESMAS E À SUA UTILIZAÇÃO COMO PRODUTOS DE DIAGNÓSTICO, COMO VACINAS E/OU COMO COMPOSIÇÕES IMUNOGÉNICAS.

**RESUMO**

**"NOVAS ESTIRPES ISOLADAS E PURIFICADAS DO VÍRUS CHIKUNGUNYA  
E SEQUÊNCIAS POLINUCLEOTÍDICAS E POLIPEPTÍDICAS, SUAS  
UTILIZAÇÕES DIAGNÓSTICAS E IMUNOGÉNICAS"**

A presente invenção refere-se a estirpes selvagens do vírus Chikungunya isoladas a partir de doentes que apresentam formas graves de infecção e que provêm de uma epidemia de arbovirose humana. A presente invenção refere-se também a sequências polipeptídicas e fragmentos das mesmas derivados do seu genoma, ao polinucleótido que codifica as mesmas e à sua utilização como produtos de diagnóstico, como vacinas e/ou como composições imunogénicas.

**DESCRIÇÃO**

**"NOVAS ESTIRPES ISOLADAS E PURIFICADAS DO VÍRUS CHIKUNGUNYA  
E SEQUÊNCIAS POLINUCLEOTÍDICAS E POLIPEPTÍDICAS, SUAS  
UTILIZAÇÕES DIAGNÓSTICAS E IMUNOGÉNICAS"**

**CAMPO DA INVENÇÃO**

A presente invenção refere-se a estirpes selvagens do vírus Chikungunya isoladas a partir de doentes que apresentam formas graves de infecção e que provêm de uma epidemia de arbovirose humana. A presente invenção refere-se também a sequências polipeptídicas e fragmentos das mesmas derivados do seu genoma, ao polinucleótido que codifica as mesmas e à sua utilização como produtos de diagnóstico, como vacinas e/ou como composições imunogénicas.

**ANTECEDENTES DA INVENÇÃO**

O vírus Chikungunya (CHIKV) é um *Alfavírus* transmitido pelo mosquito que pertence à família *Togaviridae* [1,2]. Foi isolado pela primeira vez a partir de um surto Tanzaniano em 1952 [3]. É responsável por uma infecção aguda de início abrupto, caracterizada por febres altas, artralgia, mialgia, dor de cabeça e erupção cutânea [4,5]. A poliartralgia, o sinal patognomónico da doença, é

muito dolorosa. Os sintomas são geralmente autolimitantes e duram 1 a 10 dias. No entanto, a artralgia ou os sintomas artríticos podem persistir durante meses ou anos. Nalguns doentes, foram também descritos sinais hemorrágicos secundários tais como epistaxe ou gengivorragia.

O CHIKV encontra-se geograficamente distribuído por África, Índia e Sudeste Asiático. Em África, o vírus é mantido através de um ciclo de transmissão selvático entre primatas selvagens e mosquitos tais como *Aedes luteocephalus*, *Ae. furcifer* ou *Ae. taylori* [4]. Na Ásia, o CHIKV é principalmente transmitido de humano para humano pelo *Ae. aegypti* e num menor grau pelo *Ae. albopictus* através de um ciclo de transmissão urbano. Desde o surto na Tanzânia em 1952, o CHIKV tem provocado surtos na África Oriental (Tanzânia, Uganda), na África Austral (Zimbabué, África do Sul), na África Ocidental (Senegal, Nigéria) e na África Central (República Centro-Africana, República Democrática do Congo) [4]. A reemergência epidémica mais recente foi documentada em 1999-2000 em Kinshasa, onde se estima que tenham sido infetadas 50 000 pessoas [6]. Desde o primeiro surto Asiático documentado em 1958 em Banguecoque, Tailândia, foram documentados surtos na Tailândia, Camboja, Vietname, Laos, Mianmar, Malásia, Filipinas e Indonésia [4,5]. A reemergência epidémica mais recente foi documentada em 2001-2003 em Java após 20 anos [7]. Na África ou Ásia, a reemergência foi imprevisível, com intervalos de 7-8 anos a 20 anos entre epidemias consecutivas.

Desde o final de 2004, o vírus Chikungunya (CHIKV) tem emergido nas ilhas do Oceano Índico ocidental austral. Entre janeiro e março de 2005 foram registados mais de 5 000 casos nas Comores. Mais tarde em 2005, o vírus circulou noutras ilhas, i.e. Maiote, Seicheles, Reunião e Maurícia. A partir de dezembro de 2005, a estação das chuvas deu origem a uma circulação epidémica renovada do vírus. Entre 1 de janeiro e 1 de março de 2006, foram registados 2 553, 3 471 e 4 650 casos na Maurícia, Maiote e Seicheles (12 de março de 2006). A ilha mais afetada é a Reunião com uma estimativa de 212 000 casos até 12 de março de 2006 (população total: 770 000). Mais recentemente foi documentada a circulação do vírus em Madagáscar.

Na Ilha da Reunião, os primeiros casos documentados foram doentes que regressaram das Comores em março de 2005. Foram registados mais de 3 000 casos de março a junho. A transmissão foi limitada durante a estação de inverno do hemisfério sul e foi observado um grande aumento desde meados de dezembro, com uma estimativa de 210 000 casos entre janeiro e março de 2006 [8]. Desde março de 2005, 85 doentes com uma infecção por CHIKV confirmada desenvolveram sinais clínicos graves (meningoencefalite ou hepatite fulminante) que justificaram hospitalização numa unidade de cuidados intensivos. Vários casos de meningoencefalite e síndrome álgica principal foram associados a transmissão vertical do vírus 9.

Até à data foram determinadas duas sequências de nucleótidos completas para o CHIKV, para as estirpes Ross (n.º de acesso: AF490259) e S27 [9], ambas isoladas a partir de doentes durante o surto da Tanzânia em 1952. Foi determinada outra sequência de nucleótidos completa para uma estirpe isolada de *Ae. furcifer* durante o surto do Senegal em 1983 (n.º de acesso AY726732). Khan e colaboradores [9] mostraram que o genoma S27 era semelhante na sua estrutura ao de outros alfavírus e que o vírus O'nyong-nyong (ONN) era o parente mais próximo do CHIKV. Além disso, análises filogenéticas com base em sequências El parciais de isolados Africanos e Asiáticos revelaram a existência de três grupos filogenéticos de CHIKV distintos, um que continha todos os isolados da África Ocidental, um que continha isolados da Ásia, e um que correspondia aos isolados da África Oriental, Central e Austral [10]. As estirpes isoladas em 1999-2000 na República Democrática do Congo pertenciam ao último grupo filogenético [6].

### **SUMÁRIO DA INVENÇÃO**

Um aspeto da invenção como definida pelas reivindicações 1-13 consiste em proporcionar novas ferramentas de diagnóstico e imunológicas contra doenças associadas ao vírus CHIK, tais como a arbovirose.

Um tal aspetto é particularmente conseguido proporcionando uma estirpe selvagem isolada e purificada do vírus Chikungunya (CHIK) capaz de infetar células humanas

*in vitro*; e à sua utilização para a deteção de um CHIKV associado a um arbovírus, ou para a preparação de uma composição que previne e/ou trata um arbovírus.

Outro aspeto da invenção como definida pelas reivindicações 1-13 refere-se a uma estirpe isolada e purificada de CHIKV compreendendo pelo menos uma mutação na proteína estrutural E1 e/ou proteína estrutural E2; e à sua utilização para a deteção de um CHIKV associado a um arbovírus, ou para a preparação de uma composição que previne e/ou trata um arbovírus.

Outro aspeto da invenção como definida pelas reivindicações 1-13 refere-se a um polinucleótido isolado e purificado que consiste na sequência de SEQ ID NO:1; e à sua utilização para a deteção de um CHIKV associado a um arbovírus, ou para a preparação de uma composição que previne e/ou trata um arbovírus.

Outro aspeto da invenção como definida pelas reivindicações 1-13 refere-se a um fragmento do polinucleótido da invenção que codifica o ectodomínio da glicoproteína E2 ou E1; e à sua utilização para a deteção de um CHIKV associado a um arbovírus, ou para a preparação de uma composição que previne e/ou trata um arbovírus.

Outros aspetos da invenção como definidos pelas reivindicações 1-13 referem-se a um vetor compreendendo um fragmento considerado pela presente invenção, e à célula

hospedeira compreendendo o referido vetor; e à sua utilização para a deteção de um CHIKV associado a um arbovírus, ou para a preparação de uma composição que previne e/ou trata um arbovírus.

Ainda outro aspeto da invenção como definida pelas reivindicações 1-13 refere-se a um polipéptido purificado codificado por um fragmento da invenção; e à sua utilização para a deteção de um CHIKV associado a um arbovírus, ou para a preparação de uma composição que previne e/ou trata um arbovírus.

Um outro aspeto da invenção como definida pelas reivindicações 1-13 refere-se à utilização de um anticorpo monoclonal ou policlonal ou seu fragmento que se liga especificamente a um polipéptido da invenção para a deteção de um CHIKV associado a um arbovírus.

**BREVE DESCRIÇÃO DAS FIGURAS**

**Figura 1: Localização das alterações de E1 no modelo estrutural 3D da estrutura cristalina de SFV E1 [43] [19].**

**A)** Diagrama de fita de E1, com o domínio I colorido a vermelho, domínio II a amarelo e domínio III a azul. Os tubos verde marcam as ligações dissulfureto. O péptido de fusão, na extremidade da molécula (no domínio II) está colorido de laranja e marcado. A extremidade N-

terminal e a extremidade C-terminal observadas no cristal (a qual está a 30 aa a montante da região transmembranar) estão também marcadas. As 2 únicas alterações observadas nos isolados do Oceano Índico são indicadas por inícios e marcadas: posições 226 (branca) e 284 (magenta).

**B)** Representação parcial (um oitante, ligeiramente alongado) da armação E1 icosaédrica à superfície do virião, vista ao longo de um eixo de simetria quíntuplo. Um promotor E1 encontra-se destacado a cores, como em A); todos os outros são representados a cinzento. A localização de alguns dos eixos de simetria icosaédricos é desenhada como símbolos pretos sólidos: pentágono para eixos quíntuplos, triângulo para eixos triplos, elipse para eixos duplos (os quais na rede T=4 do alfavírus são coincidentes com eixos quasi-sextuplos). Os triângulos vazios indicam grosseiramente a localização dos trímeros E2 que interagem fortemente com E1, que cobrem o domínio II e o péptido de fusão, e que apresentam os sítios antigénicos principais. Os triângulos vazios marcam também eixos de simetria quasi-triplos da rede icosaédrica superficial T=4. Uma bola magenta marca a localização da Glu 284, num sítio de contacto do promotor inter-E1. Este contacto propaga-se 240 vezes na rede da superfície (observe-se todas as bolas cor-de-rosa desenhadas nos promotores a cinzento). Refira-se que o péptido de fusão, a laranja, está a apontar para cima e a afastar-se de contactos com outros promotores E1. Isto é mais facilmente observado na periferia do virião, onde um deles se encontra marcado (FP). No virião, esta região de

E1 não está acessível, coberta, por baixo da molécula E2 [19].

**Figura 2:** Relações filogenéticas entre isolados de chikungunya com base em sequências nucleotídicas parciais de E1. Os isolados do surto do Oceano Índico (Reunião, Seicheles, Maiote, Maurícia, Madagáscar) representam um clado distinto dentro de um grande grupo filogenético da África Oriental, Central e Austral (ECSA). Os valores de reamostragem de iniciação são indicados nos nós principais. O ramo que conduz ao grupo filogenético da África Ocidental (de comprimento aprox. 15%) foi encurtado por conveniência.

**Figura 3:** Cenário evolutivo proposto dos isolados do vírus Chikungunya a partir do surto 1 do Oceano Índico. O cenário baseia-se em seis sequências de genoma determinadas por sequenciação direta de produtos de RT-PCR obtidos utilizando extratos de ARN como cadeias molde; as sequências correspondem assim às sequências consenso (Seq. Cons.) da possível mistura de genomas coexistentes (quasiespécie). Inserção: número de casos de E1-226A e E1-226V em diferentes intervalos de tempo na Ilha da Reunião, com base em sequências E1 parciais. O E1-226V foi observado nas sequências consenso 2, 3 e 4, e por conseguinte a maioria dos isolados E1-226V genotiparam com base nas sequências E1 parciais estão provavelmente relacionados com estes genótipos. No entanto, o aparecimento independente de E1-226V noutros genótipos não pode ser excluído. Int.:

sequência intermediária. A localização, tamanho e posição relativa das Ilhas e da fronteira Africana são indicativos. A sequência consenso 1 foi obtida a partir de um doente da Reunião que tinha regressado das Comores em março de 2005, e de um doente da Ilha da Reunião. As sequências 2 a 4 foram amostradas na Ilha da Reunião; a sequência 5 foi amostrada na Seicheles.

A **Figura 4** mostra a sequência de nucleótidos do genoma de uma estirpe de vírus CHIK de acordo com uma forma de realização preferida da invenção e mais especificamente para a estirpe preferida denominada 05.115 (SEQ ID NO:1).

A **Figura 5** mostra a sequência de nucleótidos do genoma de uma estirpe de vírus CHIK de acordo com a presente divulgação e mais especificamente para a estirpe preferida denominada 05.209 (SEQ ID NO:2).

A **Figura 6** mostra a sequência de nucleótidos do genoma de uma estirpe de vírus CHIK de acordo com a presente divulgação e mais especificamente para a estirpe preferida denominada 06.21 (SEQ ID NO:3).

A **Figura 7** mostra a sequência de nucleótidos do genoma de uma estirpe de vírus CHIK de acordo com a presente divulgação e mais especificamente para a estirpe preferida denominada 06.27 (SEQ ID NO:4).

A **Figura 8** mostra a sequência de nucleótidos do

genoma de uma estirpe de vírus CHIK de acordo com a presente divulgação e mais especificamente para a estirpe preferida denominada 06.49 (SEQ ID NO:5).

A **Figura 9** mostra a sequência de nucleótidos do genoma de uma estirpe de vírus CHIK de acordo com a presente divulgação e mais especificamente para a estirpe preferida denominada 05.61 (SEQ ID NO:6).

A **Figura 10** mostra uma sequência de nucleótidos de um fragmento de um vírus CHIK de acordo com a presente divulgação e mais especificamente um fragmento que codifica o ectodomínio da glicoproteína E2 da estirpe preferida denominada 06.21 (SEQ ID NO:7).

A **Figura 11** mostra uma sequência de nucleótidos de um fragmento de um vírus CHIK de acordo com a presente divulgação e mais especificamente um fragmento que codifica o ectodomínio da glicoproteína E2 da estirpe preferida denominada 06.27 (SEQ ID NO:8).

A **Figura 12** mostra uma sequência de nucleótidos de um fragmento de um vírus CHIK de acordo com a presente divulgação e mais especificamente um fragmento que codifica o ectodomínio da glicoproteína E2 da estirpe preferida denominada 06.49 (SEQ ID NO:9).

A **Figura 13** mostra uma sequência de nucleótidos de um fragmento de um vírus CHIK de acordo com uma forma de

realização preferida da invenção, e mais especificamente um fragmento que codifica o ectodomínio da glicoproteína E2 da estirpe preferida denominada 05.115 (SEQ ID NO:10).

A **Figura 14** mostra uma sequência de nucleótidos de um fragmento de um vírus CHIK de acordo com a presente divulgação e mais especificamente um fragmento que codifica uma forma solúvel da glicoproteína E2 da estirpe preferida denominada 06.21 (SEQ ID NO:11).

A **Figura 15** mostra uma sequência de nucleótidos de um fragmento de um vírus CHIK de acordo com a presente divulgação e mais especificamente um fragmento que codifica uma forma solúvel da glicoproteína E2 da estirpe preferida denominada 06.27 (SEQ ID NO:12).

A **Figura 16** mostra uma sequência de nucleótidos de um fragmento de um vírus CHIK de acordo com a presente divulgação e mais especificamente um fragmento que codifica uma forma solúvel da glicoproteína E2 da estirpe preferida denominada 06.49 (SEQ ID NO:13).

A **Figura 17** mostra uma sequência de nucleótidos de um fragmento de um vírus CHIK de acordo com uma forma de realização preferida da invenção, e mais especificamente um fragmento que codifica uma forma solúvel da glicoproteína E2 da estirpe preferida denominada 05.115 (SEQ ID NO:14).

A **Figura 18** mostra uma sequência de aminoácidos

de um polipéptido preferido do vírus CHIK de acordo com a presente divulgação e relacionado mais especificamente com o ectodomínio da glicoproteína E2 da estirpe preferida denominada 06.21 (SEQ ID NO:15).

A **Figura 19** mostra uma sequência de aminoácidos de um polipéptido preferido do vírus CHIK de acordo com a presente divulgação e relacionado mais especificamente com o ectodomínio da glicoproteína E2 da estirpe preferida denominada 06.27 (SEQ ID NO:16).

A **Figura 20** mostra uma sequência de aminoácidos de um polipéptido preferido do vírus CHIK de acordo com a presente divulgação e relacionado mais especificamente com o ectodomínio da glicoproteína E2 da estirpe preferida denominada 06.49 (SEQ ID NO:17).

A **Figura 21** mostra uma sequência de aminoácidos de um polipéptido preferido do vírus CHIK de acordo com uma forma de realização preferida da invenção, e relacionado mais especificamente com o ectodomínio da glicoproteína E2 da estirpe preferida denominada 05.115 (SEQ ID NO:18).

A **Figura 22** mostra uma sequência de aminoácidos de um polipéptido preferido do vírus CHIK de acordo com a presente divulgação e relacionado mais especificamente com uma forma solúvel da glicoproteína E2 da estirpe preferida denominada 06.21 (SEQ ID NO: 19).

A **Figura 23** mostra uma sequência de aminoácidos de um polipéptido preferido do vírus CHIK de acordo com a presente divulgação e relacionado mais especificamente com uma forma solúvel da glicoproteína E2 da estirpe preferida denominada 06.27 (SEQ ID NO:20).

A **Figura 24** mostra uma sequência de aminoácidos de um polipéptido preferido do vírus CHIK de acordo com a presente divulgação e relacionado mais especificamente com uma forma solúvel da glicoproteína E2 da estirpe preferida denominada 06.49 (SEQ ID NO:21).

A **Figura 25** mostra uma sequência de aminoácidos de um polipéptido preferido do vírus CHIK de acordo com uma forma de realização preferida da invenção, e relacionado mais especificamente com uma forma solúvel da glicoproteína E2 da estirpe preferida denominada 05.115 (SEQ ID NO:22).

**Figura 26: Elementos da Sequência de Repetição presente na região 3'NTR**

**A.** Alinhamento dos Elementos da Sequência de Repetição presentes na região 3'NTR do genoma do vírus Chikungunya. Todas as sequências formam estruturas em ansa conservadas e estáveis nas quais os nucleótidos menos conservados em torno da posição 20 constituem a volta. Em todos os genomas de chikungunya estão presentes três RSE. O primeiro (RSE1) está inserido antes da sequência poli-A interna do genoma S27 [9], enquanto os dois outros encontram-se a jusante deste motivo.

**B.** Estrutura secundária prevista para o RSE1 do isolado 05-115.

**Figura 27: Fenótipo de focos do vírus Chikungunya em células AP61 pelo Imunoensaio de Focos**

Células AP61 de mosquito em placas de 24 poços foram infetadas com culturas primárias de vírus CHIK cultivadas em células de mosquito (títulos virais 2-5 × 10<sup>8</sup> FFU. mL-1) a uma multiplicidade de infecção de 0,0001 (poço superior) ou 0,00001 (poço inferior). As células infetadas foram cobertas com CMC em meio de crescimento de Leibovitz L15 com 2% de FBS durante 2 dias para permitir o desenvolvimento de focos a 28 °C. As células foram fixadas com 3% de PFA em PBS, permeabilizadas com Triton X-100 em PBS, e os focos de replicação do vírus CHIK foram imunorrevelados com HMAF anti-CHIK de murganho (diluição 1 :2 000) e Ig anti-murganho de cabra conjugada com peroxidase (diluição 1 :100).

**Figura 28:** Preparação viral contendo pE2. Proteínas pE2 detetadas por anticorpos anti-CHIK.

**Figura 29:** Alinhamento de sequências de nucleótidos que codificam a forma solúvel da glicoproteína E2 (E2-1 a E2-361) das estirpes de vírus CHIK do Oceano Índico -21, -27, -49 e -115.

**Figura 30:** Sequências iniciadoras (SEQ ID NO:79 e 80) utilizadas para a amplificação e clonagem da forma solúvel da E2 (E2-1 a E2-364) (fragmento de ácido nucleico N-terminal e C-terminal: SEQ ID NO:81 e 82; fragmento de proteína N-terminal e C-terminal: SEQ ID NO:83 e 84) do vírus CHIK no vetor transportador pMT2/BiP/V5-HisA.

**Figura 31:** SDS-PAGE que mostra a revelação de CHIK-sE2 por azul de Coomassie.

**Figura 32:** Análise de imunotransferência da proteína sE2 de CHIK altamente purificada.

**Figura 33:** Construção do vetor TRIP que expressa a forma solúvel segregada da glicoproteína E2 (sE2) das estirpes 05 de La Reunião do vírus Chikungunya.

A **Figura 34** mostra a sequência de nucleótidos que codifica a forma solúvel segregada da glicoproteína E2 (sE2) no vetor TRIP (SEQ ID NO:85).

**Figura 35:** ensaio de imunofluorescência (IF) utilizando anticorpos anti-CHIK em células 293 transduzidas com TRIP/CHIK.sE2.

A **Figura 36** mostra um ELISA direto com  $10^{-4}$  mL de proteína pE2 enriquecida por poço. Os抗igénios foram testados respetivamente com um anti-DEN1 (diluição 1:1000), anti-WN (diluição 1:1000) e anti-CHIK (diluição 1:10 000) de murganho.

A **Figura 37** mostra uma sequência de aminoácidos da ORF 2 (proteínas estruturais) da estirpe CHIK S27 (GenBank AF339485; SEQ ID NO: 23).

A **Figura 38** mostra uma sequência de aminoácidos da ORF 2 (proteínas estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 05.61 (SEQ ID NO: 24).

A **Figura 39** mostra uma sequência de aminoácidos da ORF 2 (proteínas estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 05.209 (SEQ ID NO: 25).

A **Figura 40** mostra uma sequência de aminoácidos da ORF 2 (proteínas estruturais) de um vírus CHIK preferido de acordo com uma forma de realização preferida da invenção, nomeadamente a estirpe 05.115 (SEQ ID NO: 26).

A **Figura 41** mostra uma sequência de aminoácidos da ORF 2 (proteínas estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 06.49 (SEQ ID NO: 27).

A **Figura 42** mostra uma sequência de aminoácidos da ORF 2 (proteínas estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 06.27 (SEQ ID NO: 28).

A **Figura 43** mostra uma sequência de aminoácidos da ORF 2 (proteínas estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 06.21 (SEQ ID NO: 29).

A **Figura 44** mostra uma sequência de aminoácidos da ORF 1 (proteínas não estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 05.61 (SEQ ID NO: 30).

A **Figura 45** mostra uma sequência de aminoácidos da ORF 1 (proteínas não estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 05.209 (SEQ ID NO: 31).

A **Figura 46** mostra uma sequência de aminoácidos da ORF 1 (proteínas não estruturais) de um vírus CHIK preferido de acordo com uma forma de realização preferida da invenção, nomeadamente a estirpe 05.115 (SEQ ID NO: 32).

A **Figura 47** mostra uma sequência de aminoácidos da ORF 1 (proteínas não estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 06.49 (SEQ ID NO: 33).

A **Figura 48** mostra uma sequência de aminoácidos da ORF 1 (proteínas não estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 06.27 (SEQ ID NO: 34).

A **Figura 49** mostra uma sequência de aminoácidos da ORF 1 (proteínas não estruturais) de um vírus CHIK preferido de acordo com a presente divulgação, nomeadamente a estirpe 06.21 (SEQ ID NO: 78).

**Figura 50:** Avaliação da reatividade do Mab anti-CHIK E2 por ELISA.

**Figura 51:** Avaliação da reatividade do MAb anti-CHIK E2 em viriões CHIK por ELISA.

**Figura 52:** Análise por imunofluorescência (IF) da reatividade do Mab anti-CHIK E2 em células Vero infetadas por CHIKV.

**Figura 53:** Análise por imunofluorescência (IF) da reatividade do Mab anti-CHIK E2 em células 293A transduzidas com TRIP/CHIK.sE2.

**Figura 54:** Ligação do Mab anti-CHIK E2 na superfície celular de células Vero infetadas por vírus CHIK através de análise por FACS.

**Figura 55:** Análise de transferência de Western da expressão de CHIKsE2 em células 293A transduzidas com TRIP/CHIK.sE2.

**DESCRIÇÃO DETALHADA DA INVENÇÃO**

No presente estudo, os inventores determinaram as sequências nucleotídicas quase completas de vírus isolados a partir de seis doentes originários das Ilhas da Reunião e Seicheles. A presente divulgação permite determinar a estrutura do genoma bem como as características moleculares únicas dos isolados do surto do Oceano Índico, que os distinguem de outras sequências de CHIKV e alfavírus descritas.

Como um técnico na matéria pode entender, a originalidade da presente divulgação é a identificação de novas estirpes do vírus Chikungunya (CHIK) que se distinguem do vírus CHIK da técnica anterior, e a utilização destas estirpes de CHIK e dos polipéptidos e dos polinucleótidos que os codificam derivados do seu genoma no diagnóstico, prevenção e/ou tratamento de arbovirose.

De acordo com um primeiro aspeto, a presente invenção como definida pelas reivindicações 1-13 refere-se a uma estirpe selvagem isolada e purificada do vírus Chikungunya (CHIKV) capaz de infetar células humanas *in vitro*. Preferencialmente, a presente invenção como definida pelas reivindicações 1-13 refere-se a uma estirpe selvagem do vírus CHIK que apresenta as características do isolado 05.115. De acordo com uma forma de realização preferida, a estirpe que está no âmbito da presente invenção como definida pelas reivindicações 1-13 é caracterizada por o

seu genoma compreender pelo menos uma mutação quando comparada com a sequência do genoma da estirpe de vírus CHIK S-27 (GenBank AF339485). Também no âmbito da invenção como definida pelas reivindicações 1-13, encontra-se qualquer estirpe cultivada ou obtida por cultura de células de uma amostra de uma estirpe CHIK preferida da invenção. O genoma da estirpe preferida de acordo com a presente invenção como definida pelas reivindicações 1-13 consiste na sequência como se mostra na Figura 4 (SEQ ID NO: 1). O genoma das estirpes preferidas de acordo com a presente divulgação comprehende uma sequência como se mostra na Figura 5, 6, 7, 8 ou 9 (SEQ ID NO: 2, 3, 4, 5 ou 6).

De acordo com outro aspeto, a presente invenção como definida pelas reivindicações 1-13 proporciona uma estirpe isolada e purificada de vírus Chikungunya (CHIKV) comprehendendo pelo menos uma mutação na proteína estrutural E1 e/ou na proteína estrutural E2, e mais particularmente na sua região do ectodomínio. De acordo com uma forma de realização preferida, a estirpe da invenção como definida pelas reivindicações 1-13 é caracterizada pelo facto de o seu genoma compreender pelo menos uma mutação na proteína E2 numa posição homóloga à posição de aminoácido 382, 399, 404, 485, 489, 506, 536, 624, 637, 669, 700 ou 711 de SEQ ID NO: 23 (Fig. 37). Mais particularmente, a mutação é preferencialmente selecionada do grupo que consiste em G382K, I399M, G404E, N485T, A489T, L506M, I536T, S624N, T637M, A669T, S700T e V711A como se mostra na Tabela 6. De acordo com outra forma de realização preferida, a estirpe

da invenção como definida pelas reivindicações 1-13 é caracterizada pelo facto de o seu genoma compreender pelo menos uma mutação na E1 proteína numa posição homóloga à posição de aminoácido 1035, 1078, 1093 ou 1131 de SEQ ID NO: 23. Mais particularmente, a mutação é preferencialmente selecionada do grupo que consiste em A1035V, M1078V, D1093E e V1131A como se mostra na Tabela 6.

Como aqui utilizada, a expressão "numa posição homóloga a uma posição de aminoácido" de uma proteína, refere-se às posições de aminoácidos que são determinadas como correspondendo uma à outra com base em alinhamentos de sequência e/ou estruturais com uma proteína de referência especificada. Por exemplo, uma posição correspondente a uma posição de aminoácido de uma proteína estrutural do vírus CHIK estabelecida como SEQ ID NO: 1 pode ser determinada empiricamente alinhando as sequências de aminoácidos estabelecidas na SEQ ID NO: 1 com uma proteína estrutural particular do vírus CHIK. As posições homólogas ou correspondentes podem ser determinadas através desse alinhamento por um especialista na técnica utilizando alinhamentos manuais ou utilizando os numerosos programas de alinhamento disponíveis (por exemplo, BLASTP). As posições homólogas ou correspondentes pode ser também com base em alinhamentos estruturais, por exemplo utilizando alinhamentos simulados por computadores da estrutura da proteína. A indicação de que os aminoácidos de um polipéptido correspondem aos aminoácidos numa sequência divulgada refere-se aos aminoácidos identificados após alinhamento do polipéptido com a sequência divulgada para maximizar a identidade ou

homologia (onde os aminoácidos conservados são alinhados) utilizando um algoritmo padrão, tal como o algoritmo GAP. Como aqui utilizado, "numa posição homóloga a" refere-se a uma posição de interesse (*i.e.*, número de base ou número de resíduo) numa molécula de ácido nucleico ou proteína relativamente à posição noutra molécula de ácido nucleico ou proteína de referência. A posição de interesse relativamente à posição noutra proteína de referência pode ser, por exemplo, numa sequência de aminoácidos da mesma proteína de outra estirpe de CHIK. As posições homólogas podem ser determinadas comparando e alinhando as sequências para maximizar o número de nucleótidos ou resíduos condizentes, por exemplo, de tal forma que a identidade entre as sequências seja maior do que 95%, preferencialmente maior do que 96%, mais preferencialmente maior do que 97%, ainda mais preferencialmente maior do que 98% e muito preferencialmente maior do que 99%. À posição de interesse é então dado o número atribuído na molécula de ácido nucleico de referência.

Outro aspeto da invenção como definida pelas reivindicações 1-13 refere-se a um polinucleótido isolado e purificado que consiste na sequência como se mostra na Figura 4 (SEQ ID NO: 1).

Outro aspeto da divulgação refere-se a um polinucleótido isolado e purificado compreendendo a totalidade ou parte da sequência como se mostra nas Figuras 5, 6, 7, 8 ou 9 (SEQ ID NO: 2, 3, 4, 5 ou 6).

Outro aspeto da invenção como definida pelas reivindicações 1-13 refere-se a um fragmento do polinucleótido da invenção caracterizado pelo facto de codificar a região de ectodomínio da glicoproteína E1 ou E2. Vantajosamente, o fragmento da invenção como definida pelas reivindicações 1-13 quando codifica o ectodomínio de E2, comprehende, ou mais preferencialmente, consiste em uma sequência de nucleótidos como se mostra na Figura 13 (SEQ ID NO: 10).

Outro aspeto da divulgação refere-se a um fragmento do polinucleótido caracterizado pelo facto de codificar a glicoproteína E1 ou E2, e mais preferencialmente a sua região de ectodomínio. Vantajosamente, o fragmento da divulgação quando codifica o ectodomínio de E2 comprehende, ou mais preferencialmente, consiste em uma sequência de nucleótidos como se mostra nas Figuras 10, 11 ou 12 (SEQ ID NO: 7, 8 ou 9).

Ainda outro aspeto da invenção como definida pelas reivindicações 1-13 refere-se a um fragmento do polinucleótido da invenção caracterizado pelo facto de codificar uma forma solúvel da glicoproteína E2. De acordo com uma forma de realização preferida, o referido fragmento do polinucleótido da invenção que codifica um fragmento solúvel da glicoproteína E2 como definida pelas reivindicações 1-13, consiste em uma sequência de nucleótidos como se mostra na Figura 17 (SEQ ID NO. 14).

Ainda outro aspeto da divulgação refere-se a um fragmento do polinucleótido caracterizado pelo facto de codificar uma forma solúvel da glicoproteína E2. De acordo com uma forma de realização preferida, o fragmento do polinucleótido que codifica um fragmento solúvel da glicoproteína E2 comprehende ou, mais preferencialmente, consiste em uma sequência de nucleótidos como se mostra nas Figuras 14, 15 ou 16 (SEQ ID NO. 11, 12 ou 13).

Como um especialista na técnica pode entender, um fragmento como considerado pela presente invenção pode ser obtido por:

- utilização de enzimas de restrição em que os seus sítios de dissociação estão presentes no polinucleótido que comprehende o referido fragmento;
- amplificação com iniciadores específicos para o referido fragmento;
- transcrição *in vitro*; ou
- síntese química.

De acordo com outro aspeto, a presente invenção como definida pelas reivindicações 1-13 refere-se a um polipéptido isolado e purificado codificado por um fragmento da invenção. Como aqui utilizado, os termos "polipéptido" e "proteína" são utilizados indistintamente para denotar um polímero de aminoácidos ou um conjunto de dois ou mais polímeros de aminoácidos interatuantes ou ligados.

Por "isolado" entende-se, quando se refere a um polipéptido, que a molécula indicada é separada e discreta do organismo inteiro com o qual a molécula se encontra presente na natureza ou está presente na ausência substancial de outras macromoléculas biológicas do mesmo tipo. O termo "isolado" em relação a um polinucleótido é uma molécula de ácido nucleico desprovida, na totalidade ou parte, de sequências normalmente associadas ao mesmo na natureza; ou uma sequência, como existe na natureza, mas possuindo sequências heterólogas em associação com aquela; ou uma molécula dissociada do cromossoma.

Amplamente definidos, os termos "polipéptido purificado" ou "polinucleótido purificado" referem-se a polipéptidos ou polinucleótidos que estão suficientemente isentos de outras proteínas ou polinucleótidos, ou hidratos de carbono, e lípidos com os quais estão naturalmente associados. O polipéptido ou polinucleótido pode ser purificado através de qualquer processo pelo qual a proteína ou polinucleótido é separado de outro elementos ou compostos com base, por exemplo, na carga, tamanho molecular ou afinidade de ligação.

Os péptidos preferidos da invenção como definidos pelas reivindicações 1-13 compreendem pelo menos uma substituição de aminoácidos em comparação com a sequência de aminoácidos da estirpe S-27 (GenBank AF339485) e são derivados da sequência de uma proteína codificada por um fragmento da invenção como definida pelas reivindicações 1-

13. Preferencialmente, um polipéptido purificado da invenção como definida pelas reivindicações 1-13 consiste na sequência como definida na Figura 40 (sequência de aminoácidos da ORF2 (SEQ ID NO: 26) da estirpe de vírus CHIK 05.115).

Um polipéptido purificado da divulgação compreende a totalidade ou parte da sequência de aminoácidos de uma ORF 1 ou 2 do vírus CHIK considerada pela presente divulgação tal como uma definida em qualquer uma das SEQ ID NOS 24 a 25 e 27 a 29 (ORF 2) ou da SEQ ID NOS 30 a 34 e 78 (ORF 1). Mais preferencialmente, um polipéptido purificado da divulgação compreende a totalidade ou parte da sequência de aminoácidos de uma glicoproteína E2 considerada pela presente divulgação tal como uma definida em qualquer uma das SEQ ID NOS 15 a 18 (Figuras 18 a 21). Ainda mais preferencialmente, um polipéptido purificado da divulgação compreende a totalidade ou parte da sequência de aminoácidos de uma forma solúvel da glicoproteína E2 considerada pela presente divulgação tal como uma definida em qualquer uma das SEQ ID NOS 19 a 22 (Figuras 22 a 25).

A presente invenção como definida pelas reivindicações 1-13 refere-se também a um vetor compreendendo um fragmento de um polinucleótido da invenção como definida pelas reivindicações 1-13. Como aqui utilizado, o termo "vetor" refere-se a uma construção polinucleotídica concebida para transdução/transfeção de um ou mais tipos de

células. Os vetores podem ser, por exemplo, "vetores de clonagem" que são concebidos para o isolamento, propagação e replicação de nucleótidos inseridos, "vetores de expressão" que são concebidos para expressão de uma sequência de nucleótidos numa célula hospedeira, ou um "vetor viral" que é concebido para resultar na produção de um vírus recombinante ou partícula semelhante a vírus, ou "vetores transportadores", os quais compreendem os atributos de mais do que um tipo de vetor. Os vetores preferidos são aqueles depositados na CNCM (Collection Nationale de Cultures de Microorganismes), 28 rue du Docteur Roux, 75724 PARIS Cedex 15, França, em 15 de março de 2006 sob os números de acesso I-3587, I-3588, I-3589 e I-3590. Outro vetor preferido considerado pela presente divulgação é o plasmídeo denominado TRIP-CHIK.sE2 que foi depositado na CNCM (Collection Nationale de Cultures de Microorganismes), 28 rue du Docteur Roux, 75724 PARIS Cedex 15, França, em 14 de março de 2007, sob número de acesso I-3733. Um tal vetor comprehende um fragmento que codifica uma forma solúvel da glicoproteína E2 da invenção. Este vetor preferido foi otimizado para produção eficiente da proteína E2 recombinante em células de mamíferos. Como aqui utilizado, o termo "otimizado" significa que o vetor incorpora sequências de regulação, tal como uma sequência peptídica sinal, para proporcionar expressão adequada da proteína codificada desejada.

