(19)中华人民共和国国家知识产权局

(12)发明专利

(10)授权公告号 CN 104918952 B

(45)授权公告日 2019.01.22

(21)申请号 201380055697.6

(22)申请日 2013.09.26

(65)同一申请的已公布的文献号

申请公布号 CN 104918952 A

(43)申请公布日 2015.09.16

(30)优先权数据

12306176.4 2012.09.27 EP

(85)PCT国际申请进入国家阶段日

2015.04.24

(86)PCT国际申请的申请数据

PCT/EP2013/070137 2013.09.26

(87)PCT国际申请的公布数据

WO2014/049094 EN 2014.04.03

(83)生物保藏信息

CNCM I-2889 2002.06.12

CNCM I-3824 2007.09.06

(73)专利权人 巴斯德研究院

地址 法国巴黎

专利权人 特密斯生物科学有限责任公司

国立科学研究中心

(72)发明人 F·坦吉 S·布兰德勒

P·德普雷 A·哈贝尔

(74)专利代理机构 永新专利商标代理有限公司 72002

代理人 王健

(51)Int.Cl.

C07K 14/18(2006.01)

A61K 39/00(2006.01)

A61K 39/12(2006.01)

A61K 39/295(2006.01)

C12N 7/04(2006.01)

C12N 15/86(2006.01)

(56)对比文件

CN 102317308 A,2012.01.11,

Samantha Brandler等.Pediatric measles vaccine expressing a dengue tetravalent antigen elicits neutralizing antibodies against all four dengue viruses.《Vaccine》.2010,第28卷(第41期),第6730、6733页.

Samantha Brandler,et al..Recombinant vector derived from live attenuated measles virus: Potential for flavivirus vaccines.《Comparative Immunology, Microbiology & Infectious Diseases》.2007,第31卷(第2-3期),第271-281页.

审查员 张娜

权利要求书2页 说明书22页

序列表90页 附图10页

(54)发明名称

表达屈曲病毒多肽的重组麻疹病毒及其应用

(57)摘要

本发明涉及表达屈曲病毒多肽的麻疹病毒，特别涉及在其表面含有屈曲病毒包膜和壳体蛋白的病毒样颗粒(VLP)。这些颗粒是施用后能在宿主中复制的重组感染性颗粒。本发明提供产生这些重组感染性颗粒的工具，特别是核酸、载体、细胞和拯救系统。本发明还涉及这些重组感染性颗粒，特别是以组合物形式，更特别是以疫苗剂型用于治疗或预防屈曲病毒感染的用途。

CN 104918952 B

1.核酸构建体,其包含编码屈曲病毒(CHIKV)的C-E3-E2-6K-E1结构蛋白的多核苷酸,所述多核苷酸可操作地连接至cDNA分子,所述cDNA分子编码麻疹病毒(MV)全长、感染性反基因组(+)RNA链的核苷酸序列。

2.权利要求1的核酸构建体,特征在于所述核酸构建体符合麻疹基因组的6倍规则。

3.权利要求1的核酸构建体,其包含基因转录单元,从5'到3'包括:

(a)编码MV的N蛋白的多核苷酸,

(b)编码MV的P蛋白的多核苷酸,

(c)编码CHIKV的C-E3-E2-6K-E1结构蛋白的多核苷酸,

(d)编码MV的M蛋白的多核苷酸,

(e)编码MV的F蛋白的多核苷酸,

(f)编码MV的H蛋白的多核苷酸,和

(g)编码MV的L蛋白的多核苷酸,

所述多核苷酸可操作地连接至所述核酸构建体,并受病毒复制和转录调控序列控制。

4.权利要求3的核酸构建体,特征在于所述病毒复制和转录调控序列是MV先导和尾随序列。

5.权利要求1的核酸构建体,特征在于所述麻疹病毒为减毒病毒毒株。

6.权利要求5的核酸构建体,特征在于所述减毒病毒毒株选自施瓦兹毒株、萨格勒布毒株、AIK-C毒株和Moraten毒株。

7.权利要求1的核酸构建体,特征在于编码CHIKV的C-E3-E2-6K-E1结构蛋白的所述多核苷酸已经为Macacca密码子使用优化或为人密码子使用优化。

8.权利要求1的核酸构建体,特征在于编码CHIKV的C-E3-E2-6K-E1结构蛋白的所述多核苷酸包含至少一个选自被突变的AAAGGG、AAAAGG、GGGAAA以及GGGGAA的麻疹编辑样序列。

9.权利要求1的核酸构建体,特征在于所述屈曲病毒来自命名为06-49毒株的毒株。

10.权利要求1的核酸构建体,特征在于包膜蛋白E2的胞外域具有SEQ ID NO:11、13、15的序列。

11.权利要求1的核酸构建体,特征在于所述C-E3-E2-6K-E1结构蛋白由SEQ ID NO:21、22、23、24、25、26和28限定。

12.权利要求1的核酸构建体,特征在于所述核酸构建体包含SEQ ID NO:20、27和31限定的序列。

13.转移载体质粒,其包含权利要求1-12任一项的核酸构建体。

14.细胞,特征在于所述细胞用权利要求1-12任一项的核酸构建体或权利要求13的转移载体质粒转化。

15.权利要求14的细胞,其是真核或禽类或CEF细胞、哺乳动物细胞或酵母细胞。

16.重组感染性复制MV-CHIK病毒颗粒,其包含权利要求1-12任一项的核酸构建体作为其基因组。

17.权利要求16的重组感染性复制MV-CHIK病毒颗粒,特征在于所述病毒颗粒的基因组中包含多核苷酸序列,所述多核苷酸序列包含由SEQ ID NO:27或31限定的序列。

18.活性成分组合物或组合体,其包含与CHIKV-C-E3-E2-6K-E1VLP联用的权利要求16或17的重组感染性复制MV-CHIK病毒颗粒以及药物学可接受的载体。

19.权利要求18的活性成分组合物或组合体在制备疫苗中的用途，所述疫苗通过引发针对所述CHIKV蛋白的抗体和/或细胞免疫应答，用于在宿主中引发针对CHIK病毒的保护性免疫应答。

20.权利要求19的用途，其中所述宿主是有需要的人宿主。

21.与CHIKV-C-E3-E2-6K-E1VLP联用的权利要求16或17的重组感染性复制MV-CHIK病毒颗粒、或权利要求18的活性成分组合物或组合体在制备用于预防受试者中屈曲病毒的感染的疫苗中的用途。

22.权利要求21的用途，其中所述受试者是人。

23.拯救表达屈曲病毒(CHIKV)的C-E3-E2-6K-E1结构蛋白或CHIKV-C-E3-E2-6K-E1病毒样颗粒(VLP)的重组麻疹病毒(MV)的方法，包含以下步骤：

1)用(i)权利要求13的转移载体质粒以及(ii)编码MV L聚合酶的载体共转染稳定表达T7RNA聚合酶以及麻疹N和P蛋白的辅助细胞；

2)在能够产生MV-CHIKV重组病毒的条件下培养所述共转染的辅助细胞；

3)通过将步骤2)的所述辅助细胞与能够使病毒增殖的细胞共培养，增殖由此产生的重组病毒；

4)回收复制的MV-CHIKV重组病毒和CHIKV的C-E3-E2-6K-E1结构蛋白或CHIKV-C-E3-E2-6K-E1VLP。

24.权利要求23的方法，其中所述转移载体质粒包含由SEQ ID NO:27或31限定的序列。

25.权利要求23的方法，其中步骤1)的辅助细胞是HEK293辅助细胞。

26.权利要求23的方法，其中步骤3)能够使病毒增殖的细胞是Vero细胞。

表达屈曲病毒多肽的重组麻疹病毒及其应用

技术领域

[0001] 本发明涉及表达屈曲(Chikungunya)病毒多肽的重组麻疹病毒，特别涉及在其表面含有屈曲病毒包膜和壳体蛋白的病毒样颗粒(VLP)。这些颗粒是施用后能在宿主中复制的重组感染性颗粒。本发明提供产生这些重组感染性颗粒的工具，特别是核酸、载体、细胞和拯救系统。本发明还涉及这些重组感染性颗粒，特别是以组合物形式，更特别是以疫苗剂型用于治疗或预防屈曲病毒感染的用途。

背景技术

[0002] 屈曲病毒(CHIKV)是披膜病毒科(Togaviridae)家族甲病毒属(Alphavirus)的正链RNA病毒，于1952年在坦桑尼亚首次分离。

[0003] 这种病毒的感染引起人类疾病，其特征在于与登革热类似的症状，具有2-5天的急性发热期，随后是长期的关节痛疾病，影响四肢的关节。CHIKV是非洲、印度和东南亚地方性的，通过伊蚊(Aedes mosquitoes)经都市或森林传播循环传播。2006年，印度洋的多个岛(科摩罗、毛里求斯、塞舌尔群岛、马达加斯加、留尼汪岛…)爆发了CHIKV热，随后波及印度，在印度据报道预计140万病例。最近，欧洲报道了输入性感染，意大利报道了约200例地方性病例(Jose,J.et al.,A structural and functional perspective of alphavirus replication and assembly.Future Microbiol,2009.4(7):p.837-56)。临床上，该CHIKV流行伴随着比之前爆发更严重的症状，据报道有严重的多关节痛和肌痛、并发症和死亡。

[0004] CHIKV基因组是11.8kb的具有正极性的单链RNA分子。该病毒与西门利克森林病毒(SFV)、辛德毕斯病毒(SINV)和其他旧大陆甲病毒密切相关，而与新大陆甲病毒如委内瑞拉马脑炎病毒较为疏远(Griffin,D.E.,Alphaviruses,in Fields Virology,5th ed.,D.M.Knipe,Editor 2007,Wolters Kluwer,Lippincott Williams&Wilkins.p.1023-1067)。基因组RNA经加帽并直接翻译成名为P1234的全长非结构聚蛋白(nsP)，其由基因组5′的三分之二编码(Jose,J.,J.E.Snyder,and R.J.Kuhn,A structural and functional perspective of alphavirus replication and assembly.Future Microbiol,2009.4(7):p.837-56;Kuhn,R.J.,Togaviridae:the viruses and their replication,in Fields Virology,5th ed.,D.M.Knipe,Editor 2007,Wolters Kluwer,Lippincott Williams&Wilkins.p.1001-1022)。该前体自我切割产生P123和具有RNA依赖的RNA聚合酶活性的nsP4。这些蛋白与细胞辅因子一起装配成产生反义基因组RNA分子的复制复合物。P123随后切割成nsP1和P23，产生制备有义和反义基因组RNA的聚合酶复合物。P23进一步切割成nsP2和nsP3，产生仅制备正义基因组RNA分子的聚合酶复合物。除了复制病毒基因组，该病毒蛋白复合物从病毒基因组的3′端转录26S亚基因组RNA。该信使RNA翻译为聚蛋白前体，其由病毒和细胞酶的组合切割产生壳体蛋白(C)、两个主要的包膜蛋白(E1和E2)以及两个较小的附属肽E3和6k。一旦装配，CHIKV病毒体为直径65-70nm的球形颗粒，基本上由与壳体蛋白相关的基因组RNA分子组成，并包被在宿主来源的脂膜中，其由组装成二十面体晶格的E1-E2异二聚体装饰(Voss,J.E.,et al.,Glycoprotein organization of Chikungunya virus particles revealed by X-ray crystallography.Nature,2010.468(7324):p.709-12)。

[0005] 病毒演化和向新的地理区域传播所体现的疾病的严重性是需要应对的严重的公众健康问题。为了解决该问题,已经开发了具有减毒活病毒、具有嵌合甲病毒、具有重组DNA或具有病毒样颗粒的疫苗。

[0006] 一种福尔马林失活疫苗经证实在非人灵长类和人中具有免疫原性,但需要在BSL-3条件下制备大规模免疫所需的大量抗原限制了该策略的发展(Tiwari,M.,et al.,Assessment of immunogenic potential of Vero adapted formalin inactivated vaccine derived from novel ECSA genotype of Chikungunya virus.(Vaccine,2009.27(18):p.2513-22)。由美国陆军开发的减毒活TSI-GSD-218CHIKV疫苗具有免疫原性,但在II期临床试验中产生与毒力回复有关的副作用,引发安全问题。因此,尽管基于获得减毒活病毒的疫苗所获结果表明通过这种方式能获得有效的免疫,这些疫苗仍旧受到质疑,因为存在可能副作用的风险(Edelman R et al.,Am J Trop Med Hyg.2000Jun;62(6):681-685)。编码来自CHIKV的E1、E2和壳体蛋白的嵌合甲病毒疫苗策略在小鼠中具有免疫原性(Wang,E.,et al.,Chimeric alphavirus vaccine candidates for Chikungunya.(Vaccine,2008.26(39):p.5030-9),但甲病毒易于重组的能力产生了影响这些策略发展的安全问题(Weaver,S.C.,et al.,Recombinational history and molecular evolution of western equine encephalomyelitis complex alphaviruses.J Virol,1997.71(1):p.613-23)。

[0007] 另一种已开发的策略是设计重组DNA构建体用作疫苗。编码E1、E2和壳体蛋白的基于DNA的CHIKV疫苗已被证实在小鼠和非人灵长类中具有免疫原性(Muthumani,K.,et al.,Immunogenicity of novel consensus-based DNA vaccines against Chikungunya virus.Vaccine,2008.26(40):p.5128-34;Mallilankaraman,K.,et al.,A DNA vaccine against chikungunya virus is protective in mice and induces neutralizing antibodies in mice and nonhuman primates.PLoS Negl Trop Dis,2011.5(1):p.e928),但DNA策略不引发清除人中CHIKV所需的强烈的中和性免疫应答。DNA疫苗的缺陷在于需要大量DNA引发免疫应答,并且必须进行多次加强接种。对多次加强以及大量DNA注入许多细胞的核内的需求引发对于DNA疫苗能够整合入宿主DNA并造成插入突变的忧虑。因此,最近一项研究报道了将DNA疫苗与减毒活病毒组合使用(WO2011/082388)。尽管该技术允许减少减毒活病毒与DNA疫苗的缺点,但仍然存在提供副作用降低的疫苗的需求。

[0008] 为了避免减毒活病毒与DNA疫苗的缺点,已经开发了其他类型的疫苗,例如基于病毒样颗粒(VLP)的疫苗,其通过表达屈曲病毒结构蛋白获得。这些结构蛋白能在病毒样颗粒中自装配。在此基础上,已经开发了包含编码所有屈曲病毒结构蛋白的多核苷酸的疫苗(Akahata,W.,et al.,A virus-like particle vaccine for epidemic Chikungunya virus protects nonhuman primates against infection.Nat Med,2010.16(3):p.334-8)。然而,体外产生的VLP制备昂贵,并且完全免疫需要三次施用,因此这些疫苗价格不菲。Akata et al.中公开的CHIKV VLP策略需要用佐剂进行几次免疫以诱导保护。出于这个原因,仍然存在设计改进的疫苗的需求,所述疫苗能够使CHIKV VLP在感染的细胞,特别是感染的宿主细胞中体内产生,并由此提供有效而持久的免疫,所述疫苗尤其是仅在单次或两