Num aspeto relacionado, a presente invenção como definida pelas reivindicações 1-13 proporciona uma célula

hospedeira compreendendo um vetor como definido acima. O termo "célula hospedeira" refere-se a uma célula que tem uma nova combinação de segmentos de ácido nucleico que não estão ligados covalentemente entre si na natureza. Uma nova combinação de segmentos de ácido nucleico pode ser introduzida num organismo utilizando uma grande matriz de técnicas de manipulação de ácido nucleico disponíveis para os especialistas na técnica. Uma célula hospedeira pode ser uma célula eucariótica única, ou uma célula procariótica única ou uma célula de mamífero. A célula hospedeira pode alojar um vetor que é extragenómico. Um vetor de ácido nucleico extragenómico não se insere no genoma da célula. Uma célula hospedeira pode alojar ainda um vetor ou uma porção do mesmo que é intragenómico. O termo intragenómico define uma construção de ácido nucleico incorporada dentro de um genoma da célula hospedeira. Uma célula hospedeira preferida de *E. coli* tal como uma que contém um vetor como definido pelas reivindicações 1-13 encontra-se depositada na CNCM (Collection Nationale de Cultures de Microorganismes), 28 rue du Docteur Roux, 75724 PARIS Cedex 15, França, em 15 de março de 2006 sob os números de acesso I-3587, I-3588, I-3589 e I-3590 e em 14 de março de 2007 sob número de acesso I-3733.

A presente invenção como definida pelas reivindicações 1-13 refere-se ainda à utilização de um anticorpo monoclonal ou anticorpos policlonais, ou seus fragmentos, que se ligam especificamente a um polipeptídeo da invenção para a deteção *in vitro* de um CHIKV associado a uma

arbovirose como definida pelas reivindicações 1-13. Como aqui utilizado, o termo "liga-se especificamente a" refere-se a anticorpos que se ligam com uma afinidade relativamente alta a um ou mais epítopos de uma proteína da invenção, mas que não reconhecem substancialmente nem se ligam a moléculas diferente(s) daquela(s) de interesse. Como aqui utilizado, o termo "afinidade relativamente alta" significa uma afinidade de ligação entre o antícorpo e a proteína de interesse de pelo menos  $10^{-6}$  M, e preferencialmente de pelo menos cerca de  $10^{-7}$  M e ainda mais preferencialmente  $10^{-8}$  M a  $10^{-10}$  M. A determinação dessa afinidade é preferencialmente realizada sob condições de imunoensaio de ligação competitiva convencional que é do conhecimento comum para um especialista na técnica.

Como aqui utilizado, o termo "antícorpo" refere-se a uma glicoproteína produzida por células linfoides em resposta a uma estimulação com um imunogénio. Os anticorpos possuem a aptidão para reagir *in vitro* e *in vivo* específica e seletivamente com um determinante antigénico ou epítopo desencadeando a sua produção ou com um determinante antigénico estreitamente relacionado com o antigénio homólogo. O termo "antícorpo" entende-se como abrangendo construções que utilizam a região (variável) de ligação de um tal antícorpo, e outras modificações de antícorpo. Assim, um antícorpo útil na utilização da invenção pode compreender um antícorpo inteiro, um fragmento de antícorpo, um agregado de anticorpos polifuncional, ou em geral uma substância compreendendo um ou mais sítios de ligação

específicos de um anticorpo. O fragmento de anticorpo pode ser um fragmento tal como um fragmento Fv, Fab ou F(ab')<sub>2</sub> ou um seu derivado, tal como um fragmento Fv de cadeia simples. O anticorpo ou fragmento de anticorpo pode ser não recombinante, recombinante ou humanizado. O anticorpo pode ser de um isotipo de imunoglobulina, e.g., IgG, IgM, e assim por diante. Além disso, pode utilizar-se um agregado, polímero, derivado e conjugado de uma imunoglobulina ou um seu fragmento, quando apropriado.

Outro aspecto da invenção como definida pelas reivindicações 1-13 é a utilização de um elemento selecionado do grupo que consiste em uma estirpe, um polinucleótido, um fragmento, um vetor, uma célula hospedeira, um polipeptídeo da invenção para a deteção de um CHIKV associado a uma arbovirose, ou para a preparação de uma composição que previne e/ou trata uma arbovirose.

Outro aspecto da invenção como definida pelas reivindicações 1-13 é um anticorpo monoclonal ou policlonal ou seu fragmento que se liga especificamente a um polipeptídeo da invenção como definida pelas reivindicações 1-13 para a deteção de um CHIKV associado a uma arbovirose.

Outro aspecto da presente invenção como definida pelas reivindicações 1-13 refere-se a uma composição para tratar e/ou prevenir uma arbovirose. A composição da presente invenção como definida pelas reivindicações 1-13 comprehende vantajosamente pelo menos um elemento seleciona-

do do grupo que consiste em uma estirpe, um polinucleótido, um fragmento, um vetor, uma célula hospedeira e um polipéptido da invenção como definida pelas reivindicações 1-13. A composição da invenção como definida pelas reivindicações 1-13 pode compreender ainda um transportador aceitável. Num aspeto relacionado, a divulgação proporciona um método de tratamento e/ou prevenção de uma arbovirose. O método comprehende o passo de administração de uma composição da invenção a um indivíduo necessitado da mesma.

Como aqui utilizado, o termo "tratar" refere-se a um processo pelo qual o desenvolvimento de uma infecção por um CHIKV é afetado ou completamente eliminado. Como aqui utilizado, o termo "prevenção" refere-se a um processo pelo qual a infecção por CHIKV é obstruída ou retardada.

Como aqui utilizada, a expressão "um transportador aceitável" significa um veículo para conter os componentes (ou elementos) da composição da invenção como definida pelas reivindicações 1-13 que pode ser administrado a um animal hospedeiro sem efeitos adversos. Os transportadores adequados conhecidos na técnica incluem, mas não se limitam a, partículas de ouro, água estéril, soro fisiológico, glucose, dextrose ou soluções tamponadas. Os transportadores podem incluir agentes auxiliares incluindo, mas não se limitando a, diluentes, estabilizantes (*i.e.*, açúcares e aminoácidos), conservantes, humectantes, emulsionantes, agentes tampão de pH, aditivos de melhoria da viscosidade, cores e semelhantes.

A quantidade de componentes da composição da invenção como definida pelas reivindicações 1-13 é preferencialmente uma quantidade terapeuticamente eficaz. Uma quantidade terapeuticamente eficaz dos componentes da composição é a quantidade necessária para permitir que a mesma realize o seu papel de prevenção e/ou tratamento contra uma infecção por CHIKV sem originar efeitos excessivamente negativos no hospedeiro ao qual a composição é administrada. A quantidade exata de componentes a ser utilizada e a composição a ser administrada variará de acordo com fatores tais como o modo de administração, bem como os outros ingredientes na composição.

A composição da invenção como definida pelas reivindicações 1-13 pode ser administrada a um hospedeiro (tal como um humano) através de várias vias de administração. Por exemplo, a composição pode ser administrada na forma de preparações injetáveis estéreis, tais como suspensões aquosas ou oleaginosas injetáveis estéreis. Estas suspensões podem ser formuladas de acordo com técnicas conhecidas na matéria utilizando dispersantes ou humectantes e agentes de suspensão adequados. As preparações injetáveis estéreis podem ser também soluções ou suspensões injetáveis estéreis em diluentes ou solventes parentericamente aceitáveis, não tóxicos. Estas podem ser administradas por via parentérica, por exemplo por via intravenosa, intramuscular ou subcutânea por injeção, por infusão ou *per os*. As dosagens adequadas variarão, dependendo de

fatores tais como a quantidade de cada um dos componentes na composição, o efeito desejado (curto ou longo prazo), a via de administração, a idade e o peso do hospedeiro a ser tratado. Quaisquer outros métodos bem conhecidos na técnica podem ser utilizados para administrar a composição da invenção.

Ainda outro aspeto da invenção como definida pelas reivindicações 1-13 é a utilização de uma composição como definida como definida pelas reivindicações 1-13 para a preparação de um medicamento para tratar e/ou prevenir uma arbovirose num indivíduo necessitado do mesmo.

Ainda outro aspeto da invenção como definida pelas reivindicações 1-13 consiste em proporcionar um kit para a deteção de um CHIKV associado a uma arbovirose, compreendendo pelo menos um elemento selecionado do grupo que consiste em uma estirpe, um polinucleótido, um fragmento, um vetor, uma célula hospedeira, um polipéptido da invenção como definida pelas reivindicações 1-13 e anticorpos monoclonais ou policlonais ou seu fragmento que se ligam especificamente a um polipéptido da invenção como definida pelas reivindicações 1-13. Os kits de acordo com esta forma de realização da invenção podem compreender embalagens, contendo, cada, um ou mais dos elementos supramencionados (tipicamente na forma concentrada) que são necessários para realizar os respetivos testes diagnósticos.

**EXEMPLOS**

Os exemplos aqui a seguir destacarão outras características e vantagens da presente invenção como definida pelas reivindicações 1-13, e servirão para ilustrar o âmbito de utilização da presente invenção e não para limitar o seu âmbito. Pode utilizar-se outros métodos ou produtos equivalente àqueles que se encontram a seguir para testar ou realizar a presente invenção como definida pelas reivindicações 1-13; são descritos os materiais e métodos preferidos.

**EXEMPLO 1:** Identificação e caracterização dos vírus CHIK que originam o surto do Oceano Índico

Os inventores (como por vezes aqui referidos por "nós") descrevem a sequência quase completa do genoma de seis isolados clínicos selecionados, juntamente com as sequências parciais da glicoproteína E1 de um total de 60 doentes das Ilhas Reunião, Seicheles, Maurícia, Madagáscar e Maiote. Os atuais resultados indicam que o surto foi iniciado por uma estirpe relacionada com isolados da África Oriental, a partir dos quais se desenvolveram variantes virais após uma história de microevolução rastreável. Foram identificadas características moleculares únicas dos isolados do surto. De maneira assinalável, na região que codifica as proteínas não estruturais foram encontradas dez alterações de aminoácidos, estando três delas localizadas em posições conservadas em alfavírus da nsP2 (a qual contém

atividades helicase, protéase e ARN-trifosfatase) e da polimerase nsP4. O único isolado obtido a partir do líquido cefalorraquidiano de um doente mostrou alterações únicas nas nsP1 (T301I), nsP2 (Y642N) e nsP3 (supressão de E460). Na região da proteína estrutural foram observadas duas alterações dignas de nota (A226V e D284E) na glicoproteína de fusão da membrana E1. A modelação 3D por homologia permitiu o mapeamento destas duas alterações em regiões que são importantes para a formação do virião e para fusão da membrana. A alteração E1-A226V estava ausente nas estirpes iniciais, mas foi observada em >85% das sequências virais posteriores da Reunião, denotando um sucesso evolutivo possivelmente devido à adaptação ao vetor mosquito.

**MATERIAL E MÉTODOS**

**Doentes.** Os 60 doentes para os quais foram determinadas as sequências nucleotídicas parciais ou completas do CHIKV eram originários da Reunião (N=43), Seicheles (N=3), Madagáscar (N=7), Maiote (N=4) e Maurícia (N=3). As características dos doentes e amostras biológicas são listados na Tabela 1.

**Isolamento do vírus e extração do ARN.** Os vírus foram isolados de soro ou líquido cefalorraquidiano (CSF) (Tabela 1). Resumidamente, células de *Aedes albopictus* C6-36 foram inoculadas com 1 mL de soro ou CSF diluído a 1:10 em meio L15 (Gibco). As células foram cultivadas a 28 °C em L15 suplementado com 5% de soro fetal bovino e 10% de

fosfato de triptose. As células e sobrenadantes foram colhidos após a primeira passagem (5 dias) e a segunda passagem (7 dias). Os vírus isolados foram identificados como CHIKV por imunofluorescência indireta, utilizando líquido ascítico hiperimune de CHIKV. No caso dos isolados 05.115, 06.21, 06.27 e 06.49 cujos genomas foram sequenciados, foi confirmada a ausência de febre-amarela, dengue e vírus do Nilo Ocidental por imunofluorescência indireta utilizando soros específicos. O ARN foi extraído utilizando o Minikit Viral QIAAmp (Qiagen, França).

**Sequenciação de nucleótidos.** Os iniciadores (Tabela 4) foram concebidos com base na sequência de nucleótidos 20 da estirpe S27. A RT-PCR foi realizada utilizando o kit Titan One Tube RT-PCR (Roche, França). Os fragmentos de RT-PCR foram purificados por ultrafiltração antes da sequenciação (Millipore, França). As reações de sequenciação foram realizadas utilizando o kit de sequenciação BigDye Terminator v1.1 cycle (Applied Biosystems, EUA) e purificadas por precipitação com etanol. Os cromatogramas das sequências foram obtidos nos analisadores de sequência automáticos ABI3100 ou ABI3700 (Applied Biosystems). Todos os produtos de amplificação foram sequenciados em ambas as cadeias.

**Montagem das sequências do genoma e análise das sequências.** A montagem de sequências contíguas foi realizada independentemente por operadores e softwares distintos, utilizando BioNumerics versão 4.5 (Applied-Maths, Sint-

Martens-Latem, Bélgica) ou PhredPhrap/Consed [11]. Ambas as análises produziram exatamente a mesma sequência consenso para todas as estirpes. Foi obtida uma única sequência contígua de 11 601 nt para cinco isolados, enquanto para a estirpe 05.61, faltava uma porção da sequência, entre as posições 5 246 a 5 649 da S27 (posições 390 a 524 da nsP3). Os alinhamentos de sequências e o cálculo das tabelas de substituição foram realizados utilizando os programas BioNumerics, DNASP versão 4.10 [12] e DAMBE versão 4.2.13 [13]. Os alinhamentos das sequências de nucleótidos e aminoácidos contra sequências de alfavírus selecionados foram realizados com o software ClustalW1.7 [14]. As identidades de sequência foram calculadas com o pacote Phylip [15]. A estrutura secundária do ARN foi prevista com o servidor de estruturas secundárias de ARN Vienna [16]. As árvores de junção de vizinhos foram construídas utilizando MEGA versão 3.1 [17] com as correções de parâmetros Kimura-2 de substituições múltiplas. A fiabilidade dos nós foi avaliada por reamostragem de iniciação com 1 000 replicados. As quantidades de substituições sinónimas por sítio sinónimo (Ks) e de substituições não sinónimas por sítio não sinónimo (Ka) foram estimadas utilizando DNASP. A RDP2 [18] foi utilizada para detetar sequências mosaico putativas.

**Modelação da estrutura 3D.** A estrutura cristalográfica do ectodomínio da glicoproteína E1 do Vírus da Floresta de Semliki (SFV) a pH neutro [19]; o código 2ALA) do Banco de Dados de Proteínas foi utilizado como uma

cadeia molde para modelar e analisar as duas mutações de aminoácidos dos isolados do Oceano Índico. A Figura 2 foi preparada utilizando o programa RIBBONS [20].

**Deteção dos focos viral por revelação imunológica.** Células *Aedes pseudoscutellaris* AP61 foram cultivadas em placas de cultura de tecidos de 24 poços em meio de crescimento Leibovitz L-15 com 10% de soro fetal de vitelo inativado termicamente (FCS) durante 24 h. As monocamadas de células de mosquito foram lavadas uma vez com Leibovitz L-15 e foram adicionados 0,2 mL de Leibovitz L-15/2% de FCS. As células foram infetadas com vírus CHIK em 0,2 mL de Leibovitz L-15/2% de FCS e incubadas a 28 °C durante 1 h. Foi então adicionado meio de cobertura consistindo em 0,4 mL de Leibovitz L-15/2% de FBS e carboximetilcelulose (CMC) (1,6%) e as placas de cultura de tecidos foram incubadas a 28 °C durante 2 dias. Os focos de células infetadas foram visualizados pelo Imunoensaio de Focos (FIA). As células foram lavadas com PBS, fixadas com 3% de paraformaldeído (PFA) em PBS durante 20 min, e permeabilizadas com 0,5% de Triton X-100 em PBS durante 4 min à temperatura ambiente. As células fixas foram incubadas durante 20 min a 37 °C com diluição a 1:2 000 de líquido ascítico hiperimune de murganho (HMAF) dirigido contra CHIKV. A IgG de cabra anti-murganho, conjugada com peroxidase de rábano-silvestre foi utilizada como o segundo anticorpo (diluição a 1:100) a 37 °C durante 20 min. Os focos foram visualizados com Substrato de Peroxidase DAB (Sigma).

**1. Estrutura do genoma e assinaturas moleculares  
dos vírus Chikungunya do surto do Oceano Índico**

*Organização do genoma.* Nós determinámos as sequências quase completas dos genomas de seis isolados de CHIKV (05.115, 05.61, 05.209, 06.21, 06.27 e 06.49) que representam origens geográficas, pontos no tempo e formas clínicas distintas (Tabela 1) do surto do Oceano Índico do vírus Chikungunya. Foram determinados 11 601 nucleótidos, correspondentes às posições 52 (5'NTR) a 11 667 (3'NTR, extremidade do terceiro Elemento de Sequência de Repetição) na sequência de nucleótidos do isolado Tanzaniano de 1952 S27 (comprimento total 11 826 nt). Existiam três eventos de inserção/supressão entre os isolados de S27 e Reunião, dois dos quais foram observados na 3'NTR. Primeiro, o segmento poli-A interno de 14 nucleótidos observado na S27 (11 440-11 443) e que corresponde a um possível sítio poli-A interno [9] foi substituído por um segmento de apenas 5 A nos isolados do Oceano Índico, semelhante ao que foi observado noutro vírus Chikungunya, e.g. a estirpe Ross (n.º de acesso: AF490259). Segundo, faltava um A nos isolados do Oceano Índico num segmento de 5 A na posição 11 625 da S27. Finalmente, faltava um codão no isolado 06.27, correspondente ao codão 460 da nsP3, onde todos os outros isolados do Oceano Índico analisados e sequências alfavírus disponíveis são GAA, que codifica uma Glu.

As sequências dos genomas dos seis isolados ali

apresentados foram semelhantes às anteriormente descritas para o alfavírus [9, 21, 22]. As sequências codificantes consistiam em duas grelhas de leitura aberta (ORF) grandes de 7 422 nt e 3 744 nt que codificam a poliproteína não estrutural (2 474 aminoácidos) e a poliproteína estrutural (1 248 aminoácidos), respectivamente. A poliproteína não estrutural é o precursor das proteínas nsP1 (535 aa), nsP2 (798 aa), nsP3 (530 aa) e nsP4 (611 aa), e a poliproteína estrutural é o precursor das proteínas C (261 aa), p62 (487 aa, precursor para E3 - 64 aa - e E2 - 423 aa), 6K (61 aa), e E1 (439 aa). Os sítios de dissociação característicos da família de alfavírus nas poliproteínas não estruturais e estruturais foram conservados. Os sítios de glicosilação em E3, E2 e E1 foram também conservados. Foi identificada uma sequência junção de 65 nt entre o codão de terminação (TAG, 7499-7501) da ORF não estrutural e o codão de iniciação (7567-7569) da ORF estrutural. A região não traduzida 5' (5'NTR) terminava na posição 76. A região 3'NTR começava na posição 11 314 e continha três elementos da sequência de repetição (RSE) com estruturas secundárias previstas (Fig. 26) que eram coerentes com um trabalho anterior [9].

*Diferenças entre os isolados do surto do Oceano Índico e a estirpe S27.* Em comparação com a estirpe S27, o isolado 05.115 da Reunião apresentou 28 alterações de aa (1,13%) nas proteínas não estruturais (Tabela 5, com a proporção mais alta na nsP3 (2,26%) e a mais baixa na nsP2 (0,6%). Dez das 12 alterações de aminoácidos na nsP3 estavam concentradas entre as posições 326 e 524 (5,0% de

variação), semelhante às observações no vírus ONN [23]. Uma diferença importante com a S27 foi que os isolados do Oceano Índico apresentavam um codão de terminação opal (UGA) no codão 524 da nsP3, em vez da Arg (CGA) em S27. Este codão opal foi observado em alfavírus relacionados [9, 22, 23] e julga-se que regule a expressão de nsP4, a ARN-polimerase putativa, por um mecanismo de leitura [21, 24].

Em comparação com a S27, as proteínas estruturais apresentavam 21 (1,68%, para 05.115) a 22 (1,76%, para outros isolados) substituições de aminoácidos em isolados do Oceano Índico (Tabela 6). De forma assinalável, a proteína E2 do envelope apresentava a variação mais alta, com 14 (3,3%) alterações de aa, superior à proteína E1 do envelope (0,68%) e à proteína do capsídeo (0,38%). A relação das taxas de evolução de sítios sinónimos e não sinónimos (Ks/Ka) entre a S27 e os isolados 05.115 foi de 11,0 para a poliproteína inteira, enquanto foi apenas de 6,12 para a proteína E2, provavelmente indicativo de uma seleção positiva em favor das alterações de aminoácidos nesta proteína imunogénica. Por comparação, a Ks/Ka foi de 18,75 para a poliproteína não estrutural.

*Assinaturas moleculares do surto do Oceano Índico em proteínas não estruturais e variação fenotípica.* Dez posições (excluindo as posições polimórficas) tinham aa que eram únicos para as proteínas não estruturais dos isolados do surto, quando comparadas com outras sequências de CHIKV (Tabela 2). Primeiro, a nsP2-54 era Asn em isolados do

Oceano Índico e no SFV, mas era Ser em todas as outras sequências. Segundo, a nsP2-374 era Tyr em isolados do Oceano Índico, mas era His ou Asn noutras sequências de alfavírus (Tabela 2). Terceiro, a posição 500 na nsP4 era Leu nas sequências do Oceano Índico em vez de Gln nas outras quatro sequências de CHIKV descritas. Curiosamente, esta posição, o qual está a cerca de 30 aa do motivo catalítico "GDD", é um Glu rigorosamente conservado em todos os outros alfavírus. As restantes sete alterações ocorreram em regiões relativamente variáveis.

Foram observadas alterações específicas adicionais nos isolados 05.209 (S358P) e 06.27 (nsP1-T301I, nsP2-Y642N e nsP3-460del). De forma assinalável, os nossos ensaios fenotípicos realizados em paralelo mostraram diferenças para a estirpe 06.27. O imunoensaio de Focos mostrou que as culturas-mãe de 05.115, 06.21, 06.27 e 06.49 do CHIKV formavam misturas de focos com diferentes tamanhos em células de *Ae. Albopictus* C6/36 (dados não apresentados) e *Ae. pseudoscutellaris* AP61 (Fig. 27). Curiosamente, apenas o isolado 06-27 formou focos médios, enquanto os outros formaram focos minúsculos e pequenos. O fenótipo particular da 06-27 podia ser relacionado com as diferenças de aa observadas nas proteínas não estruturais, as quais estão envolvidas na replicação viral [21].

*Assinaturas moleculares do Oceano Índico em proteínas estruturais e modelação 3D.* Quando se analisa as sequências de aa das proteínas estruturais, verificou-se

que sete posições (quatro na E2, uma na 6K e duas na E1) são únicas para os isolados do surto do Oceano Índico (Tabela 2). Duas destas estavam localizadas no ectodomínio de E2, com a Thr 164 e a Met 312 a serem identificadas nos nossos isolados em vez das Ala e Thr, respetivamente, em todas as outras sequências de CHIKV disponíveis (Tabela 2). A primeira destas duas posições é variável nos alfavírus; ela situa-se numa região anteriormente definida como contendo epítópos neutralizantes [5, 25]. Na posição 312, a Thr está presente noutras CHIKV, no ONNV e no SFV, mas varia noutras alfavírus; situa-se numa região identificada como importante para a oligomerização E1-E2 [5, 25].

Na E1, foram observadas duas substituições cruciais, uma no resíduo 284, específica para isolados do Oceano Índico, e uma no resíduo 284, presente em 3 dos 6 isolados do Índico (06.21, 06.27 e 06.49). Ambas as mutações foram mapeadas na estrutura 3D (modelada a partir da estrutura cristalina da E1 do SFV) na Fig. 1. Curiosamente, o resíduo 226 é Ala em todas as sequências de CHIKV descritas (Tabela 2) e era também Ala no primeiro dos nossos isolados do Oceano Índico aqui sequenciados (05.61 e 05.115, obtidos no início do surto). Todos os isolados posteriores (obtidos de doentes recolhidos em novembro e dezembro de 2005) apresentaram um resíduo Val nesta posição. Embora a posição 226 seja relativamente variável entre alfavírus, observou-se que uma única mutação nesta posição (Pro para Ser) permitia o SFV adaptar-se ao

crescimento em células de insetos deprimidas de colesterol [26, 27].

O outro aa único observado na E1 de isolados do Oceano Índico era o Glu 284. Esta é uma posição altamente conservada na E1, a qual apresenta uma Asp na maioria dos alfavírus ou uma Asn no SIN (Tabela 2). Este aminoácido está localizado na interface entre os promotores de E1 na superfície do virião, participando em contactos que constituem a armação icosaédrica de E1 (Fig. 1).

## **2. Análise filogenética**

Trabalho anterior com base em sequências da proteína E1 mostrou uma estrutura filogeográfica forte das espécies do vírus Chikungunya [6, 10]. A fim de determinar o grupo filogenético progenitor a partir do qual surgiram os isolados do surto do Oceano Índico, nós comparámos uma região de 1 044 nt dentro da sequência codificante de E1 (posições 271 a 1314, i.e., codões 91 a 438) de 63 espécimes biológicos de 60 doentes das Reunião, Seicheles, Madagáscar, Maiote e Comores (Tabela 1) com outras 29 sequências de chikungunya disponíveis (Tabela 7). A análise filogenética (Fig. 2) demonstrou claramente que os atuais isolados do Oceano Índico representam um clado homogéneo dentro de um grupo lato (grupo ECSA) que compreende isolados da África Oriental, Central e Austral (ECSA, Fig. 2). Os isolados de um surto na República Democrática do Congo [6] formaram também um clado homogéneo dentro do

grupo ECSA. Não houve nenhum membro do grupo ECSA que apresentasse uma relação significativamente mais próxima com os isolados do Oceano Índico. Os isolados Asiáticos estavam menos relacionados com os isolados do Oceano Índico e constituíam o grupo irmão de grupo ECSA, enquanto os isolados da África Ocidental eram ainda mais divergentes. A inclusão de outros alfavírus, incluindo o parente mais próximo ONN, colocou a raiz dos isolados de chikungunya no ramo que conduz ao grupo filogenético da África Ocidental (dados não apresentados).

A comparação das sequências dos isolados do surto do Oceano Índico com a sequência de S27 revelou 316 (2,7%) substituições de nucleótidos no isolado 05.115 (Tabela 8). A estirpe Nagpur do clado Asiático mostrou 5,1% de divergência média dos nucleótidos com a 05.115, enquanto a estirpe Senegal 37997 do clado da África Ocidental apresentou 15% de diferença (Tabela 8). Curiosamente, a última estirpe apresentou conservação total de uma porção de 87 nucleótidos (9 958 - 10 045, na junção entre as proteínas estruturais 6K e E1) com os isolados do surto da África Oriental e do Oceano Índico. A identidade de sequência nesta porção pode espelhar um evento passado de recombinação genética entre as estirpes da África Ocidental e da África Oriental/Central. De forma distinta, nós não encontrámos suporte estatístico ( $P > 7E-2$ ) para mosaicismo de sequência ou recombinação devido à divisão entre a S27 e os isolados da Reunião, embora algumas regiões genómicas difiram na sua densidade de polimorfismos de nucleótidos.

### **3. Variação genotípica e fenotípica entre isolados do surto do Oceano Índico e cenário microevolutivo**

Foram observadas alterações de aa específicos nas proteínas não estruturais nos isolados 05.209 (S358P) e 06.27 (nsP1-T301I, nsP2-Y642N e nsP3-460del). Nas proteínas estruturais, a alteração E1-A226V foi observada nos isolados 06.21, 06.27 e 06.49, e a alteração E2-Q146R no isolado 05.209 das Seicheles. Além destas alterações não sinónimas, existiam 8 substituições silenciosas, observadas em 05.209, 06.27 e 06.49 (Tabela 3).

Uma história de provável evolução da sequência que ocorreu durante o surto (Fig. 3) foi deduzida das 14 variações de aminoácidos observadas entre os seis genomas completos (Tabela 3). O isolado 05.61 foi inicialmente selecionado para análise do genoma porque tinha sido isolado em março de 2005, no início do surto, de um doente da Reunião que regressava das Ilhas Comores, onde o surto estava em curso desde janeiro de 2005. Notavelmente, os isolados 05.61 e 05.115 (que foi o segundo isolado mais antigo analisado), o isolado S27 Africano e isolados de chikungunya anteriores não relacionados de África e Ásia foram idênticos em todos os 14 sítios polimórficos. Por conseguinte, a sequência consenso dos isolados 05.61 e 05.115 (sequência consenso 1) representa provavelmente o genótipo ancestral do surto da Reunião. A distribuição dos 14 polimorfismos sugeriu que este fundador deu origem a

três sequências consenso que provavelmente evoluíram em quatro passos. Primeiro, a substituição na posição 10 670 do genoma (que originou a alteração A226V da E1) deu origem à sequência consenso 2, representada pelo isolado 06.21 do final de novembro de 2005. Segundo, uma substituição sinónima de G para A na posição 6 547 (nsP4) levou a uma sequência intermediária, que deu ela mesma origem a duas sequências posteriores: sequência consenso 3 (isolado 06.27), após quatro substituições adicionais e supressão de um codão (Tabela 3), e sequência consenso 4 (06.49), a qual surgiu após três substituições sinónimas distintas (Tabela 3). Uma quinta sequência consenso foi representada pelo isolado 05.209 das Seicheles sozinho, o qual apresentava quatro substituições (duas delas originando alterações de aa na nsP3 - S358P e na E2 - Q146R) em comparação com a sequência consenso 1 (Fig. 3).

Uma vez que os isolados da Reunião tinham E1-226A no início do surto e E1-266V A no início do surto e E1-266V mais tarde durante a epidemia, nós comparámos o resíduo 226 em 57 sequências adicionais (57 sequências a partir de 54 soros e 3 CSF) da epidemia do Oceano Índico. Notavelmente, a natureza da E1-226 diferia completamente na Ilha da Reunião antes e após a estação de inverno. Cinco sequências de doentes amostrados entre março a junho de 2005 (incluindo a sequência originária de um viajante que regressava das Comores) tinham E1-226A. Entre setembro e finais de dezembro de 2005, 21 sequências apresentavam E1-226V. Entre as 17 sequências da Reunião de 2006, a E1-226V

foi observada 12 vezes e a E1-226A 5 vezes (Tabela 1). Nas sequências de Madagáscar e Seicheles, para as quais foram recolhidas amostras quando se suspeitou dos primeiros casos clínicos (i.e provavelmente no início dos surtos), apenas foi observada a E1-226 Ala. Nas sequências de 2006 da Maiote, apenas foi observada a E1-226 V. Nas sequências de 2006 da Maurícia foram observadas tanto a E1-226 Ala como a Val.

Até à data tinha sido completamente sequenciadas apenas estirpes laboratoriais de CHIKV, passadas muitas vezes em células de mosquitos ou mamíferos, [9]. Nós proporcionámos pela primeira vez sequências nucleotídicas quase completas de seis isolados clínicos passados *in-vitro* apenas uma vez ou duas vezes (ver secção M&M). A presença em doentes infetados de uma população viral mista, denominada quasiespécie [31-33], com genótipos coexistentes num equilíbrio governado por um balanço entre mutação e seleção natural. A presença na S27 de um codão de Arg em vez do codão de terminação opal nos isolados do Oceano Índico é provavelmente explicada pelas numerosas passagens *in-vitro* da S27, já que foi experimentalmente observada uma evolução de opal para Arg nos vírus ONN [23]. Embora possa ser vantajoso para as quasiespécies virais manter o codão opal *in-vivo*, um codão de Arg confere, provavelmente, uma vantagem seletiva *in-vitro*, como se observa para o Vírus da Floresta de Semliki estreitamente relacionado [34]. A situação das quasiespécies do vírus Chikungunya *in-vivo* pode explicar também o polimorfismo nsP1-T301I observado

para o isolado 06.27 de LCR. De facto, é provável que a seleção de um subconjunto de genótipos que tenha esta alteração possa estar associada à invasão de LCR [33]. Estes resultados sublinham que a sequência do genoma de estirpes laboratoriais de "referência" pode não espelhar com exatidão a situação natural, já que a complexidade genotípica das quasiespécies *in-vivo* é sujeita a erosão pela seleção *in-vitro*. Uma vez que os isolados do Oceano Índico aqui sequenciados foram submetidos a seleção *in-vitro* durante apenas algumas gerações, eles correspondem provavelmente mais de perto aos genótipos *in-vivo* do que as estirpes de chikungunya anteriormente sequenciadas.

As diferenças de aminoácidos (aa) detetadas entre os isolados do surto 1 podem estar relacionadas com características biológicas ou patogénicas do vírus. Embora os nossos resultados da cultura viral sejam preliminares, eles mostram claramente diferenças fenotípicas entre o isolado único de CSF (06.27), o isolado a partir de um caso de encefalopatia neonatal, e três outros isolados, associados à forma clássica da doença ou a encefalopatia. Os focos maiores observados na cultura com a 06.27 poderiam espelhar uma taxa de replicação mais alta do vírus e estar ligados a alterações específicas de aminoácidos identificadas nas nsP1, nsP2 e nsP3. Foi anteriormente demonstrado que as alterações de aminoácidos únicos na nsP1, incluindo uma alteração Thr/Ile (resíduo 538 do vírus Sindbis) [35,36] e uma supressão de 18 nt na nsP3, afetam a neurovirulência noutros alfavírus [35-37]. No entanto, na

ausência de dados estruturais da nsP1, é difícil prever o impacto estrutural ou funcional da alteração I301T observada no isolado 06.27. Deve referir-se também que todas as sequências virais determinadas a partir do soro ou dos isolados de três casos de encefalopatia neonatal e um caso de meningoencefalite adulta tinham E1-226 Val. No entanto, uma vez que este genótipo é observado também nas formas clássicas da doença, uma ligação potencial de E1-226 Val com a neuropatogénese necessita de estudos adicionais. Fatores do hospedeiro têm de ser considerados na ocorrência das formas neurológicas da doença. Por exemplo, a travessia hematoencefálica pode ser facilitada por idade jovem ou hipertensão.

As assinaturas moleculares únicas dos genomas do surto do Oceano Índico foram identificadas quando foram comparadas com todas as outras sequências de alfavírus descritas. Estas características representam alvos interessantes para futuros estudos funcionais, bem como para seguimento epidemiológico. Uma característica particularmente interessante foi o resíduo E1-226 Val (ver acima). Outra assinatura molecular interessante dos genomas do surto do Oceano Índico foi o E1-284 Asp. Embora o modelo pseudo-atómico da armação utilizado seja de resolução modesta (a resolução da estrutura cristalina é limitada - cerca de 3Å - e o modelo resulta do ajuste desta estrutura numa reconstrução de microscopia crio-eletrónica com resolução de 9Å), parece que a cadeia lateral do Asp 284 interage com a cadeia principal de um polipéptido de E1 adjacente no

virião. De facto, encontra-se numa posição compatível com a aceitação de uma ligação de hidrogénio da amida 379 da cadeia principal do promotor de E1 vizinho. Uma vez que o empacotamento é muito apertado (ver Fig. 1 B), é possível que a cadeia lateral mais comprida do ácido glutâmico (que tem um grupo CH<sub>2</sub> extra em comparação com Asp ou Asn) possa introduzir uma ligeira distorção nos sítios de contacto, um efeito que se propaga pela simetria T=4 icosaédrica do virião. Assim, um efeito cooperativo devido a esta alteração na posição Asp 284 pode desempenhar um papel quer permitindo uma formação menos eficiente de novas partículas nas células infetadas, ou um processo de desagregação mais eficiente da partícula durante a invasão de uma nova célula, ou uma combinação de ambos. Esta informação pode orientar novos estudos de mutagénese específica de um lócus, utilizando genética inversa, para testar o efeito da substituição Asp/Glu no ciclo do vírus.

**Exemplo 2: Identificação e caracterização de uma forma solúvel de E2 (sE2) do vírus CHIK.**

O plasmídeo TOPO/CHIK-21.pE2 (CNCM I-3587) contendo o ADNc que codifica a glicoproteína pE2 (E3+E2) da estirpe de vírus CHIK 21 (Schuffenecker *et al.*, *Plos Med.*, 3:1058, 2006) foi utilizado como uma cadeia molde para amplificação por PCR da sequência do ectodomínio da glicoproteína de envelope E2 (Fig. 29). O ectodomínio de gp-E2 (E2-1 a E2-364; 85% de E2) é rigorosamente conservado entre as linhas de células CHIK-21, -27, 49 e 115 isoladas

no Oceano Índico durante o surto epidémico de 2005-06 (ver Fig. 29). A forma solúvel do sE2 corresponde ao ectodomínio de gp-E2 que está suprimido na extremidade carboxílico da sua região âncora transmembranar. É interessante notar que a forma solúvel tem os epítópos principais que desencadeiam anticorpos neutralizantes do vírus. Os iniciadores de PCR são descritos na Figura 30 (SEQ ID NO:79 e 80): eles permitem, por um lado, a clonagem da sequência sE2 entre os sítios únicos *BglII* e *NotI* do vetor pMT/BiP/V5-HisA (Invitrogen), numa fase dependente do péptido sinal BiP na extremidade N-terminal, e por outro lado a junção sucessiva de marcadores V5 (His)6 na sua extremidade carboxílica.

Células de drosophila S2 foram transfetadas com o plasmídeo pMT/BiP/CHIK-sE2 recombinante na presença do plasmídeo que codifica o gene de resistência à blasticidina. A linha de células S2/CHIK-sE2 estável foi obtida por passagens sucessivas na presença de blasticidina. A linha de células foi selecionada quanto à sua capacidade para promover a secreção eficiente do vírus CHIK- sE2 após ativação do promotor metalotioneína.

As células S2/CHIK-sE2 em suspensão foram induzidas para a secreção de sE2 durante 21 dias na presença de Cu<sub>2+</sub>. O sobrenadante celular é filtrado a 0,22 µM e concentrado durante 16 horas numa coluna de afinidade de 5 mL de HiTrap Chelating HP (Amersham Biosciences) com a ajuda de uma bomba peristáltica. A proteína sE2 do CHIK é eluída da coluna de afinidade na

presença de concentrações crescentes de imidazole (50, 100 e 500 mM, pH 8). A proteína sE2 do CHIK é especificamente eluída a uma concentração de imidazole 500 mM (eluuição E<sub>3</sub>) da fração E<sub>37</sub> (Fig. 31). A proteína sE2 é detetada como sendo altamente purificada por PAGESDS após revelação com Azul de Coomassie. A proteína sE2 eluída na fração E<sub>39</sub> é imunodetectada especificamente por uma ascite (HMAF) de um murganho hiperimunizado contra o vírus CHIK que foi produzida na unidade IMFH (Fig. 32). Não foi observada reatividade cruzada com o HMAHF anti-dengue (DEN) ou anti-Nilo Ocidental (WN) e o anticorpo monoclonal 9D12 anti-DEN E. As proteínas DEN sE solúveis (DEN-3 e DEN-4) e WN sE purificadas dos sobrenadantes dos clones celulares S2 induzidos de acordo com o protocolo descrito acima são utilizadas como抗énios virais de controlo para a especificidade dos anticorpos murídeos anti-CHIK.

**Exemplo 3: Construção do vetor TRIP que expressa a forma solúvel de E2 (sE2) do vírus CHIK de acordo com a presente invenção.**

O gene que codifica a proteína sE2 de CHIK foi otimizado pela firma Genecust de forma a proporcionar um ADN sintético com um teor G+C enriquecido em comparação com o ADNc obtido a partir do ARN genómico viral. Os codões ricos em G+C (aminoácidos E2-1 a E2-364, ectodomínio de gp-E2 solúvel, sE2) foram fundidos com a sequência peptídica sinal da calreticulina humana (ssCRT) MLLSVPLLLGLLGLAA (SEQ ID NO: 77) para translocação da proteína viral para a

via de secreção. Foram adicionados sítios de restrição de enzima *BamHI* em 5' e *XhoI* em 3' nas respectivas extremidades das sequências que codificam a fusão sCRT+proteína sE2.

O gene sintético foi clonado no vetor TRIP entre os sítios *BamHI* e *XhoI* sob a transcrição do promotor ieCMV. O plasmídeo TRIP/CHIK.sE2 não replicativo e integrativo assim produzido foi validado quanto à expressão da proteína sE2 após transdução de células 293.

Como se mostra na figura 33, os inventores construíram um vetor que expressa a CHIK sE2. Como mencionado acima, a sequência CHIK sE2 original clonado no vetor TRIP foi modificada para melhorar a expressão em células de mamíferos (fig. 34).

A Figura 35 mostra células de mamíferos, tais como as células 293, transduzidas com o vetor TRIP-CHIK.sE2. A proteína sE2 expressada foi revelada por IF com anticorpos anti-CHIK.

**Exemplo 4: Produção de proteína recombinante sE2 e anticorpos monoclonais específicos.**

Os inventores geraram a linha de células S2/CHIK.sE2 induzível estável que liberta a forma solúvel da glicoproteína E2 do envelope (sE2) das estirpes de vírus CHIK da Reunião. Os inventores geraram também uma linha de

células estável 293A/CHIK.sE2 que foi transduzida pelo vetor lentiviral recombinante TRIP/CHIK.sE2. Um gene sE2 sintético que foi modificado para utilização de codão ótima nas células de mamíferos teve de ser utilizado para se obter expressão eficiente da sE2 de vírus CHIK em células fibroblásticas 293A humanas. O vetor TRIP/CHIK.sE2 é atualmente aferido pela sua capacidade para induzir imunidade protetora num modelo murídeo de infecção experimental. A suspensão viral principalmente enriquecida em CHIK pE2 (E2 precursora ou E3E2) foi obtida solubilizando viriões CHIK cultivados em células de mosquito com Triton X-100. Murganhos adultos foram hiperimunizados com CHIK pE2 na presença de adjuvante para gerar um hibridoma dirigido contra proteínas estruturais do CHIK. Os anticorpos monoclonais anti-CHIK E2 produzidos pelo hibridoma de murganho foram caracterizados através de um ensaio ELISA em virião CHIK altamente purificado e transferência de Western em sE2 segregada a partir da linha de células estável S2/CHIK.sE2. (Figs 50, 51 e Tabela 9). Foram também estabelecidos ensaios de imunodeteção de fluorescência de抗igenos virais intracelulares ou da superfície em células Vero infetadas por vírus CHIK e na linha de células transduzida com 293A/CHIK.sE2 estável (Figs 52-55). Os MAbs anti-CHIK.sE2 dos inventores têm utilização potencial no desenvolvimento diagnóstico viral inicial da doença de CHIK com base na imunocaptura de viriões CHIK no sangue virémico de doentes, e como ferramentas para estudos imunológicos e virológicos.

Tabela 1. Características dos doentes

N.º Doentes	Ilha	Região ou Ilha	Cidade ou Localidade	Amostra (b)	Data de Amostragem	Sintomas clínicos (c)	N.º do Isolado	El-226 (e)
1	Reunião (Cores) (a)	-	-	S	16-mar-05	Clássicos	05.61 (G)	A (*)
2	Reunião	Ocidental	St Gilles Les Bains	S	11-abr-05	Clássicos	05.55	A (*)
3	Reunião	Sul	Saint Pierre	S	2-maio-05	Clássicos	05.107	A (*)
4	Reunião	Ocidental	Mare Sèche Cilaos	S	4-maio-05	Clássicos	05.111	A (*)
5	Reunião	Sul	La Rivière St Louis	S	6-maio-05	Clássicos	05.115 (G)	A (*)
6	Reunião	Sul	St Louis	CSF	7-set-05	Encefalopatia neonatal	05.223	V (**)
7	Reunião	Sul	La Rivière St Louis	S	11-out-05	Clássicos	06.55	V (**)
8	Reunião	Sul	St Louis	S	21-out-05	Clássicos	06.59	V (**)
9	Reunião	Sul	La Rivière St Louis	S	21-out-05	Clássicos	06.53	V (**)
10	Reunião	Sul	La Rivière St Louis	P	26-out-05	Clássicos	n.i.	V (**)

(continuação)

N.º Doentes	Ilha	Região ou Ilha	Cidade ou Localidade	Amostra (b)	Data de Amostragem	Sinais Clínicos (c)	N.º do Isolado	El-226 (e)
11	Reunião	Sul	St Joseph	P	9-nov-05	Clássicos	n.i.	V (**)
12	Reunião	Sul	La Rivière St Louis	P	10-nov-05	Clássicos	n.i.	V (**)
13	Reunião	Sul	St Louis	P	20-nov-05	Clássicos	n.i.	V (**)
14	Reunião	Sul	La Rivière St Louis	P	21-nov-05	Clássicos	n.i.	V (**)
15	Reunião	Sul	La Rivière St Louis	S	23-nov-05	Clássicos	06.45	V (**)
16	Reunião	Sul	La Rivière St Louis (país)	S	28-nov-05	MENonatal	06.21 (G)	V (**)
17	Reunião	Sul	St Joseph	S	23-nov-05	Clássicos	06.47	V (**)
18	Reunião	Sul	La Rivière St Louis	P	24-nov-05	Clássicos	n.i.	V (**)
19	Reunião	Sul	Le Tampon	P	26-nov-05	Clássicos	n.i.	V (**)
20	Reunião	Sul	Ravine des Capris	P	25-nov-05	Clássicos	n.i.	V (**)
21	Reunião	Sul	St Joseph (país)	S	29-nov-05	MENonatal	06.25	V (**)
21	Reunião	Sul	St Joseph (país)	CSF	29-nov-05	MENonatal	06.27 (G)	V (**)
22	Reunião	Sul	St Louis	S	2-dez-05	Clássicos	06.49 (G)	V (**)

(continuação)

N.º Doentes	Ilha	Região ou Ilha	Cidade ou Localidade	Antostra (b)	Data de Amostragem	Sinais Clínicos (c)	N.º do Isolado	E1-226 (e)
23	Reunião	Sul	St Louis	P	8-dez-05	Clássicos	n.i.	V (**)
24	Reunião	Sul	Ravine des Capris	S	9-dez-05	ME	06.17	V (**)
25	Reunião	Sul	St Louis	P	13-dez-05	Clássicos	n.i.	V (**)
26	Reunião	Sul	St Pierre	P	2-jan-06	Clássicos	n.i.	A (**)
27	Reunião	Sul	St Pierre	P	4-jan-06	Síndrome álgica	n.i.	A (**)
28	Reunião	Oriental	St André	S	4-jan-06	Clássicos	n.i.	A (**)
29	Reunião	Sul	St Louis	P	29-dez-05	Síndrome álgica	n.i.	V (**)
30	Reunião	Sul	St Louis	CSF	29-dez-05	Síndrome álgica	n.i.	V (**)
31	Reunião	Sul	La Rivière St Louis	P	29-dez-05	Clássicos	n.i.	V (**)
32	Reunião	Sul	St Pierre	P	27-dez-05	Erupção cutânea vesicular grave dos membros inferiores	n.i.	V (**)
32	Reunião	Sul	St Pierre	L	28-dez-05	Erupção cutânea vesicular grave dos membros inferiores	n.i.	V (**)

(continuação)

N.º Doentes	Ilha	Região ou Ilha	Cidade ou Localidade	Amostra (b)	Data de Amostragem	Sinais Clínicos (c)	N.º do Isolado	El-226 (e)
33	Reunião	Sul	Ravine des Capris	P	4-jan-06	n.d.	n.i.	Viral(d)
34	Reunião	Sul	St Joseph	P	3-jan-06	Síndrome álgica	n.i.	A (**)
35	Reunião	Sul	St Louis	P	2-jan-06	Síndrome álgica	n.i.	V (**)
36	Reunião	Sul	St Joseph	P	5-jan-06	Síndrome álgica	n.i.	A ser determinada
37	Reunião	Sul	Ravine des Capris	P	6-jan-06	Clássicos	n.i.	V (**)
38	Reunião	Sul	St Louis	P	6-jan-06	Síndrome álgica	n.i.	V (**)
39	Reunião	Ocidental	Hospital de Saint-Paul	S	5-jan-06	n.d.	n.i.	A (**)
40	Reunião	Ocidental	St Leu	S	19-jan-06	Clássicos	n.i.	V (**)
41	Reunião	Sul	Les Aïrons	S	30-jan-06	Clássicos	06.97	V (**)
42	Reunião	Oriental	St Benoît	S	3-fev-06	Hepatite	n.i.	V (**)
43	Reunião	n.d.	n.d.	S	22-fev-06	Clássicos	n.i.	V (**)
44	Seicheles	Ilha Mahé	Anse aux Pins	S	9-agosto-05	Clássicos	05.209 (G)	A (*)
45	Seicheles	Ilha Mahé	Anse aux Pins	S	10-agosto-05	Clássicos	negativo	A (**)
46	Seicheles	Ilha Mahé	Anse aux Pins	S	10-agosto-05	Clássicos	negativo	A (**)
47	Madagascar	Oriental	Tamatrina	S	1-fev-06	Clássicos	06.103	A (**)
48	Madagascar	Oriental	Tamatrina	S	8-fev-06	Clássicos	06.99	A (**)
49	Madagascar	Oriental	Tamatrina	S	9-fev-06	Clássicos	06.101	A (**)

(continuação)

N.º Doentes	Ilha	Região ou Ilha	Cidade ou Localidade	Amostra (b)	Data de Amostragem	Sinais Clínicos (c)	N.º do Isolado	El-226 (e)
50	Madagáscar	Oriental	Toamasina	S	15-fev-06	Clássicos	n.i.	Viral(d)
51	Madagáscar	Norte	Antsirany	S	14-fev-06	Clássicos	n.i.	A (**)
52	Madagáscar	Norte	Djambanjary	S	14-fev-06	Clássicos	n.i.	A (**)
53	Madagáscar	Norte	Djambanjary	S	15-fev-06	Clássicos	n.i.	A (**)
54	Maiote	n.d.	n.d.	S	7-fev-06	Clássicos	06.111	A ser determinada
55	Maiote	n.d.	n.d.	S	11-fev-06	Clássicos	n.i.	V (**)
56	Maiote	n.d.	n.d.	S	13-fev-06	Clássicos	n.i.	V (**)
57	Maiote	n.d.	n.d.	S	13-fev-06	Clássicos	n.i.	V (**)
58	Maurícia	n.d.	n.d.	S	12-fev-06	Clássicos	06.93	V (*)
59	Maurícia	n.d.	n.d.	S	27-fev-06	Clássicos	n.i.	A (**)
60	Maurícia	n.d.	n.d.	S	1-mar-06	Clássicos	n.i.	A (**)

(a) Este pai regressou de Canóres e julga-se que teria sido lá infetado.

(b) S: soro; P: plasma; CSF: Líquido cefalorraquídiano

(c) ME: meningoencefalite

(d) Isolados marcados com (G) correspondem àqueles para os quais foi estabelecida a sequência quase completa do genoma

(e) A: Alainina; V: Valina; (\*) Sequência determinada a partir de isolados de vírus; (\*\*) Sequência determinada a partir de amostras biológicas;

(\*\*\*) Sequência determinada a partir de isolados de vírus e amostras biológicas

n.i.: não foi tentado o isolamento do vírus; n.d.: não determinado

Tabela 2. Alterações relevantes de aminoácidos identificados entre isolados do Oceano Índico versus uma seleção de sequências de Alfavírus.

	Proteínas não estruturais								Proteínas estruturais								
	mSP 1	mSP 1	mSP 2	mSP 2	mSP 2	mSP 3	mSP 3	mSP 3	mSP 4	mSP 4	E2	E2	E2	E2	6K	E1	E1
Proteína																	
Posição no polipeptídeo (a)	301	488	589	909	117	132	155	167	169	179	180	193	211	236	47	48	63
Posição na proteína (a)	301	488	54	374	642	793	217	337	358	460	471	75	254	500	14	16	31
05.115 (Genótipo 1)	T	R	N	Y	Y	V	H	I	S	E	S	A	A	L	Q	T	M
06.21 (Genótipo 2)	T	R	N	Y	Y	V	H	I	S	Na	nd	A	A	L	Q	T	M
06.27 (Genótipo 3)	I	R	N	Y	N	V	H	I	S	Del	S	A	A	L	Q	T	M
06.49 (Genótipo 4)	T	R	N	Y	Y	V	H	I	S	E	S	A	A	L	Q	T	M
05.209 (Genótipo 5)	T	R	N	Y	Y	V	H	I	P	E	S	A	A	L	R	T	M
S27	T	Q	S	H	C	A	Y	T	S	L	P	T	T	Q	Q	A	T
ROSS	T	Q	S	H	C	A	Y	T	S	L	P	T	T	Q	Q	A	T

	Proteínas não estruturais												Proteínas estruturais											
	T	K	S	H	V	A	V	T	S	P	P	T	T	Q	Q	A	T	S	V	A	A	V	D	
3797 (Grupo filogenético da África Ocidental)																								
Nagpur (Grupo filogenético Asiático)	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	nd	
OMV	S	Q	S	N	H	A	V	**	S	**	**	T	T	E	H	A*	T	S	I*	T*	A	V	D	
EEV	S	Q	S	N	E	R	N	**	**	**	**	I	V	E	E	G*	S	A	I*	D*	A	E	D	
SEV	V	S	N	H	V	A	I	**	S	**	**	T	T	E	H	V*	T	S	C*	A*	P	M	D	
RRV	V	N	S	H	V	G	S	**	**	**	**	V	T	E	H	D*	D	S	C*	A*	P	M	D	
SIV	V	M	S	H	E	R	K	**	**	**	**	I	T	E	V	A*	V	T	V*	S*	A	V	N	

(a) numeração de referência do S27

\* Posição hipervariável,  
\*\* Posição variável,

OMV: vírus o'nyong-nyong; SEV: Vírus da Floresta de Semliki; RRV: vírus de Ross River; SIV: vírus Sindbis; EEV: vírus da Encefalite Equina Oriental.

nd: não determinado. Assinale-se que o codão de terminação opal observado na nsP3-524 de isolados do surto do Oceano Índico, mas não no S27, não é representado na tabela.

Tabela 3. Polimorfismos observados entre isolados do Oceano Índico

Posição no genoma (b)	978	1378	3605	4705	5147	5452	5453-5455	6547	7045	8978	9600	10670	11295	11421
Proteína	nsP1	nsP1	nsP2	nsP3	nsP3	nsP3	nsP3	nsP4	nsP4	E2	E2	E1	E1	3'NTR
Posição na proteína (para alteração de aa)														
05.115 (Genótipo 1)	C	G	T	T	T	G	GAA (Glu)	G	C	A	T	C	G	C
05.61 (Genótipo 1)	C	G	T	T	T	n.d.	n.d.	G	C	A	T	C	G	C
06.21 (Genótipo 2)	C	G	T	T	T	G	GAA (Glu)	G	C	A	T	T (Ala>Val)	G	C
06.27 (Genótipo 3)	T (Thr>Ile)	G	A (Tyr>Ser)	A	T	A	Surpreendido	A	C	A	T (Ala>Val)	G	C	
06.49 (Genótipo 4)	C	A	T	T	T	G	GAA (Glu)	A	C	A	C	T (Ala>Val)	G	T
05.209 (Genótipo 5)	C	G	T	T	C (Ser>Pro)	G	GAA (Glu)	G	T	G (Gln>Arg)	T	C	A	C
S27 (a)	C	G	T	T	T	G	GAA (Glu)	G	C	A	T	C	G	C

(a) São representados apenas os sítios que são variáveis entre os isolados do surto do Oceano Índico; aqueles sítios que foram distintos entre isolados do surto do Oceano Índico e outros vírus são dados como Informação Suplementar

(b) numeração de S27

Tabela 4. Iniciadores utilizados para RT-PCR e sequenciação

<b>Fragmento</b>	<b>Gene</b>	<b>Iniciador</b>	<b>Sequência (5' para 3')</b>	<b>SEQ ID</b>
(a)				<b>NO.</b>
FG1	5'NC	18F	CACGTAGCCTACCAGTTCTTA	35
	nsP1	871R	ATGGAACACCGATGGTAGGTG	36
FG2	nsP1	616F	AACCCCGTTCATGTACAATGC	37
	nsP1	1435R	CGGTACCACAAAGCTGTCAAAC	38
FG3	nsP1	1317F	CACTGACCTGCTGCTGTCTATG	39
	nsP2	2130R	AGTCCTGCAGCTTCTTCCTTC	40
FG4	nsP1	1412F	CGAGTTGACAGCTTGTGGTA	41
	nsP2	2227R	ATGACTGCAATTGTATGGGC	42
FG5	nsP2	1908F	CAATCTGCCTGAAGACTTCC	43
	nsP2	2709R	TCCACTACAATCGGCTTGTG	44
FG6	nsP2	2530F	GTGCGGCTTCTTCAATATGATG	45
	nsP2	3343R	TCCAGGCCTATTATCCCAGTG	46
FG7	nsP2	2577F	AACATCTGCACCCAAGTGTACC	47
	nsP2	3504R	GTCTCCTGTTGGCCGGTATAAT	48
FG8	nsP2	3332F	TAATAGGCCTGGAGGGAAGATG	49
	nsP3	4134R	CTACGCACTCTTCATCGTTCTT	50
FG9	nsP2	3885F	GAACGAGTCATCTCGTATTGG	51
	nsP3	4725R	ATATCTCTGCCATATCCACTGC	52
FG10	nsP3	4458F	TCTTACAGCCATGGACTCGAC	53
	nsP3	5273R	CGACAGGTACGGTGCTCATTAC	54
FG11	nsP3	5065F	TGTACAGGAAGCGAGTACGACC	55
	nsP4	5874R	TCTACTTGCACGACTGATACC	56
FG12	nsP4	5630F	ACGGACGACGAGTTACGACTAG	57
	nsP4	6380R	CCCAGTATTCTGGTTGCATG	58

(continuação)

Fragmento	Gene	Iniciador (a)	Sequência (5' para 3')	SEQ ID NO.
FG13	nsP4	6184F	AAAACAGCACGCTTACCACG	59
	nsP4	6936R	AACTTGAAGCGCGTACCTGTC	60
FG14	nsP4	6732F	TCATAGCCGCACACTTAAGC	61
	nsP4	7495R	AGGACCGCCGTACAAAGTTAC	62
FG15	nsP4	7278F	GCAGGTGACGAACAAGATGAG	63
	C	8034R	CCGCTTAAAGGCCAATTG	64
FG16	C	7910F	TCGAAGTCAAGCACGAAGG	65
	E2	8670R	GTCTGTCGCTTCATTCTGATG	66
FG17	E3	8459F	TGCTTGAGGACAACGTCATGAG	67
	E2	9240R	TTTGTGATTGGTGACCGCG	68
FG18	E2	9093F	AGTCCGGAACGTAAAGATCAC	69
	6K	9861R	AAAGGTTGCTGCTCGTCCAC	70
FG19	E2	9648F	AGTTGTGTCAGTGGCCTCGTTC	71
	E1	10403R	TAAAGGACGCGGAGCTTAGCTG	72
FG20	E1	10145F	ACAAAACCGTCATCCCGTCTC	73
	E1	11158R	TGACTATGTGGCCTTCGGAGG	74
FG21	E1	10959F	CAGCAAGAAAGGCAAGTGTGC	75
	3'NC	11770R	TTTGCCAATTATGGTATTCA	76

(a) O nome do iniciador indica a sua posição e direção na sequência de nucleótidos do genoma de S27.

Tabela 5. Alterações das minerais observadas entre a essência S7 e essências do súro do Oceano Índico nos preâmbulos e estuários

Protocolo	esp1	esp2	esp3	esp4	esp5	esp6	esp7	esp8	esp9	esp10	esp11	esp12	esp13	esp14	esp15	esp16	esp17	esp18	esp19	esp20	esp21	esp22	esp23	esp24	esp25	esp26	esp27	esp28
Possível presente	172	234	301	383	384	481	483	307	34	374	642	643	733															
Possível polimórfico	172	234	301	383	384	481	483	307	34	374	642	643	733															
327	L	E	T	M	I	T	O	I	S	H	C	S	A															
0811	V	X	I	L	L	I	R	N	Y	Y	N	V																
0841	V	X	I	L	L	I	R	N	Y	Y	Y	N	V															
0862	V	X	I	L	L	I	R	N	Y	Y	Y	N	V															
08621	V	X	I	L	L	I	R	N	Y	Y	Y	N	V															
08622	V	X	I	L	L	I	R	N	Y	Y	Y	N	V															
0863	V	X	I	L	L	I	R	N	Y	Y	Y	N	V															
08630	V	X	I	L	L	I	R	N	Y	Y	Y	N	V															

esp3 + esp4

Protocolo	esp1	esp2	esp3	esp4	esp5	esp6	esp7	esp8	esp9	esp10	esp11	esp12	esp13	esp14	esp15	esp16	esp17	esp18	esp19	esp20	esp21	esp22	esp23	esp24	esp25	esp26	esp27	esp28
Possível presente	175	217	326	321	337	323	358	376	382	460	461	462	471	524	735	234	300	514	533	634								
Possível polimórfico	1498	1550	1636	1644	1720	1691	1695	1715	1724	1724	1725	1725	1725	1725	1725	1725	1725	1725	1725	1725	1725	1725	1725	1725	1725	1725	1725	
3531	I	H	S	A	I	E	K	S	I	A	L	L	S	P	N	S	Tern.	A	A	A	A	A	A	A	A	A		
08621	I	H	S	A	I	E	S	T	T	E	E	E	E	P	N	S	Tern.	A	A	A	A	A	A	A	A	A		
08622	I	H	S	A	I	E	S	T	T	E	E	E	E	P	N	S	Tern.	A	A	A	A	A	A	A	A	A		
08623	I	H	S	A	I	E	S	T	T	E	E	E	E	P	N	S	Tern.	A	A	A	A	A	A	A	A	A		
086230	I	H	S	A	I	E	S	T	T	E	E	E	E	P	N	S	Tern.	A	A	A	A	A	A	A	A	A		

As espécies com ocorrências a que foram subtraídas esteiadas das ações da Diatomais.

Tabela 6. Alterações de amioacéidos observadas nos genótipos estruturais entre as estípulas S27, Ross e do surto do Oceano Índico.

Proteína	C	E1	E2	E3	E2	E2	E2	E1	E2	E1	E1	E1	E1											
Posição de AA na proteína dissoctada	63	23	37	74	79	144	163	164	181	194	195	211	231	237	239	312	344	375	386	54	226	289	344	322
Posição de AA na poliproteína	63	284	382	395	404	471	485	486	516	519	520	538	578	592	624	827	829	739	741	802	103	187	109	113
Posição da at. no genoma [a]	775	843	871	878	877	897	902	903	923	912	912	917	938	934	943	947	957	958	959	959	957	106	107	108
S27	K	I	G	I	G	Q	Q	N	A	I	S	Q	I	R	M	S	T	A	S	V	V	I	A	
Ross	R	I	G	I	G	Q	Q	N	A	I	G	I	R	I	P	R	S	T	A	S	V	V	I	
95.115	R	T	K	M	E	Q	T	T	M	S	G	Y	R	R	N	M	T	T	A	I	V	V	A	
95.61	R	T	K	M	E	Q	T	T	M	S	G	Y	R	R	N	M	T	T	A	I	V	V	A	
95.21	R	T	K	M	E	Q	T	T	M	S	G	Y	R	R	N	M	T	T	A	I	V	V	A	
95.27	R	T	K	M	E	Q	T	T	M	S	G	Q	Y	R	R	N	M	T	T	A	I	V	V	
95.49	R	T	K	M	E	Q	T	T	M	S	G	Q	Y	R	R	N	M	T	T	A	I	V	V	
95.29	R	T	K	M	E	Q	T	T	M	S	G	Q	Y	R	R	N	M	T	T	A	I	V	V	

[a] Quando existem duas posições variáveis no mesmo codão, é indicada apenas a posição da proteína do surto.

As células a cinzento correspondem a alterações de aminoácidos entre isolados do surto do Oceano Índico.

Tabela 7. Sequência utilizada para a análise filogenética  
de sequências parciais de E1

N.º de acesso	Estirpe	Domínio genómico	Origem da Estirpe	Datas de Isolamento	Grupo Filogenético	Referência
AF192906	CAR 256	E1 parcial	Região Africana Central	Desconhecida	África Central	1
AF192907	Aq41855	E1 parcial	Uganda	1982	África Central	1
AY549583	ChikRCA	E1 parcial	DRC (b)	1996	África Central	2
AF192903	AR 18211	E1 parcial	República da África do Sul	1976	África Central-Oriental/Austral	1
AF192904	SA H2123	E1 parcial	República da África do Sul	1976	África Central-Oriental/Austral	1
AF192905	Ross	E1 parcial	Tanzânia	1953	África Central-Oriental/Austral	1
AF490259	Ross	Genoma completo	Tanzânia	1953	África Central-Oriental/Austral	na
AF369024	S27	Genoma completo	Tanzânia	1952	África Central-Oriental/Austral	3
AY549576	DRC010	E1 parcial	DRC (b)	2000	África Central	2
AY549577	DRC027	E1 parcial	DRC (b)	2000	África Central	2
AY549579	DRC1719	E1 parcial	DRC (b)	2000	África Central	2
AY549575	DRC007	E1 parcial	DRC (b)	2000	África Central	2
AY549578	DRC1718	E1 parcial	DRC (b)	2000	África Central	2
AY549581	DRC1725	E1 parcial	DRC (b)	2000	África Central	2
AY549582	DRC1728	E1 parcial	DRC (b)	2000	África Central	2

(continuação)

N.º de acesso	Estirpe	Domínio genómico	Origem da Estirpe	Datas de Isolamento	Grupo Filogenético	Referência
AY549580	DRC1720	E1 parcial	DRC (b)	2000	África Central	2
AY549584	DRC1730	E1 parcial	DRC (b)	2000	África Central	2
AF192896	644188	E1 parcial	Tailândia	1988	Asiático	1
AF192899	3412/78	E1 parcial	Tailândia	1978	Asiático	1
AF192894	RSU1	E1 parcial	Indonésia	1985	Asiático	1
AF192895	H15483	E1 parcial	Filipinas	1985	Asiático	1
AF192898	1455/75	E1 parcial	Tailândia	1975	Asiático	1
AF192901	Gibbs 63-263	E1 parcial	Índia	1963	Asiático	1
AF192902	P0731460	E1 parcial	Índia	1973	Asiático	1
AF192897	C-03295	E1 parcial	Tailândia	1995	Asiático	1
AF192900	SV045196	E1 parcial	Tailândia	1996	Asiático	1

Tabela 7. Sequência utilizada para a análise filogenética de sequências parciais de E1. (continuação)

N.º de acesso	Estirpe	Domínio genómico	Origem da Estirpe	Data de Isolamento	Grupo Filogenético	Referência
L37661	Estirpe da vacina	gene da poliproteína	Na	na	Asiático	na
AF192892	37997	E1 parcial	Na	na	África Ocidental	1
AY726732	37997	genoma completo	Senegal	1983	África Ocidental	4
AF192891	PM2951	E1 parcial	Senegal	1966	África Ocidental	1
AF192893	IbH35 E1	E1 parcial	Nigéria	1964	África Ocidental	1

Referências: (1) Powers *et al.*, Pastorino *et al.*, 2004; (3) Khan *et al.*, 2002; (4) Vanlandingham *et al.*, 2005.  
 (b) República Democrática do Congo

Tabela 8. Percentagem de semelhança da sequência com base em aminoácidos e nucleótidos (entre parênteses) para as proteínas estruturais (SP) e não estruturais (NSP) de Alfavírus selecionados.

Vírus	Estirpe	N.º de Acesso	05.115/06.49	05.115	06.49
			NSP	SP	SP
CHIKV	05.115	A ser submetida	100 (100)	100 (100)	-
	06.49	A ser submetida	100 (99,97)	99,91 (99,95)	100 (100)
	S27	AF369024	98,79 (97,3)	98,47 (97,34)	98,38 (97,33)
	37997	AY726732	95,88 (85,5)	95,82 (84,87)	95,74 (84,81)
	Nagpur	AY424803	NA	97,18 (94,85)	97,10 (94,79)
	Vacina	L37661	NA	98,92 (94,24)	96,83 (94,19)
ONNV	Gulu	M20303	85,90	87,30	87,22
SFV	genoma de ARN de 42S	X04129	70,55	65,20	65,20
RRV	NB5092	M20162	69,66	64,40	64,40
SINV	HRSP	J02363	59,25	47,40	47,31

CHIKV: vírus Chikungunya; ONNV: vírus o'nyong-nyong; SFV: Vírus da Floresta de Semliki; RRV: vírus Ross River; SINV: vírus Sindbis NA: Não Disponível.

Tabela 9: Lista de ensaios biológicos realizados para validar a reatividade de MAbs anti-CHIK E2

ENSAIOS BIOLÓGICOS
<i>ELISA em抗igénios solubilizados de viriões CHIK</i>
<i>ELISA em viriões CHIK purificados (Ilha da Reunião)</i>
<i>ELISA em viriões CHIK purificados (+ TX-100)</i>

(continuação)

<b>ENSAIOS BIOLÓGICOS</b>
<b>ELISA em viriões CHIK purificados (+ NP-40)</b>
<b>Ensaio de IF em células VERO infetadas com CHIKV</b>
<b>Análise FACS na superfície celular de células VERO infetadas com CHIKV</b>
<b>Transferência de Western em CHIK sE2 recombinante de células S2</b>
<b>Ensaio de IF de um clone estável de células 293A transduzidas com TRIP/CHIK.sE2</b>
<b>Transferência de Western em CHIK sE2 recombinante do clone de células 293A transduzidas com TRIP/CHIK.sE2</b>

**REFERÊNCIAS**

1. Strauss EG, Strauss JH (1986) Structure and replication of the alphavirus genome. In Schlesinger S, Schlesinger MJ, editors. The Togaviridae and Flaviviridae. New York: Plenum Press. pp. 35-90.
2. Porterfield JH (1980) Antigenic characteristics and classification of the Togaviridae. In: Schlesinger R, editor. The Togaviruses. New York: Academic Press. pp. 13-46.
3. Ross RW (1956) The Newala epidemic. III. The virus: isolation, pathogenic properties and relationship to the epidemic. J Hyg 54: 177-191.

4. Jupp PG, McIntosh BM (1988) Chikungunya disease. In: editors MTP, editor. *The Arboviruses: epidemiology and ecology*. Boca Raton, Florida: CRC Press. pp. 137-13 157.
5. Johnston RE, Peters CJ (1996) Alphaviruses associated primarily with fever and polyarthritits. In: Fields BN, Knipe DM, Howley PM, editors. *Fields Virology*. pp. 16 843-898.
6. Pastorino B, Muyembe-Tamfum JJ, Bessaud M, Tock F, Tolou H, et al. (2004) Epidemic resurgence of Chikungunya virus in democratic Republic of the Congo: identification of a new central African strain. *J Med Virol* 74: 277-282.
7. Laras K, Sukri NC, Larasati RP, Bangs MJ, Kosim R, et al. (2005) Tracking the re-emergence of epidemic chikungunya virus in Indonesia. *Trans R Soc Trop Med Hyg* 99: 128-141.
8. Paquet C, Quatresous I, Solet JL, Sissoko D, Renault P (2006) Chikungunya outbreak in Reunion: epidemiology and surveillance, 2005 to early January 2006. *Eurosurveillance weekly* 11: 2.
9. Khan AH, Morita K, Parquet Md Mdel C, Hasebe F, Mathenge EG, et al. (2002) Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. *J Gen Virol* 83: 3075-3084.

10. Powers AM, Brault AC, Tesh RB, Weaver SC (2000) Re-emergence of Chikungunya and O'nyong-nyong viruses: evidence for distinct geographical lineages and distant evolutionary relationships. *J Gen Virol* 81: 471-479.
11. Gordon D AC, Green P. (1998) Consed: a graphical tool for sequence finishing. *Genome Res* 8: 195-202.
12. Rozas J, Sanchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA 2 polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19: 2496-2497.
13. Xia X, Xie Z (2001) DAMBE: software package for data analysis in molecular biology and evolution. *J Hered* 92: 371-373.
14. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Research* 22: 4673-4680.
15. Felsenstein J (1989) PHYLIP - Phylogeny Interfeme Package (version 3.2). *Cladistics* 5: 164-166.
16. Hofacker IL (2003) Vienna RNA secondary structure server. *Nucleic Acids Res* 31: 3429-3431.

17. Kumar S, Tamura K, Nei M (2004) MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* 5: 150-163.
18. Martin DP, Williamson C, Posada D (2005) RDP2: recombination detection and analysis from sequence alignments. *Bioinformatics* 21: 260-262.
19. Roussel A, Lescar J, Vaney MC, Wengler G, Wengler G, et al. (2006) Structure and interactions at the viral surface of the envelope protein E1 of semliki forest virus. *Structure* 14: 75-86.
20. Carson M (1987) Ribbon models of macromolecules. *J Mol Graph* 5: 103-106.
21. Strauss JH, Strauss EG (1994) The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev* 58: 491-562.
22. Lavergne A, Thoisy BD, Lacoste V, Pascalis H, Pouliquen JF, et al. (2005) Mayaro virus: Complete nucleotide sequence and phylogenetic relationships with other alphaviruses. *Virus Res* in press.
23. Lanciotti RS, Ludwig ML, Rwaguma EB, Lutwama JJ, Kram TM, et al. (1998) Emergence of epidemic O'nyong-nyong fever in Uganda after a 35-year absence: genetic characterization of the virus. *Virology* 252: 258-268.