次施用步骤后诱导持久的免疫。

[0009] 发明概述

[0010] 为此，本发明人实现了产生基于重组感染性复制麻疹病毒的疫苗，所述麻疹病毒由编码屈曲病毒抗原的多核苷酸重组，当重组病毒在施用后特别是在宿主中复制时回收所述疫苗。因此，本发明涉及基于广泛使用的施瓦兹(Schwarz)小儿麻疹疫苗的活CHIKV疫苗活性成分。在优选实施方案中，该重组活MV-CHIKV疫苗通过在感染细胞中复制产生CHIK病毒样颗粒。

[0011] 麻疹病毒是副粘液病毒科(Paramyxoviridae)家族麻疹属(Morbilivirus)的不分节单链反义包被RNA病毒。该病毒分离于1954年(Enders,J.F.,and T.C.Peebles.1954.Propagation in tissue cultures of cytopathogenic agents from patients with measles.Proc.Soc.Exp.Biol.Med.86:277-286.)，减毒活疫苗来源于该病毒，由此产生疫苗株，并且特别来自施瓦兹株。过去30年麻疹疫苗已施用于数亿儿童，其效力和安全性已经证实。其在许多国家大规模生产并低价分配。出于所有这些原因，本发明人使用减毒麻疹病毒产生稳定表达屈曲病毒结构抗原的重组麻疹病毒颗粒，特别是VLP。

[0012] 因此，本发明涉及核酸构建体，其包含编码至少一种屈曲病毒(CHIKV)结构抗原的多核苷酸，所述多核苷酸可操作地连接、特别是克隆入编码麻疹病毒(MV)全长、感染性反基因组(antigenomic)(+)RNA链的cDNA分子中。

[0013] 本发明的核酸构建体特别是纯化的DNA分子，由可操作共同连接的各种不同来源的多核苷酸重组获得或可获得。

[0014] 表述“可操作地连接”指本发明核酸构建体的不同的多核苷酸之间存在的功能性连接，由此所述不同的多核苷酸和核酸构建体有效转录，并且在适当的情况下翻译，特别是在细胞或细胞系中，尤其是在用作拯救系统的一部分产生本发明的嵌合感染性MV颗粒的细胞或细胞系中或宿主细胞中。

[0015] 在本发明的具体实施方案中，通过将编码一个或多个CHIKV结构蛋白的多核苷酸克隆入编码麻疹病毒全长反基因组(+)RNA的cDNA分子中而制备构建体。可选择地，本发明的核酸构建体可以使用核酸片段合成或由模板聚合，包括通过PCR的步骤制得。

[0016] 在本发明的具体实施方案中，编码至少一个CHIKV蛋白的多核苷酸或每个这些多核苷酸被克隆入插入到麻疹病毒cDNA中的ATU(额外转录单元)中。ATU序列是技术人员已知的，用于克隆入MV的cDNA的步骤，包括转基因的MV依赖的表达所必需的顺式作用序列，例如在MV cDNA中位于插入物之前的基因启动子，所述插入物由掺入多克隆位点盒的编码CHIKV蛋白的多核苷酸代表。

[0017] 当用于进行本发明时，ATU有利地定位于编码MV反基因组全长(+)RNA链的cDNA分子的N末端序列中，并且尤其定位于该病毒的P和M基因之间或H和L基因之间。经观察，MV的病毒RNA转录遵循从5'向3'端的梯度。这解释了当插入cDNA编码序列的5'端时，ATU会使得其含有的异源DNA序列(例如编码至少一种CHIKV结构蛋白的多核苷酸)更有效的表达。

[0018] 因此编码至少一种CHIKV结构蛋白的多核苷酸可以插入麻疹病毒cDNA分子任意的基因间的区域中，特别是ATU中。本发明具体的构建体如实施例中所例示。

[0019] 在具体实施方案中，当几种不同的多核苷酸存在于DNA构建体中时，这些编码至少一种CHIKV结构蛋白的多核苷酸的每一个可以插入MV cDNA的不同位点中，可能是麻疹病毒

cDNA的不同ATU中。

[0020] 在本发明的优选实施方案中，编码至少一种CHIKV结构蛋白的多核苷酸插入麻疹病毒cDNA分子的P和M基因的基因间的区域中，特别是ATU中。

[0021] 用于本申请的表述“编码”限定核酸分子在选择的细胞或细胞系中转录以及(如果适合)翻译以进行产物表达的能力。相应地，本发明的核酸构建体可以包含控制编码序列转录的调控元件，特别是转录启动子和终止序列，以及可能的增强子和其他顺式作用元件。这些调控元件可以是与CHIK多核苷酸序列异源的。

[0022] 术语“蛋白”与术语“抗原”或“多肽”可互换使用，限定由氨基酸残基连接产生的分子。特别地，本申请中公开的蛋白源自CHIKV并且是结构蛋白，其可与天然蛋白相同或可选地可由天然蛋白经突变衍生，包括取代(特别通过保守氨基酸残基)或添加氨基酸或翻译后二次修饰，或缺失部分天然蛋白产生相对于作为参考的天然蛋白具有缩短尺寸的片段。本发明包含具有适于在宿主特别是人宿主中引发免疫应答的天然蛋白的表位的片段，优选该应答能够防止CHIKV感染或CHIKV相关疾病。表位特别是B型表位，通过在蛋白施用的宿主中或其在施用本发明的感染性复制颗粒后表达该蛋白的宿主中激活抗体产生，参与引发体液免疫应答。可选地，表位可以是T型表位，参与引发细胞介导的免疫应答(CMI应答)。片段可以具有代表多于50％的CHIKV天然蛋白氨基酸序列尺寸的尺寸，优选至少90％或95％。可选地，片段可以是具有至少10个氨基酸残基的短多肽，其具有天然蛋白的表位。就这一点来说，片段也包括本文限定的多表位。

[0023] 在本发明的具体实施方案中，在核酸构建体中编码MV全长感染性反基因组(+)RNA链的核苷酸序列的cDNA符合麻疹病毒基因组的6倍规则(rule of six)。

[0024] 麻疹病毒基因组的组构及其复制和转录过程已被现有技术充分确认，尤其是公开于Horikami S.M.和Moyer S.A.(Curr.Top.Microbiol.Immunol.(1995)191,35-50或Combredet C.et al(Journal of Virology,Nov 2003,p11546-11554)的病毒施瓦兹接种毒株，或Neumann G.et al(Journal of General Virology(2002)83,2635-2662)中广泛认定的反义RNA病毒。

[0025] “6倍规则”是指代表MV(+)链RNA基因组的核酸或包含MV(+)链RNA基因组的核酸构建体中存在的核苷酸总数是6的倍数。“6倍规则”已被本领域认为是麻疹病毒基因组中有关核苷酸总数的要求，其使得MV基因组RNA有效或优化复制。在限定满足6倍规则的核酸构建体的本发明的实施方案中，所述规则适用于限定编码全长MV(+)链RNA基因组的cDNA的核酸构建体。就此，6倍规则单独适用于编码麻疹病毒全长感染性反基因组(+)RNA链的核苷酸序列的cDNA，但不必适用于克隆入所述cDNA并编码一或多种CHIKV蛋白的多核苷酸。

[0026] 根据本发明的具体方面，核酸构建体包含以下基因转录单元，由5’到3’包含：

[0027] (a)编码MV的N蛋白的多核苷酸，

[0028] (b)编码MV的P蛋白的多核苷酸，

[0029] (c)编码至少一种CHIKV结构蛋白的多核苷酸，

[0030] (d)编码MV的M蛋白的多核苷酸，

[0031] (e)编码MV的F蛋白的多核苷酸，

[0032] (f)编码MV的H蛋白的多核苷酸，和

[0033] (g)编码MV的L蛋白的多核苷酸，

[0034] 所述多核苷酸和核酸构建体可操作地连接，并受病毒复制和转录调控序列如MV先导序列和尾随序列控制。

[0035] 表述“N蛋白”、“P蛋白”、“M蛋白”、“F蛋白”、“H蛋白”和“L蛋白”分别指麻疹病毒的核蛋白(N)、磷蛋白(P)、基质蛋白(M)、融合蛋白(F)、血凝素蛋白(H)和RNA聚合酶大蛋白(L)。这些成分在现有技术中已经鉴定，并特别公开于Fields，Virology(Knipe&Howley，2001)。

[0036] 在本发明的优选实施方案中，编码麻疹病毒全长感染性反基因组(+)RNA链的cDNA分子具有MV减毒株的特征或获自MV减毒株。

[0037] 麻疹病毒的“减毒株”定义为在相同宿主中无毒或比母株低毒的毒株，施用于宿主时保留免疫原性以及可能的佐剂性，即保留免疫显性的T和B细胞表位以及可能的佐剂性例如诱导T细胞共刺激蛋白或细胞因子IL-12。

[0038] 麻疹病毒的减毒株相应地指在选取的细胞上经系列传代并且可能地，适应于其他细胞以产生适于制备疫苗株的种子毒株的毒株，其具有不允许回复到病原性也不整合入宿主染色体的稳定基因组。作为具体的“减毒株”，经确证的用于疫苗的毒株是适于本发明的减毒株，其满足FDA(美国食品和药物管理局)限定的规则，即经过对实验室和临床数据严格的审阅后，其满足安全性、效力、质量和重复性规则(www.fda.gov/cber/vaccine/vacappr.htm)。

[0039] 可用于实施本发明并尤其获得核酸构建体的MV cDNA的具体毒株是施瓦兹MV毒株、萨格勒布(Zagreb)毒株、AIK-C毒株和Moraten毒株。所有这些毒株均已在现有技术中记载并可获得，特别是用作商业化疫苗。

[0040] 根据本发明的具体实施方案，cDNA分子位于异源表达控制序列的控制下。

[0041] 当在具有cDNA的天然控制序列不能使该cDNA完全转录的细胞类型中寻求该cDNA的表达时，插入该cDNA表达控制是有利的。

[0042] 根据本发明的具体实施方案，异源表达控制序列包含T7启动子和T7终止子序列。这些序列各自定位于MV的全长反基因组(+)RNA链的编码序列的5’和3’并与该编码序列周围的序列相邻。

[0043] 在本发明的具体实施方案中，如上限定的cDNA分子经修饰，即包含额外的核苷酸序列或基序。

[0044] 在优选实施方案中，本发明的cDNA分子在其5’端(与编码MV确证疫苗株的全长反基因组(+)RNA链的核苷酸序列的第一个核苷酸相邻)进一步包含GGG基序，其后是锤头状的核酶序列，并且在其3’端(与编码全长反基因组(+)RNA链的所述核苷酸序列的最后一个核苷酸相邻)包含核酶序列。肝炎delta病毒核酶(δ)适于进行本发明。

[0045] 位于5’端，与上述编码序列的第一个核苷酸相邻的GGG基序改善所述cDNA编码序列的转录效率。为了麻疹病毒颗粒的适当装配，编码反基因组(+)RNA的cDNA应符合6倍规则，当添加GGG基序时，在cDNA编码序列的5’端还添加核酶，其位于GGG基序的3’端，从而能够在MV的全长反基因组(+)RNA链的第一编码核苷酸处切割转录物。

[0046] 在本发明的具体实施方案中，为了制备本发明的核酸构建体，通过已知方法实现编码现有技术中公开的麻疹病毒全长反基因组(+)RNA的cDNA分子的制备。所述cDNA插入载体例如质粒中时，特别提供基因组载体。

[0047] 适于制备本发明核酸构建体的具体cDNA分子是使用麻疹病毒施瓦兹毒株获得的cDNA分子。相应地，用于本发明的cDNA可以如WO2004/000876中公开获得，或由CNCM的巴斯德研究所保藏的质粒pTM-MVSchw(保藏号I-2889，保藏日期2012年6月12日)获得，其序列于WO2004/000876中公开，通过引用并入本文。质粒pTM-MVSchw获自Bluescript质粒并包含位于T7RNA聚合酶启动子控制下的编码施瓦兹毒株全长麻疹病毒(+)RNA链的多核苷酸。其具有18967个核苷酸，序列如SEQ ID NO:1所示。来自其他MV毒株的cDNA分子(为了方便，也命名为麻疹病毒cDNA或MV cDNA)可类似地由例如本文所述的减毒MV的病毒颗粒纯化的核酸获得。

[0048] 本发明的核酸构建体适合并用于制备重组感染性复制麻疹-屈曲病毒(MV-CHIKV)，并且相应地，所述核酸构建体用于插入转移基因组载体，由此其包含麻疹病毒尤其是施瓦兹毒株的cDNA分子，用于产生所述MV-CHIKV病毒和产生CHIKV结构蛋白，特别是CHIKV VLP。pTM-MVSchw质粒适于制备转移载体，通过插入CHIKV结构蛋白，特别是CHIKV VLP表达所必需的CHIKV多核苷酸。

[0049] 本发明因此涉及转移载体，当从辅助细胞拯救(rescue)时，其用于制备重组MV-CHIKV颗粒。本发明的转移载体最好是质粒，特别是获自Bluescript质粒的质粒。

[0050] 本发明还涉及转移载体转化适于拯救病毒MV-CHIKV颗粒的细胞的用途，特别是分别使用质粒或含有本发明核酸构建体的病毒载体转染或转导所述细胞，根据表达所需麻疹病毒蛋白的能力选择所述细胞，所述麻疹病毒蛋白用于病毒重组基因组的适当复制、转录和壳体化，其对应于重组感染性复制MV-CHIKV颗粒中的本发明的核酸构建体。

[0051] 本发明还涉及由本发明的转移载体和提供辅助功能和蛋白的其他多核苷酸转化的细胞或细胞系。

[0052] 因此多核苷酸位于所述细胞中，其编码特别包括麻疹病毒N、P和L蛋白的蛋白(即能够形成核糖核蛋白(RNP)复合物的天然MV蛋白或其功能性变体)，优选作为稳定表达的蛋白，至少在重组病毒MV-CHIKV颗粒的转录和复制中起作用的N和P蛋白。N和P蛋白可以在细胞中由包含其编码序列的质粒表达，或可以由插入细胞基因组中的DNA分子表达。L蛋白可以由不同的质粒表达。其可以暂时表达。辅助细胞也能表达RNA聚合酶，其适于合成来自本发明的核酸构建体的重组RNA，可能作为稳定表达的RNA聚合酶。RNA聚合酶可以是T7噬菌体聚合酶或其核形式(nlsT7)。

[0053] 在一个实施方案中，麻疹病毒的cDNA克隆来自与N蛋白和/或P蛋白和/或L蛋白相同的麻疹病毒毒株。在另一个实施方案中，麻疹病毒的cDNA克隆来自与N蛋白和/或P蛋白和/或L蛋白不同的病毒毒株。

[0054] 本发明因此涉及制备重组感染性麻疹病毒颗粒的方法，包含：

[0055] 1)在辅助细胞系中转移，特别是转染本发明的核酸构建体或含有该核酸构建体的转移载体，所述辅助细胞系还在病毒颗粒能够装配的条件下由MV的cDNA表达蛋白，所述蛋白是MV的反基因组(+)RNA序列转录、复制和壳体化所必需的；以及