24. Strauss EG, Levinson R, Rice CM, Dalrymple J, Strauss JH (1988) Nonstructural proteins nsP3 and nsP4 of Ross River and O'Nyong-nyong viruses: sequence and comparison with those of other alphaviruses. *Virology* 164: 265-274.
25. Griffin DE (2001) Alphaviruses. In: Knipe DM, Howley PM, editors. *Fields Virology*. Philadelphia: Lippincott Williams & Wilkins. pp. 917-962.
26. Vashishtha M, Phalen T, Marquardt MT, Ryu JS, Ng AC, et al. (1998) A single point mutation controls the cholesterol dependence of Semliki Forest virus entry and exit. *J Cell Biol* 140: 91-99.
27. Ahn A, Schoepp RJ, Sternberg D, Kielian M (1999) Growth and stability of a cholesterol-independent Semliki Forest virus mutant in mosquitoes. *Virology* 262: 452-456.
28. Williams MC, Woodall JP, Corbet PS, Gillett JD (1965) O'nyong-Nyong Fever: An Epidemic Virus Disease In East Africa. 8. Virus Isolations From Anopheles Mosquitoes. *Trans R Soc Trop Med Hyg* 59: 300-306.
29. Weaver SC, Barrett AD (2004) Transmission cycles, host range, evolution and emergence of arboviral disease. *Nat Rev Microbiol* 2: 789-801.
30. Lu YE, Cassese T, Kielian M (1999) The cholesterol

requirement for sindbis virus entry and exit and characterization of a spike protein region involved in cholesterol dependence. *J Virol* 73: 4272-4278.

31. Holland J, Spindler K, Horodyski F, Grabau E, Nichol S, et al. (1982) Rapid evolution of RNA genomes. *Science* 215: 1577-1585.
32. Domingo E, Holland JJ (1997) RNA virus mutations and fitness for survival. *Annu Rev Microbiol* 51: 151-178.
33. Vignuzzi M, Stone JK, Arnold JJ, Cameron CE, Andino R (2006) Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature* 439: 344-348.
34. Kim KH, Rumenapf T, Strauss EG, Strauss JH (2004) Regulation of Semliki Forest virus RNA replication: a model for the control of alphavirus pathogenesis in invertebrate hosts. *Virology* 323: 153-163.
35. Heise C, Kirn DH (2000) Replication-selective adenoviruses as oncolytic agents. *J Clin Invest* 105: 847-851.
36. Heise MT, White LJ, Simpson DA, Leonard C, Bernard KA, et al. (2003) An attenuating mutation in nsP1 of the Sindbis-group virus S.A.AR86 accelerates non-structural

protein processing and up-regulates viral 26S RNA synthesis. *J Virol* 77: 1149-1156.

37. Suthar MS, Shabman R, Madric K, Lambeth C, Heise MT (2005) Identification of adult mouse neurovirulence determinants of the Sindbis virus strain AR86. *J Virol* 79: 4219-4228.

38. Condon RJ, Rouse IL (1995) Acute symptoms and sequelae of Ross River virus infection in South-Western Australia: a follow-up study. *Clin Diagn Virol* 3: 273-284.

39. Selden SM, Cameron I ron AS (1996) Changing epidemiology of Ross River virus disease in South Australia. *Med J Aust* 165: 313-317.

40. Mazaud R, Salaün JJ, Montabone H, Goube P, Bazillio R (1971) Troubles neurologiques et sensoriels aigus dans la dengue et la fièvre a Chikungunya. *Bull Soc Pathol Exot* 64: 22-30.

41. Nimmannitya S, Halstead SB, Cohen SN, Margiotta MR (1969) Dengue and chikungunya virus infection in man in Thailand, 1962-1964. I. Observations on hospitalized patients with hemorrhagic fever. *Am J Trop Med Hyg* 18: 954-971.

42. Gratz NG (2004) Critical review of the vector status of *Aedes albopictus*. *Med Vet Entomol* 18: 215-227.

43. Lescar J, Roussel A, Wien MW, Navaza J, Fuller SD, et al. (2001) The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. *Cell* 105: 137-148.

Lisboa, 27 de julho de 2015

**REIVINDICAÇÕES**

1. Estirpe selvagem 05.115 isolada e purificada do vírus Chikungunya (CHIKV) capaz de infetar células humanas *in vitro* e em que o seu genoma consiste na sequência mostrada na Figura 4.
2. Polinucleótido isolado e purificado que consiste na sequência como se mostra na Figura 4.
3. Fragmento do polinucleótido de acordo com a reivindicação 2, em que codifica o ectodomínio da glicoproteína E2.
4. Fragmento do polinucleótido de acordo com a reivindicação 2, em que codifica uma forma solúvel da glicoproteína E2.
5. Fragmento do polinucleótido de acordo com a reivindicação 2, em que codifica o ectodomínio da glicoproteína E1.
6. Vetor compreendendo um fragmento como definido em qualquer uma das reivindicações 3 a 5.
7. Célula hospedeira compreendendo um vetor como definido na reivindicação 6.

8. Polipéptido purificado codificado por um fragmento de polinucleótido como definido em qualquer uma das reivindicações 3 a 5.

9. Polipéptido purificado consistindo na sequência como definida na Figura 40.

10. Utilização de um elemento selecionado do grupo que consiste em:

- uma estirpe 05.115 do vírus Chikungunya que consiste no genoma definido na Figura 4;
- um polinucleótido que consiste na sequência que se mostra na Figura 4;
- um fragmento como definido em qualquer uma das reivindicações 3 a 5;
- um vetor como definido na reivindicação 6;
- uma célula hospedeira como definida na reivindicação 7;
- um polipéptido como definido em qualquer uma das reivindicações 8 ou 9; e
- anticorpos monoclonais ou policlonais, ou fragmento dos mesmos, que se ligam especificamente a um polipéptido como definido em qualquer uma das reivindicações 8 e 9; para a deteção *in vitro* de um CHIKV associado a uma arbovirose.

11. Utilização de um elemento selecionado do grupo que consiste em:

- uma estirpe 05.115 do vírus Chikungunya que comprehende um genoma como definido na Figura 4;
  - um polinucleótido que consiste na sequência que se mostra na Figura 4;
  - um fragmento como definido em qualquer uma das reivindicações 3 a 5;
  - um vetor como definido na reivindicação 6;
  - uma célula hospedeira como definida na reivindicação 7; e
  - um polipéptido como definido em qualquer uma das reivindicações 8 ou 9;
- para a preparação de uma composição que previne e/ou trata uma arbovirose.

12. Composição compreendendo pelo menos um elemento selecionado do grupo que consiste em:

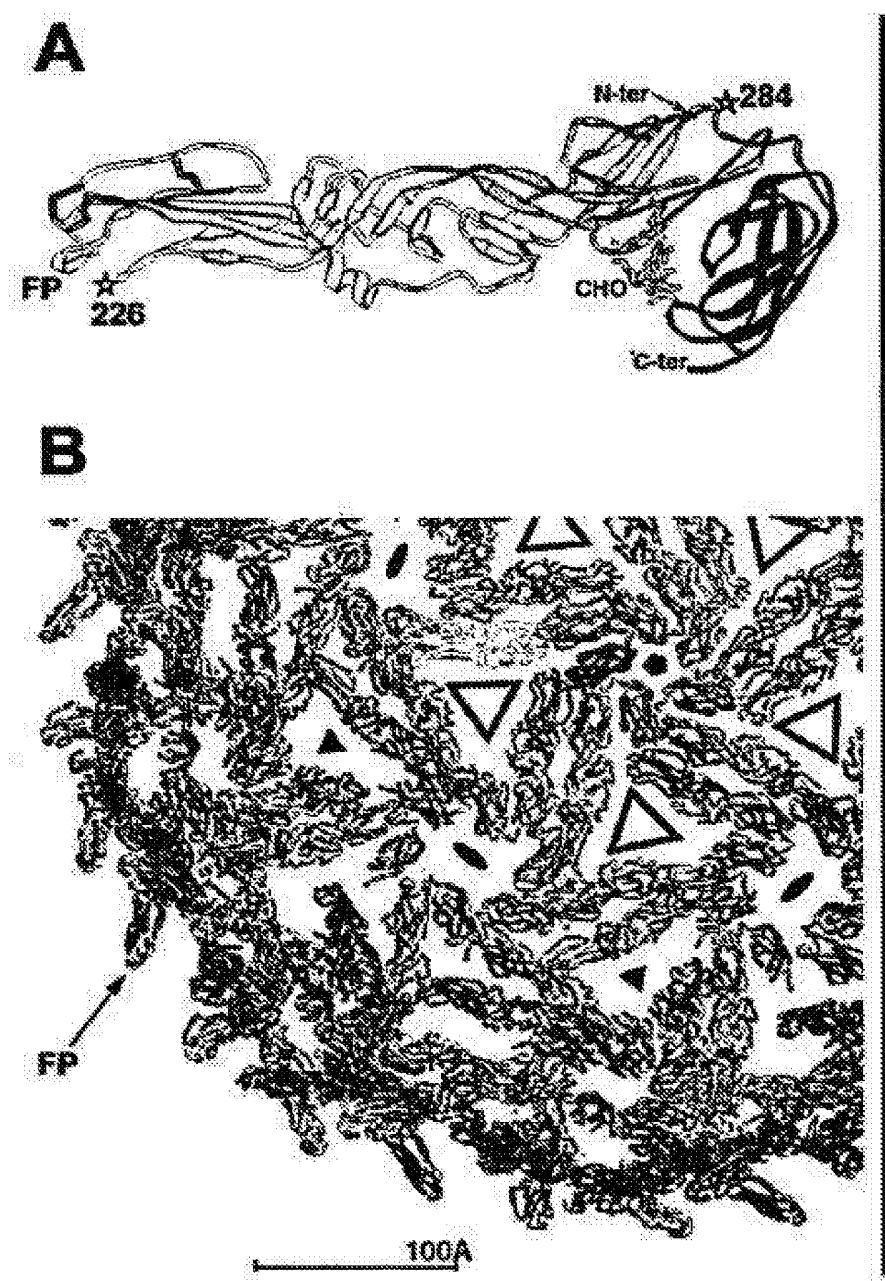
- uma estirpe 05.115 do vírus Chikungunya que comprehende um genoma como definido na figura 4;
- um polinucleótido que consiste na sequência que se mostra na Figura 4;
- um fragmento como definido em qualquer uma das reivindicações 3 a 5;
- um vetor como definido na reivindicação 6;
- uma célula hospedeira como definida na reivindicação 7; e
- um polipéptido como definido em qualquer uma das reivindicações 8 ou 9, e que comprehende ainda um transportador aceitável.

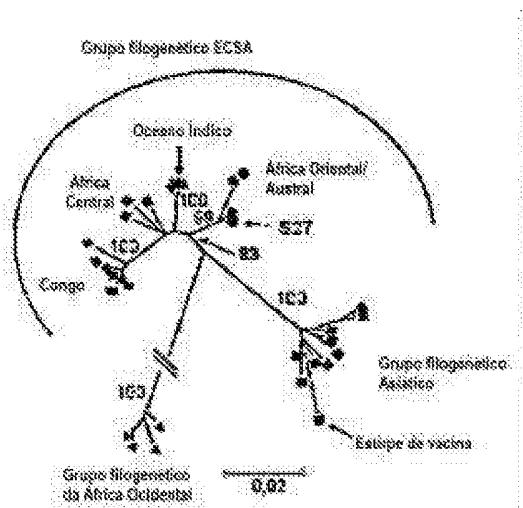
13. Kit para a deteção de um CHIKV associado a

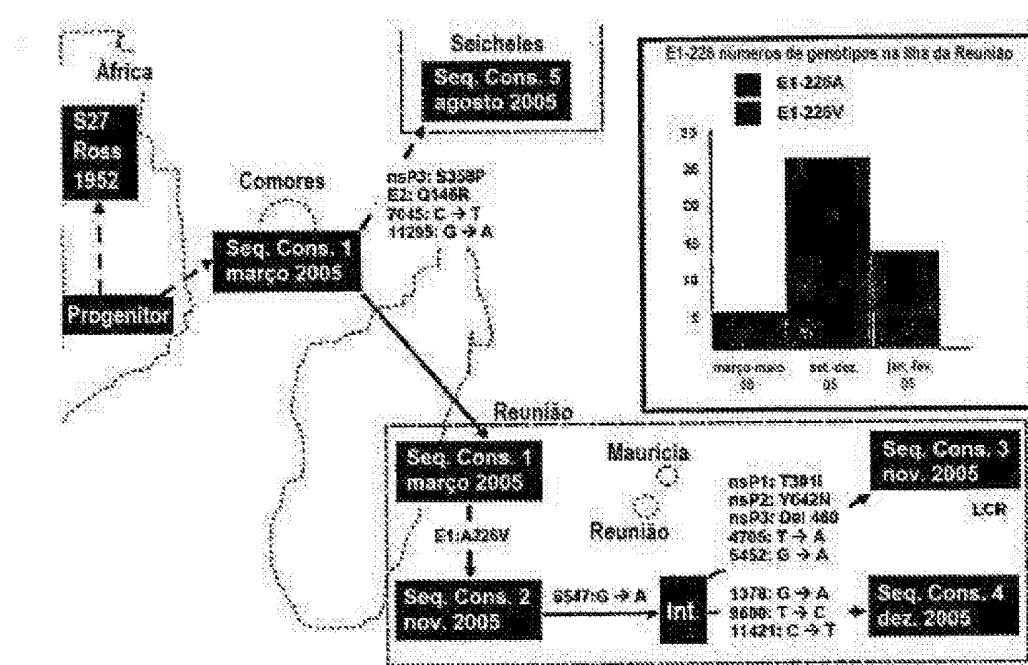
uma arbovirose, que compreende pelo menos uma embalagem compreendendo pelo menos um elemento selecionado do grupo que consiste em:

- uma estirpe 05.115 do vírus Chikungunya que consiste no genoma como definido na figura 4;
- um polinucleótido que consiste na sequência que se mostra na Figura 4;
- um fragmento como definido em qualquer uma das reivindicações 3 a 5;
- um vetor como definido na reivindicação 6;
- uma célula hospedeira como definida na reivindicação 7;
- um polipéptido como definido em qualquer uma das reivindicações 8 ou 9; e anticorpos monoclonais ou policlonais, ou fragmento dos mesmos, que se ligam especificamente a um polipéptido como definido em qualquer uma das reivindicações 8 e 9.

Lisboa, 27 de julho de 2015

**FIGURA 1**

**FIGURA 2**

**FIGURA 3**

**FIGURA 4**

**FIGURA 4 (cont.)**

**FIGURA 4 (cont.)**

**FIGURA 4 (cont.)**

**FIGURA 4 (cont.)**

aggatatagt gtcccccataag agacacacttg tacatcggca ataatctata gatccaaaaagg	11340
ctaccccaaccc ctgttaatagt acaaaatcac ccataatcta aaatttataa aaaccggaaa	11400
atacatcaaat aggataatgt gtcccccataag agacacacttg tatgttagtg ataaatgtatcg	11460
atccaaagggc cgsataaaccc ctgtttatgtt acaaaatataa aaatctataa aaatctataa	11520
atataaaaaaa ccataaaaaatg aatgtttttt aacggctataa aaacccttga atatgtttttt	11580
aaatctataat ttatccaaat t	11601

## FIGURA 5

**FIGURA 5 (cont.)**

ttatggaaatcc	ccggcttcacgg	catacaacgtc	agggccacgtc	aaatgttacttc	ttaacgcgtac	2880
ggaaaggataaa	cttcgtatggc	agacacatccc	cgggggaccgg	tggatataaga	ccgtgcgaaa	2840
ctttacccgaaa	ggaxaccttc	aaagaaatctt	ttgggggttgg	ggggggggggc	atggccatct	3000
aatggggggc	atctggggcc	ccccaaatggc	cttcgtatcc	ttcccaatctt	aaaggccatgt	3060
tttttgttgggt	aaagggtttgg	tccatcatctt	ccatccacgg	ggatataaaa	ttaaatgtatc	3120
ggatgttgggt	ccatataatcc	aaggcccttc	aaaggacaaa	gtatctccat	cttgatgttgc	3180
cttggatgtat	atctgttccgc	gtatgtatccg	gttggatctt	gatccggggcc	tatkttttctaa	3240
aaatgttgggt	tctgtgtttat	aaegeggatcc	ccatctgggg	aaatagggttgc	gaggggaaaat	3300
gtttggatgtt	aaaccccccgg	caggatcatcc	tcttggaaaa	aaatgttccat	ttccatccatgg	3360
aaatgttggat	aaatccatgg	aaatgttggat	ctttttccgg	aggatggaaa	acccatccccc	3420
ttccatccatgg	atctatccgg	ccaaacccgg	atccatccac	ttttttttgg	ccggatccccc	3480
ttccatccatgg	aaayggaaagaa	ttggaaatgg	ttttttccgg	ataaaaacggcc	acccatccgtgt	3540
cttggatgtat	gtttatccat	ttttttccgg	tatccatcc	ttttttttgg	ttttttttgg	3600
aaatgttgggt	ggatgttccgc	aaatccatcc	ttttttccgg	ttttttttgg	ttttttttgg	3660
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	3720
gttccatccatgg	ccatccatcc	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	3780
aaatgttggat	ccatccatcc	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	3840
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	3900
aaatgttggat	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	3960
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4020
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4080
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4140
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4200
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4260
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4320
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4380
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4440
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4500
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4560
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4620
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4680
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4740
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4800
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4860
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4920
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	4980
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5040
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5100
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5160
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5220
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5280
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5340
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5400
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5460
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5520
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5580
tttttgtatgt	tttttgtatgt	ttttttccgg	ttttttccgg	ttttttttgg	ttttttttgg	5640

### FIGURA 5 (cont.)

FIGURA 5 (cont.)

**FIGURA 5 (cont.)**

```

atacatatat aggtatacgt gtccctttagg agacacatttgg tatgttttgtt ataaatgttttt 11460
atcaatgggtt cgtatatacc ctgtatgtta acaaaaaatggt aaacaccatgtt aaaaatgttt 11550
tatggatttttt cttatatacgtt aggttttgtttt aagggttttata aaaccccttggt atatgttttt 11580
atccatataat ttatataat c 11601

```

**FIGURA 6**

**FIGURA 6 (cont.)**

taatgaaaaac	cgcgtctatcg	cataaacgtc	agagcacgtc	aacgttacttc	taacgcgtac	2880
ggaaatgtaaa	ctggtatggc	agacactttc	ccgcacccgg	tgtataaaggc	ccgtgcacaa	2940
ccccacggaaa	ggaaaccttc	aagcaactat	ttaggagtgg	gagggtggagc	atgcacatcat	3000
aatggggggc	atctgcagtc	accaaatgac	cttgcataca	ttccaaaaata	aaggcaacgt	3060
tttgttggct	aagagcttgg	tcccttactct	cgaaacacgc	gggataaacc	taatgtatacg	3120
gcacgttgtct	cagataatcc	aagcttccaa	agaagacaaa	gcataactcac	ctgaagtgae	3180
ctgtatgtaa	atatgtacgc	gcatgtatgg	ggtggatcta	gacagccggcc	tattttctaa	3240
accgttggtg	tcgtgttatt	acgcccgttca	ccactgggtat	aataggccgt	gaggggaaat	3300
gttggattt	macccccggg	cagcatccat	tctagaaaga	aagtatccat	tcacaaaaagg	3360
gaaggggggac	atcaacaaggc	agatctgtt	gacttccatgg	gggtatggaaq	actttttcccc	3420
taccacccaa	atctatccccc	ccaaacaggag	ataccacac	tcaatgttgg	ccggaaacaccy	3480
cccaatggaa	ggggggaaagaa	tggatggct	gtttaaccaag	atzaacccycc	accacgtgt	3540
cttgttgcgt	ggcttataacc	tttgcacttgc	tactaagaga	gttcaatgggg	tagegeccgtt	3600
agggtgttgc	ggagccggact	acacatccaa	cttagatgttgc	ggtgttgcac	caacgttgg	3660
taggtatgtac	ctatggtca	taaaatccaa	ccacaccccc	ccgtatccac	attaccacca	3720
gttgtgtggcc	cacgtccatgt	aactgcacat	gtctgggggt	gacttcattgc	gactgttcaa	3780
acccggggcc	tccatccatgt	tcacggcata	ttgttacccca	gttgcacccca	gttgcacgggt	3840
atctgtgtgt	ttggggacgc	atgtttagat	gtcttagagcc	tttgcacccac	catgtgtcac	3900
ccggccacact	gagatgtttt	tccttccatgg	ccacccgttc	estggccggaa	ggaaattttcac	3960
ccactatgttc	atgtacatcc	asctgtatgc	ggcccttccgt	ggccacgttca	ccggggccgg	4020
atgtgcacccg	tcgttacccgg	taaaacccat	ggacatccgc	aagaaacgttgc	aaagatgtgt	4080
atgtcaacgc	gttacacccctt	gggggtttaac	gggttgcacgt	gttgcacagg	caatatacaa	4140
aaaatggccgg	gagtttctta	agaaacagtgc	accacccgttgc	ggaaacccgttca	aaacatgttat	4200
gttgcggccgt	tatccatgtaa	tccacgttgc	tggaccggaa	tttctttaatt	attttggatgc	4260
ccggggggcc	ggggaaatgg	tcgttgcata	tcggggaaatgc	ccaaaggaaat	tcacttgggtt	4320
gggggttataat	atgttgcgtt	taccccttct	tcgttgcacgt	gttatactcc	gaggggaaaga	4380
ccggctgtac	tcgttgcacgt	accacccctt	tacccgttgc	gacccgttgc	atgttgcacgt	4440
gttccatccat	tcggggggcc	gggttgcgttgc	tcgtatccat	tttgcacccgtt	tcacatgttgc	4500
ggcccaatgtt	gagcttccgttgc	atgttgcacat	tcggatattgc	tttgcgttgc	ccctggccat	4560
ccctggccatgc	agcttggccat	tcggggggcc	tcgtatccat	tttgcgttgc	ccggggccat	4620
atatccatgt	ggggcccccgtt	tccatccatgt	ggcttgcgttgc	atggggggaa	tatctatctat	4680
gtggccacccgg	ccaaacccatgg	tcgttgcacat	tcgtatccat	tttgcacccgtt	ggggaaatgtt	4740
tgtatgttgc	ggggggatgtt	ccatccatgttgc	tcgtatccat	tttgcacccgtt	tcacatgttgc	4800
tgtccatccat	ttttggccgtt	tcgttgcacat	tcggatattgc	tttgcgttgc	ccctggccat	4860
ccatccatgttgc	agatccatgtt	tgtttccatccat	tttgcacccgtt	ccaaatgttgc	aaatagaagg	4920
atgttgcacccat	gttcaatgttgc	tcgtatccat	tttgcacccgtt	ccaaatgttgc	catgtgtgtt	4980
ccatccatgttgc	ccatccatgttgc	tcgtatccat	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5040
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5100
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5160
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5220
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5280
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5340
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5400
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5460
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5520
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5580
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5640
ccatccatgttgc	tcgtatccat	tttgcacccgtt	tttgcacccgtt	ccatccatgttgc	ccatccatgttgc	5700

**FIGURA 6 (cont.)**

agtccacgag gagaatgttt acccacctaa gctggatgc gcaaggagcc aacttattact  
taaagaaactc cggggatgtt catccatggc caacagaagc aggtatcagt 5760  
aaaaaaatcg aaxggagcaa tcataccatgg actaaagaga cyggccaaatgt 5820  
gtccatggacc ccataatgtcc ctatcttacgg tataacttataat ggtgttagac 5880  
tcggatccaaat gtcogatgtt cccatcccgaa atcaatggatc tataacttataat 5940  
agcttggaaaat tatccatcttg tcataccatgg gactacatgtt gtcactcgcc 6000  
agacatgttgc ggggttgcgg agatgttgcgt tgactacatgtt gtcactcgcc 6060  
caggatgttgc cggaaatccgg acgttaccaat gtcactcgcc 6120  
ccatcttacgg aataacatccatgg agaaatgttactt gtcactcgcc 6180  
caccatggatg agggatttacccatgttggaa atcaatggatc ttcggatccaaat 6240  
aaatatttgcg tgcacccaaatgg aataacttggaa atcaatggatc ttcggatccaaat 6300  
sactgaaatgtt ttagcaacctt atgttacttacgg ctcggatccaaat 6360  
cycaaaaatccatggatcataatctac tggccactaca ttcggatccaaat 6420  
tatgttggatc gacgttgcgg ttcacttccgg gtcactcgcc 6480  
ggatgttgcggatc cgggggttgcgg accatcttgcgtt ttcacatggatc 6540  
gttgcgttggatc aggttgcggatc ctcgttccgg gtcactcgcc 6600  
ttccatgtgcgc tcatatggatc ttcacatggatc 6660  
aaatggatcataatggatc gtcgttccgg gtcactcgcc 6720  
gcttgcgttgcggatc ttccatgtgcgc ttcacatggatc 6780  
gttgcgttggatc ttccatgtgcgc ttcacatggatc 6840  
gttgcgttggatc ttccatgtgcgc ttcacatggatc 6900  
atcgatgttgcgc ttcacatggatc 6960  
atgttgcgttgcgc ttcacatggatc 7020  
atgttgcgttgcgc ttcacatggatc 7080  
gttgcgttggatc ttccatgtgcgc ttcacatggatc 7140  
ttccatgtgcgc ttcacatggatc 7200  
ttccatgtgcgc ttcacatggatc 7260  
ttccatgtgcgc ttcacatggatc 7320  
ttccatgtgcgc ttcacatggatc 7380  
ttccatgtgcgc ttcacatggatc 7440  
ttccatgtgcgc ttcacatggatc 7500  
ttccatgtgcgc ttcacatggatc 7560  
ttccatgtgcgc ttcacatggatc 7620  
ttccatgtgcgc ttcacatggatc 7680  
ttccatgtgcgc ttcacatggatc 7740  
ttccatgtgcgc ttcacatggatc 7800  
ttccatgtgcgc ttcacatggatc 7860  
ttccatgtgcgc ttcacatggatc 7920  
ttccatgtgcgc ttcacatggatc 7980  
ttccatgtgcgc ttcacatggatc 8040  
ttccatgtgcgc ttcacatggatc 8100  
ttccatgtgcgc ttcacatggatc 8160  
ttccatgtgcgc ttcacatggatc 8220  
ttccatgtgcgc ttcacatggatc 8280  
ttccatgtgcgc ttcacatggatc 8340  
ttccatgtgcgc ttcacatggatc 8400  
ttccatgtgcgc ttcacatggatc 8460  
ttccatgtgcgc ttcacatggatc 8520  
ttccatgtgcgc ttcacatggatc 8580

**FIGURA 6 (cont.)**

**FIGURA 6 (cont.)**

straaacggc cgaaatcaccc ctgtatgttc acaaaatcttg aaaaatccataa aaatctataa 11520  
 aatgtaaaaaa ccataazacag aagtatgttca aagggtatata aaacccttga atgtatccaa 11580  
 aacataaaaat ttataaaaat c 11601

FIGURA 7

**FIGURA 7 (cont.)**

FIGURA 7 (cont.)

tcaatttacaaa	cagaaggtcag	tacggccatgc	agtgtctcccg	gtggaaaccccc	tggaggaaat	5700
ccacgaggag	aagtgttacc	cacctaagct	ggatyaagca	aaaggacycaac	tattacttta	5760
aaaaactccag	gagagtgcgt	ccatggccaa	cagaaggcagg	tatcagtcgc	gcgggggtags	5820
aaaaatgtaaa	gcagcaatca	tccagagact	aaagagagggc	tttagactat	acttaatgtc	5880
agagacccca	aaagtccata	cttacggac	tatcataccg	gcgttgttgc	atccycttc	5940
gttccatcgt	cgattgtcc	atccggatc	cycsgtggca	gcgtgtttgc	agttcttgc	6000
tagaaactat	ccaaatgtct	catcatacc	atffaccyac	gatgtatgtg	catacttgcg	6060
cattatcgac	gggttcggaga	tttgccttgg	cogaqgaca	ticaatccgt	caaaactca	6120
ggatcaccgg	aaacaycacy	tttaccacgc	gcctccatc	agaaaggcgt	taccgttccc	6180
tttccatggac	acactacaga	atgtacttgc	agcayccacg	aaaggaaact	gcacacgtcac	6240
acagatgggg	gaatttccca	cttggactc	agcagtatc	aaatgtgggt	gtttccaaaa	6300
atccatgtc	aaacaaagat	actggggaaa	atttgtctcc	agccatattt	ggataaaaaac	6360
tggaaatttt	ycacccatgt	ttatcttacc	aaaaaygcac	aaaggcgtcg	cyclatrtgc	6420
aaaaacccat	aatatctgtc	cactacggat	agtaccaatg	gatgggttca	cagtatgtat	6480
aaaaaggggac	gtaaagggtg	ctctgtgtac	aaaycataca	gaggaaazqac	cttgggtgca	6540
ggtttatacag	gggggtggaa	ccttggcgac	agcatatcta	tgtggatcc	acagagatc	6600
ggtttggggg	ctggaaacccg	cccttcttacc	caatgtatc	atccatatttgc	aaatgttgc	6660
cgaggatttc	gatgcataca	tagccycaca	ctttaegcca	ggagacactg	ttttggaaaa	6720
ggacatagcc	tccttggata	agggccaaag	tgatttcaat	geyctttatg	ctttgtatgt	6780
tttagggatg	tttgggggtgg	atccatcttgc	gttggacttc	atagggcttg	ttttccggaga	6840
ttttttccgg	tgtcacatc	cgacaggatc	ggttttccgg	ttagggggca	tgtgtaaatc	6900
agggtatgttc	cttacttctgt	tgttcaacac	atgtttaatc	atccatctgtc	ccggccgggt	6960
gttggaaatg	cgctctgacaa	atccggcggt	cgggcccttc	atggggggatc	acaaacataat	7020
acatggggtc	gttcccgatg	atttgttgc	agccatgtgt	qcacttgg	tgacatgg	7080
agtgtaaatc	atagatgtc	ttttatectt	gaatggccct	tatittttg	gagggitttt	7140
actgtcacat	acrgtgcacag	gaacggatc	caaggtggca	gacgggtttaa	aaagggtttt	7200
taaaactgggc	aaacacgtcg	cggtgggtga	cgaaacatgt	gaatggatgg	gacgagcgt	7260
ggctgtacgca	gtgtatcgat	ggcaacggaa	agggtttttt	gtatggctgg	aaaggccgt	7320
atactcttgc	tacgtaaatgc	gggtgtatcc	aggtgtatcc	atgttccatgg	ccaccccttgc	7380
aaatgtccaga	tccaaatcccg	agaatgttc	aggacccgtc	ataaatccgt	acgggggtcc	7440
ttttatgggtt	cgccatccat	ctatcttattt	tgcacaaatcc	gacggccatgt	atctaaacac	7500
ttatctatgt	caatggatgt	caccccaac	tttttttttt	aaaaatgggg	gtatccatgt	7560
cgaccccttgg	ctccggcccc	tacttccaa	gttatacgcc	ccggacccgg	ccctccatgg	7620
aaaggctggcc	aaatgttccca	gttgcatttc	aaatgtatcc	gggggggttc	ccgggggggt	7680
cccccaaccc	agccacccgg	gaatggaaat	aaaaatggaa	aaaaacggaaa	ccacacggcc	7740
ccacaaaaaca	acacaaatca	aaaaaaggcg	ccacccaaaa	aaaaacccgg	ttttttggaa	7800
aaqaatggcc	ggccggccgg	ggggatgtgc	atgtttatcg	aaaaatgttg	tattttggaa	7860
gttcaatggcc	aaatgtatgt	aaatgggttc	gggtgtttcc	ttttttggaa	atgtatgtaa	7920
ccaggccatcc	ttttatgggg	catcgatcc	ggggccctgg	ccacccatgg	ttttttgggg	7980
ttatcttcaat	atgtatccat	atggccggcg	atccatgttc	acatgtatgc	ccggccgttc	8040
aaatgttccatc	atggggatcc	gggggggttc	tatccatgtc	acccatggcc	atgtatgttc	8100
tcaggggggcc	gtttccatgt	ccctatcggt	ccggggatcc	ccggggatcc	ccggggatcc	8160
atcttccatca	aaatggggcc	ccgtgttgc	atgttttttg	ggggggatcc	ccggggatcc	8220
cgtagccccc	tctccgtgt	gacccatgt	aaaaatgggg	tttttttttt	ccggggatcc	8280
ggggccggaa	atgtggatct	tgtccatcccc	tttatagtgg	tgttgcacaa	ccacccgttc	8340
ccctgttccca	aggcccccctt	ccggccatgt	ttgtatccaa	ggggccgggg	ggggccccc	8400
cycatgttttgc	aggacaaacgt	catggatcc	gggttacttcc	atgtgtatca	agcatccat	8460
acatgttttc	cccaacccggcc	ggggccggcc	accaaggaca	ttttttatgt	cttataaagcc	8520

FIGURA 7 (cont.)

acaagaccat	acttagctca	ctgtcccgac	tgtggagaag	ggcactcg	tgccatagtccc	8580
gtggcactag	aaacgcattcag	aaatggaaagcc	aaagacgg	ccgtgaaaat	ccagggtctcc	8640
ttggaaatcg	gaataaagac	ggatgacagd	cacgattg	ccaadgtgg	ttatatggac	8700
aaacmacatcg	wagcagacgc	agagggcc	gggttattt	taa	aaacccatcg	8760
acgttactcg	gaaacatgg	acatccatc	ctggcccgat	gtccaa	aaaaatctcg	8820
acgggtggat	tcactgacag	tagaaagatt	atgttactat	gtacgcaccc	atttcaccac	8880
gaccccttcgt	tgatagtg	ggaaaaattt	cattcccgac	cgaa	ggatggatgt	8940
acttgcggca	cttacgttgc	gagcaccggc	gaaactcc	aggagataga	gttacacatg	9000
ccccccggca	ccccctgtatcg	cacattaatg	tracacatg	ccggccacgt	aaagatccaa	9060
gtcaatggcc	ggacgggtgc	gtacaaatgt	atgttgcgt	gttcaatgt	aggacttaca	9120
actacacgaca	aaatgtttat	taactgc	gttgcgtat	gttgcgtat	gttcaaccaat	9180
cacacaaatcg	ggcgttat	tttttttttgc	gttgcgtat	atgttgcgt	ttggggccca	9240
aaaggaaatcg	tttccatccc	gttcccg	ccaaatgtaa	catgcggat	gttcaatgt	9300
ggggccccc	ccgttgcgt	ccggaaaaac	caatgtatca	tttgcgtat	ttttttttcc	9360
ccaaacactcc	tgtttttatcg	aaatgtggg	aaatgtatca	aaatgtatca	aaatgtgggt	9420
atgcataaaga	aggaaatgt	gtttttatcg	ccgttgcgt	gggttgcgt	ccatgtggcc	9480
aaucacccgg	ccgttgcgt	tttgcgtat	ttatcttcc	aaatgtatca	ccatggccac	9540
ccgttgcgtat	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	9600
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	9660
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	9720
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	9780
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	9840
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	9900
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	9960
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10020
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10080
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10140
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10200
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10260
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10320
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10380
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10440
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10500
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10560
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10620
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10680
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10740
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10800
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10860
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10920
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	10980
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	11040
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	11100
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	11160
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	11220
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	11280
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	11340
tttttttttgc	tttttttttgc	tttttttttgc	ttatcttcc	ttatcttcc	ttatcttcc	11400

FIGURA 7 (cont.)

cataaaatgg tatacggtgc ccctaagaga cacatgtat gtatggata agtatacata aaaggggccgaa ataaacccctg aatagtttca aaatatgaaa atcaataaaa atcataaaaaat agaaaaaaacca taaaccaaaag tagtttcaaag ggctataaa cccctgtataa gtaacaaaaac ataaaaattaa taataatc	11460 11520 11580 11598
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----------------------------------

**FIGURA 6**

caaaggcaaga	gatctaataaac	ccatcatggaa	tcctgtgtac	gtggacatag	acgtgcacag	60
cgccttttg	aaggccctgtc	aacgtgtgtt	ccccatgttt	gagggtggaa	caaggcagg	120
cacaccgaat	gaccatgtctt	atgtcgatgtc	gttcttcgtat	ctagctataa	aaactaataga	180
gcaggaaatt	gaccctgtat	caaccatctt	ggatatcgcc	agtgcggcc	caaggaggat	240
gtatgtcgcc	aggasgtacc	actgcgtctg	cccgtatgtcc	agtgcggasg	atccccggat	300
actgcgtat	tatgtcgatgt	actgtatgtc	ttccgtatgg	aaaggcttgg	acagaaaact	360
ctctggaaag	atcggtttact	tacaatgtt	tttttttttt	ccagacacgg	agacgcccac	420
ttttttgttt	ccatcatgtt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	480
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	540
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	600
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	660
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	720
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	780
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	840
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	900
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	960
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1020
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1080
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1140
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1200
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1260
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1320
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1380
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1440
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1500
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1560
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1620
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1680
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1740
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1800
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1860
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1920
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	1980
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2040
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2100
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2160
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2220
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2280
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2340
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2400
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2460
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2520
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2580
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2640
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2700
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2760
tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	2820

**FIGURA 8 (cont.)**

taatgaaaaac cccgtctatg catccacgtc  
ggaaaggtaaa ctggatggc agacacttc  
ccccacccaaa ggcacattca aagcaacatc  
aatggggggc atctgcagtc accaaatgac  
ttgttgggtc aagagtttg tcctatctt  
ccagtgtct cagatatttc aagccatcc  
cttgcatacgat atatgttgcgatgtatgg  
acgttgttgc tctgtgttatt acggggatcc  
gttcggatcc aaccggatgg cgtatccat  
ggaaagtggaaat atcaacaagg aagatcgt  
taccaccaac atcaatccgg cccatccatgg  
cccagtaaaaa cgggaasggaa tggatggct  
cttggtcagt ggtatataccc ttgcactggc  
agggtgcggc ggagoggact acacatccaa  
taggtatgtac ctatgtgtca taacatccca  
gtgtgtgtac caccggatgg aactgtcaat  
acccggggcc ttcctatgg tcagggatcc  
catctggatc ttggggatgg agttttgc  
cggcaacact gagatgtttt tccttattcc  
aactcatgc atgaccaatc aactgtatgc  
atgtgcaccc bggttacccggg taacacgtatc  
agtcacccgc gctaacccttcc gggggtttcc  
aaaatggccg ggttcccttta agaaacgttgc  
gtgtgtgtac tttttatgttttccatgttgc  
tggaaaggccg cggaaattttc cagttgtcc  
ggggatgtttt agtgtatgtca tacettcttct  
cagggtgtcc cagtctatgtca accatcttcc  
ggtcatctac tgeggccaca aagatgttgc  
gaccraagta gacgtgtgtcc atggugcacat  
ccctgtacccg agtttggggatc gcaaaaaagg  
statcttagaa gggacccgtt ttcatcagac  
gtggccaaag caaacaggcc ccaatggcc  
tgaatccatc aacccaaatc gccccgggtgg  
tgcocccgtgc ctgttgcgtt acgttatgtac  
ccacgttcaaa aacatattttt tggtgttcc  
aagtccaaagg ctcaatgttgc ttaagggtat  
aagtccaaagg gaaatataatgtt ctccatgg  
actgtacccat aacatcaatgg acctaaatgt  
cctggatgtc gacggccccccg ccctgtaccc  
gcacatccaca aacggaaaacc ttgggggggt  
cggccggcc aacggaaaggcc gggggggatc  
gaataatccatc cccatggatc ggttgcgtt  
aagaaacggcc gagaacgttgc acacacatcc  
ggaaacggatcc catccggccat ttccttcc  
ttggggatcc aacggggatcc aatccatgg  
cttttacccaa gggggatgtgg atgttttgc  
ggacggacggatcc ttatgtatcc aacggggatcc

**FIGURA 8 (cont.)**

**FIGURA 8 (cont.)**

FIGURA 9 (cont.)

aggatatatgt gtccccctaag agacacatcg tacatgcgg ataatctata gatcggaggc	11349
ctacgcaccc cctgaaatgt aacaaaatat aaaaatcacta aaaaattataa aaacagaaaa	11400
atacataaat aggatatacg gtccccctaag agacacatcg tatgttaggtg ataaatgtatag	11459
atcaaaaggcc cgaataaccc ctgtatagta acacaaatgt aaaaatcaata aaaaatcataa	11520
atatggaaaa ccataacacag aatgtttca aagggttata aaacccctgtt atatgttacaa	11589
aacatataat tataaaaaat c	11601

**FIGURA 9**

cagtttctta	ctgcgttact	ctgcggcggca	agagattat	aaccatcat	ggatctgtg	60
tacgtggaca	tagacgttga	cggcgcccttt	ttaagggucc	tguacgttc	gtaccatgt	120
tttgggtgg	aarcaggca	ggtcacacgg	atggccatgt	ctaattgttg	agcytfctcg	180
catrtaagttt	taaaaactsat	agaycaggaa	atggccccc	actccaccat	cctggatata	240
ggcagtgcgc	cagaaggag	qatgtatgtcg	gacaggaaat	accactgtt	ctggccatgt	300
ccgcgtgggg	asgatccccg	gagactcgcc	atttatgtgg	gaaagcttgc	atctggcgca	360
ggaaasgtcc	tggacagss	catatcttgg	magatccccg	acttacaaac	sgtaatggcc	420
gtgcacggaa	cgggacggcc	aacattttgc	ttacacacag	acgttcttgc	tagacagaga	480
gcgacgtcg	cttatatact	aaaggtttat	ytgtatccac	gtttatccat	gtttatccat	540
caggccatgt	aaagggttcc	agtgggttac	tgggttgggt	togacacaaac	ccggttcatg	600
tacatgttca	tggcgggttgc	ctatccatca	tactcgaaat	actggcaga	tyggcgggtt	660
ctggaaatgtt	agaacatagg	attatgttca	acaaacatgt	eggaaggatg	acggggcaag	720
tttgttattt	tgagagggaa	aaagctaaaa	ccgtggggace	gtgtgtgttt	ctcgttgggg	780
tcacatgtct	acccggaaag	ccggcggctt	ctttagggct	ggccatctgc	atrggtgttc	840
catttaaagy	gcacactcg	cttacatgtgc	ccgtgtgtat	cagtggtttc	gtgttggggc	900
tacgtcggtt	aaagggatcac	gtatggccca	ggccctttat	gaaaaaccac	agggatgtcg	960
gtaaacccacc	acggacggcc	atccatgtat	tcacacacta	ccggacgggt	tgcggggcga	1020
agagtgttat	tcgggtgtt	ccatctgttg	ccggggacaa	ttttgtatgt	aatggccgtc	1080
atccatgttca	tcggatgttc	ggccggggat	gtatggaaage	tgtttgttgg	gtgttgggg	1140
agaatagtgg	ttaacggcag	aaacycaacgg	ataatggaaaca	ccatgtaaaa	ttatctgttt	1200
cccggtggcc	ccccggccctt	cgttacatgttg	gtttttttgt	gtttttttgt	gtttttttgt	1260
gaaaaactcc	tgggggttca	agaaaagacca	ctgacccgtt	gttgttctat	ggccatttcaag	1320
aaggcggaaaa	cacacacgg	cttacaaaggg	cctgtatcc	agtcacatca	gaagggttccag	1380
ggcgagttt	acagttttgt	ggtacccgt	ctgttgggt	ccgggttgc	aatcccttgc	1440
aggacttggaa	tcaatgtttt	gttacggcaag	gtgtccaaaaaa	ccggccctgtat	cccatccatgc	1500
ggagatggccc	gayaaggcc	ggacggccaa	aaaggacgg	agggggaaacg	gyaaggcggaa	1560
ctgacttgc	asggccatcc	acccatctac	ccggccatgg	aaatgttttc	gttggggatcc	1620
gacgtggaa	agcttgcgg	ccggccggcc	ccggggatata	ttatggactcc	ggtggggatcc	1680
atccaaatgtt	ctgcggccat	aaacggaccac	gttggggatgg	agtaactgtt	actcttcccg	1740
cggccgtatc	taatgttgc	gaaatgttgc	cttggggatgg	gtttttttgt	gtttttttgt	1800
acgtgcacgc	acaaacggacy	ggccgggggg	ttatggggatgg	aaatgttgc	ccggccggatc	1860
ctatgtccct	caggctatgt	atccatgttgc	aaatgttgc	aaatgttgc	aaatgttgc	1920
acgtatgtgt	ataacggaaa	agagtgttgc	aaatgttgc	aaatgttgc	aaatgttgc	1980
ggacccggcc	tcaatccatca	cgaaaggatcg	tatggatgttgc	tggggggcaga	ggggccggaa	2040
cacggatgtc	tcatgtatgt	gtatggatgt	agatgttgc	aaatgttgc	aaatgttgc	2100
ctgggtactgg	ttgggtgtt	gtatgtatgt	ccatccacat	ttatggatgt	ttatggatgt	2160
aaaaatccggc	ctgtatgttgc	atccatgttgc	ccatccatgt	ttatggatgt	ttatggatgt	2220
tctggcaatgt	cgttattat	caagaaatcc	ccatccatgt	ttatggatgt	ttatggatgt	2280
aaaggaaatgg	actggccatgg	aaatccatcc	ccatccatgt	ttatggatgt	ttatggatgt	2340
tctgcacgtt	gggttgcactc	gttgcgttgc	aaatggatgt	aaatggatgt	aaatggatgt	2400
taatgttgc	gggggttttgc	gttgcgttgc	ccatccatgt	ttatggatgt	ttatggatgt	2460
ggcccaatggc	agaaaatgttgc	acttgcgttgc	ccatccatgt	ttatggatgt	ttatggatgt	2520
atgcgtatgt	aaatgttgc	taatgttgc	ccatccatgt	ttatggatgt	ttatggatgt	2580
tcggccggatgt	gtatgttgc	gtatgttgc	ccatccatgt	aaatgttgc	aaatgttgc	2640
atgtgtatgt	caaataatgt	aaatgttgc	ttatggatgt	ttatggatgt	ttatggatgt	2700
cctgacccgt	gatgttgc	catgtatgt	ttatggatgt	ttatggatgt	ttatggatgt	2760
gactatgttgc	gatgttgc	catgtatgt	ccatccatgt	aaatgttgc	aaatgttgc	2820

**FIGURA 9 (cont.)**

gtttacgttag	ttagacaaaa	agttaatgaa	aaccgcgtct	atgcattcaac	gtcaggagcac	2880
gtcaacgtac	tccttaacgcg	tacggaaagt	aaactggtat	ggaaggacact	tttccggcgac	2940
ccgtggataa	agacgtgtca	gaacccuaccg	aaagguaact	tcaaaggcaac	tatthaaggag	3000
tcggaggtgg	egcatgtcata	aataatggcg	gycatctgca	gtcaccasat	gaacctttagat	3060
acatccccaa	ataaaagccaa	cgttttgtgg	gtcaagaaat	tggtccctat	cctcgtaaaaaa	3120
gcggggataa	aactaaatgt	tagcactgg	tctcagataa	ttccaaaggccc	caaagaagac	3180
aaagcataact	cacctgaagt	agccctgtat	aaaaatgtta	cgccgtatgt	tgggggtggat	3240
ctagacsgcg	ggctattttc	taaaccgttg	gtgtctgtgt	attacggga	taaccactgt	3300
gataatggc	ctggggggaa	aatgttccgg	ttttaaccctcg	aggcggccatc	cattctagaa	3360
agaaaagtatc	cattcacaaa	aggggggtgg	aacatcaaca	agccatgtcg	cgtgactacc	3420
aggaggatag	aagactttaa	cortaccacc	aacatcatac	cgccccacag	ggggactacca	3480
cactcattag	tggcogaaca	ccggcccgat	aaagggaaa	gaatggaaatg	gtgggttaac	3540
agatataacg	gccaccacgt	gtctctggtc	agtggctata	accttgcact	gcctactaaag	3600
agatgtactt	gggttagcgc	gttaggtgtc	ccggggagccg	actbacacata	caacccatagaa	3660
ttaggtctgc	cagcaacgt	tggtaggtat	gacccatgtgg	tcataaataat	ccacacaccc	3720
tttgcatacc	accattacca	acatgtcgat	gaccacgcac	tgaaatgtca	aargctcgaa	3780
gtgtactcat	tgagactgt	aaaaacgggc	ggctcttata	tgatcagagc	atatgtttac	3840
cccgatagaa	ccaggtaacg	gtgtatctgc	gtattggcac	gcasgttttg	atcgctctaga	3900
ccgttggaaac	caccatgtgt	caccaggcaac	actgtggatgt	tttttcttatt	cagcaacttt	3960
gacaaatggca	gaagggattt	caaaactcat	gtcatgaaac	atcaactgtaa	tgcggcccttc	4020
gttggacagg	tcaccccgac	aggatgtgca	ccgtcgtaac	gggttaaaaa	catggacatc	4080
ccggaaacg	atggggatgt	cgtagtcaac	gccgtcaacc	ctcgggggtt	acgggggtgc	4140
gttgttgcg	aggcgatata	aaaaaaatgg	ccggggatct	ttaagaaacgg	tgcacacacca	4200
gttggggaccy	caaaaaaaatgt	tatgtgggtt	acgtatccatg	taatccacgg	tgttggacca	4260
acttttctca	attattcgg	gtctgggggg	gaccgggat	tgccagctgc	ctatcgagaa	4320
gttgcgggg	aaatgttgcg	gttgcgggg	tttttttttt	ctataatctt	cttctccaca	4380
gttgtataact	ccggggggaa	agacagggtt	accgggttac	tgatcaccat	tttttacagcc	4440
atggacttoga	ccggatgcaga	cgtgggtatc	tactggcccg	acaaaagaatg	ggggggggaaa	4500
atatatgtgg	ccatatacgat	ggggggccaa	gttagatgtc	tggatgagca	cattttccata	4560
gtactggatata	ttgttgcgt	gcaccctgtac	ggcggatgtgg	caggcggaaa	gggatccacg	4620
accacccggaa	ggccgtatgt	ctatatactca	ggggggaccc	gtttttatca	gggggggtgg	4680
atatacgccg	agatatactac	tatgtggcca	ggggggaaatgg	ggggccatgt	gcgggtctgc	4740
ctatatacc	tggggggaaaag	tatgtgtatcg	atcaggccaga	aatggcccggt	ggatgtatca	4800
gacgcataat	ctcccccacaa	aaatgtcccg	tgccctttggc	gttangctat	gactccggaa	4860
ccgttgcaccc	gggttgcgtat	gaccggccgt	acaaatgtaa	ttgtgtgttc	ttrgtttcc	4920
ctcccaaaatgt	acaaaaataga	aggaggatgtaa	aaatgtcaat	gttcttaatgt	aatgtcttattt	4980
gaccacaaacg	tgccatcgcc	cgttggatgtca	aggggatata	gttctccca	ggagggtctgc	5040
ccggggggcc	gttacatccac	gttactgtacg	catagtcaat	tgcacccatgg	cggtgtatggc	5100
gagatactgc	ccgtcccgat	agacgtggat	gttgcacgccc	ccggccatgtaa	accggccatca	5160
ccggccgggg	ccggccatccac	gttgcattcc	acaaatgggg	accttgcggc	cggtgttccaa	5220
ccggggccgggt	gggttataat	tctcggtgg	ccgggggttca	gttcttttac	acacggaaatgc	5280
atgtacgtttat	tcagtgtgtc	ccgttggatad	ccggggggaa	gttccacgggg	gggggtgttt	5340
cccaacccaa	ctggatgtgg	caaaaggacca	acttattactt	aaatgtttatc	ggggggatgtgc	5400
atccatgtcc	aaacggaaatgc	ggtatacgat	ggggggaaatgt	gatggatgttt	ggggggatccat	5460
catccatggaa	ctaaagagaa	gttggatgtact	atacttaatgt	tcggggaccc	caaaaggatcc	5520
tacttacgg	actacatatac	ggggggatgt	gtactgtatc	ccggatcaacg	tcggattgttc	5580
caatcccgag	tcggggatgt	ccggatgtatc	gttggatgttt	gttggatgttt	atccatgtatc	5640

FIGURA 9 (cont.)

ctcatacatac caaaatcccc accggatgtca	tgcataatcta gacatggtgg accgggtcgaa	5700
gaggttccctg gaccggatgtca cattcaatcc	gtccaaaactc aggaggtacc cgaaacacgca	5760
cgttttccacg gageccctccca taaaagaacgc	tgttacegttc ccattccaga acacactaca	5820
gaatgtactg ggacggacca cgaanagaaa	ctgcaacgtc acacagatgt gggatttacc	5880
cactttggac ttagcagtat tcaacgttgg	gtgtttcaaa aaatttgcgtt gcaaccasga	5940
atactggaa gaatttgttg ccacccctat	ttaggttacca actgagaatt tagccacata	6000
tgttactaaa ctaaaatggc caaaaggcgc	agcgcttacc gcaaaaaccc ataatctact	6060
ggccatcacc gaatgttccaa tggatagggt	cacatgtat atgaaaaggg acgtgtgggt	6120
gacttctgtgtt acacggatgtt cagaggaaag	acctaaggts cagggttatac agggggctgt	6180
accccttgggg Adagcataacc tatgttggat	tcacagagag ctgggttagga gggtgtgtgc	6240
tgttactata cccatgttac atacactatt	tgacatgtct gcccgggatt tggatgtccat	6300
catatccyca cacttttccac caggagacac	tgttittggaa acggacatag ctttcttttgc	6360
taaaggccaa gatgtatccat ttgtgtttac	tgttttgcgtt ctgttagagg atttaggggt	6420
ggatccatcc ctgtctggact tgatagggc	tgttttgcgtt gagatttcca gctgttaccc	6480
acccggacgggtt acgggttccaa agtggggcgc	catgtatggaa tccggatgtt tccataactct	6540
tttgcgttccaa acatgttacatccat	cccccggccgt gtgttggggat atcgatgtgc	6600
aaaatccyca tgcgggttccatccatggc	ccacaaacata atacatgggg togtttccga	6660
tgaatgtatg gtagccatgtt gggccacttg	gtgttgcgtt gatgtatgtt tccatgtgc	6720
atgttatacc ttgttggccat tttatcttttgc	tgggggggtt atactgttccy atactgtgtc	6780
ggggaaatgtt tgcggatgtt cagggccgtt	aaaaagggtt tttaaactcc gcaaaacccgt	6840
egccggccgtt gacgttccaa atgttgcgtt	aaacgttccgtt tttttttttt gatgtatccgt	6900
atggccacgtt acgggggttcc ttgtgtgtgt	ggggaaagggtt gtataacttca ggtacgttgc	6960
ggatgggtata tccgttgggtt tttgttccat	ggccacccctt gcaatgttcc gatccaaactt	7020
cgagaaatgtt acggggccgtt tccatgttcc	gtatgggggtt tccatgttcc tccgttccat	7080
atgttacccat ttttgcgttccaa cccacacgtt	gtatctaaatc actatgttccy tccatgttgg	7140
tttcatccaa cccaaatctt tttatgttccgtt	ttttttttccy tccatgttccy tccatgttgg	7200
cttactatcc aatgttccat gcccggatccg	cccccgttccy tccatgttccy tccatgttgg	7260
ttgtgtgtatc tccatgttccaa taaaatgttcc	atgttccatccy tccatgttccy tccatgttgg	7320
aggaaatggaa agatataatgtt gcaaaatgtt	aaaaaaatccgttccy tccatgttccy tccatgttgg	7380
ccaaatggaa acggccatccaa aaaaatgttcc	gttccatccatccy tccatgttccy tccatgttgg	7440
ggatgggtatgtt gatgtatgtt cccatgttcc	aaaaatgttccy tccatgttccy tccatgttgg	7500
gttccatccatccy tccatgttccy tccatgttcc	aaaaatgttccy tccatgttccy tccatgttgg	7560
acccatgttccy tccatgttccy tccatgttcc	gttccatccatccy tccatgttccy tccatgttgg	7620
ggatgttccgtt tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	7680
gttccatccatccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	7740
atgttccatccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	7800
ccatgttccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	7860
gttccatccatccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	7920
atgttccatccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	7980
tttgcgttccatccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	8040
tttgcgttccatccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	8100
ccatgttccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	8160
ccatgttccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	8220
ccatgttccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	8280
ccatgttccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	8340
ccatgttccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	8400
ccatgttccy tccatgttccy tccatgttcc	ccatgttccy tccatgttccy tccatgttgg	8460

**FIGURA 9 (cont.)**

**FIGURA 10**

FIGURA 11

ataccgggtc	acatggaaatc	cgacgttcc	aaggttcaacc	atggaaaacc	gggggggtac
tacaacttgtc	accacgggc	agtacagta	tcaggaggcc	ggtttccat	ccctacsggt
gctggcaaac	caggggacag	ccggacacog	atcttgcaca	acaaagggaacg	ctgtgttgccc
atagtcttag	gaggagctaa	tgaaggagcc	egtacagcccc	tctcggtegt	gacctggaaat
aaagacattt	tcactaaaaat	caccccgag	ggggccggaa	egtggagttt	tgccatccccaa
gttatgtccc	tgttggcaaa	caccacyttc	ccctgttccc	agcccccttg	cacggcccttgc
tgctacgaaa	aggaaacccy	ggaaacccctt	cgcatgttgc	aggccaaacgt	cttgagaccc
gestactact	agctgttaca	agcatcotta	acatgttttc	ccacccggca	ggggggggggc
accaaggaca	acttcaatgt	ctataaaggc	acaaagccat	acttggccat	ctgtttccggat
tgtggggaa	gycacttgc	ctatagttcc	gttgcactag	aaacgtttcag	aaatggggcc
acagacggg	cgctgaaaat	ccaggattec	tttgaaatcg	gaataaagac	ggatggacagc
cacgatting	ccaaatgttgc	ttatatggac	saccacatgc	cggccggccgc	aggggggggg
gggttatttgc	taaaggacata	agttttttgt	acgattactg	aaacaaatggg	acacttccatc
ctggcccgat	gtccaaaagg	ggaaacttgc	acggtggat	tcactgacag	tagggatgtt
agtcactctat	gtargcaccc	ttttcccccac	gaccccttgc	tgataggtcg	ggaaaaatcc
catteccgac	cycagccacgg	taaaggacata	cttttgcggcc	cgtaatgtca	gaggccacggcc
gcaactacc	xogagataga	gttacacatg	cccccggac	cccttgcatacg	ccatattttatg
tcacaaatgt	coggccaaatgt	aaagatataa	gtttttgtgt	aaacgggttgc	gtttttttttgt
aatttgcgtt	gttccaaatgc	aggacttca	actacagaca	aaatgtttttaa	ttttttttttgt
gttgcataat	gttcatgtccgc	ggtcaccaat	cacaaaaatgt	ggccgttataaa	ctccccccctgt
gtccccggcgt	atgttgcact	ttggggccgg	aaaggaaaaaa	tttcacatcc	ttttttttttgt
gcaaaaatgtaa	cattgggggt	gtttttttttgt	aggaaacccca	ccgttgacgta	ccccccccaaaa
caagtcatca	tgctactgtt	tccctggccac	ccaaacactcc	tgttccatcc	gaatatggga
gaagaaccaa	actatcaaaa	gggttgggttgc	atgtcaataaa	aggaaagtgt	gttcaacogtg
ccgactggaa	ggcttccgggt	cacccggggcc	acaaacccggc	cgttataatgt	ttggcccccgg
tttatotatcaa	acggatcage	ccatggggccac	ccgtatgtgg	taattttgttgc	tttattttatgt
ctgttccccc	cttttttttgt	atgtttttgttgc	ttttttttttgt	ttttttttttgt	ttttttttttgt
gttgggtatgg	cacggggggat	gttgcattttgt	ccatggggcc	gttgcattttgt	ttttttttttgt
ctgtacaccag	gagttttttgt	ccctttttttgt	ttttttttttgt	ttttttttttgt	ttttttttttgt
aaatggggggca	cataccaaat	ttttttttttgt	ttttttttttgt	ttttttttttgt	ttttttttttgt

**FIGURA 12**

ataccgggtgc	acatggaaagtc	cgacgcttgg	aaggttccaccc	tttgagaaaaac	gggggggtac	60
taccaactggc	accacggggc	agtacatgtac	tcggggggcc	ggttccaccaat	ccttacagtt	120
gttggdtaaac	caggggcacag	ccggcagaccc	atctttggaca	aaaaaggaaag	ctgtggggcc	180
atagtcttag	ggggggctaa	tggggggggcc	cgatcagccc	tctcgatgtt	gaccctggaaat	240
aaaggacattt	tcactaaat	ccatccccgg	gggggggggg	agggggggttt	tgcctatccca	300
gttatgtgcc	tgttgttgc	ccccccggcc	ccctgtcccc	agecccccgg	cacggccctgc	360
tgtctacvaaa	aggaaacccgg	ggggggggcc	cgatcggttt	aggacaaatgt	catggagaccc	420
gggttactatc	agetgttca	agccatccca	acatgttcc	ccccccggcc	gtggacggcc	480
acccaggaa	atccatgt	cttataaagtt	aaaaaggacc	atcttgcctca	ctgtggccgg	540
tytgggggggg	ggccatcg	ccatgtcc	gtggggggcc	aaacggatcc	aatgggggg	600
ccacggggga	cggtggaaat	ccgggttcc	tttggggggcc	gatggggggcc	ggatggccgg	660
cacggatgg	ccggatggcg	tttataggac	sacccatcc	ccggggggcc	ggggggggcc	720
ggggcttttg	tttggggggcc	acggatccgt	acggatccgg	ggacacatgg	acactttccat	780
cttggcccgat	gtccggaaagg	ggggggccgt	acgggtgggg	ttactgtacac	ttggggggat	840
agtcaatccat	gttccggcc	atttccaccc	ggccctccgt	tgtatgggtt	ggggggggatcc	900
cattccggcc	ccggggccgg	taaaggatgt	cccttggggcc	cgatcggtca	ggggccggcc	960
ggccatctcc	aggggataga	gggtacatcc	ccccccggcc	cccttggatcc	ccatggatcc	1020
ccatggatgt	ccgggggggt	aaaggatcc	gtttttttcc	ggggggggcc	gttccgggtt	1080
atgggggggt	gttccggatgt	aggacttaca	acttacggcc	aaatgttttt	ttttttttcc	1140
gtttgtatcaat	gttccggcc	ggggggccat	ccccccggcc	ggggggggcc	ttttttttcc	1200
gttccggcc	atgtctggact	ttggggggcc	aaaggggaaa	ttttttttcc	ttttttttcc	1260
ggccatgtaa	catggccgg	gtttttttcc	ggggggccca	ccggggggcc	ccggggggcc	1320
ccggccatcc	ttttttttcc	ccccccggcc	ccggggggcc	ccggggggcc	ccggggggcc	1380
ggggggccaa	acttacggcc	ggggggcc	ttttttttcc	ttttttttcc	ttttttttcc	1440
ccggccatcc	ggggggcc	ttttttttcc	ttttttttcc	ttttttttcc	ttttttttcc	1500
ttttttttcc	acggggccgg	ccggggggcc	ttttttttcc	ttttttttcc	ttttttttcc	1560
ccggccatcc	ttttttttcc	ttttttttcc	ttttttttcc	ttttttttcc	ttttttttcc	1620
gtttttttcc	atgtatgtt	ttttttttcc	ttttttttcc	ttttttttcc	ttttttttcc	1680
ttttttttcc	ttttttttcc	ttttttttcc	ttttttttcc	ttttttttcc	ttttttttcc	1740
aaaggggggcc	cataccaacg	ggggggccat	ttttttttcc	ttttttttcc	ttttttttcc	1797

**FIGURA 13**

aatcccggtgc	acatgaaagtcc	cggacgttccg	aaatgtccccc	atggaaaaacc	ggaggggtac	60
taccaactggc	accacggagc	agtacagtac	tcaggagcc	ggttcaccat	ccctacagggt	120
gttggccaaac	caggggacag	cgccagaccc	atcttcgaca	acaaggggacg	cgtgggtggc	180
atagtcttag	ggggagctaa	tgaaggggcc	cgtacagcc	tctcggttgt	gacctggaaat	240
aaagacatcc	tcactaaat	cacccccggg	ggggccgaaag	agtgggggtt	tgccatcccc	300
tttatgtgc	tgttggccaa	cacccatcc	ccctgtcc	agcccccttg	cacccccctgc	360
cgctacgaa	aggaaaccgg	ggaaaccctt	ccgtatgttgc	aggacaaatgt	catgagaccc	420
gggttactatc	acgttcttca	acgttcttca	acatgttcc	ccraccggcc	gegacggcgc	480
accaaggaca	acttcaatgt	ctataaagcc	acaagacccat	acttagtcca	ctgtccccgac	540
tgtggagaag	ggcactcggt	ccatagtc	ctagcactag	aaacgtatcg	aaatyaagcy	600
acacacccgg	ccgttcaat	ccatgttcc	ttgcataatcg	aaataaagac	ggatgacage	660
cacgatggca	ccaaatgttgc	ttatgttgc	accacatgc	cagcagacgc	agagagggc	720
gggttatattt	taaagacatc	acgttccgtgt	acgattactc	aaacatgggg	acacitccatc	780
ttggcccgat	gttccaaatgg	ggaaatcttg	ccgggtggat	tcactgacag	taggaagatt	840
atgtcaatcat	gtacqccaccc	atttcaaccac	gaccccttc	ttgtatggcc	ggaaaaatcc	900
cattcccgac	cggcggccgg	taaaggatata	ccttgcagcc	ctgtacgtgc	gagcaccycc	960
gcaactaccg	aggagataga	ggtacacatg	ccccccayac	ccccctgtatc	ccatattatcg	1020
tcacaacatgt	cggccacgt	aaagatccca	gtcaatggcc	agacgtgtgc	gtacaatgtt	1080
atttgcgggt	gttcaatgt	aggacttaca	actacggaca	aaatgtatcaa	taactgtcaas	1140
gttgcataat	gtcatggcc	ggtcacccat	cacaaaaatgt	ggcgtatata	ctccccccttc	1200
gtccccgggt	atgttcaat	tggggaccga	aaaggaaaaat	ttccacatccc	gtttcccgctg	1260
gttgcataat	catgcacccgt	gccttaaagca	aggaaacccca	ccgtgacgt	ccccggaaaaac	1320
caagtcatca	tgtctactgt	tcctgaccac	ccaaacactcc	tgtcttccac	aaatatggga	1380
ggagacccaa	cttataaaga	agatgtggty	atgtatata	aggaaggtgt	gtcaacccgt	1440
ccgtactgtaa	ggctcgaggt	cacgtggggc	aaacaaacggcc	ctgtatata	ttggccugcag	1500
ttatgtatca	acgttacatgc	ccatggccac	ccgtatgtat	taattttgt	ttattatgtg	1560
ctgtacccca	ctatgactgt	atgtatgtt	tcagtggcc	cttttataat	octgttggat	1620
gtgggtatgg	cacggggat	gtgtatgtgt	ggacggaccc	gtatgtatcc	accgtatgtaa	1680
ctgtacccca	gagcttccgt	ccctttccgt	cttgcgttcc	tatgtgtat	ccggacatgt	1740
aaagccggcc	cataccaaaga	ggctggccata	tacctgttgc	acggacggcc	acccttta	1797

FIGURA 14

gacaacttca atgtctataa agccacaaga ccatacttag ctcaactgtcc cgactgtgga	60
ggaggcact cgtgcatacg tccogtagca ctagaacgca tcagaaatga agcgcacagac	120
gggacgctga aaabccaggc atcccttgcac atcgaaataa agacggatga cagccacat	180
tggaccsacg tgcgttatat ggacaaccac atgcacagcag acgcacagag gggggggcta	240
tttgttaaaaa catcagdaco gtgtacgatt actggaaacca tgggacactt catccgtcc	300
ccatgtccaa aaggggaaac tctgacggtg ggatttcactg acatgtggaa gattagtccac	360
ttcatgtacgc accatattca ccacgaccct ctgtgtatag gtggggaaaa attccattcc	420
cgacccgacgc acggtaaaa ctcccttgc acgcacgttgc tggggggccac ccggccacat	480
ccggggggta tagaggtaca catggccccc gacacccctgc atgcacat aatgtccacaa	540
cagtccggcca acgttaaaaat cacgtcaat ggddagacgg tggggatcaa gtgtatgtc	600
gggtggctca atgaaggact aacaaatca gacaaatgtca ttataactg caagggttcat	660
aatgttcatg cccgggttccaa ccataaaaaa aatgtggcaat ataaactcccc tctggtcccc	720
cgtaatgtctg aacttggggaa cggaaaaaggaa aaaaatccca tccgttccc gtggggaaat	780
gttaacatgca ggggtgcataa agcaaggaaac ccacccgtca cgtacgggaa aaacccaaatc	840
atcatgttac tttatcttgc ccaccccaaa ctccgttccc acggggatata gggggggaa	900
ccaaacttcc aaaaaggatgttggatgtccaaac gagggggatata agtattggcc gcagttatet	960
acaaacggta cagccatgg ccacccgtat gagataatcc ttgtattatata tgagctgtac	1020
cccaactatga ct	1080
	1092

FIGURA 15

gacaacttca atgtctataa agccacaaga ccatacttag ctcaactgtcc cgactgtgga	60
ggaggcact cgtgcatacg tccogtagca ctagaacgca tcagaaatga agcgcacagac	120
gggacgctga aaabccaggc atcccttgcac atcgaaataa agacggatga cagccacat	180
tggaccsacg tgcgttatat ggacaaccac atgcacagcag acgcacagag gggggggcta	240
tttgttaaaaa catcagdaco gtgtacgatt actggaaacca tgggacactt catccgtcc	300
ccatgtccaa aaggggaaac tctgacggtg ggatttcactg acatgtggaa attccattcc	360
cgacccgacgc acggtaaaa ctcccttgc acgcacgttgc tggggggccac ccggccacat	420
ccggggggta tagaggtaca catggccccc gacacccctgc atgcacat aatgtccacaa	480
cagtccggcca acgttaaaaat cacgtcaat ggccacgggg tggggatcaa gtgtatgtc	540
gggtggctca atgaaggact aacaaatca gacaaatgtca ttataactg caagggttcat	600
aatgttcatg cccgggttccaa ccataaaaaa aatgtggcaat ataaactcccc tctggtcccc	660
cgtaatgtctg aacttggggaa cggaaaaaggaa aaaaatccca tccgttccc gtggggaaat	720
gttaacatgca ggggtgcataa agcaaggaaac ccacccgtca cgtacgggaa aaacccaaatc	780
atcatgttac tttatcttgc ccaccccaaa ctccgttccc acggggatata gggggggaa	840
ccaaacttcc aaaaaggatgttggatgtccaaac gagggggatata agtattggcc gcagttatet	900
acaaacggta cagccatgg ccacccgtat gagataatcc ttgtattatata tgagctgtac	960
cccaactatga ct	1020
	1080
	1092

**FIGURA 15**

ggccacttca	atgtctataaa	agccacaaga	ccatacttag	ctcactgtcc	cgactgtgga	60
gaaggggcact	cgtccccatag	tcccgtagca	cttagaaacca	tccgaaatgt	agcgacagac	120
gggcacgtca	aaaccagggt	ctctttgcac	atcgatataa	agacggatgt	cagccacgtat	180
tggaccaggc	tgcgttatat	ggacaaccac	atgcacggcag	acgcacagag	ggcggycta	240
tttgtsages	catacgtccac	gtgtacgtat	actggaaacaa	tggcacactt	catactggcc	300
cgatgtccaa	aaaggggaaaac	tctgtacgttg	gttatttcactt	acagttagaa	gtttagtccac	360
tctgttgcac	accatattca	ccacgacccat	ctctgtatag	gtccccggaaat	atccatattcc	420
ccacggcggc	acggtaaaga	ttttttttgc	acgcacgtacg	tgcacggcac	cccccgcact	480
cccgaggggaa	tagaggttaca	catacccccac	gacacccctt	atcgacattt	aatgttcacaa	540
cagtcccgca	acgttaaaggat	cacagtcaat	ggccragcgg	tgcgggttacaa	gtgttaatttgc	600
gttgcgttca	atgaaaaggat	aacaacttaca	gacaaatgtt	tttaataactt	caagggttcat	660
caatgttcatg	ccggggtttttac	caatcacaaa	aaatggccgt	ataactcccc	tctggtcccg	720
cgtaatgtt	aaatgggggtt	ccggaaaaggaa	aaattttccaa	tcccggtttcc	gttggcaaat	780
gttaacatgtca	gggtgtccat	agcgggggaa	ccacccgttg	ctgtacggggaa	aaaccgggtt	840
ttccatgttac	tgtatgttca	ccaccccaaca	cttcgttctt	acggaaatat	ggggagaagaaa	900
ccaaacatct	aaagaaaggat	gtgtatgtat	aaagganggg	tgcgttttttt	ctgtccggact	960
gggggggtctc	aggtcacgttg	ggggcaaaaac	gagccgtata	atgtttttttt	ccgtttttatct	1020
acaaacgtta	cagccatgtt	ggggcaaaaac	gagatataat	tgttatttttt	tgagctgtac	1080
cccaactatgt	ct					1092