[0056] 2)回收表达至少一种CHIKV结构蛋白的重组感染性MV-CHIKV病毒。

[0057] 根据具体实施方案，该方法包含：

[0058] 1)用转移载体转移将根据本发明的核酸构建体转染辅助细胞系，其中所述辅助细胞能够表现辅助功能以表达RNA聚合酶，并表达MV病毒的N、P和L蛋白；

[0059] 2)将步骤1)的所述转染的辅助细胞与适于cDNA来源的MV减毒株传代的传代细胞共培养；

[0060] 3)回收表达至少一种CHIKV结构蛋白的重组感染性MV-CHIKV病毒。

[0061] 根据本发明另一具体实施方案，产生重组感染性MV-CHIKV病毒的方法包含：

[0062] 1)用本发明的核酸构建体和包含编码麻疹病毒RNA聚合酶大蛋白(L)的核酸的载体重组细胞或细胞培养物，其稳定产生RNA聚合酶、麻疹病毒的核蛋白(N)和麻疹病毒的聚合酶辅因子磷蛋白(P)；以及

[0063] 2)从所述重组细胞或重组细胞培养物中回收感染性MV-CHIKV病毒。

[0064] 根据所述方法的一个具体实施方案，产生重组MV，其表达CHIKV结构蛋白，特别是CHIKV VLP，其中所述颗粒表达抗原组合，例如CHIK病毒的CE3E26KE1抗原。作为示例，拯救表达CHIKV结构蛋白，特别是CHIKV VLP的重组MV的方法包含以下步骤：

[0065] 1)用(i)转移载体，特别是质粒，其包含编码麻疹病毒的全长反基因组(+)RNA的cDNA，与编码CHIKV结构蛋白，例如编码CHIKV-CE3E26KE抗原的至少一种多核苷酸重组，以及(ii)载体，特别是编码MV L聚合酶cDNA的质粒，共转染稳定表达T7RNA聚合酶以及麻疹N和P蛋白的辅助细胞，特别是HEK293辅助细胞；

[0066] 2)在能够产生MV-CHIKV重组病毒的条件下培养所述共转染的辅助细胞；

[0067] 3)通过将步骤2)的所述辅助细胞与能够使病毒增殖的细胞例如Vero细胞共培养，增殖由此产生的重组病毒；

[0068] 4)回收复制的MV-CHIKV重组病毒和CHIKV结构蛋白，特别是CHIKV病毒样颗粒，特别是CHIKV-CE3E26KE1VLP。

[0069] 该方法与所用构建体和条件一同例示于图1B。

[0070] 用于本文，“重组”指将至少一种多核苷酸引入细胞，例如以载体形式，所述多核苷酸整合(整体或部分)或不整合入细胞基因组(例如上述)。

[0071] 根据具体实施方案，重组可用第一多核苷酸获得，其为本发明的核酸构建体。重组还能或可选择地包括引入多核苷酸，其为编码麻疹病毒RNA聚合酶大蛋白(L)的载体，其定义、性能和表达稳定性已在本文描述。

[0072] 根据本发明，稳定产生RNA聚合酶、麻疹病毒的核蛋白(N)和麻疹病毒的聚合酶辅因子磷蛋白(P)的细胞或细胞系或细胞培养物是本说明书中限定的细胞或细胞系或本说明书中限定的细胞培养物，即也是重组细胞，其经导入上述一或多种多核苷酸而修饰。在本发明的具体实施方案中，稳定产生RNA聚合酶、N和P蛋白的细胞或细胞系或细胞培养物不产生麻疹病毒的L蛋白或不稳定产生麻疹病毒的L蛋白，例如暂时表达或产生L蛋白。

[0073] 本发明MV-CHIKV病毒的产生可涉及如本文所述转化细胞的转移。本文所用“转移”指将重组细胞置于不同类型的细胞上，尤其是置于不同类型细胞的单层上。后面这些细胞能够支持感染性MV-CHIKV病毒的复制和产生，即分别在胞内形成感染性病毒，以及可能的将这些感染性病毒释放到胞外。该转移造成本发明的重组细胞与前句限定的感受态细胞共培养。当重组细胞并非有效的病毒产生培养物时，即当感染性MV-CHIKV不能从这些重组细胞有效回收时，上述转移可以是额外即任选步骤。该步骤在本发明的重组细胞与本发明的核酸构建体，以及任选地包含编码麻疹病毒RNA聚合酶大蛋白(L)的载体进一步重组之后引入。

[0074] 在本发明的具体实施方案中，需要转移步骤，因为通常因其便于重组的能力所选择的重组细胞并不足以有效支持并产生重组感染性MV-CHIKV病毒。在所述实施方案中，上述方法步骤1)的细胞或细胞系或细胞培养物是根据本发明的重组细胞或细胞系或重组细胞培养物。

[0075] 适于制备本发明的重组细胞的细胞是原核或真核细胞，特别是动物或植物细胞，更特别是哺乳动物细胞例如人细胞或非人哺乳动物细胞或禽类细胞或酵母细胞。在具体实施方案中，在其基因组重组前，细胞分离自原代培养物或细胞系。本发明的细胞可以是分裂或不分裂细胞。

[0076] 根据优选实施方案，辅助细胞来自人胚肾细胞系293，该细胞系293保藏于ATCC，保藏号CRL-1573。具体的细胞系293是在WO2008/078198中公开的细胞系，并在后续实施例中引用。

[0077] 根据该方法的另一个方面，适于传代的细胞是CEF细胞。CEF细胞可由受精鸡蛋制得，所述鸡蛋获自EARL Morizeau，8rue Moulin，28190Dangers，France，或任意其他受精鸡蛋生产商。

[0078] 本发明公开的方法有利地用于生产适于用作免疫组合物的感染性复制MV-CHIKV病毒。

[0079] 本发明因此涉及免疫原性组合物，其活性成分包含从本发明的核酸构建体拯救的感染性复制MV-CHIKV病毒，并且特别地由所公开的方法获得。

[0080] 如本文所限定，本发明的核酸构建体和本发明的MV-CHIKV病毒编码或表达至少一种CHIKV结构蛋白。

[0081] “屈曲病毒结构蛋白”是指本文所限定的“蛋白”，其序列与CHIKV毒株中的对应物相同，包括多肽，其为天然成熟的CHIKV结构蛋白或CHIKV结构蛋白前体，或如本文所限定的其片段或其突变体，特别是与天然存在的屈曲病毒壳体或包膜蛋白具有至少50％、至少80％、特别有利地至少90％或优选至少95％氨基酸序列相同性的片段或突变体。氨基酸序列相同性可由本领域技术人员使用手动比对或使用各种可用的比对程序(例如，BLASTP-http://blast.ncbi.nlm.nih.gov/)经比对确定。本发明CHIKV结构蛋白的片段或突变体可相对于本文例示的具体氨基酸序列来限定。

[0082] 根据本发明，编码至少一种CHIKV结构蛋白的多核苷酸编码一或多种以下蛋白：结构糖蛋白、结构多肽和壳体蛋白。

[0083] 在具体实施方案中，糖蛋白包含包膜糖蛋白E1、E2和E3包膜糖蛋白。多肽包含6K多肽和壳体蛋白。在以下段落中，术语蛋白或糖蛋白可互换用于命名E1、E2或E3糖蛋白或其组合。

[0084] 根据本发明的具体实施方案，编码CHIKV结构蛋白的多核苷酸选自以下组：

[0085] -编码E1、E2、E3、6K或C蛋白之一的多核苷酸；

[0086] -编码选自E1、E2、E3、6K和C蛋白的几个蛋白的融合多核苷酸；

[0087] -编码E3-E2-6K-E聚蛋白的多核苷酸；

[0088] -编码C-E3-E2-6K-E1聚蛋白的多核苷酸以及特别是来自CHIKV基因组或编码所述聚蛋白的对应cDNA的开放读码框(ORF)；

[0089] -任意这些多核苷酸，其经修饰以编码一或多种这些蛋白的突变形式，特别是E2蛋

白的突变形式。

[0090] 就这一点来说，具体的多核苷酸编码可溶形式的E2蛋白(sE2)或编码E2蛋白的胞外域(ectodomain)或其可溶形式(sE2Δstem)。在具体实施方案中，多核苷酸编码以下多肽之一：E3-sE2-6K-E1、E3-sE2Δstem-6K-E1、C-E3-sE2-6K-E1和C-E3-sE2Δstem-6K-E1。

[0091] 用于本文，术语“胞外域”指糖蛋白E2的结构域，其延伸出病毒颗粒，负责在CHIKV病毒颗粒感染过程中附着并进入细胞。

[0092] 根据具体实施方案，多核苷酸编码CHIKV结构蛋白的单个表位或编码由一或多个CHIKV结构蛋白的重复表位(具有相同或相似序列)或多个不同表位表达产生的多表位。

[0093] 举例来说，多表位由编码定位于E2糖蛋白N末端、与弗林蛋白酶E2/E3切割位点临近的E2EP3单个表位的重复多核苷酸融合形成。E2EP3表位的氨基酸序列在Kam Y.W et al.(EMBO Mol Med 4,330-343)中公开，即SEQ ID NO:33。

[0094] 根据本发明的具体实施方案，几种多核苷酸(其中每种多核苷酸编码至少一种CHIKV结构蛋白)组合或融合以形成编码几种CHIKV结构蛋白的多核苷酸。这些多核苷酸可以彼此不同，其编码不同毒株的CHIKV的蛋白。由此，多核苷酸编码本文所述多表位，例如单个E2EP3表位的多表位。

[0095] 编码至少一种CHIKV结构蛋白的多核苷酸克隆在cDNA分子(编码麻疹病毒的全长感染性反基因组(+)RNA链)中，可能在不同位点，以产生本发明的核酸构建体。

[0096] 根据本发明的一个方面，编码至少一种屈曲病毒(CHIKV)结构蛋白的多核苷酸来自分离和纯化的野生CHIKV毒株的基因组。野生CHIKV毒株可以是例如Ross毒株(GenBank:AF490259.3)，或S27毒株(GenBank:AF339485.1)，二者均在1952年坦桑尼亚爆发期间分离自患者，或在1983年塞内加尔爆发期间分离的名为Ae.furcifer(GenBank:AY726732.1)的毒株。

[0097] 根据本发明另一方面，编码至少一种CHIKV结构蛋白的多核苷酸来自以下纯化和分离的野生CHIKV毒株：05.61、05.115、05.209、06.21、06.27和06.49，其在WO2007/105111中有记载。这些代表屈曲病毒印度洋爆发的不同地理起源、时间点和临床形式的CHIKV分离物的几乎完整的基因组序列已经测序并在WO2007/105111中公开。确定了11,601个核苷酸，对应于用作参考的1952年坦桑尼亚分离物S27的核苷酸序列(全长11,826nt)中的位置52(5’NTR)至11,667(3’NTR，3个重复序列片段的末端)。

[0098] WO2007/105111中出现的分离物05.61、05.115、05.209、06.21、06.27和06.49的基因组序列以及本发明的具体实施方案中可来源于此的根据本发明的多核苷酸如下组构。编码序列由分别编码非结构聚蛋白(2,474个氨基酸)和结构聚蛋白(1,248个氨基酸)的7,422nt和3,744nt的两个大开放读码框(ORF)组成。非结构聚蛋白是蛋白nsP1(535aa)、nsP2(798aa)、nsP3(530aa)和nsP4(611aa)的前体，结构聚蛋白是蛋白C(261aa)、p62(487aa,E3(64aa)和E2(423aa)的前体)、6K(61aa)和E1(439aa)的前体。作为甲病毒家族特征的非结构和结构聚蛋白中的切割位点是保守的。E3、E2和E1中的糖基化位点也保守。公开的基因组序列通过引用并入本文。

[0099] 根据一个实施方案，编码CHIKV结构蛋白的多核苷酸来自如前所述名为06.115、06.21、06.27和06.49的野生CHIKV毒株的基因组。

[0100] 出现在多核苷酸定义中的术语“来自”仅表明所述多核苷酸的序列可以与CHIKV毒

株中的对应序列相同，或可以不同，以编码满足本发明“蛋白”定义的CHIKV结构蛋白。相应地，该术语不限制多核苷酸的产生模式。

[0101] 本发明的多核苷酸和核酸构建体可以根据本领域任意已知的方法制备，并且特别地可通过聚合作用尤其是使用PCR方法克隆、获得，或可以合成。

[0102] 本发明的核酸构建体进一步限定为包括以下编码至少一种CHIKV结构蛋白的多核苷酸之一。

[0103] 根据具体实施方案，编码一或几种CHIKV结构蛋白的多核苷酸编码可溶形式的糖蛋白E2。作为实例，该多核苷酸包含具有SEQ ID NO:2、4、6、8序列的编码结构域。

[0104] 在具体实施方案中，编码一或几种CHIKV结构蛋白的多核苷酸编码E2胞外域。作为实例，该多核苷酸包含具有SEQ ID NO:10、12、14序列的编码结构域。

[0105] 在本发明的另一个实施方案中，编码可溶形式的糖蛋白E2的多核苷酸编码一种具有选自SEQ ID NO:3、5、7、9的氨基酸序列的多肽。

[0106] 在本发明的另一个实施方案中，编码糖蛋白E2的胞外域的多核苷酸编码一种具有选自SEQ ID NO:11、13、15的氨基酸序列的多肽。

[0107] 根据本发明的具体实施方案，编码CHIKV结构蛋白的多核苷酸包含以上限定的核苷酸序列之一，其编码可溶形式的E2糖蛋白或所述蛋白的胞外域，并进一步包含编码E3、E1、6K或C蛋白之一的多核苷酸，或编码其任意组合的多核苷酸。

[0108] 相应地，编码分离自弗雷瑞斯的France/2010毒株的结构聚蛋白E2-6K-E1的本发明的具体的多核苷酸选自：包含具有SEQ ID NO:16(GenBank:CCA61130.1)和20(GenBank:CCA61131.1)的序列的编码结构域的多核苷酸。

[0109] 本发明的优选多核苷酸编码来自CHIKV毒株的结构蛋白C-E3-E2-6K-E1。在具体实施方案中，所述多核苷酸编码名为S27或06.49的毒株之一的聚蛋白C-E3-E2-6K-E1，并分别具有序列SEQ ID NO:20和27。

[0110] 在本发明另一个实施方案中，所用的多核苷酸编码E2EP3表位或由该表位的重复形成的或包含该重复的多表位。该多核苷酸特别具有SEQ ID No:32公开的核苷酸序列。

[0111] 根据优选实施方案，本发明还涉及多核苷酸的修饰和优化，以允许在宿主中MV-CHIKV嵌合感染性颗粒表面有效表达屈曲病毒蛋白。

[0112] 根据该实施方案，可进行多核苷酸序列的优化以避免核酸分子的顺式活性结构域：内部TATA盒、chi位点和核糖体进入位点；富含AT或富含GC序列段；ARE、INS、CRS序列元件；重复序列和RNA二级结构；隐藏的拼接供体和受体位点，分支点。