## FIGURA 17

gacaacttca atgtctataa sgccacacaa	ccatacttag ctcaactgtcc cgactgt gga	68
ggaaaggcact cgtccccatag tcccccgtggca	ctagaaaaaaca tcagaaaatgaa agggacagac	120
ggggacgtga aaatcccaagt ctcccttgca	atccggataaa agggggatgg cagccacat	180
tggaccacaa tgccgttatat ggacaaccac	atgcacaggcag acgcacgggg gggggggcta	240
tttgttaaaga catccacccgg gtgtatcgatt	actggaaacaa tggggacactt catccctggcc	300
cgtatgtccaa aaggggaaatc ttgtatcgatg	ggatteacctg acacgttggaa gattatgcac	360
tcatgtatcc accatccatcc ccacggccat	atctggatag gtccggaaaaa attccatcc	420
cgacccggcc acggataaaa gatacccttg	acgcacgtacg tgcacggaccc cyccggcaat	480
acccggggaa tagayygtaca catgtccccca	gacacccctg atccgtacatt atgttcacaa	540
cagtcccgca acgtaaaaatg cacaatgtcaat	ggcccaacgg tgcggatccaa gttgtatcgc	600
ggtcggctcaa atgaaaggact aacccatccat	gtccaaaaatgtaa ttatataactg caagggttgc	660
caatgtcatg ccgcgggtcac caatccaaaa	aaatgttgcgt ataaatcccc tctggtcccc	720
cgtaatgtcg aacttggggaa cccggaaaaatgg	aaaaatccaca tccggtttcc gtcggcaat	780
gttaatcatca ggggtgtttaa agccaggaaat	cccaatgtga cgttccggaa aaacccaaatgc	840
atccatgttac tggatatccatca ccaccccaat	ctccatgttcc acggatataat gggggaaatggaa	900
ccaaacttata aagaatggatgt ggtgtatgtat	aaagaaaggaaatg tggatgttccac cgttccggact	960
ggagggttcgg aaggccatgtg gggccaccaac	gagccgtata atgtttggcc gcaatgttact	1020
ccaaatgttca cgcacccatgg ccacccgtat	gagatataatgttattatgttggatgttac	1080
ccatccatgtaa ct		1092

**FIGURA 18**

IPVHNKSDAS	KFTHEKPGY	YNNHNGAVQY	SGGRFTIPTG	AGKPGDSGRP	IFINKGRVVA	60
IVLGGANEKA	RTALSVVTNN	KDIVTKITPS	GAENSLAIP	VNCILLANTTF	PCSQPFCTPC	120
CYERKEPESTL	EMLEINVMRP	GYYQLLQASL	TCSPIRQRRS	TKDNFNVYKA	TRPYLANCSD	180
CGEGHSCHSF	VALERIRNEA	TDTGLKIQVS	LQIGIKTIDS	HNTWLRYMD	NHMPADREKA	240
GLFVRTEAPC	TITGIMGHFI	LARCPKGTL	TVGFTDSEKI	SHSCTHPPFH	DPPVIGRKF	300
NSRFQHCKEL	PCSTYVOSTA	ATTEEIEVNM	FPTDTPDTLM	SQSGCNVKIT	VNGQTVRYKC	360
NCGGSNNEGLT	TIDEVINNCK	VDOQCHAATVN	HKKNQYNPL	VPRNAELGDR	KGKIHIPPFL	420
ANVTCRVFEA	RNPTVITYGN	QVIMLLYPIN	PTLLSYRNNG	EFPNYQEENV	MHKEEVVLTV	480
PTEGCLEVTNG	NNEPFYKYNPQ	LSTINGTAHGN	PHBIIILYYE	LYPTMTVVVV	SVATFILLSM	540
VGMAAGMCME	ARRRCITPYE	LTPGATVPFL	LSLICCIINTA	KAATYQEAAI	YLNNECQPL	599

**FIGURA 19**

IPVHNKSDAS	KFTHEKPGY	YNNHNGAVQY	SGGRFTIPTG	AGKPGDSGRP	IFINKGRVVA	60
IVLGGANEKA	RTALSVVTNN	KDIVTKITPS	GAENSLAIP	VNCILLANTTF	PCSQPFCTPC	120
CYERKEPESTL	EMLEINVMRP	GYYQLLQASL	TCSPIRQRRS	TKDNFNVYKA	TRPYLANCSD	180
CGEGHSCHSF	VALERIRNEA	TDTGLKIQVS	LQIGIKTIDS	HNTWLRYMD	NHMPADREKA	240
GLFVRTEAPC	TITGIMGHFI	LARCPKGTL	TVGFTDSEKI	SHSCTHPPFH	DPPVIGRKF	300
NSRFQHCKEL	PCSTYVOSTA	ATTEEIEVNM	FPTDTPDTLM	SQSGCNVKIT	VNGQTVRYKC	360
NCGGSNNEGLT	TIDEVINNCK	VDOQCHAATVN	HKKNQYNPL	VPRNAELGDR	KGKIHIPPFL	420
ANVTCRVFEA	RNPTVITYGN	QVIMLLYPIN	PTLLSYRNNG	EFPNYQEENV	MHKEEVVLTV	480
PTEGCLEVTNG	NNEPFYKYNPQ	LSTINGTAHGN	PHBIIILYYE	LYPTMTVVVV	SVATFILLSM	540
VGMAAGMCME	ARRRCITPYE	LTPGATVPFL	LSLICCIINTA	KAATYQEAAI	YLNNECQPL	599

**FIGURA 20**

IPVHNKSDAS	KFTHEKPGY	YNNHNGAVQY	SGGRFTIPTG	AGKPGDSGRP	IFINKGRVVA	60
IVLGGANEKA	RTALSVVTNN	KDIVTKITPS	GAENSLAIP	VNCILLANTTF	PCSQPFCTPC	120
CYERKEPESTL	EMLEINVMRP	GYYQLLQASL	TCSPIRQRRS	TKDNFNVYKA	TRPYLANCSD	180
CGEGHSCHSF	VALERIRNEA	TDTGLKIQVS	LQIGIKTIDS	HNTWLRYMD	NHMPADREKA	240
GLFVRTEAPC	TITGIMGHFI	LARCPKGTL	TVGFTDSEKI	SHSCTHPPFH	DPPVIGRKF	300
NSRFQHCKEL	PCSTYVOSTA	ATTEEIEVNM	FPTDTPDTLM	SQSGCNVKIT	VNGQTVRYKC	360
NCGGSNNEGLT	TIDEVINNCK	VDOQCHAATVN	HKKNQYNPL	VPRNAELGDR	KGKIHIPPFL	420
ANVTCRVFEA	RNPTVITYGN	QVIMLLYPIN	PTLLSYRNNG	EFPNYQEENV	MHKEEVVLTV	480
PTEGCLEVTNG	NNEPFYKYNPQ	LSTINGTAHGN	PHBIIILYYE	LYPTMTVVVV	SVATFILLSM	540
VGMAAGMCME	ARRRCITPYE	LTPGATVPFL	LSLICCIINTA	KAATYQEAAI	YLNNECQPL	599

**FIGURA 21**

IPVHMGESDAS KPTHEKPEGY YNNNGAVOY EGGRTFTIFTG AGKFGDSGRP IFDNEGRVVVA	60
IVLSGANEKA RTALSVVTKN KDIVTKITPE GREENSLAIF VMCLLANTTF PCSQPFCTEC	120
CYERKEPEMTL RMLEINWPRP QYVOLLOQASL TCSFHRORRS TKDNNNVYKA TREVLAHCSD	180
CGEHSISCHSP VALERIRNEA TUGILKIQVS LOIGIKTDSR HDTKLRYMD NNMFDADERRA	240
GLFVETSAPC TITGTMCHFI LARCPKGSTL TVGFTOSKIJ SHSCTHPPMH DEPVIGREKIF	300
HSRPQNGKEI PCSTIVQSTA ATIEELIEVM EPOTFORTIM SOQSGRNKIT VNGOTVRYKC	360
NCGRSNESLT TIDKVINMCK VDQCHAATVN HKKNQYNPL VERNAEGLDR KGKIHIPPFL	420
ANVICRVPKA RNPITVYTKN QVDNLLYPDH FTLLSYRNG EEPNYQSERV MNKKEWLTV	480
PTEGLEVING NNSPYKYNWPO LSINGTANGH PHEIIILYYVE LYPINTVYVV SVNTFILLSM	540
VGMADGMCNC ARERCTPFE ITPGATVFFL ISLICCIRTA KAATYQENAI YINNEQOPL	599

**FIGURA 22**

DENVYKATE PYLANCPDG EGHSCNSPVA LERIRNEATD GTLKIQVSLQ IGKTDLSHD	60
WTKLRYMINH MPADAMRAGL FVTSAPCTI TGTMGHTFILA RCPKGESTLV GFTDSERKISH	120
SCTHPPHNDP PVIGREKFHS RFQHGKELPC STYVQSTAAT TEEIEVHNPP DTPORTLMSQ	180
QSGNVKITVN QUTVRYXCNQ GGSNEGLTTF DKVINNCKVD QCHAATVNHK KNQYNSPLV	240
RNAKLGIRKG KINIPPLAN VTCRVPKARN PTWVYGENQV IMLLYFDNPT LLSYERUNGEE	300
PNYQEEWYMH KKEVVLTVFT EGLEEVINGNN EPYKYNPQLS TNGTNGKPH EILILYYVELY	360
PTMT	364

**FIGURA 23**

DENVYKATE PYLANCPDG EGHSCNSPVA LERIRNEATD GTLKIQVSLQ IGKTDLSHD	60
WTKLRYMINH MPADAMRAGL FVTSAPCTI TGTMGHTFILA RCPKGESTLV GFTDSERKISH	120
SCTHPPHNDP PVIGREKFHS RFQHGKELPC STYVQSTAAT TEEIEVHNPP DTPORTLMSQ	180
QSGNVKITVN QUTVRYXCNQ GGSNEGLTTF DKVINNCKVD QCHAATVNHK KNQYNSPLV	240
RNAKLGIRKG KINIPPLAN VTCRVPKARN PTWVYGENQV IMLLYFDNPT LLSYERUNGEE	300
PNYQEENVMH KKEVVLTVFT EGLEEVINGNN EPYKYNPQLS TNGTNGKPH EILILYYVELY	360
PTMT	364

**FIGURA 24**

DNFNVYKATR PYLANCPDCG EGHSCHSPVA LERIRNEPATD GILKIQVSLQ IGIKTIDGND	60
WTKLRYMDNN MPAADAEKAGL PVNTSAEPTI TGTMGHETLA RCPKGESTLTV GFTDSRKISH	120
SCTHPPPHHPF PVIQEEKPHS RPQNGKEIFC STYVQSTTAT TEELEVNHMP DTFORTLMSQ	180
QSGNVKITVN GQTVRYKNCG GGSENEGLTII DVNINNCIVD QCHARAVTNHK KWQVNSPLVF	240
RNAELGDRRG KIHIPFFPLAN VTCRVEPNAM PTVTYGRNQV IMLLYFDMPt LLSSYRNNGEE	300
PTVQSEWVMH KREVVLTVPT EGLLEVWGMN EPYKYNPQLS TNGTANGMPH EILLYYYELY	360
PTVQ	364

**FIGURA 25**

DNFNVYKATR PYLANCPDCG EGHSCHSPVA LERIRNEPATD GILKIQVSLQ IGIKTIDGND	60
WTKLRYMDNN MPAADAEKAGL PVNTSAEPTI TGTMGHETLA RCPKGESTLTV GFTDSRKISH	120
SCTHPPPHHPF PVIQEEKPHS RPQNGKEIFC STYVQSTTAT TEELEVNHMP DTFORTLMSQ	180
QSGNVKITVN GQTVRYKNCG GGSENEGLTII DVNINNCIVD QCHARAVTNHK KWQVNSPLVF	240
RNAELGDRRG KIHIPFFPLAN VTCRVEPNAM PTVTYGRNQV IMLLYFDMPt LLSSYRNNGEE	300
PTVQSEWVMH KREVVLTVPT EGLLEVWGMN EPYKYNPQLS TNGTANGMPH EILLYYYELY	360
PTVQ	364

**FIGURA 26**

1

13

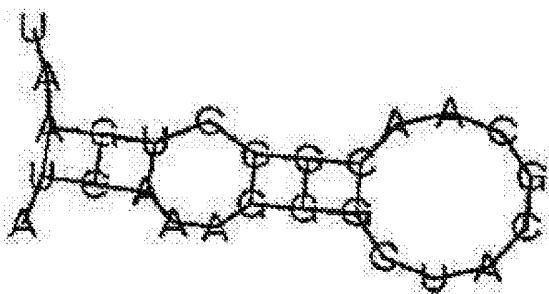
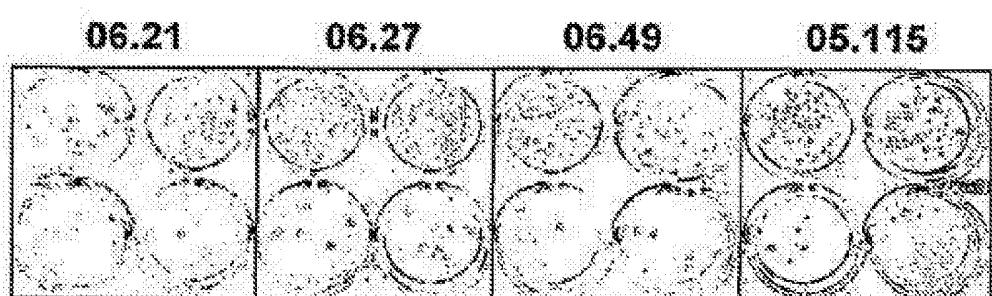


FIGURA 27

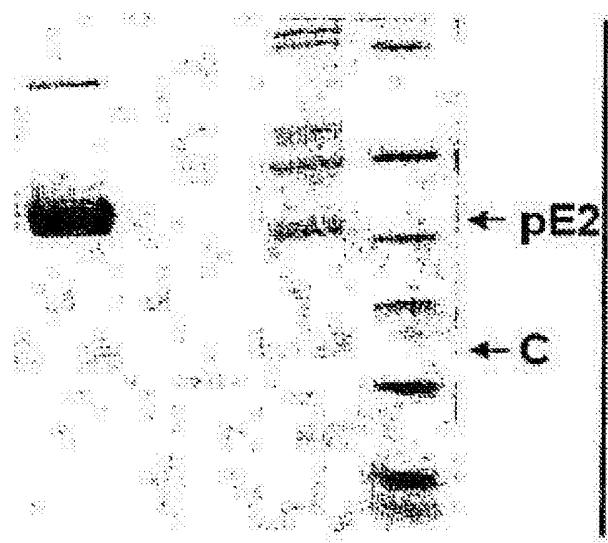
**FIGURA 28**

FIGURA 29

**FIGURA 29 (cont.)**

06-27	GTTATTCGGGCGGTCTTAACTGAGCAACTTACACACACAGTCAC
06-49	GTTATTCGGGCGGTCTTAACTGAGCAACTTACACACAGTCAC
05-115	GTTATTCGGGCGGTCTTAACTGAGCAACTTACACACAGTCAC
06-21	GTTATTCGGGCGGTCTTAACTGAGCAACTTACACACAGTCAC
06-27	AAGTTGATCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG
06-49	AAGTTGATCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG
05-115	AAGTTGATCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG
06-21	AAGTTGATCTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG
06-37	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
06-49	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
05-115	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
06-21	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
06-27	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
06-49	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
05-115	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
06-21	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
06-27	ACCAATTCATCTTGCTACTTGTCTTCCCTCCCTCCCTCCCTCC
06-49	ACCAATTCATCTTGCTACTTGTCTTCCCTCCCTCCCTCCCTCC
05-115	ACCAATTCATCTTGCTACTTGTCTTCCCTCCCTCCCTCCCTCC
06-21	ACCAATTCATCTTGCTACTTGTCTTCCCTCCCTCCCTCCCTCC
06-27	GGGAAATGAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
06-49	GGGAAATGAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
05-115	GGGAAATGAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
06-21	GGGAAATGAACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT
06-27	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
06-49	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
05-115	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
06-21	TGGCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG
06-27	ATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
06-49	ATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
05-115	ATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
06-21	ATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
06-27	ACCTTCTTACCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
06-49	ACCTTCTTACCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
05-115	ACCTTCTTACCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT
06-21	ACCTTCTTACCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT

FIGURE 3B

Iniciador para a sequência terminal de CHIK S2-EP:

22 sítios dissociados de humanas D<sup>1</sup>, E<sup>2</sup>, F<sup>3</sup>, H<sup>4</sup>, Y<sup>5</sup>, R<sup>6</sup>, P<sup>7</sup>, X<sup>8</sup>, A<sup>9</sup>, T<sup>10</sup>, K<sup>11</sup>, S<sup>12</sup>, N<sup>13</sup>, C<sup>14</sup>, F<sup>15</sup>, 3'-OHC-2'G-3'CC-3'G-C-3'G-C-3'G-C-3'G-C-3'G-C-3'G-C-3'.

卷之三

terminal de relojamiento de E2 12 Chik.

卷之三

卷之三

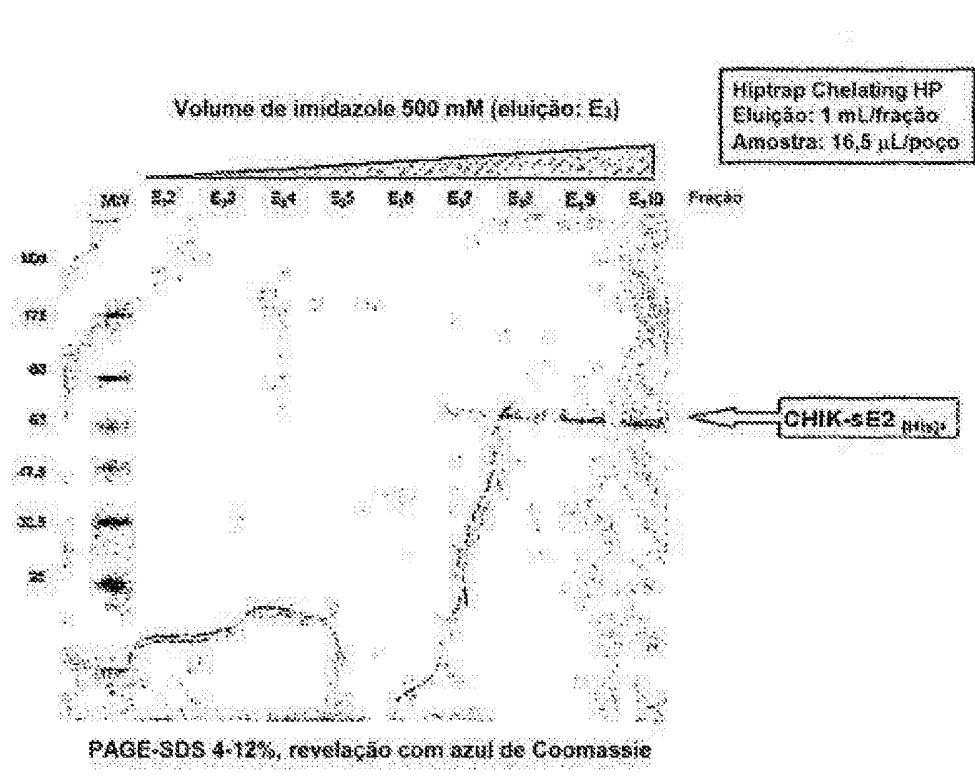
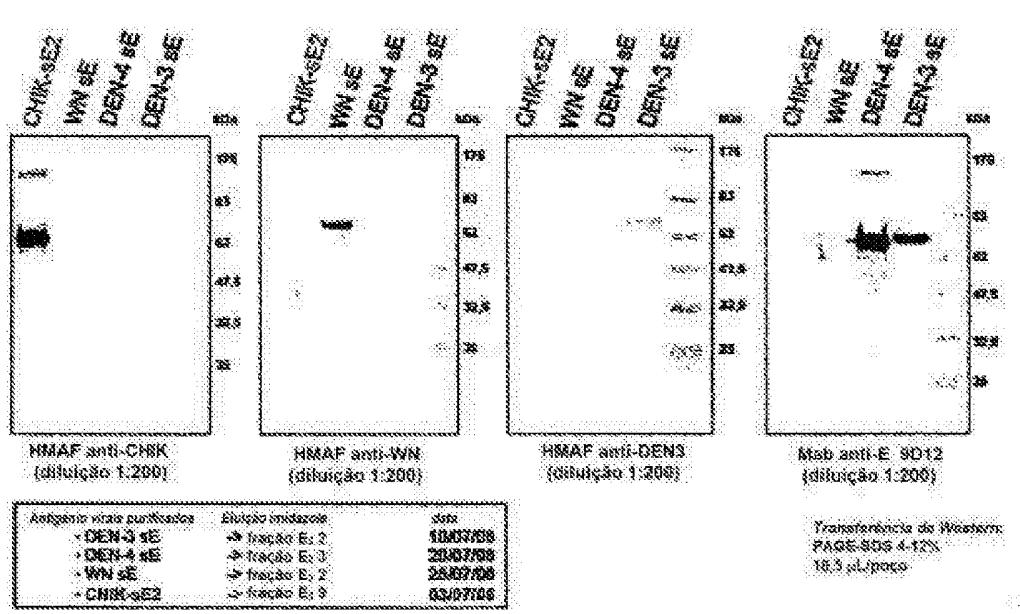
FIGURA 31

FIGURA 32



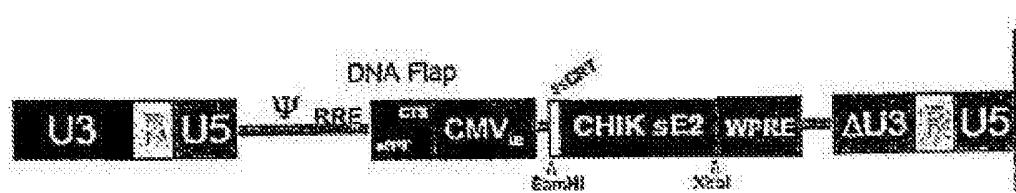
**FIGURA 33**

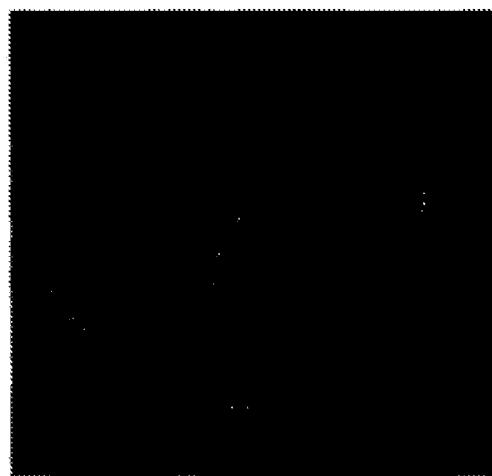
FIGURA 34

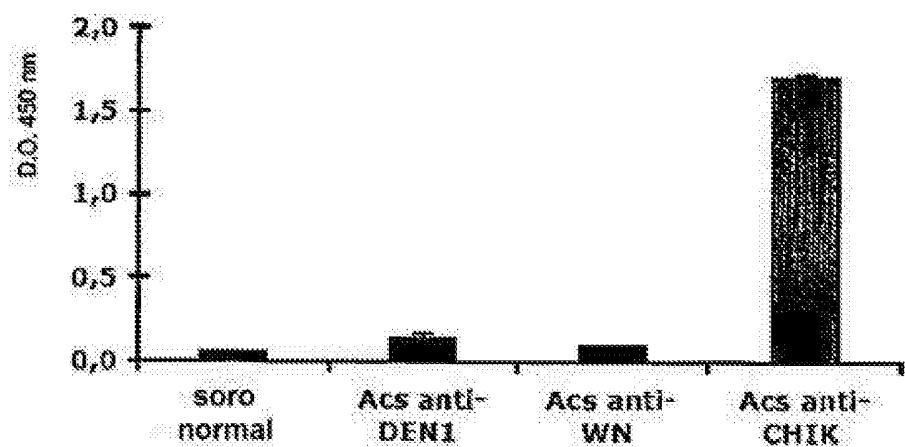
FIGURA 34 (cont.)

**PE2001900**

**53/70**

**FIGURA 35**



**FIGURA 36**

**FIGURA 37**

MEFIPTQTFY NERYQPFKWT PRPTIQVIRP RPRPQRQAGG LAQLISAVNQ LTMRAVFOOK	60
RRERDRKDKKQ KOKQDQDQIN TNGKQDFPK KIAQKXKPG KDRERMCNKIS NOCLFEVKHE	120
GEVTVGACLV GOKVMKPAHV KOTIOMADIA KLAPEKSSKY DLECAQIPVH MKSDASKFTH	180
SEPEGYVWIK HGAVQVQSGR FTIPTGAKGP GDGCRPIFON KGRVVAIVLG GANEGAPTA	240
SVVTNNKDIV TKITPEGAEE WSLAIPVMCL LANITPPCQ FPCIPCCYEK EPEETLWIE	300
DNVVRPGYYQ LLQASLTCSF HRQERSTKIN FIVYKATRPFY LAKCPDCCBG NSCHSPVALE	360
KIRNEATDGT LKIQVSLQIG IGTIDESHQWT KLYMYMINHF AIRGRAGLFV RTSAPCTTG	420
TMGHFLILAC PROGETLTVGF TUSRKISMHC TFPFHNDSPV IGREEKFHRP QNGKELPCT	480
YQGSDAPTAE EIEVEMPPDT FURTILLSCGS ENVKITYMGR TVRYKCNCGG ENEGLITTDK	540
VINNCKVUOC HAAVTRKHW QYNEFLVPMN AELGERKQKI HIFPPLANVT CMVPKAHNPT	600
VTYGEQNQVIM LYYPDNPTLL SYRSNSHEPN YOENWTRKX EVVLTVPTEG LEVTNGNNEP	660
YKVWPQLSAN GTANGHPEHI ILVYYELYPT MTYVVVVSAS FILLSMVGNA VGNCMCARRE	720
CITPYELPG ATVFFPLSLI CCINTAKAAT YQEAAYVLMN EQQNLFWLQA LIPLAALIVL	780
CCNCLRLIPOC CKTLIAFLAVM SIGNATVSAV ENVTIVIPNTV CVPYETLUNR PGYSPMVLEN	840
ELLSVTLEPT LSLOVITTCY KTVIPSPFYVK CGGTAECKDE NLPDYSCIEVF TGVYPPMNG	900
AYCFIDAEKT QLSRANVERKS ESCTEPASA YKANTASAKA KIEVLYQGNM ITVTAYANG	960
NAVTIVEDAKF IVGPMESANT PFDINKIVVYK QDVYINMTPP FGAGRPQGFG DIQERTPESK	1020
DVYNTNOLVL QRPAAGTVAV PYSCAPSGFK YWLKERGASL QNTAPPCCOI ATNPVRHANC	1080
AVGAMPISID IEDAAAPTHVV DAESELTMSC EPVACTHESD FCGVAIIKYA VSKXGKCAVN	1140
SMTRAVTIRK ARIEVEGNGSQ LQIESPSTALA SAEFRVQVCS TOVHCAARCH PPKOHIVNYP	1200
ASHITLGVQD ISATAMSWQ XITIGVGLVV AVAALTLIVV LCVSPSRH	1248

**FIGURA 38**

MESFIPTOTFY	NREYQPRPNT	PRFTIQLVIRP	RPRPQHQACQ	LAQLISAVNK	LITRAVPOQK	60
PRERNRKQKQ	KQKQOQAFQNM	TNQKQPPKK	KPAQKXKKPG	RERRMCKKIE	NDCLPEVKHS	120
GKVIGYACLV	GDGVNMKANV	KGTIDNADIA	KLAJGRSSKY	DLECAQILPVH	MKSDAASKFTH	180
EKPEGTYNNH	HGAVOVSQGR	FTIPTSGKIP	GDGSRPFIDN	KGRVVAIVLG	GANSGARTAL	240
SVVTWNKEDIV	TKITPEGAEF	WSLAIFVNL	LANTTIPPCSQ	PRCTPCCYER	EPFESTLMLK	300
DNVNRPGYYQ	LLOASLTCSF	NRQRSTKIN	TVNYKATRPy	LAHCPOCGEG	HSCHSFVALE	360
RIRNEATDGT	LKIQVSLQIG	IKEODSHONI	KLRYMQNSMP	ADAEFRAGLV	RTSEAPCTITG	420
TMKHFILARC	PKGETLTIVGF	TDSRKISHSC	THPPHDPPV	IGREKPHSRP	RHGRELPCST	480
YVQSTAATTE	EIEVHMPDGT	PORTLMSQCS	GVVKITVNGQ	TVRYKCNCGQ	SNBCLTTTDK	540
VINNCVNDQC	HAAVINHKW	QYNSPFLVHN	AELCDRUGKII	HIPPLANVT	CRVPKARNPT	600
VTVGENOVIM	LLYPFIKPTLL	SYRNCHENPN	YQEMVPHKK	BVLTVPTEG	LEVTRGHNNSP	660
YKWPOLSTN	GTANGHPIREI	ILLYVYELYP	MTVWVWSVAT	FILLSMVNA	AGCMCICARR	720
CITYKLITPG	ATVPLLSLI	CCIRKAKAAT	YQEAAYILNN	EQQFLFWLQA	LIPFLAALIVL	780
ELLSVTLEPT	LSLOYTCEY	KTVIPSMYVK	CCGTABCOK	HLEDYSCKVP	TGVYPPMNGG	840
AYCFCAHENT	QLSAEAHVEKS	ESCRTEFASA	YRAMPASASA	KLAVLNQGRN	ITVTRAYANGD	900
HAVTVKIDAF	IVGPNSSANT	PFINKIVYTK	GOVYDMDYPP	PGAGRGQGQG	DIQSRTYPESK	1020
DVYANTQVL	QRFAAGTVHV	PYQAFSGFK	YLKENGASL	QHTAPPGCQI	ATNEVRAVNC	1080
AVGNMPISID	IPRAAAPTEVV	DAPSLLTMSC	EVFACTHSD	PGGVALIKYA	ASKGKCAVN	1140
SNTNAVTIKE	AKIEVEGNSQ	LOISFSTALA	SAEFRVQVCS	TQVHCAARCH	PPFDHIVNYP	1200
ASHTTLGVQD	ISATAMSWQ	KITGGVGLVV	AVARLILIVV	LCVSPSRH		1248

**FIGURA 39**

MESFIPTOTFY	NREYQPRPNT	PRFTIQLVIRP	RPRPQHQACQ	LAQLISAVNK	LITRAVPOQK	60
PRERNRKQKQ	KQKQOQAFQNM	TNQKQPPKK	KPAQKXKKPG	RERRMCKKIE	NDCLPEVKHS	120
GKVIGYACLV	GDGVNMKANV	KGTIDNADIA	KLAJGRSSKY	DLECAQILPVH	MKSDAASKFTH	180
EKPEGTYNNH	HGAVOVSQGR	FTIPTSGKIP	GDGSRPFIDN	KGRVVAIVLG	GANSGARTAL	240
SVVTWNKEDIV	TKITPEGAEF	WSLAIFVNL	LANTTIPPCSQ	PRCTPCCYER	EPFESTLMLK	300
DNVNRPGYYQ	LLOASLTCSF	NRQRSTKIN	TVNYKATRPy	LAHCPOCGEG	HSCHSFVALE	360
RIRNEATDGT	LKIQVSLQIG	IKEODSHONI	KLRYMQNSMP	ADAEFRAGLV	RTSEAPCTITG	420
TMKHFILARC	PKGETLTIVGF	TDSRKISHSC	THPPHDPPV	IGREKPHSRP	RHGRELPCST	480
YVQSTAATTE	EIEVHMPDGT	PORTLMSQCS	GVVKITVNGQ	TVRYKCNCGQ	SNBCLTTTDK	540
VINNCVNDQC	HAAVINHKW	QYNSPFLVHN	AELCDRUGKII	HIPPLANVT	CRVPKARNPT	600
VTVGENOVIM	LLYPFIKPTLL	SYRNCHENPN	YQEMVPHKK	BVLTVPTEG	LEVTRGHNNSP	660
YKWPOLSTN	GTANGHPIREI	ILLYVYELYP	MTVWVWSVAT	FILLSMVNA	AGCMCICARR	720
CITYKLITPG	ATVPLLSLI	CCIRKAKAAT	YQEAAYILNN	EQQFLFWLQA	LIPFLAALIVL	780
ELLSVTLEPT	LSLOYTCEY	KTVIPSMYVK	CCGTABCOK	HLEDYSCKVP	TGVYPPMNGG	840
AYCFCAHENT	QLSAEAHVEKS	ESCRTEFASA	YRAMPASASA	KLAVLNQGRN	ITVTRAYANGD	900
HAVTVKIDAF	IVGPNSSANT	PFINKIVYTK	GOVYDMDYPP	PGAGRGQGQG	DIQSRTYPESK	1020
DVYANTQVL	QRFAAGTVHV	PYQAFSGFK	YLKENGASL	QHTAPPGCQI	ATNEVRAVNC	1080
AVGNMPISID	IPRAAAPTRVV	DAPSLLTMSC	EVFACTHSD	PGGVALIKYA	ASKGKCAVN	1140
SNTNAVTIKE	AKIEVEGNSQ	LOISFSTALA	SAEFRVQVCS	TQVHCAARCH	PPFDHIVNYP	1200
ASHTTLGVQD	ISATAMSWQ	KITGGVGLVV	AVARLILIVV	LCVSPSRH		1248

**FIGURA 40**

MESPIPTQTFY	NREYQPRFWT	PRPTIQVIRP	RPRPQRQAGQ	LAQLISAVNK	LTMRAVFOQK	60
PRERERENKEQ	KOKCQAPDN	TNQKQPKPK	KPAQKKRKGKPG	REERRMCMKIE	NDICIFEVKHE	120
GRVTGTYACLV	GDKVMKPAHV	KGTIDMADIA	KLAFKRSSKY	DLECAQIFVH	MKSQDASKPTH	180
EPKPEGYINNH	NGAVQYSGGF	PTIPTGACKF	CDOSGRPIFDN	KGRVVAVIYLG	GANGSARTAL	240
SVVVWNEKIV	TKITFEGARE	WSLAIPIVNL	LANTTFCPSQ	PFCTPCCYRK	EPEPTLRLML	300
DNVTRPGYYO	IHQASLTCS	PKRQRSTKDN	PNVYKATRPFY	LANCFDCGEG	HSCHSFVALE	360
RIENEATDGT	LKIQVSLQIG	IETDISHDWT	KLRYMDNSMP	ADAKRAGLFV	RTSAPCTITG	420
TMCHFILARC	PRGETLTVGF	TDGKIKSRSC	THPFHNDPPV	IGREKPHSRP	QHGKELFCST	480
YVOSTAATTE	EIEVHMPEDT	FORTILMSQGS	GNVKITVNGQ	TVRYKCNCGG	SNEGLTTTOK	540
VINNCXVOQC	KAATVNHKHN	QYNSPFLVFN	AELGDRKGKI	HIPPLANVY	CRVPKARNPY	600
VTVYGRNQVIN	LILYDINPTLL	SYRNQHSEPN	YQEMWNSIK	BVVLTVTPEC	LEVTVGNNEP	660
YKYNWQPLSTN	GTANCHPHEI	ILYYYELYPT	MTVWVSVAT	FILLEMVOMA	AGNCMCARRR	720
CITPYRLTPG	ATVPEFLSLI	CCIRTAKAAAT	YQSAATYLWN	EQQPLFWLQA	LIPLAALIVL	780
CNCLRLLPCC	CKTLAFLAVM	SVGANTVSA	EHVTIVIENTV	GVYVITLVR	PGYSPMVLEN	840
ELLSVTILEPT	LSLDYITCEY	KTIVIPSPYVK	CCGTAECKDK	NLFDTYSCKV	TGTVYPPWNG	900
AYCPFDAENT	QISeAHVKEK	ESCKTEFASA	YRANTASASA	KRVLVYQGNN	ITVTAYANGD	960
HAVTVKDAKF	IVGPMSSANT	PFINKIVVYK	GDVYKNDYPF	FGASGRPSQFQ	DIQSRTIPESK	1020
DVYANTQLVL	QRPAVGTVHV	PYSQAPSQGF	YWLKENGASL	QHTAPPCCQI	XINNFRAVNC	1080
AVGNMPISID	IPERAFTRVV	DAPSILTDMS	SVPACTHESD	FGGVVATKYA	ASKGGKCAWH	1140
SVTNNVTLRE	AEIEVEGNSQ	LOISFSTALA	SAEPFVQNGC	TQVHCAFSCH	PPIDHIIYYP	1200
ASHTTLCVQD	ISATAMSWVQ	KITSGVGLVV	AVVALLIIVV	LCVSPFSR		1248