[0113] 优化的多核苷酸也可以为在特定细胞类型中表达而经密码子优化，特别是可以为Maccaca密码子使用或人密码子使用而修饰。该优化使得嵌合感染性颗粒在细胞中的产生效率增加，而不影响表达的蛋白。

[0114] 特别地，编码CHIKV蛋白的多核苷酸的优化可以通过密码子中的摆动位置(wobble position)的修饰进行，不会影响从所述密码子翻译的氨基酸残基相对于原始氨基酸残基的相同性。

[0115] 还进行优化以避免来自麻疹病毒的编辑样序列。麻疹病毒转录物的编辑是特别发生在由麻疹病毒P基因编码的转录物中的过程。该编辑，通过在P转录物中的特定位点插入额外的G残基，产生与P蛋白相比截断的新蛋白。仅添加单个G残基导致V蛋白的表达，其含有

独特的羧基末端(Cattaneo R et al.,Cell.1989Mar 10;56(5):759-64)。

[0116] 在根据本发明的该具体实施方案的多核苷酸中,来自麻疹病毒的以下编辑样序列可被突变:AAAGGG、AAAAGG、GGGAAA、GGGGAA,及其互补序列:TTCCCC、TTTCCC、CCTTTT、CCCCTT。例如,AAAGGG能被突变为AAAGGC,AAAAGG能被突变为AGAAGG或TAAAGG或GAAAGG,GGGAAA能被突变为GCGAA。

[0117] 根据本发明的修饰和优化的多核苷酸的一个实施方案由SEQ ID NO:29所限定。该多核苷酸编码无干区(stem region)的可溶形式的包膜蛋白E2。

[0118] 编码所有结构蛋白C-E3-E2-6K-E1的修饰和优化的多核苷酸的一个实施方案由SEQ ID NO:31所限定。

[0119] 根据SEQ ID NO:29和31所限定的本发明的该具体实施方案的这些优化的多核苷酸在BsiWI和BssHII位点的序列中但非在序列末端存在突变,以保留该位点用于克隆目的。

[0120] 因此,根据该具体实施方案,本发明提供核酸构建体,其包含增加嵌合MV-CHIKV感染性颗粒产生效率的多核苷酸。

[0121] 包含编码CHIKV结构蛋白的序列并也适用于本发明核酸构建体的其他优化的多核苷酸为优化的融合多核苷酸,其编码以下结构蛋白的组合之一:E3-E2-6K-E1、E3-sE2-6K-E1、E3-sE2Δstem-6K-E1、C-E3-E2-6K-E1、C-E3-sE2-6K-E1和C-E3-sE2Δstem-6K-E1。

[0122] 本发明还涉及核酸构建体,其中编码至少一种CHIKV结构蛋白的多核苷酸编码以下多肽之一。

[0123] 根据本发明优选实施方案的CHIK病毒结构蛋白以及与来自命名为05.115、06.21、06.27和06.49毒株的可溶形式的糖蛋白E2相关的氨基酸序列的实例由SEQ ID NO:3、5、7、9所限定。

[0124] 根据本发明优选实施方案的CHIK病毒结构蛋白以及与来自命名为05.115、06.21、06.27和06.49毒株的糖蛋白E2的胞外域相关的氨基酸序列的实例由SEQ ID NO:11、13、15所限定。

[0125] 在具体实施方案中,本发明涉及由分离自弗雷瑞斯的France/2010毒株的结构蛋白E2-6K-E1构成的融合蛋白。这些蛋白分别来自具有序列SEQ ID NO:16(GenBank:CCA61130.1)和18(GenBank:CCA61131.1)的多核苷酸的表达。它们的序列分别由SEQ ID NO:17和19所限定。

[0126] 在特别优选的实施方案中,本发明涉及由所有CHIKV结构蛋白组成的融合蛋白。这些蛋白来自编码所有结构蛋白C-E3-E2-6K-E1的多核苷酸的表达,并由SEQ ID NO:21、22、23、24、25、26和28所限定。

[0127] 获得的结构蛋白C-E3-E2-6K-E1能够在CHIKV-MV颗粒中自装配成CHIKV病毒样颗粒(VLP)。

[0128] 用于本文,术语“病毒样颗粒(VLP)”指至少一个属性类似病毒但未被证实同样为感染性的结构。根据本发明的病毒样颗粒不携带编码病毒样颗粒蛋白的遗传信息,总体上,病毒样颗粒缺少病毒基因组,因此是非感染性和非复制性的。根据本发明,病毒样颗粒能大量产生并与CHIKV-MV重组颗粒一起表达。

[0129] 在具体实施方案中,本发明涉及来自名为06.49的CHIKV毒株、无干区的可溶形式的糖蛋白E2。该糖蛋白来自SEQ ID NO:29限定的修饰和优化的多核苷酸的表达。其序列由

SEQ ID NO:30限定。

[0130] 根据另一方面，本发明涉及重组CHIKV-麻疹病毒颗粒，其表达本文限定的屈曲病毒结构蛋白，特别参考其核酸和多肽序列。重组CHIKV-MV病毒有利地表达CHIKV结构蛋白为VLP。

[0131] 本发明还涉及CHIKV结构蛋白的病毒样颗粒，特别是C-E3-E2-6K-E1的VLP在组合物中与CHIKV-MV感染性复制病毒颗粒的联合。

[0132] 根据本发明的优选实施方案，重组麻疹病毒载体以该方式设计，生产过程涉及细胞，由此在用所述载体转染或转化的辅助细胞中生产、来自适于接种的麻疹病毒毒株的病毒颗粒使得产生重组麻疹-屈曲感染性和复制病毒，并产生CHIKV-VLP，用于免疫原性组合物、优选保护性或甚至疫苗组合物。

[0133] 有利地，本发明的重组麻疹-屈曲感染性病毒的基因组是能够复制的。“能够复制”是指核酸转导入表达MV的N、P和L蛋白的辅助细胞系时能够被转录和表达，从而产生新病毒颗粒。

[0134] 使用制备MV-CHIKV重组基因组的MV cDNA获得的本发明重组病毒的复制也能在宿主体内实现，特别是施用重组MV-CHIKV的人宿主。

[0135] 本发明还涉及活性成分的组合物或组合体，其包含与CHIKV结构蛋白的VLP，例如CE3E26KE1蛋白的VLP联用的重组麻疹-屈曲复制病毒。这些组合物或组合体诱导针对屈曲病毒的免疫应答，特别是保护性免疫应答，并特别引发针对屈曲病毒结构蛋白的抗体产生和/或引发针对CHIKV感染的细胞免疫应答。这些组合物相应地可包含对宿主、尤其是人宿主施用的适合的载体，例如药物学可接受的载体，并且可进一步包含但非必须的佐剂以增强宿主中的免疫应答。本发明人事实上已证实本发明活性成分的施用可引发免疫应答而无需佐剂化。

[0136] 本发明特别涉及向儿童施用的组合物。

[0137] 本发明还涉及免疫原性组合物，特别是疫苗组合物，并且特别涉及向儿童施用的疫苗组合物。所述组合物或疫苗用于在预防疗法中防止CHIKV感染。该疫苗组合物有利地具有活性成分，其包含拯救自载体的重组麻疹-屈曲感染性复制病毒颗粒，所述载体已在本文中限定为与CHIKV结构蛋白的VLP，例如CE3E26KE1蛋白的VLP联用。

[0138] 在本发明的上下文中，术语“联用”或“关联”指MV-CHIKV重组病毒颗粒和CHIKV结构蛋白在独特组合物中的同时出现，特别是作为VLP，通常以物理分离的实体出现。

[0139] 本发明还涉及与CHIKV病毒结构蛋白、特别是表达CHIKV结构蛋白的CHIKV病毒样颗粒、特别是CHIKV-CE3E26KE1VLP联用的重组MV-CHIKV感染性复制病毒颗粒，或本发明的组合物，其用于治疗或预防受试者、特别是人中屈曲病毒的感染。

[0140] 本发明还涉及MV-CHIKV感染性、复制病毒以及联用的CHIKV病毒结构蛋白，特别是联用的CHIKV结构蛋白VLP，例如CE3E26KE1蛋白VLP，用于施用方案，以及根据给药方案使用，其引发针对CHIKV病毒感染或诱导性疾病的免疫应答，有利地为保护性免疫应答，特别是在人宿主中。

[0141] 施用方案和给药方案需要独特施用选择剂量的MV-CHIKV感染性、复制病毒以及联用的CHIKV病毒结构蛋白，特别是联用的CHIKV结构蛋白VLP，例如CE3E26KE1蛋白VLP。

[0142] 可选地，在免疫-加强方案中需要多剂量施用。可使用相同的活性成分实现免疫和

加强,所述活性成分由MV-CHIKV感染性、复制病毒以及联用的CHIKV病毒结构蛋白,特别是联用的CHIKV结构蛋白VLP,例如CE3E26KE1蛋白VLP组成。

[0143] 可选地,可使用不同的活性成分实现免疫和加强施用,所述活性成分在至少一个施用步骤中包含MV-CHIKV感染性、复制病毒以及联用的CHIKV结构蛋白,特别是联用的CHIKV结构蛋白VLP,例如CE3E26KE1蛋白VLP,在其他施用步骤中包含CHIKV的其他活性免疫原,例如表达C-E3-E2-6K-E1聚蛋白的CHIKV蛋白或VLP。

[0144] 本发明还涉及不同活性成分的组合体,所述活性成分包括以下活性成分之一:MV-CHIKV感染性、复制病毒以及联用的CHIKV结构蛋白,特别是联用的CHIKV结构蛋白VLP,例如CE3E26KE1蛋白VLP。活性成分的组合体有利地用于宿主、特别是人宿主的免疫。

[0145] 本发明人已经证实MV-CHIKV感染性、复制病毒以及联用的CHIKV结构蛋白VLP的施用引发免疫应答,并且尤其引发与不同CHIKV毒株,至少是ECSA基因型毒株交叉反应的抗体。相应地,已经证实施用本发明的活性成分(用特定CHIKV毒株的编码序列制备时)能引发针对CHIKV毒株群、特别是ECSA基因型毒株群、特别是包含CHIKV印度毒株、CHIKV刚果毒株、CHIKV泰国毒株和CHIKV留尼汪毒株的毒株群的免疫应答。

[0146] 考虑到关于适于其他病原体(例如HBV或HPV)的疫苗(其涉及施用病毒样颗粒(VLP))以及还适于已知人MV疫苗的剂量的可用知识,本发明人确定用重组MV-CHIKV病毒回收CHIKV-VLP使得能够尝试施用有效低剂量的活性成分。事实上,考虑到重组MV-CHIKV病毒每个MV-CHIKV复制颗粒能够产生约10^4CHIKV-VLP,并且考虑到目前已知的人MV疫苗剂量为10^3至10^4pfu,待施用的重组MV-CHIKV病毒的适合剂量为0.1至10ng,特别是0.2至6ng,并且可能低至0.2至2ng。作为比较,在HBV或HPV疫苗中施用的VLP剂量在10μg范围内,这意味着重组MV-CHIKV疫苗的剂量可以包含少约2000或多达5000至10000倍的VLP。

[0147] 根据本发明的具体实施方案,本文限定的疫苗的免疫原性组合物还可用于防止麻疹病毒感染。

附图说明

[0148] 图1:示出屈曲病毒结构蛋白和MV基因组骨架的ORF-包括麻疹病毒表达的所述CHIK病毒抗原的MV-CHIKV构建体的示意图。

[0149] 图1B:表达CHIKV VLP的重组MV的拯救

[0150] 图2:在由MOI 0.1的重组MV-CHIKV感染24h的Vero细胞中免疫荧光检测E2抗原。

[0151] 使用抗E2单抗3E4(1/100稀释)检测E2,二抗经1/5000稀释使用。

[0152] 图3:MV-CHIKV载体表达E2和壳体蛋白。

[0153] MV-sE2Δstem和MV-CE3E26KE1感染24h的Vero细胞的细胞裂解物(细胞)和上清(SN)经western印迹分析。用3E4单抗探测E2,使用抗壳体单抗(来自P.Desprès,1/100稀释)检测C蛋白,二抗经1/5000稀释使用。

[0154] 图4:电子显微镜检查分析在由MOI 0.1的MV-CE3E26KE1重组病毒感染的Vero细胞上清中分泌的CHIKV VLP。比例尺为200nm(左)和100nm(右)。红箭头代表在颗粒表面上的刺突的特定排列以及在颗粒内的壳体蛋白的二十面体对称。

[0155] 图5:由MV-sE2重组病毒表达的截短的sE2(156aa,19kDa)的序列。

[0156] 图6:重组MV-sE2Δstem和MV-CE3E26KE1与标准MV相比在Vero细胞上的生长动力

学(MOI 0.01)。以TCID50表明细胞相关病毒滴度。

[0157] 图7:实施例2的免疫和攻击计划。

[0158] 图8:用MV-CE3E26KE1重组病毒两次免疫后,以100PFU CHIKV-06-49致死攻击的小鼠的存活曲线。

[0159] 图9:实施例3的免疫和攻击计划。

[0160] 图10:用MV-CE3E26KE1重组病毒单次免疫后,以100PFU CHIKV-06-49致死攻击的小鼠的存活曲线。

[0161] 图11:实施例4的免疫和攻击计划。

[0162] 图12:用不同剂量的MV-CE3E26KE1重组病毒免疫后,以100PFU CHIKV-06-49致死攻击的小鼠的存活曲线。

[0163] 图13:实施例5的被动转移免疫血清和攻击计划。

[0164] 图14:被动转移MV-CE3E26KE1免疫血清后,以100PFU CHIKV-06-49致死攻击的小鼠的存活曲线。

[0165] 图15:通过106TCID50的MV-CHIKV单次注射免疫的CD46-IFNAR小鼠脾细胞内引发的细胞介导的免疫应答。

[0166] 图16:实施例6的免疫和攻击计划。

[0167] 图17:用MV-CE3E26KE1免疫后,以100PFU CHIKV-06-49致死攻击的预免疫小鼠的存活曲线。

[0168] 图18:在首次免疫前第90天(加强之前)和第111天(加强之后第21天)进行针对CHIK的PRNT测定。

实施例

[0169] 表达屈曲病毒蛋白的重组麻疹病毒载体的构建和特征

[0170] 本发明人基于屈曲病毒06-49毒株天然蛋白的肽序列设计3种屈曲病毒抗原。能够制备这些肽序列的天然蛋白为5种结构蛋白,其由壳体(C)包膜和附属蛋白E1、E2、E3和6K组成。

[0171] 第一结构涉及可溶形式的包膜蛋白E2(sE2)的表达,第二结构涉及无干区的sE2(sE2Δstem)的表达,第三结构涉及所有病毒结构蛋白(C-E3-E2-6K-E1)的表达(图1)。本文根据该后一构建体描述实验方案。