**FIGURA 41**

MESPIPTQTFY	NREYQPRFWT	PRPTIQVIRP	RPRPQRQAGQ	LAQLISAVNK	LTMRAVFOQK	60
PRERERENKEQ	KOKCQAPDN	TNQKQPKPK	KPAQKKRKGKPG	REERRMCMKIE	NDICIFEVKHE	120
GRVTGTYACLV	GDKVMKPAHV	KGTIDMADIA	KLAFKRSSKY	DLECAQIFVH	MKSQDASKPTH	180
EPKPEGYINNH	NGAVQYSGGF	PTIPTGACKF	CDOSGRPIFDN	KGRVVAVIYLG	GANGSARTAL	240
SVVVWNEKIV	TKITFEGARE	WSLAIPIVNL	LANTTFCPSQ	PFCTPCCYRK	EPEPTLRLML	300
DNVTRPGYYO	IHQASLTCS	PKRQRSTKDN	PNVYKATRPFY	LANCFDCGEG	HSCHSFVALE	360
RIENEATDGT	LKIQVSLQIG	IETDISHDWT	KLRYMDNSMP	ADAKRAGLFV	RTSAPCTITG	420
TMCHFILARC	PRGETLTVGF	TDGKIKSRSC	THPFHNDPPV	IGREKPHSRP	QHGKELFCST	480
YVOSTAATTE	EIEVHMPEDT	FORTILMSQGS	GNVKITVNGQ	TVRYKCNCGG	SNEGLTTTOK	540
VINNCXVOQC	KAATVNHKHN	QYNSPFLVFN	AELGDRKGKI	HIPPLANVY	CRVPKARNPY	600
VTVYGRNQVIN	LILYDINPTLL	SYRNQHSEPN	YQEMWNSIK	BVVLTVTPEC	LEVTVGNNEP	660
YKYNWQPLSTN	GTANCHPHEI	ILYYYELYPT	MTVWVSVAT	FILLEMVOMA	AGNCMCARRR	720
CITPYRLTPG	ATVPEFLSLI	CCIRTAKAAAT	YQSAATYLWN	EQQPLFWLQA	LIPLAALIVL	780
CNCLRLLPCC	CKTLAFLAVM	SVGANTVSA	EHVTIVIENTV	GVYVITLVR	PGYSPMVLEN	840
ELLSVTILEPT	LSLDYITCEY	KTIVIPSPYVK	CCGTAECKDK	NLFDTYSCKV	TGTVYPPWNG	900
AYCPFDAENT	QISeAHVKEK	ESCKTEFASA	YRANTASASA	KRVLVYQGNN	ITVTAYANGD	960
HAVTVKDAKF	IVGPMSSANT	PFINKIVVYK	GDVYKNDYPF	FGASGRPSQFQ	DIQSRTIPESK	1020
DVYANTQLVL	QRPAVGTVHV	PYSQAPSQGF	YWLKENGASL	QHTAPPCCQI	XINNFRAVNC	1080
AVGNMPISID	IPERAFTRVV	DAPSILTDMS	SVPACTHESD	FGGVVATKYA	ASKGGKCAWH	1140
SVTNNVTLRE	AEIEVEGNSQ	LOISFSTALA	SAEPFVQNGC	TQVHCAFSCH	PPIDHIIYYP	1200
ASHTTLCVQD	ISATAMSWVQ	KITSGVGLVV	AVVALLIIVV	LCVSPFSR		1248

**FIGURA 42**

MEPFIPTQIIFY	NKRYQPRPWT	PRPTIQVIRP	RFRPQRQAGQ	LAQLISAVN	LTMRAVPOQK	60
PERRNRKDKKQ	KOKQQAQPQN	TNQCKQFFPK	KPAQKKEKFG	RERNCMKIE	MDCLPEVKHS	120
GEVIGYACLV	GDRVMKPARV	KGTIDNADLA	KLAFKRSKY	DLECAQIPVH	MRSDAKPFTH	180
EKPEGYIWNH	HGAVOYSGR	FTIPTQGKPF	GDSGRPIFDN	KGRVVAIVLG	GAMEGARTAL	240
SVVTWNKDIV	TKITPEGAEF	WSLAIFVNCI	LANTIPCSQ	PFCTPCCYEK	EPEETLMLK	300
DNVMPGYYQ	LJQASLTCSP	HRQRRESTRK	ENVYKATREY	LANCPCCGBG	HSCHSFVALS	360
RIRNEATDGT	LKIQVELQIG	IETEDSHENT	KLRYMDNSMP	AIRERAGLJFV	RTSAPCTITC	420
TNGHFLARC	PKHRTLTIVGP	TDSRKISHSC	THPPHNEFPV	IGREYFHSRF	QHGKELPCST	480
YVQSTAATTE	EIEVNPEDT	PORTLMSQS	GVVXKITVNGQ	TVRYEYCNG	SNEGLTTDK	540
VINNCKVQOC	HAAVINHKKW	QYNSPLVFN	AKLCDEKGK	HIFFLANVT	CRVEKARNT	600
VTVYGNQVIM	LLYPENMILL	SYRNGEEMN	YQESNMVMHK	EVVLTVPTEG	LEVTVGNNEP	660
YKYNPQLSLN	GTANCHMHEI	ILYYTYLYPT	MTVVVSVAT	FILLSMVGMA	AGNCMCARRR	720
CUTPYELTPG	ATVFFLISLI	CCIRTAKAAT	YQESAIYLN	EQQPLFWLQA	LIELAALIVL	780
CNCILRLLPCC	CKTLAFLAVM	SVGANTVSEY	BNVTVLNTV	CVYTKILVNN	PGYSEMULEM	840
ELLSVTLEPT	LSLDVITC	RTVIPSIPVYK	COCTASCKDK	NLSDYNSCKVY	TGVPYPPNGS	900
AYCPDCAENT	QLSERANVEKS	ESCXTEFASA	YRANTASAGA	KLAVLYQNN	ITVTVAYANGD	960
HAVTVDAK	IVGFMGSANT	PFINKIIVYK	GDVIMMNPV	PGACKPGQFG	DIOSGRTPEEK	1020
DVYTANTOLVL	QRPANGTIVH	FYSQAFSGFK	YWLNGKGMEL	QHTAPPCCQI	ATNPMRAVNC	1080
AVGNMPISID	IPERAFTIVV	DAPSLTIMSC	EVFACTHSSD	FGGVATIKEYA	ASEKOKCAVH	1140
SVTNAVNTIRE	AEIEVGENSQ	LQISFSTALA	SAPFRVQNC	TQVHCAECH	PPHDHIVNYP	1200
ASNTTILGVQD	ISATAMSWSQ	KITGGVGLVV	AVAILILIVW	LCVSFSRH	1248	

**FIGURA 43**

MEPFIPTQIIFY	NKRYQPRPWT	PRPTIQVIRP	RFRPQRQAGQ	LAQLISAVN	LTMRAVPOQK	60
PERRNRKDKQ	KOKQQAQPQN	TNQCKQFFPK	KPAQKKEKFG	RERNCMKIE	MDCLPEVKHS	120
GEVIGYACLV	GDRVMKPARV	KGTIDNADLA	KLAFKRSKY	DLECAQIPVH	MRSDAKPFTH	180
EKPEGYIWNH	HGAVOYSGR	FTIPTQGKPF	GDSGRPIFDN	KGRVVAIVLG	GAMEGARTAL	240
SVVTWNKDIV	TKITPEGAEF	WSLAIFVNCI	LANTIPCSQ	PFCTPCCYEK	EPEETLMLK	300
DNVMPGYYQ	LJQASLTCSP	HRQRRESTRK	ENVYKATREY	LANCPCCGBG	HSCHSFVALS	360
RIRNEATDGT	LKIQVELQIG	IETEDSHENT	KLRYMDNSMP	AIRERAGLJFV	RTSAPCTITC	420
TNGHFLARC	PKHRTLTIVGP	TDSRKISHSC	THPPHNEFPV	IGREYFHSRF	QHGKELPCST	480
YVQSTAATTE	EIEVNPEDT	PORTLMSQS	GVVXKITVNGQ	TVRYEYCNG	SNEGLTTDK	540
VINNCKVQOC	HAAVINHKKW	QYNSPLVFN	AKLCDEKGK	HIFFLANVT	CRVEKARNT	600
VTVYGNQVIM	LLYPENMILL	SYRNGEEMN	YQESNMVMHK	EVVLTVPTEG	LEVTVGNNEP	660
YKYNPQLSLN	GTANCHMHEI	ILYYTYLYPT	MTVVVSVAT	FILLSMVGMA	AGNCMCARRR	720
CUTPYELTPG	ATVFFLISLI	CCIRTAKAAT	YQESAIYLN	EQQPLFWLQA	LIELAALIVL	780
CNCILRLLPCC	CKTLAFLAVM	SVGANTVSEY	BNVTVLNTV	CVYTKILVNN	PGYSEMULEM	840
ELLSVTLEPT	LSLDVITC	RTVIPSIPVYK	COCTASCKDK	NLSDYNSCKVY	TGVPYPPNGS	900
AYCPDCAENT	QLSERANVEKS	ESCXTEFASA	YRANTASAGA	KLAVLYQNN	ITVTVAYANGD	960
HAVTVDAK	IVGFMGSANT	PFINKIIVYK	GDVIMMNPV	PGACKPGQFG	DIOSGRTPEEK	1020
DVYTANTOLVL	QRPANGTIVH	FYSQAFSGFK	YWLNGKGMEL	QHTAPPCCQI	ATNPMRAVNC	1080
AVGNMPISID	IPERAFTIVV	DAPSLTIMSC	EVFACTHSSD	FGGVATIKEYA	ASEKOKCAVH	1140
SVTNAVNTIRE	AEIEVGENSQ	LQISFSTALA	SAPFRVQNC	TQVHCAECH	PPHDHIVNYP	1200
ASNTTILGVQD	ISATAMSWSQ	KITGGVGLVV	AVAILILIVW	LCVSFSRH	1248	

FIGURA 44

KCPVYVVDIDA	DSAFKLALQR	AIPMPEVEPR	QVTPNDHNRN	RAFSHLAIKL	IEGSILPDT	69
TLDIGSAAAR	RIMSDORKYRC	VCPKESARDP	ERLANYAKRL	ASRAAGRVLDR	NISGKIGDLO	129
AVMAVPOVET	PTPCLIHTDVS	CRQRADVAIY	QDVTAVHAPT	SLYNAIKGV	RVAYWVGFDT	189
TCPPMDAAMNG	AIFSYCTINWA	DEQVLKAEI	GLCSTDTES	NRCKLSIMRG	XELKNTDORVL	249
PSVGSTLYPS	SRELLKSNRL	PSVPHLEGEL	SFTCKCIVVV	SCSGYVVERI	TNSPGLYGKF	309
TGTAUTHAC	GFLNCKTTDT	VDGEERVSFV	CTYVPAATICD	QMTGILATEV	TPEDRACKLLV	369
GENQRIVVNG	RTQRRNTTMR	NYILPVVAQD	FSKNAKECRK	LNEDSKILLGV	KERTLICCCCL	429
WAFERQKTHI	VYKRFETOSI	OKVQAEPDSY	WPZELM22GL	SIFLRTRIKW	LLSKVPKTEL	489
IFVSOQDREA	RDAEKAABEE	REAELETRAL	PFLQNAQEDV	QVEIDVEOLE	DRAGAGLLET	549
PRGALXWTAQ	PIONWVGEYL	VLSPOTVLRS	QKLSLHALA	EQVKICHTING	RASEYAVEAY	609
IGRVLVTSGGY	AISPEPQSI	SEKATMVNE	REFVNRRNLH	IAMNGPAINT	DRESYELVRA	669
ERTEHEYYVYD	VOQRCCKKE	ERAGLVLVGD	LTNPPYHEFA	YEGLAKIRPAC	FYKLAVIGV	729
GWPGESEKSAI	IENLWTRQDL	VTSGKHNCO	KEITDVRNDR	GLEISARTVD	SILLNECNRP	789
VIVIAVDEAP	ACHSGTLLAL	IAIUVPRQKV	VLCHDOPKCG	FFNMNMQMKVN	YNNHNICTQVY	849
INNSISRECTL	PVIAITVSSLH	YEGDMRITNE	INNEPIVWDTT	GSTKDPGCOL	VLTCPFRGWWX	909
OLQIDNVRGYS	VMTAARSQGL	TRKGUYAVRO	KVNENPLIAS	TERNWVLLT	RTEGKLWVKT	969
LSEPPWIXTL	QNEPKGNPKA	TIEKENNEVHA	SIMAGICSHQ	MITDFTQNEA	NUCNWAKSELVP	1029
ILETMGIKLN	DNQWSQIIQA	FEEDKAYSPE	VALNEICHTN	YGVOLUSLDF	SEPLVSVVYTA	1089
UNHNEFEGE	IMFCGPNEAA	SILERNYPFT	NGKNNNNKQI	CVTTRRIEDF	SPPTINIIIFAN	1149
HLPNHSVLAS	HFPVKGERME	MLVNRKINGHH	VLLVSGYNLA	LPTKRVTVWA	FLGVRGADYT	1209
YNNELGLPAT	LGHYIILVWIN	IHTPPRISHM	QCYCWHAMKL	QNLGHSISKL	LKPQGSLLIR	1269
AVGYADRTSE	XVTCVLGRKF	RSSRALEPPC	VTENTEMPL	FSRPFINGRNN	FTTHVMENOL	1329
NAAPVGCVTR	ACCAPSTRVK	RMDIARNDDE	CYVIAANPRG	LFODGTCKAV	YKXWPFESPIN	1389
SATPVSTAKT	VMECTYPIH	AVGPNPSNYS	ESEGDRELLA	AYREVAKEVT	SLGNHNSVAIP	1449
LISTCVYSGG	KIRLTIQSINH	LPTAMKSTIA	CVVYICRKEE	WEKKISKEAQ	MRTQVELLDE	1509
KISINDCIVK	VHPESSELAGR	KYSTTEGAL	YSYLEGTRFH	QJAVIMARAH	TMWFRQITERN	1569
DOVCLVALOS	SIESIRONCP	VDEADNSSPP	KTIVNCLCRYA	MTPENTRILR	MNVVTEIIVC	1629
SSPPPLPKXII	SGVQKVKCSK	VMLFDHNVRG	RVSFREIRSS	QESAGECASTI	TSLTHSQFIL	1689
SVNKEILPWF	SOLDACRPL	EPALDODGATH	TLFSTTMENLA	AVSOWVMSTV	FVAPPFERRG	1749
ERLTVTCDER	EENITPMASV	RFFRAELCPV	VQRTAETRUE	AMSLQAPPST	ATEPKHFFIS	1809
PGASSETTPPI	TFGDFNNEGEI	ESLSSSELLTF	HOFLGEVDD	LTDGSINSTCS	DTIDELIJRA	1869
GGYIFSSDTG	PGHLOOKSVR	QSVLPVNTL	EVHEEKCYFP	KLDANEROLL	LKKLQESAEM	1929
AMRSRYQGRK	VIENNEKAATIQ	RLKRGCRLL	MEETPKVPTY	RTTYFAPVYS	PPINVRLEMP	1989
ESRVAACNEF	LARNYFTVSS	YQTIDETDAY	LIMVINGRSC	LDRATINPSK	LRSYPKQHAY	2049
NAFFSIRSAVF	SPPQNTLQN	LAATKERNCN	VTQMRRELPTL	DSAVFVNECF	EKFACNQEIN	2109
SEPARASPIR	ITENLATIVT	KLKIPKAAM	FAKTHNLLPL	QEVPMDEFTV	BNKRDIVKVT	2169
GTRNTEERFK	VQVIOQAEP	ATAYLQSIHR	KLVERLQAVL	LINVTILSEM	SAGDFIALIA	2229
ANPXPQDTVL	ETUJIASFOKS	QDLSLALTAL	MLLEDQWCH	SLLDLIEAAP	GEISSCHLPT	2289
GTRKPGAM	ESGMQFLTLFV	NTLINITINS	RYLKDRLTKS	ACAAFIGDGN	LIHGVVSDKL	2349
NEARCATWM	MEVKIIDAVV	SLKAPYFCG	FTLHDOTVTST	ACRVADPLER	LFLKGKPLAA	2409
QDEQDIEKRE	ALADEVINMQ	PTGLIDELER	AVYSKYEVQG	ISVYVWMSMAY	FASSEKSFEX	2469
LRGFVITLYG	GPK					2473

**FIGURA 45**

MDFVYVDIQA DSAFLAKALQR AYPMFEVEPR	GVTPTDHNNA RAYSHLAIKL IEQRIQFDST	60
ILDGISAPAR RMMSIDSKYHC VCPWRSASDIP	ERLANYARKI ASAAACKVLDR NISGKIGDLO	120
ANNGAVPDTET PTFCIHTIVS CRQADVAIT	GOVYAVHAPT SLYHQAINGV RVATWVGPFIT	180
TPFVWVNAIMC AYPSYNTMIA DECVLKAMNI	GLOCSTIDTNE RECKLSINHG XXLXNCTIRVL	240
FSVGSTLYPE SKILLASWNL PSVTHLKGKL	SFTCRCDTIVV SCRGYVVRRV TNSPGLYGKT	300
TGYAVTHED GELNCTTET VENSERVSFSV	CTYVPATICD QMTGILATEV TPEHAOKLLV	360
GLBQRIWVNG STQRTNTIME NYILLPVVAQD	FSKNAKECEK IWEDEKILGV RERTLTCCLL	420
SAPEKQXTIT VYERPDQSI QKVQAEFDOSF	VVPSLNSSGL SIPLRTRIKW LLSKVPKTOL	480
JFYZCDAREA REAEKPAESE REAELTREAL	PPFLORAQSDV QNEIDVEOLE DRAGASLIET	540
FRGAIRYTAQ PTDSHVGEYL VLSQPTVLR	QKLSLHALA EJVKXCTHNG RACRYAVEAY	600
IGEVNLVPSGY AISPEDFQSL SESATMVYNE	REFVNEELHH IASHGPAINT DERSYELVRA	660
ERTEHEVYTD VOQARCCKEE BAAGLVLVGD	LTSUPPYMEFA YRGKIRPAC PYKIAVIGVF	720
GWPGSGNRAK IKNLVTRQSL VTSQKXENQ	KETTEVNRQD CLEISARTVO SLLNGCNRP	780
WDVIXVTERAF ACHSGTILLAL TALAVPHQKV	VLGGOPKOCG PFRMMSQMEVN YNNHINCTQVY	840
HKSISRCIL PVTATVSSLR YEGKMRITME	INXPIVWDTT GSTKPDPGDL VLTCPRGNVK	900
OLQIOUROYE VMTRAAASQGL TRNQVYAVRO	KVNNENPLAS TSEHVNVLITY RTGOKLUVWT	960
LSGDFWIKL QNPSPKRNPKA TIEKEVEENA	SEIMAGICSQH MTFOTQONKA NVCMAESLVP	1020
ILRTNGTKEIN DRONEKITION FKEKDKEYSPE	VALMEICRM YGVILDSCLP SKPLIVEVYIA	1080
ENHNNINPOOG KMFQFNPBEA SILERKYPFT	EJENNNINQI CVTTRIAED NFPTINII PAN	1140
ERLNSLIVAE HEPVNGEME NLWVKINGHN	VLLVSEGYLN LPTKRNVINVA PLGVRGADYT	1200
YNLHLGLPAT LGHYDRUVIN INTFRINHY	QOCVHENAKL OMLGGDSIRL LKKGGSLLIR	1260
AVGYADORTSE RVICULSRKF RSSRALKEPC	VTSNTEMFPL FENFINGREN FTIHVMINQL	1320
KAFAWQOYTR ACCAPSYTRVZ KMDIANDEE	CVVMIAANPFG LPGCGVCAV YKINWESPKM	1380
BATTNGTAKT VMCSTYPPVH AVSINFENYV	ESEGDREKLA AYREVAXEVN RIGVNSVAIP	1440
LLSTGUVYSGG KIRLICOSLH LPTAMISTDA	DVVIYCYRKE WEKKISEATO MRTQVELLOR	1500
MISIDXDIVN VHPDSSLAGK EGYSTTREGAL	YSTYLEGTRFH QTAVIMAKIN TMWPKQTEAN	1560
EQVILVALQE SIESIRQSCP VDDGASSEPP	XIVPCLCRYA NTPERVTRIL MNHVTISIIVC	1620
EEFFPLKTKC EGVCKYKPCSK VMLFDHNVPS	EVSPLEYRES QESAQESTI TSLTHSOFEL	1680
SVDGEILPVP FUDADADAPAL BEALDQGATH	TLPSTTGCNL AUSCIVVMSTV PVAPPERRNG	1740
NNLTUTCINER EGNITEMASV RFFRAELCPV	VOETAETROT AMSIQAPPST ATFPNNPPIS	1800
PGASSETPTI TPGDNNEGSI ESLSSELLTP	CEPPIGEVDD LTDISINSTCS DTIDRILLRA	1860
OCTYIPESDNG PEHLQCESVR QSVLPUVTLR	EVBHEECCYPP KLOEAKEQLLI LKXLOESASH	1920
MURERYQSHN VENMKMATTQ RINGCRRLYL	MSBTIFKVFTY RTTYVAPVYS PFINVERLSNP	1980
EEAVAAACNEP LARNYFTVSE YQTIDETDAY	LDMVGSSESC LIRATPNPSK LRSYPIQHAY	2040
RAFSIRSARVP SFPONTLQNV LAATKRNCH	VTQNRFLPTL DSAVFNVRCP KRFACINQEW	2100
ERFAASPIRI UTRNLATVTV KLGKPKAAAL	PAKTHNLLPL QEVPNQRFV DMRKIVKVT	2160
GTKEKEENPK VOVIQKASPL ATAYLOGIHE	ELVNRLLAVL LRVNVTLFDM SAEDPFIATA	2220
ANFKPSGTVL ETDIASFDKS QDDSLALTAL	MLEEDLCVHN SILLOLJERAF GEISSCHLPT	2280
GTREPKQDM XSGMPLTLSV NTLLNITIAS	PVLEDRITIS ACAAFISHON IINGVMSIEL	2340
SAARCATWNN MEVKIIDAVV SLKAFYFCGG	FILHDTVIGT XCRVADPLR LFKLGEPLAA	2400
GRGQDIEHRR ALADEVENWQ RTGLIDELEX	AVTSKYEVOG ISVVVMSMRT FASSERNTEN	2460
LRGPVTTYG GPK		2473

**FIGURA 46**

MDFVYVYVIIIA	DSAFIKALQR	AYPHFEVEVPR	QVTPEMHANA	RAFEHLAIKL	IHQKIDPDST	60
ILDIGSARAE	RISMSIRKWC	VCFMRSAGDP	ERILANTYAREL	ASAAAGWVLDS	MSISKKICDLO	120
AVMAVPITST	PTFCCLTDVS	CRIADWATY	QDVIZAWHFT	SLYHQAIKGV	BVAYW/GFDT	180
TEPPNRYMAMAG	AYPSYTTWA	DSJWIXXAKI	GLCCTESTLTER	REGKLSIMNG	KXLRPCDRL	240
PSVGSTLYPE	SRELLKSNHL	PSWVHLKGKL	SFTCRCTTVV	SCGYVVERI	TMSPOLYGET	300
TSYAVTHRAED	GFLACKTIDT	VIGERNSFEV	CTYVATACIQ	QMTGELATEV	TFEDRQNLIV	360
GLNQRTIVVNG	FTORNTNTNK	NYLLPVVAQD	FSKMKKCBRK	IPMSDEKLLOV	REFTLTCCL	420
WAPKEQKHTT	YVYRFDTQSI	QNVQAEFDSP	VVPEELNSEGL	SIFLKTRIKW	LLSKVPKTL	480
IPYSGDAREA	RDAKKAERE	REAKLTTRAL	PFPLQACEDV	QVEILDNEQLE	DRAGAGIIST	540
SEGAIKVTAQ	FISIVVGEYL	VI3GPQTVLS	QRLSLMLA	SOVTKCTING	HAGRYAVEZ	600
D38VLVIESSY	AISPEDOTISL	HESAHVMVNE	REFVNNELHH	IAMHGPAINT	DRERSVHLVA	660
ERTENEVYVD	VIQEQCCYCE	EMANGLIVED	LINSPYHESFA	YEGLKIRPAC	PYKIAVIGVY	720
GT/PGSGKSGAI	TKNLVTRQDL	VTSGKXENQ	EITCDVWQG	GLBLSRATTW	SILLNCINRP	780
VOVLYVDEAF	ACHSSTLLAL	IALMRPQEV	VLCEDPKOOG	FFNMNCMEWS	YMININCTQVY	840
HKSISRECTL	PVTAJIVSSLH	YEGKOMPTNS	YKCPVWVDTT	GSTKEDPGEL	VLTCFRGAVV	900
QLQIDYEGYE	VNTAAASQQL	TRKQVYAVRQ	ICVNNENPLAS	TSEHVNWL/T	KTRBNLWNR	960
LSSEDPNIKTEL	QNPBKGNFNA	TIKNEVISHA	SIMPAGCSHQ	MTDPTQMC	NUCKAKELVP	1020
ILSTAGIKLM	DRONSQLIQA	FREDRAYSPP	VALMEICTRM	YQVDEECLP	SKPLVSVTYA	1080
DNHWDNRPQG	KMFGFNPEAA	SILARSKYFT	KGSNINNKCI	CVTTEREDF	NFTTNILPAM	1140
RRLPHSILVAE	HRPVKGSHME	WLWNLKINGH	VLLVSGVNL	LPFTRTVWYA	FLGVRGADYT	1200
XNLELGIPAT	LGRYDLVWIN	INTFHFRHY	QOCVUHANKL	QMLQGDSLRL	LKPKAGELLR	1260
AVGYAIDTEE	EVICVLGKPF	RSSEPAKCP	WTSMTEMEPL	FSNFDRGERN	FTTHUNNINOL	1320
NAAFVGQVTR	ASCAFSYRVK	RMDIAJMDP	CVVGIAWRS	LFGDGVCEN	YKNPPEEFIN	1380
SATPVGIAKT	VNCQTYPVH	AVGDNFSNYS	ESBACRELAA	AYREVAKHET	BLGVNSVAIP	1440
LLSTGCVYSGG	KIRLJHSILH	LFTAMDISTN	DVVIYCRDKH	WEKKSJEAO	MRTTOVELL	1500
HISIUCDIVE	VHFDSPLAGR	KOYPTTNGAL	YSVLGRKRFH	QTAVINMAEH	TNSWEPQTBAN	1560
EQVNLVALGE	ZIESIRQKCP	VENDASSPP	RTVNECLARYA	MTSERVYTLA	MMRUTGJIIVC	1620
SSFPLPKYKI	EGVQNVKCSK	VMLQDHNVS	RVSREXRS	QESAOGRASFI	TSLLTHSQFIL	1680
SVDGSLTPV	SILIDAFAL	RFALDDGATH	TLPSSTTOLIA	AUSDWJMSTV	IVAPPKRRNG	1740
RNLTVTCDER	ECHITPPMASV	RFFRAELCPV	VOSTRTRITD	AMSLOAPPST	ATEPKHPPIS	1800
XGASSETVPI	TNGDFNREGI	RSLSSELLTT	GDPLAGEVDO	LTDGSLNTHCS	DTIKKIJLDR	1860
GGYIFSSDTG	PGHLGOKSVR	QSVLPVNTLQ	EVHEEKCYPP	ELDEAKROLL	LELQCESARM	1920
ANHRSYQSK	VERNQAAIIQ	RLXRGCRAL	MEETSKQVFT	RTTYDAPVYS	BDINVELSNP	1980
ESAVAACNEF	LARNYPTVSS	YQITDNYDAY	LDNTDGSERSC	LURATTNPCK	LASYPKQHAY	2040
HAPSIRSAVP	SPFONTLQNV	LAANTIONCN	VTQMEELPTL	DSAVENVECF	KRFAKQEQW	2100
SEFPAASPIRI	TTENLATYVT	KLXGPKAAAL	FRATRNLL	QSVPMEDRPTV	DSKRDWVKVTF	2160
GTKITEERFX	VQVIAQAEPL	ATAYLOGINA	ELVPLINAVL	LPNQNTLPTM	SANDPDAIA	2220
ANFKPGDTVL	ETDIOFPIKS	QDQGLALITAL	MLEDLGVDN	SLLDLJEAAP	GRISSECHLPT	2280
GTTRPKFAMM	KSGMFLTLV	NTLNNTIAS	EVLEDRDONG	ACRAGEDON	LIHGVVDEL	2340
MAARCATWWI	MEVKIIIDAVV	SLKAFYFCG	FILHDTWIST	ACRVADPLX	LFLGLSRPLAA	2400
GEQDQDNERK	ALAOEVIRWQ	RTSGLJDLSK	AVISYFPEVQ	ISVYVMEPAT	FASSEKSFPER	2460
LRGFVITLYG	CPK					2520

FIGURA 47.

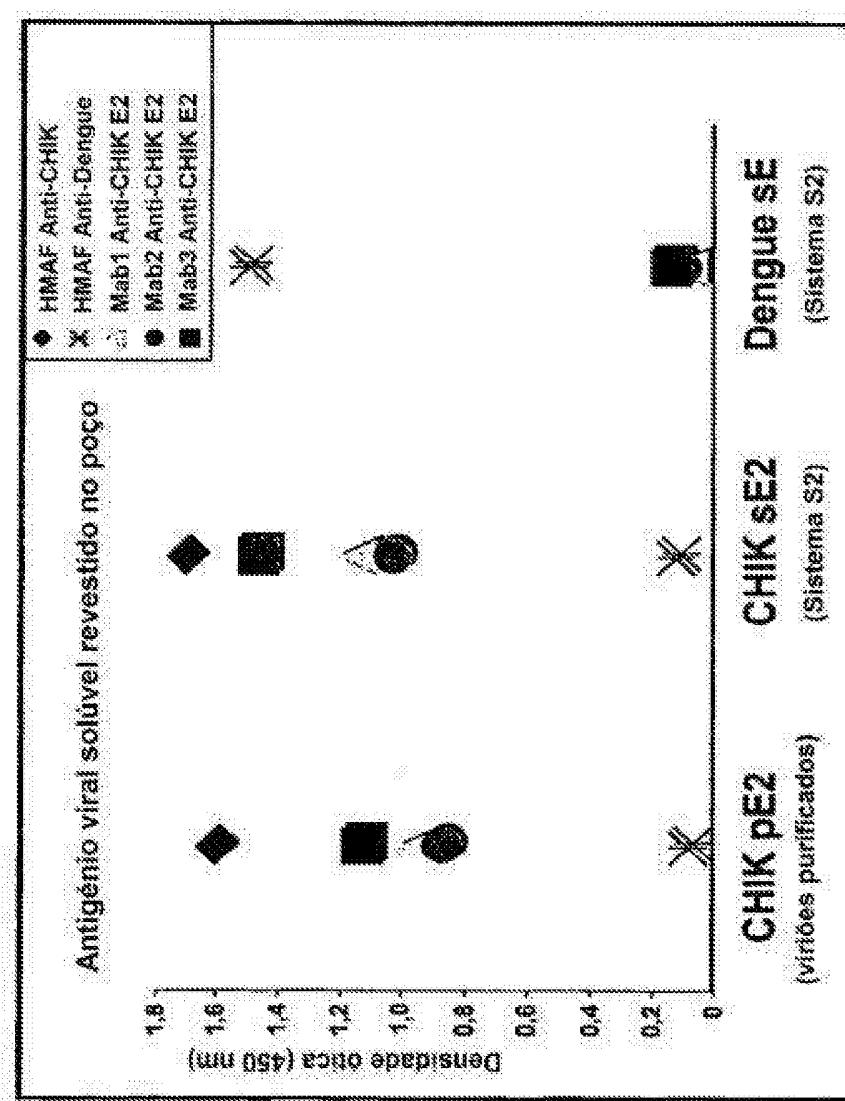
NEPVYVVDIEA	DSAFIKALQR	AYPEFVEPR	QVTRMHNAAA	RAPSHLAIKL	IEQEIEIPDST	60
ILSIGSAPAR	RNSMEDRKYNG	VCPMRSAEDP	SELLNYARNL	ASANGWVLDS	NISSSKIGDLQ	120
AVMVNPITET	FTFCIATDVN	CRCRADWVAY	QDVIAVHAPT	SLYHOAIRGV	EVAYWVGFDT	180
TEPMYXNAMG	AYFSYSTINWA	DECYLNKNI	SLCETDETEG	RNGKLSIMRG	KKIMPCLAVL	240
PSVGOSTLYPE	SRKLKSMHL	PSVFNLKKGKL	STFCRCDTYY	SCGYVVERI	TMEPGLYGET	300
TYVANTERAD	GFLMKCTDT	VDCGERSFSV	CTYVPRACID	QMGSLATEV	TREDRQCLIV	360
GLNQHIVVNG	RTORNNTNMK	HYLLMWWAQ	FENWAKCRX	DMEDEWLGLV	RENTLTOCCL	420
WAFKQSTHTY	YVFRPDQSI	OKVQASFDSF	VVPSELWSGL	SIFLNTRIKR	LLSKVPTEL	480
IPYSGDAREA	RDAEKEAKER	REAEKTREAL	PPLQNAQCV	QVEIDUQQLS	DRAGAGIIET	540
FRGAIKVIAQ	FTDHYWGSYL	VLSPOQTIVLES	QHLSLHHLA	BQVKTCTHNG	RAGRYAVRAY	600
EGRVILVPSCY	AISFEDFOSL	SESATHWVNE	REFVNUHJH	IAMPICALMT	DEBESVELVRA	660
ERTEHEYVYD	VIQRECCCNK	RANGLVLEED	LTPNPFYHETR	TESLKRIPAC	PYKIAVIGVY	720
QVPGSGESAK	JKNLVTNRQJL	VTSGKKNQ	KITNDVHQ	GLEISARTVD	SILLNGCNRF	780
VENLYVNEAF	ACHSSTILAL	IALVLRPROJV	VLCXEDPROCG	PPNNSCMYDN	YNNINCTQVY	840
HRSISERCTL	PVTAIVSSLR	YECRMETTSNE	YNEKLVWVTT	GSTKEDPGOL	VIATCPROMVK	900
QLOIDYRGSYE	VMIRAAQNL	TRNGVINVRO	KVNESTPLVAS	TSEHUVNLLT	RTERKLJWWT	960
LSGEFWINTL	QNPFGNFKA	TIKNEVEHA	SIMPAGCSHQ	MTFDTFQICA	NVCWAKSLVP	1020
ELSTAGIKNL	DRQNSQIQA	PKEDKAYSP	VALNRICTM	YGVBLISGLP	SKEPNVSVYTA	1080
INNNWNEKPGG	KNPGFNPRAA	SILERKYPFT	NGHONJNQOI	CVITTRKIEDP	NPTTTRLIPIAN	1140
BRLENSIVAS	HRFVNGRME	WLUNKINGHH	VLLVSGYHNA	LPTKRVYHVA	PLAVRGAET	1200
XNLELGLPAT	LGRYDLVVIN	INTPFRINN	QNCVNDKML	OMGGOSLRL	LRPGGSSLIR	1260
AYGVADRTSE	XVICVLGRKP	RSRRAKPC	VSNTENPFL	PSNPINGRRN	FTTHUNSNQ	1320
RAAFVGCVTRX	ACCAPSTRVX	RMDIAINDR	CYVNGAANPRG	LPGDGCKAV	YKIMFESFKN	1380
SATFVGTAKT	VMCGTYFVIH	AVGPNPSNYS	ESBENDRILAA	AYREVAKEVT	BLGVNSVAIR	1440
LLSTGTVYSGG	KIRLTOQLNH	LPTAMISTRA	DNVIYICRDK	WEKKISRAIQ	MRTQVELLIS	1500
HISIDICRIVR	VHPDSSLAGR	KSYSTTEGAL	YSYLDRTRPH	QTAIVGARSH	TMWPKQTEAN	1560
EQVCLYALGS	SEISSIRQNF	VTDAGASSPF	XTVPLCKPVA	MTPERVTRLE	MNATTSIIVC	1620
SSFFPLPKYKL	EGVQNVXCSK	VNLFDRHNVPS	IVGSPREYNESS	QSSAQASTI	TSLTNSQFOL	1680
SVDEGILPVP	SULEADAFAL	EPALEGATH	TLPSTTGNLA	AUSIDWMSIV	FVAPPKREKNG	1740
BNLTVICDER	EIGNITPMASV	EPPRAELCFV	VQESTASTRID	ANSLQAPEST	ATEPNNHPPIS	1800
PGASSETPTI	TFGDNNGEI	ESLSESELLTF	COFLNCEVDO	LTDSEWSTCS	UTIOKLLDNA	1860
QSYIPSSDTG	PGHLQQKSVR	QSVLPINTLE	EVHEEKCYPP	ELDEAKEQLL	LKXLQESASM	1920
ANRSKYQSRK	VERNKAATIQ	RKRGCRILY	MESETPXYPT	EITYPAPVYH	SPPINVRLENP	1980
ESAVARACNEF	LARNYPTVSS	YQITDYEYDAY	LOSWDGSRSC	LDRATNTPSK	LRSTYPQHAY	2040
NAFGRISAVP	SPPQNTLQNV	LAATIRNCN	VPQMRSLPTL	DSAPVNPVEC	KKFACNQEW	2100
REPTAASPIR	TTCNLATYVT	KLKGPKAAAL	FAXTNLLPL	QEVPNDRPTV	DMKGDWVYTP	2160
CTRTHTEEREM	VOVQIAAEPL	ATAYLOGIHE	ELVVERLNAVL	LENVNTLFD	SASDFDIA	2220
ANPFDGDTVL	STDIASFOKS	QDQSLALTAL	MILEDLQVDR	SLLDLIEAAF	GRISSECHLPT	2280
CTRFLYGMAM	KGSMFLTLFV	NTLINITLRS	RVLEDRLTKS	ACRAFIQION	IIHGVVSDEL	2340
NAARCATWN	REVKIICAVV	SLRAPHFCGG	FILHETVYGT	ACRVADPLXR	LFKLGRPLAA	2400
CDEDEDEERR	ALADEVIENQ	ETGLLUSLX	AVYSKYEVUG	ISVWVWSEHMT	PASSENPEK	2460
			LRGIVVITYG	GPX		2473