[0172] 细胞培养。Vero(非洲绿猴肾)细胞培养于DMEM GlutaMAX^{TM}(Gibco-BRL)中,添加5%热灭活胎牛血清(FCS,Invitrogen,Frederick,MD)。用于重组麻疹病毒拯救的HEK-293-T7-MV辅助细胞(WO2008/078198)培养于添加10%FCS的DMEM中。

[0173] pTM-MVSchw-CE3E26KE1构建。pTM-MVSchw质粒已在别处描述(Combredet,C.,et al.,A molecularly cloned Schwarz strain of measles virus vaccine induces strong immune responses in macaques and transgenic mice.J Virol,2003.77(21):p.11546-54),其含有对应于施瓦兹MV疫苗株的反基因组的感染性MV cDNA。编码结构CE3E26KE1CHIKV抗原的cDNA由化学合成制备(GenScript,USA)。其含有来自CHIKV毒株06-49的病毒结构蛋白C-E3-E2-6K-E1的序列(WO2007/105111)。完整序列遵守“6倍规则”,其规定进入MV基因组的核苷酸数目必须是6的倍数,并在5’端含有BsiWI限制酶切位点,3’端含

有BssHII限制酶切位点。序列经优化用于麻疹病毒在哺乳动物细胞中的表达。该cDNA插入BsiWI/BssHII消化的pTM-MVSchw-ATU2中，其在施瓦兹MV基因组的磷蛋白(P)和基质(M)基因之间含有额外的转录单元(ATU)(Combredet,C.,et al.,A molecularly cloned Schwarz strain of measles virus vaccine induces strong immune responses in macaques and transgenic mice.J Virol,2003.77(21):p.11546-54)。所得质粒命名为pTM-MVSchw-CE3E26KE1。

[0174] 重组MV-CE3E26KE1的拯救。使用前述拯救系统如前所述拯救来自质粒pTM-MVSchw-CE3E26KE1的重组施瓦兹MV-CHIKV(Radecke,F.,et al.,Rescue of measles viruses from cloned DNA.J Virol,1995.14(23):p.5773-84;WO2008/078198)。通过在Vero细胞上的端点极限稀释测定确定病毒滴度，并使用Kärber法计算TCID50。

[0175] 免疫荧光。如别处所述在感染细胞上进行免疫荧光染色(Lucas,M.,et al.,Infection of mouse neurons by West Nile virus is modulated by the interferon-inducible 2'-5'oligoadenylate synthetase 1b protein.Immun.Cell Biol.,2003.81:p.230-236)。用鼠抗E2(3E4)和抗壳体抗体探查细胞。Cy3缀合的山羊抗鼠Cy3缀合IgG抗体(Jackson Immunoresearch laboratories)用作二抗。

[0176] Western印迹测定。来自重组病毒感染的Vero细胞的蛋白裂解物经SDS-PAGE凝胶电泳分离，并转移至纤维素膜(Amersham Pharmacia Biotech)。用鼠抗E2单抗3E4和抗壳体抗体探查印迹。山羊抗鼠免疫球蛋白G(IgG)-辣根过氧化物酶(HRP)缀合物(Amersham)用作二抗。使用增强型化学发光检测试剂盒(Pierce)对过氧化物酶活性进行显影。

[0177] 通过电子显微镜分析VLP产生。Vero细胞(3x T-150细颈瓶)用MOI1的MV-CHIKV重组病毒感染。感染后36h收集的上清于3000rpm离心30min澄清，铺于PBS中20％的蔗糖垫层上，在SW41转子中于41,000rpm离心2h。用含有1％BSA的PBS重悬沉淀，并用电子显微镜分析。用2％乙酸双氧铀在涂覆碳的铜载网上进行负染色，并在使用前辉光放电。用Jeol JEM1200(Tokyo,Japan)透射电子显微镜在80kV观察样品。使用Eloise Keenview相机和Analysis Pro-software 3.1版本(Eloise SARL,Roissy,France)记录图像。

[0178] 小鼠实验。如前所述制备对MV感染易感的CD46-IFNAR(Combredet,C.,et al.,A molecularly cloned Schwarz strain of measles virus vaccine induces strong immune responses in macaques and transgenic mice.J Virol,2003.77(21):p.11546-54)。小鼠在特定的无病原体条件下饲养于巴斯德研究所动物设施中。对于免疫，6周龄CD46-IFNAR小鼠用10^5TCID50重组MV-CE3E26KE1或MV经腹膜内(i.p.)接种。对于保护测定，免疫小鼠用100pfu CHIKV 06-49毒株经腹膜内接种，并跟踪死亡率2周。所有实验均获批准并按照巴斯德研究所试验动物护理办公室的准则进行。对于被动转移研究，用20μl来自以10^5TCID50MV-CE3E26KE1免疫的6只小鼠的混合血清经腹膜内接种CD46-IFNAR小鼠。对照小鼠接受20μl来自以10^5TCID50空MVSchw免疫的小鼠的混合血清或20μl抗CHIKV HMAF。血清在被动转移前24h、用100pfu CHIKV 06-49毒株攻击前16h、然后在攻击后12h稀释于总体积100μl的PBS中，以模拟感染动物中的抗体存在。分析小鼠死亡率2周，以确定保护。

[0179] 体液免疫应答分析。为了评价特定的抗体应答，小鼠在免疫后不同时间通过眼窝途径采血。血清在56℃热灭活30min，通过ELISA(ENZYGNOST-Siemens)检测抗MV抗体。HRP缀合的抗鼠免疫球蛋白(Jackson Immuno Research)用作二抗。用特定ELISA检测抗CHIKV抗

体。简言之，用在大肠杆菌中产生的重组CHIKV-E2蛋白包被96孔板。HRP缀合的抗鼠免疫球蛋白用作二抗。混合血清的端点滴度以产生来自作为阴性对照的MV接种小鼠的血清两倍吸光度的最终稀释的倒数计算。使用噬菌斑减少中和试验(PRNT)测量抗CHIKV中和抗体。将Vero细胞种植于12孔板24h。在DMEM Glutamax/2％FCS中系列稀释血清样品。100μl稀释液与含100pfu 06-49毒株的等体积CHIKV于37℃、温和搅拌下孵育2h。然后在覆盖含有最终0.8％(重量/体积)羧甲基纤维素的DMEM GlutaMAX™/2％FCS的Vero细胞单层上测定剩余感染性。孵育3天后，固定细胞并用结晶紫染色，进行噬菌斑计数确定。端点中和滴度以测试的减少噬菌斑数量至少50％(PRNT50)的最高血清稀释计算。

[0180] 细胞介导的免疫应答分析。6周龄CD46+/-IFNα/βR-/-小鼠腹膜内接种10^6TCID50的MV-CHIKV重组病毒。对照小鼠用10^6TCID50空MV载体免疫。感染后7天对小鼠进行安乐死，并收集脾脏。来自免疫小鼠的脾细胞于RPMI，10％FCS和10IU重组人白介素-2(rh-IL-2；Boehringer Mannheim)中孵育。其刺激时分泌IFN-γ的能力通过酶联免疫印迹(ELISPOT)测定进行检测。细胞用刀豆素A(5μg/ml；Sigma)作为阳性对照，RPMI-IL-2(10U/ml)作为阴性对照刺激18h，CHIKV(MOI 1)，或MV(MOI 1)。多筛选-HA96孔板用处于PBS中的5μg抗鼠IFN-γ/ml(R4-6A2；Pharmingen)于4℃包被过夜，清洗，然后用100μl RPMI和10％FCS于37℃孵育1h。用100μl细胞悬浮液(每孔5x10^5脾细胞，一式三份)和100μl刺激剂替换培养基。于37℃2h后，添加加热的FCS(10％)，并将板在37℃孵育18h。清洗后，添加生物素化的抗鼠IFN-γ抗体(XMG1.2；Pharmingen)，并将板在室温孵育2h。链霉亲和素-碱性磷酸酶缀合物(Roche)用作第二步骤。用BCIP/NBT(Promega)形成斑点并计数(ELISpot Reader；Bio-Sys)。

[0181] 重组MV载体表达CHIKV病毒样颗粒。CHIKV VLP已证实引发针对CHIKV感染的保护性免疫(Akahata,W.,et al.,A virus-like particle vaccine for epidemic Chikungunya virus protects nonhuman primates against infection.Nat Med,2010.16(3):p.334-8)。为了受益于该能力，本发明人设计了能够诱导CHIKV VLP分泌的重组MV载体。为了该目的，编码CHIKV VLP产生所需的C-E3-E2-6K-E1结构蛋白的cDNA经化学合成(Genscript)并针对在哺乳动物细胞中的麻疹病毒表达进行优化，然后导入施瓦兹MV疫苗感染性cDNA的额外的转录单元(ATU)(图1)。通过将该质粒转染入HEK-293辅助细胞并在Vero细胞上增殖以获得MV-CHIKV。病毒母液在Vero细胞上生长，并确定滴度。

[0182] 由免疫荧光、western印迹和电子显微镜证实，高含量的CHIKV VLP分泌到感染细胞的培养基中。

[0183] 该策略提供了在每轮复制分泌CHIKV VLP的活的重组MV疫苗病毒。并未预期天然甲病毒颗粒的装配会发生，并且这些VLP不会妨碍副粘病毒的同时复制。这在此首次得到证实。由于MV疫苗以粗病毒提取物工业化制备，批量的重组MV-CHIKV含有活MV病毒和非复制CHIKV VLP。该策略允许得益于VLP上展示的多聚体抗原有利的免疫原性能，无需苛刻且昂贵的纯化和浓缩过程。而且，由于VLP得益于活疫苗有利的免疫原性特征，例如平衡的Th1应答和长期记忆，无需佐剂。

[0184] CHIKV E2和壳体抗原的表达在感染的Vero细胞中通过免疫荧光得到证实，其使用针对CHIKV的E2蛋白的特异性抗体(单抗3E4)(图2)。为了寻找分泌VLP的存在，感染细胞的培养基经低速离心澄清，然后铺于20％蔗糖垫层，并通过在SW41转子中于41,000rpm离心2h

浓缩。沉淀溶解于含1％BSA的PBS中。从细胞裂解物和浓缩的培养基中提取的蛋白经SDS-PAGE凝胶电泳分离，并转移至纤维素膜。用检测E2的3E4小鼠单抗和抗壳体单抗对印迹进行标记，所述3E4小鼠单抗由2007年9月6日以巴斯德研究所的名义保藏在CNCM(Collection Nationale de Cultures de Microorganismes,Paris,France)、编号I-3824的杂交瘤产生。在细胞裂解物和浓缩的感染细胞上清中均发现正确大小(46KDa)的E2蛋白，表明MV-CHIKV病毒诱导包含E2蛋白的高密度颗粒的分泌。高密度颗粒中也发现了壳体蛋白，证实CHIKV-VLP的形成。C和E2蛋白二者在感染细胞的浓缩上清中的存在表明CHIKV VLP的形成。为了观察其物理存在，发明人通过电子显微镜分析了从MV-CHIKV感染细胞的上清中浓缩的沉淀。图像表明高含量颗粒的存在，其具有与野生型CHIKV感染后所述的颗粒相似的尺寸和表型(Pletnev,S.V.,et al.,Locations of carbohydrate sites on alphavirus glycoproteins show that E1forms an icosahedral scaffold.Cell,2001.105(1):p.127-36;Zhang,W.,et al.,Placement of the structural proteins in Sindbis virus.J Virol,2002.76(22):p.11645-58)(图4)。观察到的颗粒呈现65nm的外部直径和40nm核心直径。表面组构表明VLP表面上存在刺突，与其他甲病毒类似排列(Pletnev,S.V.,et al.,Locations of carbohydrate sites on alphavirus glycoproteins show that E1forms an icosahedral scaffold.Cell,2001.105(1):p.127-36;Zhang,W.,et al.,Placement of the structural proteins in Sindbis virus.J Virol,2002.76(22):p.11645-58)。该发现证实重组MV-CHIKV病毒对Vero细胞的感染使得分泌高含量的自装配的CHIK VLP。

[0185] E2蛋白还由麻疹病毒-sE2Δstem重组病毒，以及麻疹病毒-CE3E26KE1重组病毒以正确大小(46KDa)表达并分泌。

[0186] 不幸的是，对MV-sE2感染细胞的分析再次证实截短形式的E2蛋白的表达。发明人在RT-PCR扩增感染细胞后对MV-sE2病毒产生的E2mRNA进行测序。分析证实产生STOP密码子的突变的存在，由此导致截短(图5)。

[0187] 发明人继而使用低MOI(0.01)比较了麻疹病毒-sE2、麻疹病毒-sE2Δstem以及麻疹病毒-CE3E26KE1重组病毒在Vero细胞上相对于标准麻疹病毒母液产生的复制速率(图6)。MV-sE2的生长与对照MV类似。麻疹病毒-sE2Δstem和麻疹病毒-CE3E26KE1重组病毒的生长稍微延缓，但其最终滴度与空麻疹病毒范围相同。

[0188] MV-sE2和MV-CE3E26KE1的免疫原性以及在CD46-IFNAR小鼠中的保护

[0189] 对MV感染易感的CD46-IFNAR用于评价重组MV-CHIKV病毒的免疫原性及其保护效力。这些小鼠表达人CD46基因，其具有人样组织特异性并缺乏I型干扰素受体。小鼠在特定的无病原体条件下饲养于巴斯德研究所动物设施中，所有实验均获批准并按照巴斯德研究所试验动物护理办公室的准则进行。6周龄CD46-IFNAR小鼠用10^3至10^5TCID50剂量的MV-CHIKV重组病毒经腹膜内(i.p.)接种并在1个月后用相同剂量的重组病毒加强。对照小鼠用相同剂量的空MVSchw载体免疫。对于抗体确定，在初次接种后1个月，然后在增强后的2或4周通过眼窝途径收集血样。

[0190] 之前的研究证实，IFNAR小鼠对致死的屈曲病毒感染易感，显示感染的病理表现并提供评价保护的免疫机制的模型(Couderc et al;2008)。

[0191] 对麻疹病毒感染易感的CD46-IFNAR用于评价重组麻疹-屈曲病毒的免疫原性及其

保护效力。这些小鼠表达人CD46基因，其具有人样组织特异性并缺乏I型干扰素受体。小鼠在特定的无病原体条件下饲养于巴斯德研究所动物设施中，所有实验均获批准并按照巴斯德研究生试验动物护理办公室的准则进行。

[0192] 实验1分析麻疹病毒-CE3E26KE1重组病毒在CD46-IFNAR小鼠中的免疫原性和保护效力

[0193] 对麻疹病毒感染易感的6周龄CD46-IFNAR小鼠用2.10^4TCID$_{50}$麻疹病毒-屈曲病毒重组病毒经腹膜内接种，并在1个月后用相同剂量的重组病毒加强。对照小鼠用相同剂量的空麻疹病毒施瓦兹载体(MV Schw)免疫。对于抗体确定，在初次接种后1个月，然后在增强后的2周通过眼窝途径收集血样。然后通过腹膜内注射100pfu屈曲病毒06-49毒株对小鼠进行攻击，以评价保护(免疫和攻击计划如图7所示)。

[0194] 为了评价特异性抗体应答，在接种后不同时间点对小鼠采血。血清在56℃热灭活30min，并通过ELISA检测抗屈曲病毒抗体。用在大肠杆菌中产生的重组屈曲病毒-E2蛋白包被96孔板。HRP缀合的抗鼠免疫球蛋白用作二抗，抗屈曲病毒小鼠抗体用作阳性对照。使用50PFU屈曲病毒06-49(在Vero细胞上产生)在Vero细胞上通过噬菌斑减少中和试验(PRNT)检测抗屈曲病毒中和抗体(Warter L et al.JIM 2011(D4enclosed)and Russell PK et al.JIM 1967)on Vero cells using 50PFU of Chikungunya virus-06-49(produced on Vero cells)。端点滴度以测试的减少PFU数量至少50％(PRNT50)或90％(PRNT90)的最高血清稀释计算。

[0195] 单次注射麻疹病毒-CE3E26KE1重组病毒诱导高抗体滴度，其被二次注射强烈加强(表1-图8)。在两次免疫后，诱导高中和滴度(PRNT50＝450-4050,PRNT90＝50-450)。用10^4或10^5TCID50免疫的所有动物都免受100PFU CHIKV-06-49的CHIKV致死攻击，而用较低剂量(10^3TCID50)免疫保护83％的动物。

[0196] 表1.CD46-IFNAR小鼠对MV-sE2和MV-CE3E26KE1免疫的抗体应答(在混合小鼠血清中确定)

[0197]

	Elisa 1 剂量	Elisa 2 剂量	PRNT50	PRNT90
MV	<100	<100	<50	<50
MV-CHIK.sE2	450	4000	<50	<50
MV-CHIK.CE3E26KE1	4000	>12000	1350	150
抗 CHIKV HMAF	ND	ND	4050	450

[0198] 在两次免疫后，诱导高中和滴度(PRNT50＝1350,PRNT90＝150)，其保护小鼠免受100PFU屈曲病毒-06-49的致死攻击。

[0199] 实验2分析单剂量麻疹病毒-CE3E26KE1重组病毒在CD46-IFNAR小鼠中的免疫原性和保护效力

[0200] 6周龄CD46-IFNAR小鼠用10^5TCID$_{50}$麻疹病毒-CE3E26KE1重组病毒经腹膜内接种。对照小鼠用相同剂量的空麻疹病毒Schw载体免疫。在免疫后2周通过眼窝途径收集血样用于抗体确定，然后通过腹膜内注射100pfu屈曲病毒06-49对小鼠进行攻击(免疫和攻击计划

如图9所示)。

[0201] 在IFNAR小鼠中单次注射后麻疹病毒-CE3E26KE1重组病毒诱导高抗体滴度(表2-图10),以及足以给予对100pfu屈曲病毒06-49致死攻击进行保护的中和滴度。

[0202] 表2.CD46-IFNAR小鼠中用MV-CE3E26KE1病毒单次免疫后引发的抗体应答

[0203]

	MV Elisa 滴度	CHIKV Elisa 滴度	CHIKV PRNT50	CHIKV PRNT90
MV	10 000	<100	<50	<50
MV-CHIKV	10 000	4 050	150	50
抗 CHIKV HMAF	ND	ND	12150	450

[0204] 实验3确定CD46-IFNAR小鼠中麻疹病毒-CE3E26KE1重组病毒的保护剂量

[0205] 6周龄CD46-IFNAR小鼠用10^3至10^5TCID50剂量的MV-CHIKV重组病毒经腹膜内(i.p.)接种,并在1个月后用相同剂量加强。对照小鼠用相同剂量的空MVSchw载体免疫。最后一次免疫后1个月,通过腹膜内注射100pfu CHIKV 06-49对小鼠进行攻击(免疫和攻击计划如图11所示)。对于抗体确定,在初次接种后1个月,然后在增强后1个月,攻击之前通过眼窝途径收集血样。进行特异性Elisa’s以检测抗MV和抗CHIKV结合抗体。在Vero细胞上通过噬菌斑减少中和试验(PRNT)确定抗CHIKV中和抗体滴度。

[0206] 结果示于表3和图12。

[0207] 表3.用不同剂量的MV-CE3E26KE1接种后的抗体应答

[0208]

	首次免疫				2 次免疫			
	MV Elisa 滴度	CHIKV Elisa滴度	CHIKV PRNT50	CHIKV PRNT90	MV Elisa 滴度	CHIKV Elisa滴度	CHIKV PRNT50	CHIKV PRNT90
MV 10^5	10000	<100	<50	<50	300000	<100	<50	<50
MV-CHIKV 10^3	1000	1 350	50	<50	3000	2700	450	50
MV-CHIKV 10^4	3000	4 050	150	50	30000	12150	1350	150
MV-CHIKV-10^5	3000	12 150	450	150	300000	48600	4050	450

[0209] 当重组MV剂量增加时,抗MV和抗CHIKV抗体的滴度均增加。MV-CE3E26KE1病毒单次免疫诱导高抗体滴度,其被二次注射增强。在两次免疫后,诱导高中和滴度(PRNT50＝450-4050,PRNT90＝50-450)。用10^4或$10^5$$TCID_{50}$免疫的所有动物都免受100PFU CHIKV-06-49的CHIKV致死攻击,而用较低剂量(10^3TCID50)免疫保护83％的动物(图12)。

[0210] 实验4评价来自由重组MV-CE3E26KE1病毒免疫小鼠的血清的被动转移给予的保护

[0211] 6周龄CD46-IFNAR小鼠用20μl来自10^5TCID50的重组MV-CE3E26KE1免疫小鼠的混合血清经腹膜内(i.p.)接种。对照小鼠接受20μl来自以10^5TCID50空麻疹病毒施瓦兹免疫小鼠的混合血清或20μl抗CHIKV病毒HMAF。血清在总体积100μl的PBS中稀释。血清在用

100pfu屈曲病毒06-49攻击前24h以及16h，然后在攻击后12h转移，以模拟感染动物中的抗体存在。分析小鼠死亡率2周以确定保护(免疫和攻击计划如图13所示)。

[0212] 由MV-CE3E26KE1病毒免疫小鼠的免疫血清的被动转移保护83％的受体小鼠免受致死屈曲病毒攻击，而接受抗屈曲病毒HMAF的小鼠全部受到保护。相比之下，接受来自由空麻疹病毒免疫小鼠的免疫血清的小鼠全部死亡。这些结果表明，由麻疹病毒-CE3E26KE1重组病毒诱导的体液免疫应答在CD46-IFNAR小鼠中给予对屈曲病毒的保护(图14)。

[0213] 实验5诱导特异性细胞介导的免疫应答

[0214] 为了确定MV-CHIKV免疫是否引发细胞介导的免疫应答，我们通过ELISPOT测定测量了来自免疫小鼠的脾细胞在特异性离体刺激时分泌IFN-γ的能力。单次免疫后7天收集脾细胞，对MV-特异性和CHIKV-特异性应答进行评价。1MOI的CHIKV和MV用于脾细胞刺激。检测到显著数量的CHIKV-特异性细胞(多达300/10^6脾细胞，平均150/10^6)(图15)，其代表相似刺激条件下三分之一的MV-特异性应答(多达600/10^6脾细胞，平均500/10^6)。所有用MV-CHKV免疫的八只小鼠除一只外均具有对CHIKV显著的CMI应答。相比之下，用空MVSchw免疫的对照小鼠具有相似的MV-特异性应答，但不具有CHIKV-特异性应答。这些结果表明，单次接种MV-CHIKV在免疫小鼠的脾脏中诱导高水平的CHIKV和MV特异性细胞免疫应答。

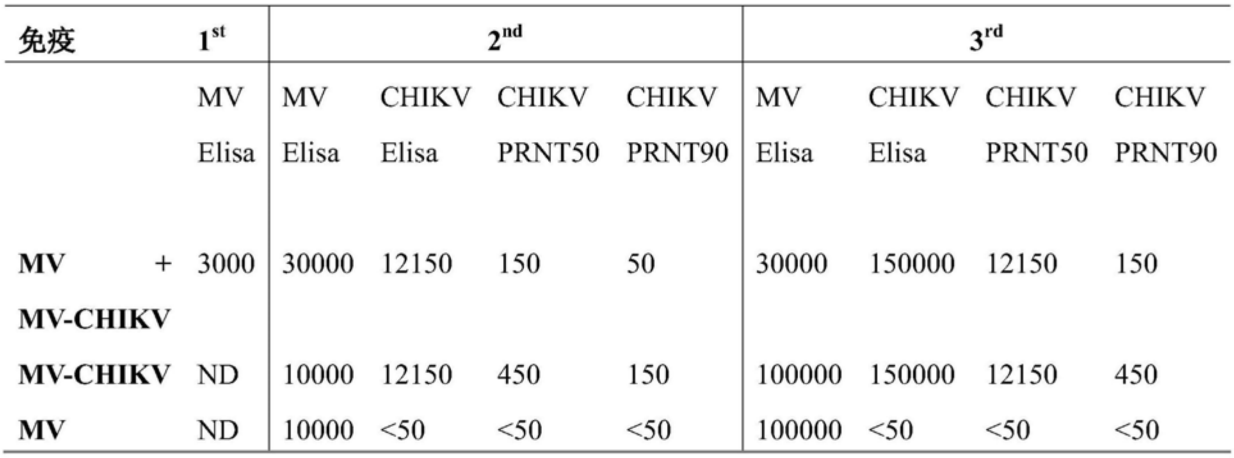
[0215] 实验6分析麻疹病毒预免疫对重组MV-CE3E26KE1病毒在CD46-IFNAR小鼠中的免疫原性和保护效力的影响

[0216] 6周龄CD46-IFNAR小鼠用5.$10^3$$TCID_{50}$的空麻疹病毒施瓦兹经腹膜内(i.p.)接种(图16组1)以模拟预免疫。3个月后，这些小鼠两次注射$10^5$$TCID_{50}$麻疹病毒-CE3E26KE1重组病毒，间隔1个月。对照小鼠用$10^5$$TCID_{50}$麻疹病毒-CE3E26KE1重组病毒(组2)或$10^5$$TCID_{50}$空麻疹病毒Schw(组3)免疫。对于抗体确定，如图16所示通过眼窝途径收集血样，然后通过腹膜内注射100pfu屈曲病毒06-49毒株对小鼠进行攻击，用于保护测定。

[0217] 该实验证实由5.10^3TCID50空麻疹病毒预先免疫的CD46-IFNAR小鼠能够在用麻疹病毒-CE3E26KE1重组病毒免疫后增加保护性屈曲病毒免疫应答。ELISA和PRNT滴度(表4-图17)保持较高，并与天然小鼠中诱导的滴度(ELISA滴度不变，PRNT滴度减少1倍稀释)处于相同范围。在用麻疹病毒-CE3E26KE1重组病毒免疫的预免疫和天然组的动物中，100％的接种动物免受屈曲病毒致死攻击。

[0218] 表4.CD46-IFNAR小鼠在对MV载体预免疫存在下对MV-CE3E26KE1的抗体应答

[0219]



免疫	1st	2nd				3rd			
	MV Elisa	MV Elisa	CHIKV Elisa	CHIKV PRNT50	CHIKV PRNT90	MV Elisa	CHIKV Elisa	CHIKV PRNT50	CHIKV PRNT90
MV + MV-CHIKV	3000	30000	12150	150	50	30000	150000	12150	150
MV-CHIKV	ND	10000	12150	450	150	100000	150000	12150	450
MV	ND	10000	<50	<50	<50	100000	<50	<50	<50

[0220] 实验7:由接种引发的抗体交叉反应

[0221] 为了确定MV-CHIKV免疫是否引发对不同CHIKV原代分离物的交叉反应性抗体应答,测试获自实验4动物的血清中和不同CHIKV原代分离物的能力。选择属于ECSA基因型的4种毒株。

[0222] •CHIKV印度毒株,临床分离物n°3710(NRC for Arbovirus,France),分离于2011年。在Vero细胞上传代1次(2011年11月),病毒滴度6.3log PFU/ml

[0223] •CHIKV刚果毒株,临床分离物n°525(NRC for Arbovirus,France),分离于2011年。在Vero细胞上传代1次(2011年6月),病毒滴度6.5log PFU/ml

[0224] •CHIKV泰国毒株,临床分离物n°1499(NRC for Arbovirus,France),分离于2009年。在C6/36细胞上传代(2009年12月),病毒滴度6.3log PFU/ml

[0225] •CHIKV留尼汪毒株,临床分离物2006.49(NRC for Arbovirus,France),分离于2006年。在Vero细胞上传代3次(2011年4月4日),病毒滴度7.3log PFU/ml

[0226] 使用噬菌斑减少中和试验(PRNT)检测抗CHIKV中和抗体。将Vero细胞种植于12孔板24h。在DMEM Glutamax/2%FCS中系列稀释血清样品。100μl稀释液与含100pfu 06-49毒株的等体积CHIKV于37℃、温和搅拌下孵育2h。然后在覆盖含有最终0.8%(重量/体积)羧甲基纤维素的DMEM GlutaMAX^{TM}/2%FCS的Vero细胞单层上测定剩余感染性。孵育3天后,固定细胞并用结晶紫染色,进行噬菌斑计数确定。端点中和滴度以测试的减少噬菌斑数量至少50%($PRNT_{50}$)的最高血清稀释计算。

[0227] 表5.来自MV-CHIKV免疫小鼠的中和抗体的交叉反应性

[0228]

	1^{st}后的PRNT50		2nd后的PRNT50		攻击后的PRNT50	
	MV-CHIKV	MV	MV-CHIKV	MV	MV-CHIKV	MV
06-49毒株	50	<50	1350	<50	450	ND
印度毒株	150	<50	1350	<50	13250	ND
刚果毒株	<50	<50	1350	<50	1350	ND
泰国毒株	150	<50	4050	<50	12150	ND

[0229] 结果表明,MV-CHIKV免疫小鼠诱导交叉反应性抗体,其能够中和来自不同国家的CHIKV的几种原代分离物。有趣的是,用06-49留尼汪病毒攻击动物导致甚至扩展应答,并将对印度和泰国分离物的中和增强至极高水平。仅对ECSA基因型进行测试,因为其在巴斯德研究所可获得。

[0230] 实验8:MV-CHIKV在食蟹猴中的免疫原性

[0231] 在非人灵长类中测试重组麻疹病毒疫苗对屈曲的免疫原性。预先选择的对黄病毒和麻疹病毒血清阴性的2组4只食蟹猴(macaca fascicularis)于第0天皮下接种10^4或10^5 $TCID_{50}$的MV-CHIKV,然后在第90天用相同剂量加强。收集血清和血浆并储存于-20℃用于后续分析。通过使用PRNT测定检测对屈曲病毒的中和抗体。

[0232] 表6

[0233]