**FIGURA 48**

MDFVYVDIDA DSAFLKALOR AYPMFEVEFR QVTENHANAA RAFSHLAIKL IEQEIDPDST	60
ILGIGSAPAN RNMNSORKYIC VCPMRSAEOP KRLANYARKL ASAAGKVILR HISGGKIGDLQ	120
AVMGRVPIET PTFCLHTDVS CRORADWAIY QDVYAVRPT SLYWQAIKGV RVAYWVGFDI	180
TPFMYNAMAG ATPSYSTHNA DRDNLAKEIN GRCSTDLTEG ARSKLSEINERG KXJXFCIRVL	240
PSVCGSTLYPF SRKLLKSWHL PSVPHLKQKL SEPTCERCTVV SCGYVVKI TMSPGLYCKT	300
IGNANTHAD CFLMKITDT VDCHVSPSV CTYVPATCQ QMTGILATIV TPEDAQKLAV	360
GLNQHIVNG HTOJNTNTK NYLPLWQAQ FSGKNAKEKRC INSEDELLGV BERTLTCCCL	420
WAFFKQETHT VYRFOTOSI OKVDAEFDSE VVESLWSSGL SIPLRTREW LLSRWPXIL	480
IPYSGDAREA ROAKKEAEE REARLYTHEAL PPLQAAQEDV QWEIDUEQLE DRAGINGLIET	540
PRGAJIKVTAQ PTENVUGEYL VLSPQTVLRS QKLSSLHAA KQVKTCTHNG RAGRYAVEAY	600
DSPVVLVPGCY AISPEDFQSL SESATHWYNE PSFVNDRPLH IAMEIGEALAT DEESYKLARA	660
ERTERHEVYTD VDQRESCKEE KARPLVLED LTNPFFYHRPA YRGKTRPAC PVKIAVIGVP	720
CWPGSGESEAL INNLVTRQD VTSQKXENIQ EJTTVWRAQW QKASPARTVD SLLANGKRP	780
UDVLYVIEAP ACHSGTIAL IAIWVPRQW VLCCGDPNQCS FPNMSQMVN YNNHNICIQVY	840
NKHSISRKCTL PVTATVSESLH YBEMRMITTE YHKPIVMDTT GSTKRPQGOL VLTCPFGNVX	900
OLQIDYROYE VNTAAASQQL TRNGVYAVRQ KUNENPLYAS TEEHVNVILLE RTESKLVWKT	960
LAGEDPWIKTL QNPPKGNTKA TINENEVESA SIMAGICSHQ MIPDTTQNKA NYCWAKSLVP	1020
ILSTAGIKNL DRONSOIIQA FKEDKAYSFR VALMEICTRM VGVDLISGLP SKPLVSVYKA	1080
CSHHWNUHNGG KMPGENPEAA SILLSKYPFT KQONHMKOI CWPTRKIEOF NFTTNIJIFAN	1140
ERLPHSIVAE HRPVYGERME WLWINKHGT VLLVSGNMLA LPTKEVTWVA PLVVRGADYT	1200
YNLELGLPAT LCRYOLVVIN INTFRHIHY QMVTNDPMKL QMLGCCSLNL LKNGCISLLIR	1260
AYCVYDRTSE EVICVLGRPF RSSRALKFFC VTNTRCPFL FESFNGRNH FTTHVNRNQL	1320
KRAAFVGQVTR AGCRPSYHVK RMIAIENDKE CVVNAAMPNG LKEDGCKAV YKINPESPKH	1380
SATPVGTAKT VMCGITYPVH AVGFWTSNYS ESBDODRILAA AYREVAKEVY KLGVNSVAIP	1440
LLSTGIVYSGG KDRLTQSLAH LPTAMOSTA DVUTVYCHNE WERKLSRATQ MRTQVELLDS	1500
HISIDCDIVR VNPOSSLAGK KSYSTTEGAL YSYLEGTRFH QTAVERAELIH TMWFKOTEAR	1560
EQVCLVALGE SIESIREQCF VELIADASSFP KTVFCLCKYA MTFERVTELK MSHVTSILIVE	1620
GGFFLPKXKI EGONVNCSEK VNGJFDHSTVPS RVSPRETYASS QESACASTI TSIJTHSQFDL	1680
SDVGEILPVE SLDQADAFAL SFLDQXGAIN TLPSCTQNLIA AVGDWVMSTV FVANSHRRRG	1740
RNLTVTCIDEK EGNIITFMSV RPPRAELCPV VQETATRHTD AMELQAFGST ATENHPPILP	1800
GASSETPPIT PGDFNEGEIN SLSSELLTRG DFLPGCEVDEL TDSDWSTCED TDDELLDRAG	1860
GYIPESSITCP CHLOOKSVRQ SYLVPNTLKS VREKCYTPK LOZAGQOLLK XKLQESASMA	1920
NRSEKYOQSKV ENKKAACIQR LERSCRLYLM SETPENETTA TTYPXAVVSP SIVWRELSIPE	1980
SAVARACNEFL ARNYPTVSSY QITISBYDAYL BTWEDGSECL DEATPNPSKL RSYPKOHATH	2040
APSIRSAYPS PFQNTLQNL AAATHRNQW TQNSRELYTLD SAVTIVRCPTK EPACMQNTWE	2100
EPRASFINIT TEILATFVTK LKSPKMALE JXTTINLILQ EYVMIKPTVW MRGDIVVYTG	2160
TKHTEERPKV QVIOAASPLA TAYLGSITRN LVRALNAILV PIVHTLPTMG AEDPDAIIA	2220
NFKRGDTVLE TDIASPNSQ DRSALTAIM LLELILWODIS LLELILRAGC EISSCHLPTG	2280
TRFKFGGPNK SONFTLTFVN TLLINITASH VLEINRJKSA CRAFTGDDNI INGVVSEOKLM	2340
AARCATWMMF EVKITIRAVVS LKAPYFCGGF ILHETVISTA CRWADLERL FKIGKPLAAG	2400
DEQDURKDKA LADEVINMOR TGLIDELSEA VYSRYEVQCI SVVUMEMATE ASERSNPEKL	2460
RGFVITLYGG PK	2472

FIGURA 49

MPEVEPFIQVT PNEKANARAF SHIAKILIEG EIDPDESTLID TOSAPARRMM SDRKYHCVCP	60
MRSAAEDPERL ANYARKLASA AGKVLOOMIS GRIGDLOAVM AVPDITETPTF CLNTIDVSCRO	120
ENQWAIYQOV YAHIAPTSLY HOATEGVNA YWVGFDTTFF MYHAMAGAYP SYSTINWADEQ	180
VLEKAKNIGLC STILTREERG KLSITMRGKEL KNCIDKVLFSV GSTLYPESRK LLKSNHLPBV	240
PHLNOKRILSFT CRCUTVVSCE GYVKRITMS POLYGRITSY AVTHBADSPL MCXFTTDIVEG	300
ENVSFSVCTY VSTATICQNT GILATEVTPE DAQNLVGLN QRIVVNGRTQ RNTNTMGYL	360
LPVVAQAFSK NAKECRIMM DEKLJFVRP TLTCCOCMAF KIQKINTVYK RNUIQSIQKV	420
QAKFISPVVUH SLSSEHLSIP LRTRKINLIS KVFETDILPV SCDAEANRDA EKEARNEREA	480
KLTIBALPPL QAAQEDVQVE IDNEQLEDRA CGGIJSTPSG AJKYTAOFID HVVSEYLNLS	540
EQTVLRLSQL SLIHALAQV KTCYHNGRAG RYAVEAVDGR VLVPSCGYAIS PEDPQELSES	600
ANWVYNEREF VNKKLAHILAM HGSALNTDSS SYBLVRAEST ENKYYVVDVQ KROCKKRRAA	660
SLALVUGDLTN PPYHEFAYEGL KIRPACPFY IAVIGVNTF GSGLSATIKN LVTBQOLVTS	720
OKKESICQETT TIVMRQRGLL ISARTVDSLL LNSCQFVIV LYVDEAFACH SGTLLALIAL	780
VRPPQKCNVLC GEPEKQCGFPN KQKQKVNINN MICKQNTHQS ISKRNLWVU AIVSSLNTEG	840
SKKFTSNEYWK PIVVTTGST KFDGRGLVLT CFRGNVYKQLQ IDWRSYEVMT AAASQGLTRK	900
QWVAVRQEVM ENPLYACTSE HVNVLLTSTE GXLVWKTLSG DFWIATLCPN PKGNFRATIK	960
ENEVENASIM AGICSHQMFT DTQNQHEANVC WAKSLVFILE TAGTQHNRQ NSQJTOAFKE	1020
DKAISPEVAL NEICTRYQV DLSQGLFSKP LVSVVYYADNM WENRPOSSMF CPNPEAASIL	1080
BRNYPPFTKGK WNIINKQICVT TERRIDOFNPT TNLIRANRRL PHSLAVASHRP VRGERMENLV	1140
NSKINHHVLL VSGYNALAPT KRVINVAPLG VSGADTYTIL KLGFLATLG YDLYVVININT	1200
PPNIRHHYQOC VHNAMKQLM. CGCISLRLLF GCSLILRAGV YADRTTERRV CVLSEKEXPRESS	1260
HALEPPCVTS NTENMFPLSN FINGRRNFTT HVWNNQLNAA FVGQVTEAGC APSYRVKEND	1320
TAINDEECVV NAAMPRGLPG DGVKEAVYKK WPFESPENSAT PVGTTAKTVMC GTYVPIHANG	1380
PMFSSTSYSESE GORELAAAYR EVAKEVTRIG VNSVVAIFLIS TGVYSSGGDR LTQGLOHLFT	1440
AMDSTDADVV TYCRDREWK KISEAIIQNET QVELLDENIS IDCIVRVMF DSSLACRRTY	1500
STTEGALSY LEOSTRFHTA VINAETINTM PROTEANEQV CLYALGESIE SIROKCPVDO	1560
ADAESSPPETV PCLCRYAMTP ERVTRLRMNS VTSIIIVSSF PLPKYKLGKV QKVKCSKVL	1620
PDRHNVPSRVS FREYRSSQES AQESTASTSI THSQDILSVD GRILFVPSIL DAIJAPALEKA	1680
LDDGATHTLF STTGMLAAVS DWVNSTVPAF PFSRERCRNL TVTCDEREGN ITIPMASVRFF	1740
RAELCPVWQS TAETRDNTMS LQAPPSTATE FMUPPFISFGA SSETFPPITPG DFNRGEIERSL	1800
SSELLITFGDF LPGEVODDLTD SINSTCSDTD DRILLDRAGGY IFSSSITGFGH LQQKSVRQSV	1860
LPVNTLBEVH EKCYCFPKLD RAENQLLIZK LOEBSASKMR SRYQSRREVEN MKAATTORLK	1920
RGCRLYLNSE TPKVPTTYRTT YPAPVYSPPI NVRLISNPESA VAACNRFIAR NYFTVSESYOI	1980
TDEYDAYLDM VDGSESCLOR ATFPNSPSKLRS YPKONAYHAP SIRSAPVSPFF QNTLQNVLA	2040
ATKRNNCNVTQ MRELPFTLGA VFNVNECPKIF ACNQRYWIEF AASFIRITTE NLATYVTKLK	2100
GPKAAALPAK THNLLPLOEV PMERPTVIMK RQVIVTPTGX HTERZERPVQV IQAAEPLATA	2160
YLGGINRRELW RHLNAVLLNN VHTLFIMSAE DFDALLARIF KIGINIVLETD IASFDKSQD	2220
SLALTALMIL BOLGDVHSLL SLLRNAGEI SSCHLPTGTR FKPGNMMEBG MFLCLFVNTL	2280
INITIASRVL EDRLTKSACA AFISGDNILH GVVSDELMAA KCTTWMEV KIIDAVVSLK	2340
APYFCGGFIL KDTVYGTACK VADPLKRLFK LOKPLAMSON QEDRDRNALA DEVIRWQRTG	2400
LIDELEKAVY SRYEVQGIVS VVMSMATEAS SRSNFTEGLRG PVTTLXGGPK	2460

**FIGURA 50**

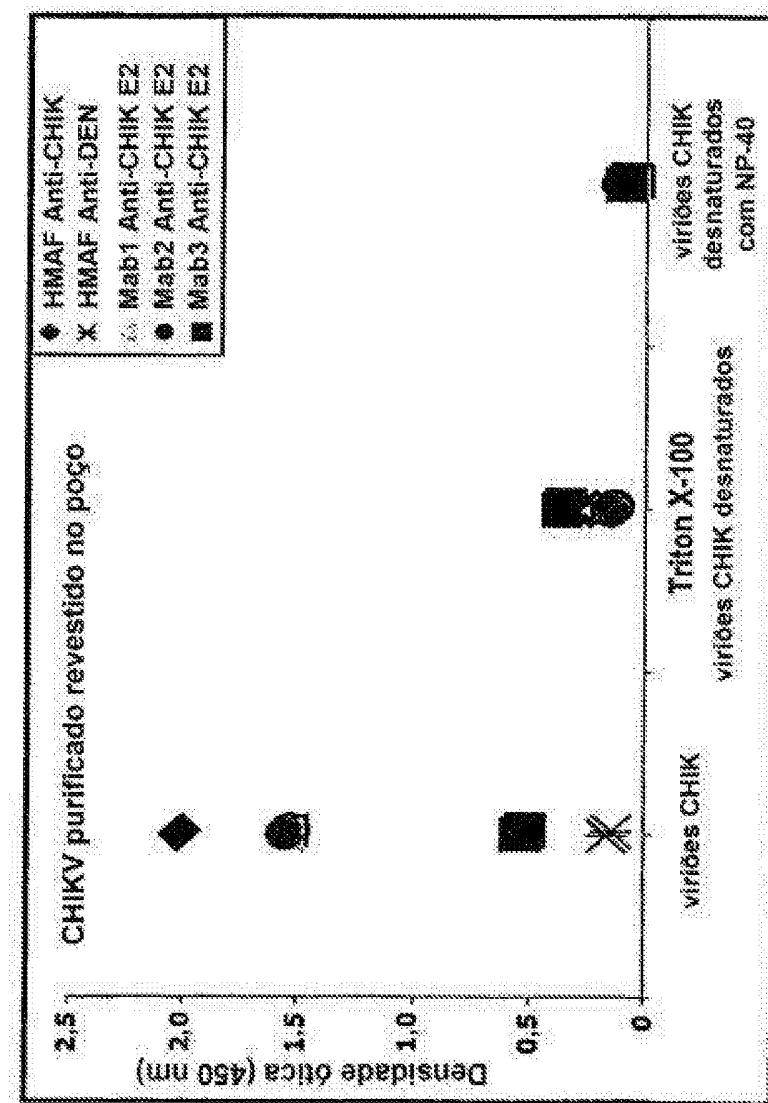
**FIGURA 51**

FIGURA 52

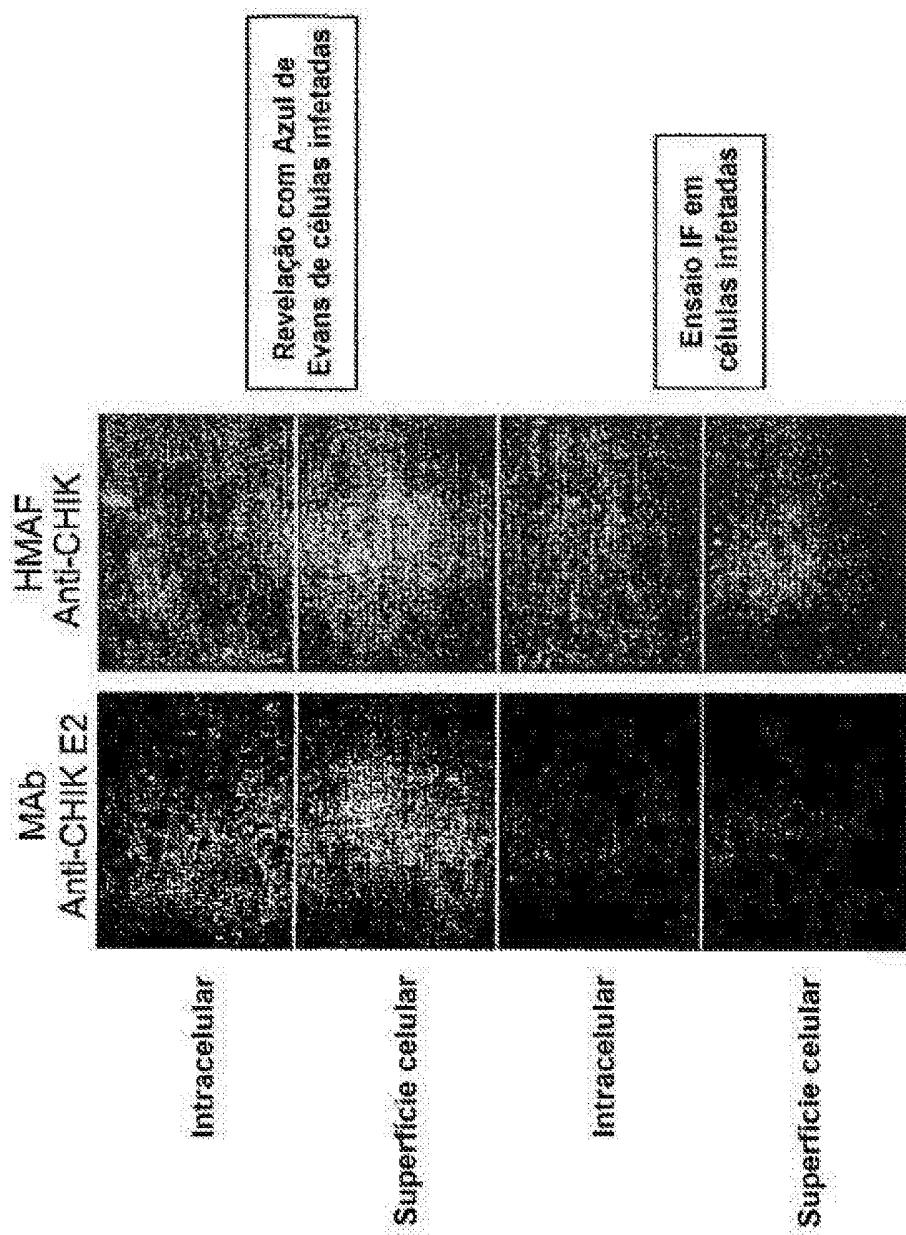


FIGURA 63

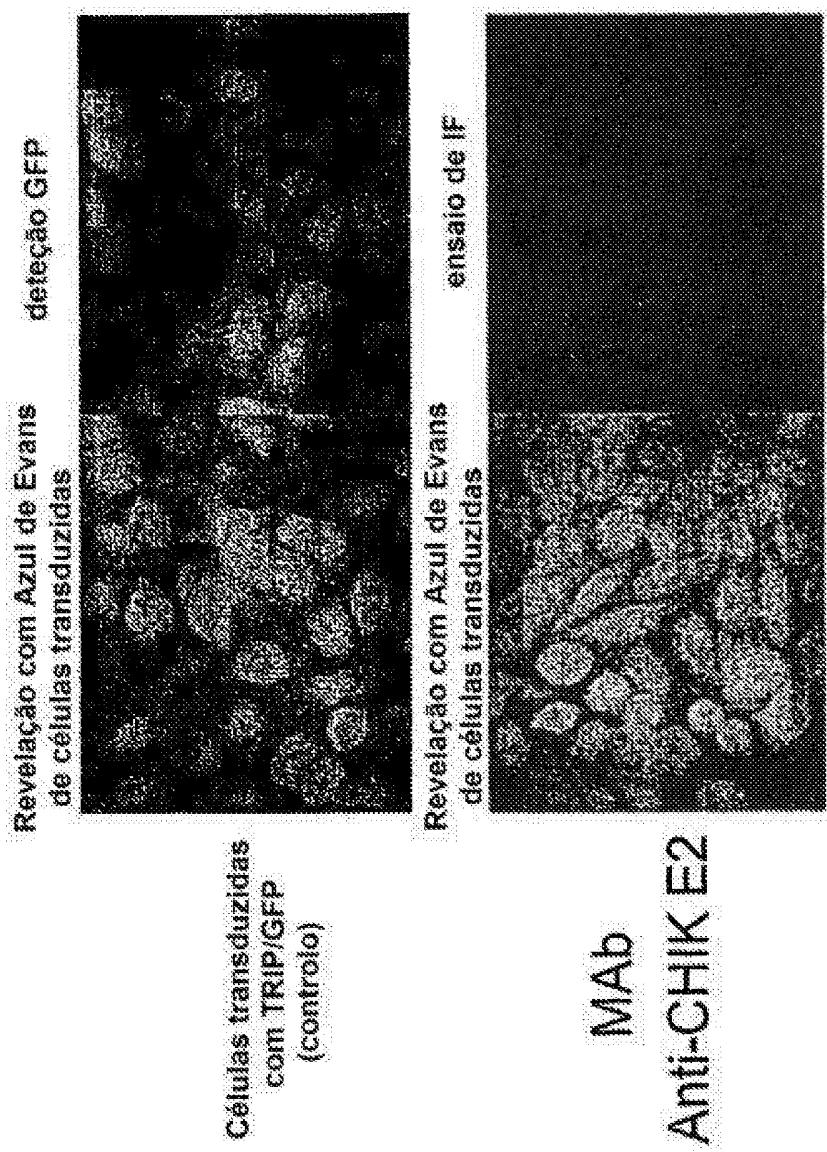
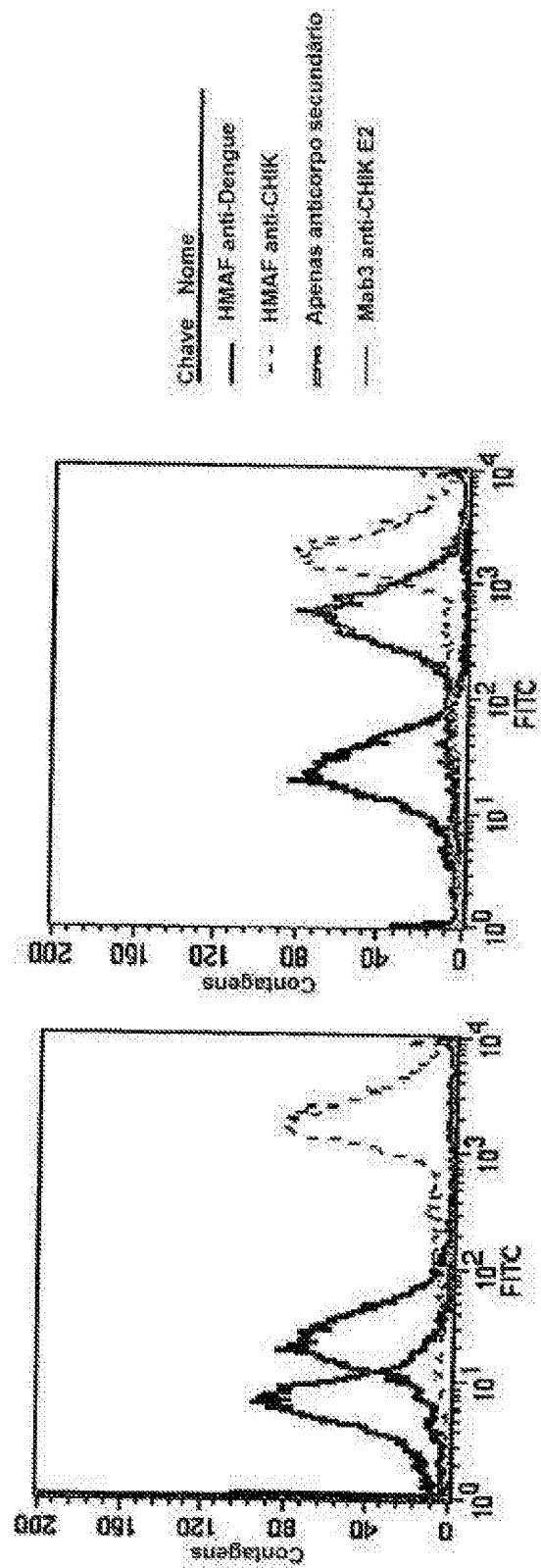
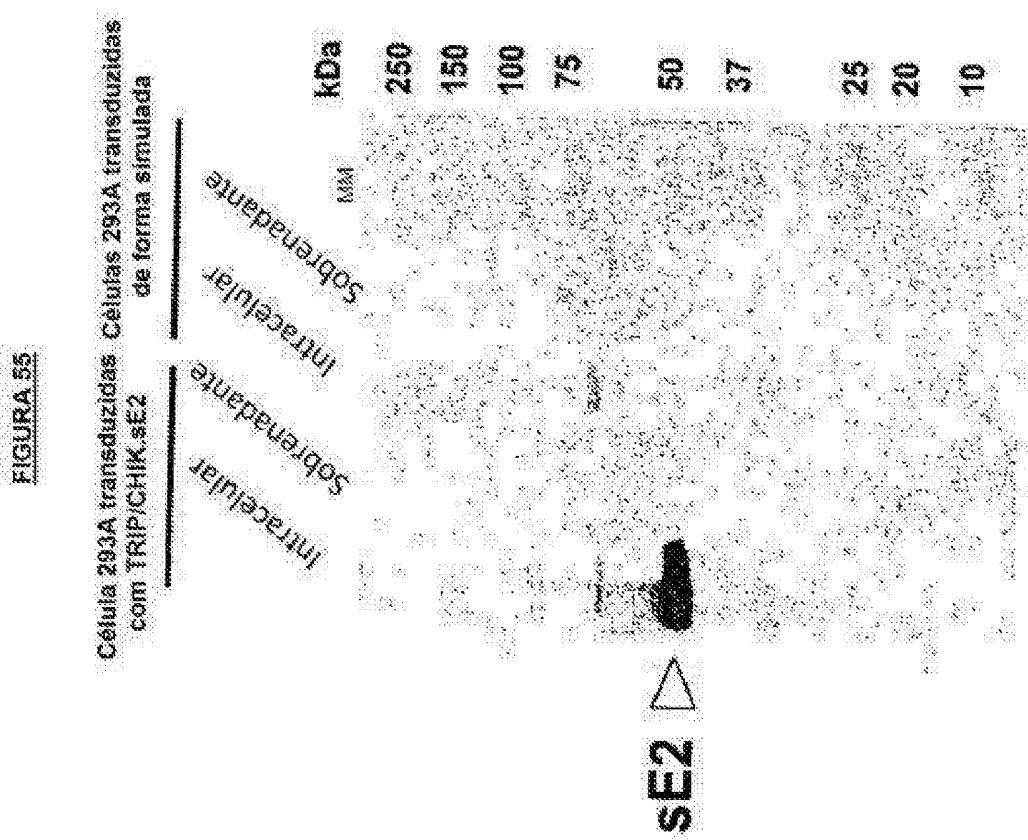


FIGURA 54





## REFERÊNCIAS CITADAS NA DESCRIÇÃO

Esta lista de referências citadas pelo requerente é apenas para conveniência do leitor. A mesma não faz parte do documento da patente europeia. Ainda que tenha sido tomado o devido cuidado ao compilar as referências, podem não estar excluídos erros ou omissões e o IEP declina quaisquer responsabilidades a esse respeito.

### Literatura que não é de patentes citada na Descrição

- \* Structure and replication of the alphavirus genome. STRAUSS EG ; STRAUSS JH. *The Togaviridae and Flaviviridae*. Plenum Press, 1986, 35-90
- \* Antigenic characteristics and classification of the Togaviridae. PORTERFIELD JH. *The Togaviruses*. Academic Press, 1980, 13-48
- \* ROSS RW. The Newala epidemic. III. The virus: isolation, pathogenic properties and relationship to the epidemic. *J Hyg*, 1960, vol. 54, 177-191
- \* Chikungunya disease. JUPP PG ; MCINTOSH BM. *The Arboviruses: epidemiology and ecology*. CRC Press, 1988, 137-13 157
- \* Alphaviruses associated primarily with fever and polyarthritides. JOHNSTON RE ; PETERS CJ. *Fields Virology*, 1996, vol. 16, 843-898
- \* PASTORINO B ; MUYEMBE-TAMFUM JJ ; BES-SAUD M ; TOCK F ; TOLOU H et al. Epidemic resurgence of Chikungunya virus in democratic Republic of the Congo: identification of a new central African strain. *J Med Virol*, 2004, vol. 74, 277-282
- \* LARAS K ; SUKRI NC ; LARASATI RP ; BANGS MJ ; KOSIM R et al. Tracking the re-emergence of epidemic chikungunya virus in Indonesia. *Trans R Soc Trop Med Hyg*, 2005, vol. 99, 128-141
- \* PAQUET C ; QUATRESONS I ; SOLET JL ; SIS-SOKO D ; RENAULT P. Chikungunya outbreak in Reunion: epidemiology and surveillance, 2005 to early January 2006. *Eurosurveillance weekly*, 2006, vol. 11, 2
- \* KHAN AH ; MORITA K ; PARQUET MD MDEL C ; HASEBE F ; MATHENGE EG et al. Complete nucleotide sequence of chikungunya virus and evidence for an internal polyadenylation site. *J Gen Virol*, 2002, vol. 83, 3075-3084
- \* POWERS AM ; BRAULT AC ; TESH RB ; WEAVER SC. Re-emergence of Chikungunya and O'nyong-nyong viruses: evidence for distinct geographical lineages and distant evolutionary relationships. *J Gen Virol*, 2000, vol. 81, 471-479
- \* GORDON D AC ; GREEN P. Consed: a graphical tool for sequence finishing. *Genome Res*, 1998, vol. 8, 195-202
- \* ROZAS J ; SANCHEZ-DELBARRO JC ; MESSEG-UEX X ; ROZAS R. DnaSP: DNA 2 polymorphism analyses by the coalescent and other methods. *Bioinformatics*, 2003, vol. 19, 2486-2497
- \* XIA X ; XIE Z. DAMBE: software package for data analysis in molecular biology and evolution. *J Hered*, 2001, vol. 92, 371-373
- \* THOMPSON JD ; HIGGINS DG ; GIBSON TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 1994, vol. 22, 4673-4680
- \* FELSENSTEIN J. PHYLIP - Phylogeny Inference Package (version 3.2). *Cladistics*, 1989, vol. 5, 164-168
- \* HOFACKER IL. Vienna RNA secondary structure server. *Nucleic Acids Res*, 2003, vol. 31, 3429-3431
- \* KUMARS ; TAMURA K ; NEI M. MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform*, 2004, vol. 5, 158-163
- \* MARTIN DP ; WILLIAMSON C ; POSADAS D. RDP: recombination detection and analysis from sequence alignments. *Bioinformatics*, 2005, vol. 21, 260-262
- \* ROUSSEL A ; LESCAR J ; VANAY MC ; WENGER G ; WENGLER G et al. Structure and interactions at the viral surface of the envelope protein E1 of semliki forest virus. *Structure*, 2003, vol. 14, 75-86
- \* CARSON M. Ribbon models of macromolecules. *J Mol Graph*, 1987, vol. 5, 103-106
- \* STRAUSS JH ; STRAUSS EG. The alphaviruses: gene expression, replication, and evolution. *Microbiol Rev*, 1994, vol. 58, 491-562
- \* LAVERGNE A ; THOISY BD ; LACOSTE V ; PAS-CALISH ; POULIQUEN JF et al. Mayaro virus: Complete nucleotide sequence and phylogenetic relationships with other alphaviruses. *Virus Res* in press, 2006
- \* LANCIOTTI RS ; LUDWIG ML ; RWAGUMA EB ; LUTWAMA JJ ; KRAM TM et al. Emergence of epidemic O'nyong-nyong fever in Uganda after a 35-year absence: genetic characterization of the virus. *Virology*, 1998, vol. 252, 258-268
- \* STRAUSS EG ; LEVINSON R ; RICE CM ; DALRYMPLE J ; STRAUSS JH. Nonstructural proteins nsP3 and nsP4 of Ross River and O'Nyong-nyong viruses: sequence and comparison with those of other alphaviruses. *Virology*, 1988, vol. 164, 265-274
- \* Alphaviruses. GRIFFIN DE. *Fields Virology*. Lippincott Williams & Wilkins, 2001, 917-932

- VASHISHTHA M ; PHALEN T ; MARQUARDT MT ; RYU JS ; NG AC et al. A single point mutation controls the cholesterol dependence of Semliki Forest virus entry and exit. *J Cell Biol*, 1998, vol. 140, 91-99.
- AHN A ; SCHOEPP RJ ; STERNBERG D ; KIELIAN M. Growth and stability of a cholesterol-independent Semliki Forest virus mutant in mosquitoes. *Virology*, 1999, vol. 262, 452-456.
- WILLIAMS MC ; WOODALL JP ; CORBET PS ; GILLETT JD. O'nyong-Nyong Fever: An Epidemic Virus Disease in East Africa. 8. Virus Isolations From Anopheles Mosquitoes. *Trans R Soc Trop Med Hyg*, 1965, vol. 59, 300-306.
- WEAVER SC ; BARRETT AD. Transmission cycles, hostrange, evolution and emergence of arboviral disease. *Nat Rev Microbiol*, 2004, vol. 2, 789-801.
- CASSESE T ; KIELIAN M. The cholesterol requirement for sindbis virus entry and exit and characterization of a spike protein region involved in cholesterol dependence. *J Virol*, 1999, vol. 73, 4272-4278.
- HOLLAND J ; SPINDLER K ; HORODYSKI F ; GRABAU E ; NICHOL S et al. Rapid evolution of RNA genomes. *Science*, 1982, vol. 215, 1577-1585.
- DOMINGO E ; HOLLAND JJ. RNA virus mutations and fitness for survival. *Annu Rev Microbiol*, 1997, vol. 51, 151-178.
- VIGNUZZI M ; STONE JK ; ARNOLD JJ ; CAMERON CE ; ANDINO R. Quasispecies diversity determines pathogenesis through cooperative interactions in a viral population. *Nature*, 2006, vol. 439, 344-348.
- KIM KH ; RUMENAPP T ; STRAUSS EG ; STRAUSS JH. Regulation of Semliki Forest virus RNA replication: a model for the control of alphavirus pathogenesis in invertebrate hosts. *Virology*, 2004, vol. 323, 153-163.
- HEISE C ; KIRN DH. Replication-selective adenoviruses as oncolytic agents. *J Clin Invest*, 2000, vol. 105, 847-851.
- HEISE MT ; WHITE LJ ; SIMPSON DA ; LEONARD C ; BERNARD KA et al. An attenuating mutation in nsP1 of the Sindbis-group virus SAAR86 accelerates non-structural protein processing and up-regulates viral 26S RNA synthesis. *J Virol*, 2003, vol. 77, 1149-1156.
- SUTHAR MS ; SHABMAN R ; MADRICK K ; LAMBETH C ; HEISE MT. Identification of adult mouse neurovirulence determinants of the Sindbis virus strain AR86. *J Virol*, 2005, vol. 79, 4219-4228.
- CONDON RJ ; ROUSE IL. Acute symptoms and sequelae of Ross River virus infection in South-Western Australia: a follow-up study. *Clin Diagn Virol*, 1995, vol. 3, 273-284.
- SELDEN SM ; CAMERON IRON AS. Changing epidemiology of Ross River virus disease in South Australia. *Med J Aust*, 1996, vol. 165, 313-317.
- MAZAUD R ; SALAUN JJ ; MONTABONE H ; GOUBE P ; BAZILLIO R. Troubles neurologiques et sensoriels aigus dans la dengue et la fièvre à Chikungunya. *Bull Soc Pathol Exot*, 1971, vol. 64, 23-30.
- NIMMANNITYA S ; HALSTEAD SB ; COHEN SN ; MARGIOTTA MR. Dengue and chikungunya virus infection in man in Thailand, 1962-1964. I. Observations on hospitalized patients with hemorrhagic fever. *Am J Trop Med Hyg*, 1969, vol. 18, 954-971.
- GRATZ NG. Critical review of the vector status of *Aedes albopictus*. *Med Vet Entomol*, 2004, vol. 18, 215-227.
- LESCAR J ; ROUSSEL A ; WIEN MW ; NAVAZA J ; FULLER SD et al. The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenek activation at endosomal pH. *Cell*, 2001, vol. 105, 137-148.