组	动物	编码	疫苗	剂量	倒数 CHIK-PRNT			
					第 0 天	第 21 天	第 90 天	第 111 天
D	D1	12/9	MV-CHIK	10^4	<10	574	<10	14
	D2	12/4	MV-CHIK	10^4	<10	142	73	97
	D3	12/5	MV-CHIK	10^4	<10	16	17	121
	D4	12/10	MV-CHIK	10^4	<10	70	83	118
E	E1	12/16	MV-CHIK	10^5	<10	247	101	382
	E2	12/23	MV-CHIK	10^5	<10	103	178	355
	E3	12/26	MV-CHIK	10^5	<10	6173	151	652
	E4	12/22	MV-CHIK	10^5	<10	97	386	228

[0234] 表6所示结果表明，所有猴子出现中和CHIKV的高滴度抗体。最高剂量比较低剂量更有效。如图18所示，在大多数动物中，增强是有效的(第90天，该天进行增强，对比第111天，增强后第21天)。这些结果证实MV-CHIKV疫苗候选物在非人灵长类中的免疫原性。

[0235] 总体结论

[0236] 发明人产生了稳定表达CHIKV毒株06.49的完整结构蛋白CE3E26KE1的重组MV-CHIK病毒。被该重组病毒感染的Vero细胞表达高水平的CHIKV蛋白并且分泌高密度VLP。该重组病毒的生长动力学稍微延迟，但产生与空MV载体相似的滴度。经在对MV感染易感的CD46-IFNAR小鼠中的评价，该疫苗候选物根据施用的剂量和次数诱导高水平的CHIKV中和抗体(PRNT50＝450-4050;PRNT90＝150-450)。所有免疫小鼠均可重复地受到保护，即使是在单次施用后，表明该疫苗候选物强烈的免疫能力。在天然动物中被动转移免疫血清给予对致死攻击的保护，即使是在这些高度易感的小鼠中。最后，发明人证实CD46-IFNAR小鼠中对MV载体预先免疫的存在并不妨碍在用MV-CE3E26KE1疫苗候选物免疫后诱导保护性免疫。基于这些结果，由此获得的重组载体值得在可靠的非人灵长类感染模型中进行评估。

序列表

<110> 巴斯德研究院
国立科学研究中心
特密斯生物科学有限责任公司

<120> 表达屈曲病毒多肽的重组麻疹病毒及其应用

<130> B09995A - AD/DP/AM

<140> PCT/ EP XXXXXXX
<141> 2013-09-26

<150> EP 12306176.4
<151> 2012-09-27

<160> 33

<170> PatentIn version 3.5

<210> 1
<211> 18967
<212> DNA
<213> Artificial Sequence

<220>
<223> pTM-MVSchw (GenBank: FW366202.1)

<400> 1
gcggccgcta atacgactca ctatagggcc aactttgttt ggtctgatga gtccgtgagg 60

acgaaacccg gagtcccggg tcaccaaaca aagttgggta aggatagttc aatcaatgat 120

catcttctag tgcacttagg attcaagatc ctattatcag ggacaagagc aggattaggg 180

atatccgaga tggccacact tttaaggagc ttagcattgt tcaaaagaaa caaggacaaa 240

[0001] ccacccatta catcaggatc cggtggagcc atcagaggaa tcaaacacat tattatagta 300

ccaatccctg gagattcctc aattaccact cgatccagac ttctggaccg gttggtgagg 360

ttaattggaa acccggatgt gagcgggccc aaactaacag ggcactaat aggtatatta 420

tccttatttg tggagtctcc aggtcaattg attcagagga tcaccgatga ccctgacgtt 480

agcataaggc tgttagaggt tgtccagagt gaccagtcac aatctggcct taccttcgca 540

tcaagaggta ccaacatgga ggatgaggcg gaccaatact tttcacatga tgatccaatt 600

agtagtgatc aatccaggtt cggatggttc gggaacaagg aaatctcaga tattgaagtg 660

caagaccctg agggattcaa catgattctg ggtaccatcc tagcccaaat ttgggtcttg 720

ctcgcaaagg cggttacggc cccagacacg gcagctgatt cggagctaag aaggtggata 780

aagtacaccc aacaaagaag ggtagttggt gaatttagat tggagagaaa atggttggat 840

gtggtgagga acaggattgc cgaggacctc tccttacgcc gattcatggt cgctctaatc 900

ctggatatca agagaacacc cggaaacaaa cccaggattg ctgaaatgat atgtgacatt 960

gatacatata tcgtagaggc aggattagcc agttttatcc tgactattaa gtttgggata 1020

gaaactatgt atcctgctct tggactgcat gaatttgctg gtgagttatc cacacttgag 1080

tccttgatga acctttacca gcaaatgggg gaaactgcac cctacatggt aatcctggag 1140

aactcaattc agaacaagtt cagtgcagga tcatacccct tgctctggag ctatgccatg 1200

ggagtaggag tggaacttga aaactccatg ggaggtttga actttggccg atcttacttt 1260

gatccagcat attttagatt agggcaagag atggtaagga ggtcagctgg aaaggtcagt 1320

tccacattgg catctgaact cggtatcact gccgaggatg caaggcttgt ttcagagatt 1380

gcaatgcata ctactgagga caagatcagt agagcggttg gacccagaca agcccaagta 1440

tcatttctac acggtgatca aagtgagaat gagctaccga gattgggggg caaggaagat 1500
aggagggtca aacagagtcg aggagaagcc agggagagct acagagaaac cgggcccagc 1560
agagcaagtg atgcgagagc tgcccatctt ccaaccggca cacccctaga cattgacact 1620
gcaacggagt ccagccaaga tccgcaggac agtcgaaggt cagctgacgc cctgcttagg 1680
ctgcaagcca tggcaggaat ctcggaagaa caaggctcag acacggacac ccctatagtg 1740
tacaatgaca gaaatcttct agactaggtg cgagaggccg agggccagaa caacatccgc 1800
ctaccatcca tcattgttat aaaaaactta ggaaccaggt ccacacagcc gccagcccat 1860
caaccatcca ctcccacgat tggagccaat ggcagaagag caggcacgcc atgtcaaaaa 1920
cggactggaa tgcatccggg ctctcaaggc cgagcccatc ggctcactgg ccatcgagga 1980
agctatggca gcatggtcag aaatatcaga caacccagga caggagcgag ccacctgcag 2040
ggaagagaag gcaggcagtt cgggtctcag caaaccatgc ctctcagcaa ttggatcaac 2100
tgaaggcggt gcacctcgca tccgcggtca gggacctgga gagagcgatg acgacgctga 2160
aactttggga atcccccaa gaaatctcca ggcatcaagc actgggttac agtgttatta 2220
cgtttatgat cacagcggtg aagcggttaa gggaatccaa gatgctgact ctatcatggt 2280
tcaatcaggc cttgatggtg atagcaccct ctcaggagga gacaatgaat ctgaaaacag 2340
cgatgtggat attggcgaac ctgataccga gggatatgct atcactgacc ggggatctgc 2400
tcccatctct atggggttca gggcttctga tgttgaaact gcagaaggag gggagatcca 2460
cgagctcctg agactccaat ccagaggcaa caactttccg aagcttggga aaactctcaa 2520
tgttcctccg cccccggacc ccggtagggc cagcacttcc gggacaccca ttaaaaaggg 2580

[0002]

cacagacgcg agattagcct catttggaac ggagatcgcg tctttattga caggtggtgc 2640
aacccaatgt gctcgaaagt caccctcgga accatcaggg ccaggtgcac ctgcggggaa 2700
tgtccccgag tgtgtgagca atgccgcact gatacaggag tggacacccg aatctggtac 2760
cacaatctcc ccgagatccc agaataatga agaaggggga gactattatg atgatgagct 2820
gttctctgat gtccaagata ttaaaacagc cttggccaaa atacacgagg ataatcagaa 2880
gataatctcc aagctagaat cactgctgtt attgaaggga gaagttgagt caattaagaa 2940
gcagatcaac aggcaaaata tcagcatatc caccctggaa ggacacctct caagcatcat 3000
gatcgccatt cctggacttg ggaaggatcc caacgacccc actgcagatg tcgaaatcaa 3060
tcccgacttg aaacccatca taggcagaga ttcaggccga gcactggccg aagttctcaa 3120
gaaacccgtt gccagccgac aactccaagg aatgacaaat ggacggacca gttccagagg 3180
acagctgctg aaggaatttc agctaaagcc gatcgggaaa aagatgagct cagccgtcgg 3240
gtttgttcct gacaccggcc ctgcatcacg cagtgtaatc cgctccatta taaaatccag 3300
ccggctagag gaggatcgga agcgttacct gatgactctc cttgatgata tcaaaggagc 3360
caatgatctt gccaagttcc accagatgct gatgaagata ataatgaagt agctacagct 3420
caacttacct gccaacccca tgccagtcga cccaactagt acaacctaaa tccattataa 3480
aaaacttagg agcaaagtga ttgcctccca aggtccacaa tgacagagac ctacgacttc 3540
gacaagtcgg catgggacat caaagggtcg atcgctccga tacaacccac cacctacagt 3600
gatggcaggc tggtgcccca ggtcagagtc atagatcctg gtctaggcga caggaaggat 3660
gaatgcttta tgtacatgtt tctgctgggg gttgttgagg acagcgattc cctagggcct 3720

ccaatcgggc gagcatttgg gttcctgccc ttaggtgttg gcagatccac agcaaagccc 3780
gaaaaactcc tcaaagaggc cactgagctt gacatagttg ttagacgtac agcagggctc 3840
aatgaaaaac tggtgttcta caacaacacc ccactaactc tcctcacacc ttggagaaag 3900
gtcctaacaa cagggagtgt cttcaacgca aaccaagtgt gcaatgcggt taatctgata 3960
ccgctcgata ccccgcagag gttccgtgtt gtttatatga gcatcacccg tctttcggat 4020
aacgggtatt acaccgttcc tagaagaatg ctggaattca gatcggtcaa tgcagtggcc 4080
ttcaacctgc tggtgaccct taggattgac aaggcgatag gccctgggaa gatcatcgac 4140
aatacagagc aacttcctga ggcaacattt atggtccaca tcgggaactt caggagaaag 4200
aagagtgaag tctactctgc cgattattgc aaaatgaaaa tcgaaaagat gggcctggtt 4260
tttgcacttg gtgggatagg gggcaccagt cttcacatta gaagcacagg caaaatgagc 4320
aagactctcc atgcacaact cgggttcaag aagaccttat gttacccgct gatggatatc 4380
aatgaagacc ttaatcgatt actctggagg agcagatgca agatagtaag aatccaggca 4440
gttttgcagc catcagttcc tcaagaattc cgcatttacg acgacgtgat cataaatgat 4500
gaccaaggac tattcaaagt tctgtagacc gtagtgccca gcaatgcccg aaaacgaccc 4560
ccctcacaat gacagccaga aggcccggac aaaaaagccc cctccgaaag actccacgga 4620
ccaagcgaga ggccagccag cagccgacgg caagcgcgaa caccaggcgg ccccagcaca 4680
gaacagccct gacacaaggc caccaccagc caccccaatc tgcatcctcc tcgtgggacc 4740
cccgaggacc aacccccaag gctgcccccg atccaaacca ccaaccgcat ccccaccacc 4800
cccgggaaag aaacccccag caattggaag gcccctcccc ctcttcctca acacaagaac 4860
[0003] tccacaaccg aaccgcacaa gcgaccgagg tgacccaacc gcaggcatcc gactccctag 4920
acagatcctc tctccccggc aaactaaaca aaacttaggg ccaaggaaca tacacaccca 4980
acagaaccca gaccccggcc cacggcgccg cgcccccaac ccccgacaac cagagggagc 5040
ccccaaccaa tcccgccggc tcccccggtg cccacaggca gggacaccaa cccccgaaca 5100
gacccagcac ccaaccatcg acaatccaag acgggggggc ccccccaaaa aaaggccccc 5160
aggggccgac agccagcacc gcgaggaagc ccacccaccc cacacacgac cacggcaacc 5220
aaaccagaac ccagaccacc ctgggccacc agctcccaga ctcggccatc accccgcaga 5280
aaggaaaggc cacaacccgc gcaccccagc cccgatccgg cggggagcca cccaacccga 5340
accagcaccc aagagcgatc cccgaaggac ccccgaaccg caaaggacat cagtatccca 5400
cagcctctcc aagtcccccg gtctcctcct cttctcgaag ggaccaaaag atcaatccac 5460
cacacccgac gacactcaac tccccacccc taaaggagac accgggaatc ccagaatcaa 5520
gactcatcca atgtccatca tgggtctcaa ggtgaacgtc tctgccatat tcatggcagt 5580
actgttaact ctccaaacac ccaccggtca aatccattgg ggcaatctct ctaagatagg 5640
ggtggtagga ataggaagtg caagctacaa agttatgact cgttccagcc atcaatcatt 5700
agtcataaaa ttaatgccca atataactct cctcaataac tgcacgaggg tagagattgc 5760
agaatacagg agactactga gaacagtttt ggaaccaatt agagatgcac ttaatgcaat 5820
gacccagaat ataagaccgg ttcagagtgt agcttcaagt aggagacaca agagatttgc 5880
gggagtagtc ctggcaggtg cggccctagg cgttgccaca gctgctcaga taacagccgg 5940
cattgcactt caccagtcca tgctgaactc tcaagccatc gacaatctga gagcgagcct 6000
ggaaactact aatcaggcaa ttgagacaat cagacaagca gggcaggaga tgatattggc 6060

tgttcagggt gtccaagact acatcaataa tgagctgata ccgtctatga accaactatc 6120
ttgtgattta atcggccaga agctcgggct caaattgctc agatactata cagaaatcct 6180
gtcattattt ggccccagtt tacgggaccc catatctgcg gagatatcta tccaggcttt 6240
gagctatgcg cttggaggag acatcaataa ggtgttagaa aagctcggat acagtggagg 6300
tgatttactg ggcatcttag agagcggagg aataaaggcc cggataactc acgtcgacac 6360
agagtcctac ttcattgtcc tcagtatagc ctatccgacg ctgtccgaga ttaagggggt 6420
gattgtccac cggctagagg gggtctcgta caacataggc tctcaagagt ggtataccac 6480
tgtgcccaag tatgttgcaa cccaagggta ccttatctcg aattttgatg agtcatcgtg 6540
tactttcatg ccagagggga ctgtgtgcag ccaaaatgcc ttgtacccga tgagtcctct 6600
gctccaagaa tgcctccggg ggtacaccaa gtcctgtgct cgtacactcg tatccgggtc 6660
ttttgggaac cggttcattt tatcacaagg gaacctaata gccaattgtg catcaatcct 6720
ttgcaagtgt tacacaacag gaacgatcat taatcaagac cctgacaaga tcctaacata 6780
cattgctgcc gatcactgcc cggtagtcga ggtgaacggc gtgaccatcc aagtcgggag 6840
caggaggtat ccagacgctg tgtacttgca cagaattgac ctcggtcctc ccatatcatt 6900
ggagaggttg gacgtaggga caaatctggg gaatgcaatt gctaagttgg aggatgccaa 6960
ggaattgttg gagtcatcgg accagatatt gaggagtatg aaaggtttat cgagcactag 7020
catagtctac atcctgattg cagtgtgtct tggagggttg atagggatcc ccgctttaat 7080
atgttgctgc agggggcgtt gtaacaaaaa gggagaacaa gttggtatgt caagaccagg 7140
[0004] cctaaagcct gatcttacgg gaacatcaaa atcctatgta aggtcgctct gatcctctac 7200
aactcttgaa acacaaatgt cccacaagtc tcctcttcgt catcaagcaa ccaccgcacc 7260
cagcatcaag cccacctgaa attatctccg gcttccctct ggccgaacaa tatcggtagt 7320
taatcaaaac ttagggtgca agatcatcca caatgtcacc acaacgagac cggataaatg 7380
ccttctacaa agataacccc catcccaagg gaagtaggat agtcattaac agagaacatc 7440
ttatgattga tagaccttat gttttgctgg ctgttctgtt tgtcatgttt ctgagcttga 7500
tcgggttgct agccattgca ggcattagac ttcatcgggc agccatctac accgcagaga 7560
tccataaaag cctcagcacc aatctagatg taactaactc aatcgagcat caggtcaagg 7620
acgtgctgac accactcttc aaaatcatcg gtgatgaagt gggcctgagg acacctcaga 7680
gattcactga cctagtgaaa ttaatctctg acaagattaa attccttaat ccggataggg 7740
agtacgactt cagagatctc acttggtgta tcaacccgcc agagagaatc aaattggatt 7800
atgatcaata ctgtgcagat gtggctgctg aagagctcat gaatgcattg gtgaactcaa 7860
ctctactgga gaccagaaca accaatcagt tcctagctgt ctcaaaggga aactgctcag 7920
ggcccactac aatcagaggt caattctcaa acatgtcgct gtccctgtta gacttgtatt 7980
taggtcgagg ttacaatgtg tcatctatag tcactatgac atcccaggga atgtatgggg 8040
gaacttacct agtggaaaag cctaatctga gcagcaaaag gtcagagttg tcacaactga 8100
gcatgtaccg agtgtttgaa gtaggtgtta tcagaaatcc gggtttgggg gctccggtgt 8160
tccatatgac aaactatctt gagcaaccag tcagtaatga tctcagcaac tgtatggtgg 8220
ctttggggga gctcaaactc gcagcccttt gtcacgggga agattctatc acaattccct 8280
atcagggatc agggaaaggt gtcagcttcc agctcgtcaa gctaggtgtc tggaaatccc 8340

caaccgacat gcaatcctgg gtccccttat caacggatga tccagtgata gacaggcttt 8400
acctctcatc tcacagaggt gttatcgctg acaatcaagc aaaatgggct gtcccgacaa 8460
cacgaacaga tgacaagttg cgaatggaga catgcttcca acaggcgtgt aagggtaaaa 8520
tccaagcact ctgcgagaat cccgagtggg caccattgaa ggataacagg attccttcat 8580
acggggtctt gtctgttgat ctgagtctga cagttgagct taaaatcaaa attgcttcgg 8640
gattcgggcc attgatcaca cacggttcag ggatggacct atacaaatcc aaccacaaca 8700
atgtgtattg gctgactatc ccgccaatga agaacctagc cttaggtgta atcaacacat 8760
tggagtggat accgagattc aaggttagtc cctacctctt cactgtccca attaaggaag 8820
caggcgaaga ctgccatgcc ccaacatacc tacctgcgga ggtggatggt gatgtcaaac 8880
tcagttccaa tctggtgatt ctacctggtc aagatctcca atatgttttg gcaacctacg 8940
atacttccag ggttgaacat gctgtggttt attacgttta cagcccaagc cgctcatttt 9000
cttactttta tccttttagg ttgcctataa agggggtccc catcgaatta caagtggaat 9060
gcttcacatg ggaccaaaaa ctctggtgcc gtcacttctg tgtgcttgcg gactcagaat 9120
ctggtggaca tatcactcac tctgggatgg tgggcatggg agtcagctgc acagtcaccc 9180
gggaagatgg aaccaatcgc agatagggct gctagtgaac caatcacatg atgtcaccca 9240
gacatcaggc atacccacta gtgtgaaata gacatcagaa ttaagaaaaa cgtagggtcc 9300
aagtggttcc ccgttatgga ctcgctatct gtcaaccaga tcttataccc tgaagttcac 9360
ctagatagcc cgatagttac caataagata gtagccatcc tggagtatgc tcgagtccct 9420
cacgcttaca gcctggagga ccctacactg tgtcagaaca tcaagcaccg cctaaaaaac 9480
[0005] ggattttcca accaaatgat tataaacaat gtggaagttg ggaatgtcat caagtccaag 9540
cttaggagtt atccggccca ctctcatatt ccatatccaa attgtaatca ggatttattt 9600
aacatagaag acaaagagtc aacgaggaag atccgtgaac tcctcaaaaa ggggaattcg 9660
ctgtactcca aagtcagtga taaggttttc caatgcttaa gggacactaa ctcacggctt 9720
ggcctaggct ccgaattgag ggaggacatc aaggagaaag ttattaactt gggagtttac 9780
atgcacagct cccagtggtt tgagcccttt ctgttttggt ttacagtcaa gactgagatg 9840
aggtcagtga ttaaatcaca aacccatact tgccatagga ggagacacac acctgtattc 9900
ttcactggta gttcagttga gttgctaatc tctcgtgacc ttgttgctat aatcagtaaa 9960
gagtctcaac atgtatatta cctgacattt gaactggttt tgatgtattg tgatgtcata 10020
gaggggaggt taatgacaga gaccgctatg actattgatg ctaggtatac agagcttcta 10080
ggaagagtca gatacatgtg gaaactgata gatggtttct tccctgcact cgggaatcca 10140
acttatcaaa ttgtagccat gctggagcct ctttcacttg cttacctgca gctgagggat 10200
ataacagtag aactcagagg tgctttcctt aaccactgct ttactgaaat acatgatgtt 10260
cttgaccaaa acgggttttc tgatgaaggt acttatcatg agttaactga agctctagat 10320
tacattttca taactgatga catacatctg acaggggaga ttttctcatt tttcagaagt 10380
ttcggccacc ccagacttga agcagtaacg gctgctgaaa atgttaggaa atacatgaat 10440
cagcctaaag tcattgtgta tgagactctg atgaaaggtc atgccatatt ttgtggaatc 10500
ataatcaacg gctatcgtga caggcacgga ggcagttggc caccgctgac cctccccctg 10560
catgctgcag acacaatccg gaatgctcaa gcttcaggtg aagggttaac acatgagcag 10620
tgcgttgata actggaaatc ttttgctgga gtgaaatttg gctgctttat gcctcttagc 10680

ctggatagtg atctgacaat gtacctaaag gacaaggcac ttgctgctct ccaaagggaa 10740
tgggattcag tttacccgaa agagttcctg cgttacgacc ctcccaaggg aaccgggtca 10800
cggaggcttg tagatgtttt ccttaatgat tcgagctttg acccatatga tgtgataatg 10860
tatgttgtaa gtggagctta cctccatgac cctgagttca acctgtctta cagcctgaaa 10920
gaaaaggaga tcaaggaaac aggtagactt tttgctaaaa tgacttacaa aatgagggca 10980
tgccaagtga ttgctgaaaa tctaatctca aacgggattg gcaaatattt taaggacaat 11040
gggatggcca aggatgagca cgatttgact aaggcactcc acactctagc tgtctcagga 11100
gtccccaaag atctcaaaga aagtcacagg ggggggccag tcttaaaaac ctactcccga 11160
agcccagtcc acacaagtac caggaacgtg agagcagcaa aagggtttat agggttccct 11220
caagtaattc ggcaggacca agacactgat catccggaga atatggaagc ttacgagaca 11280
gtcagtgcat ttatcacgac tgatctcaag aagtactgcc ttaattggag atatgagacc 11340
atcagcttgt ttgcacagag gctaaatgag atttacggat tgccctcatt tttccagtgg 11400
ctgcataaga ggcttgagac ctctgtcctg tatgtaagtg accctcattg cccccccgac 11460
cttgacgccc atatcccgtt atataaagtc cccaatgatc aaatcttcat taagtaccct 11520
atgggaggta tagaagggta ttgtcagaag ctgtggacca tcagcaccat tccctatcta 11580
tacctggctg cttatgagag cggagtaagg attgcttcgt tagtgcaagg ggacaatcag 11640
accatagccg taacaaaaag ggtacccagc acatggccct acaaccttaa gaaacgggaa 11700
gctgctagag taactagaga ttactttgta attcttaggc aaaggctaca tgatattggc 11760
[0006] catcacctca aggcaaatga gacaattgtt tcatcacatt tttttgtcta ttcaaaagga 11820
atatattatg atgggctact tgtgtcccaa tcactcaaga gcatcgcaag atgtgtattc 11880
tggtcagaga ctatagttga tgaaacaagg gcagcatgca gtaatattgc tacaacaatg 11940
gctaaaagca tcgagagagg ttatgaccgt taccttgcat attccctgaa cgtcctaaaa 12000
gtgatacagc aaattctgat ctctcttggc ttcacaatca attcaaccat gacccgggat 12060
gtagtcatac ccctcctcac aaacaacgac ctcttaataa ggatggcact gttgcccgct 12120
cctattgggg ggatgaatta tctgaatatg agcaggctgt ttgtcagaaa catcggtgat 12180
ccagtaacat catcaattgc tgatctcaag agaatgattc tcgcctcact aatgcctgaa 12240
gagaccctcc atcaagtaat gacacaacaa ccggggggact cttcattcct agactgggct 12300
agcgaccctt actcagcaaa tcttgtatgt gtccagagca tcactagact cctcaagaac 12360
ataactgcaa ggtttgtcct gatccatagt ccaaacccaa tgttaaaagg attattccat 12420
gatgacagta aagaagagga cgagggactg gcggcattcc tcatggacag gcatattata 12480
gtacctaggg cagctcatga aatcctggat catagtgtca caggggcaag agagtctatt 12540
gcaggcatgc tggataccac aaaaggcttg attcgagcca gcatgaggaa gggggggtta 12600
acctctcgag tgataaccag attgtccaat tatgactatg aacaattcag agcagggatg 12660
gtgctattga caggaagaaa gagaaatgtc ctcattgaca aagagtcatg ttcagtgcag 12720
ctggcgagag ctctaagaag ccatatgtgg gcgaggctag ctcgaggacg gcctatttac 12780
ggccttgagg tccctgatgt actagaatct atgcgaggcc accttattcg gcgtcatgag 12840
acatgtgtca tctgcgagtg tggatcagtc aactacggat ggtttttgt cccctcgggt 12900
tgccaactgg atgatattga caaggaaaca tcatccttga gagtcccata tattggttct 12960

accactgatg agagaacaga catgaagctt gccttcgtaa gagccccaag tcgatccttg 13020
cgatctgctg ttagaatagc aacagtgtac tcatgggctt acggtgatga tgatagctct 13080
tggaacgaag cctggttgtt ggctaggcaa agggccaatg tgagcctgga ggagctaagg 13140
gtgatcactc ccatctcaac ttcgactaat ttagcgcata ggttgaggga tcgtagcact 13200
caagtgaaat actcaggtac atcccttgtc cgagtggcga ggtataccac aatctccaac 13260
gacaatctct catttgtcat atcagataag aaggttgata ctaactttat ataccaacaa 13320
ggaatgcttc tagggttggg tgttttagaa acattgtttc gactcgagaa agataccgga 13380
tcatctaaca cggtattaca tcttcacgtc gaaacagatt gttgcgtgat cccgatgata 13440
gatcatccca ggatacccag ctcccgcaag ctagagctga gggcagagct atgtaccaac 13500
ccattgatat atgataatgc acctttaatt gacagagatg caacaaggct atacacccag 13560
agccatagga ggcaccttgt ggaatttgtt acatggtcca caccccaact atatcacatt 13620
ttagctaagt ccacagcact atctatgatt gacctggtaa caaaatttga gaaggaccat 13680
atgaatgaaa tttcagctct catagggat gacgatatca atagtttcat aactgagttt 13740
ctgctcatag agccaagatt attcactatc tacttgggcc agtgtgcggc catcaattgg 13800
gcatttgatg tacattatca tagaccatca gggaaatatc agatgggtga gctgttgtca 13860
tcgttccttt ctagaatgag caaaggagtg tttaaggtgc ttgtcaatgc tctaagccac 13920
ccaaagatct acaagaaatt ctggcattgt ggtattatag agcctatcca tggtccttca 13980
cttgatgctc aaaacttgca cacaactgtg tgcaacatgg tttacacatg ctatatgacc 14040
tacctcgacc tgttgttgaa tgaagagtta gaagagttca catttctctt gtgtgaaagc 14100
[0007] gacgaggatg tagtaccgga cagattcgac aacatccagg caaaacactt atgtgttctg 14160
gcagatttgt actgtcaacc agggacctgc ccaccaattc gaggtctaag accggtagag 14220
aaatgtgcag ttctaaccga ccatatcaag gcagaggcta tgttatctcc agcaggatct 14280
tcgtggaaca taaatccaat tattgtagac cattactcat gctctctgac ttatctccgg 14340
cgaggatcga tcaaacagat aagattgaga gttgatccag gattcatttt cgacgccctc 14400
gctgaggtaa atgtcagtca gccaaagatc ggcagcaaca acatctcaaa tatgagcatc 14460
aaggctttca gacccccaca cgatgatgtt gcaaaattgc tcaaagatat caacacaagc 14520
aagcacaatc ttcccatttc agggggcaat ctcgccaatt atgaaatcca tgctttccgc 14580
agaatcgggt tgaactcatc tgcttgctac aaagctgttg agatatcaac attaattagg 14640
agatgccttg agccagggga ggacggcttg ttcttgggtg agggatcggg ttctatgttg 14700
atcacttata aagagatact taaactaaac aagtgcttct ataatagtgg ggtttccgcc 14760
aattctagat ctggtcaaag ggaattagca ccctatccct ccgaagttgg ccttgtcgaa 14820
cacagaatgg gagtaggtaa tattgtcaaa gtgctcttta acgggaggcc cgaagtcacg 14880
tgggtaggca gtgtagattg cttcaatttc atagttagta atatccctac ctctagtgtg 14940
gggtttatcc attcagatat agagaccttg cctgacaaag atactataga gaagctagag 15000
gaattggcag ccatcttatc gatggctctg ctcctgggca aaataggatc aatactggtg 15060
attaagctta tgcctttcag cggggatttt gttcagggat ttataagtta tgtagggtct 15120
cattatagag aagtgaacct tgtataccct agatacagca acttcatctc tactgaatct 15180
tatttggtta tgacagatct caaggctaac cggctaatga atcctgaaaa gattaagcag 15240
cagataattg aatcatctgt gaggacttca cctggactta taggtcacat cctatccatt 15300

aagcaactaa gctgcataca agcaattgtg ggagacgcag ttagtagagg tgatatcaat 15360

cctactctga aaaaacttac acctatagag caggtgctga tcaattgcgg gttggcaatt 15420

aacggaccta agctgtgcaa agaattgatc caccatgatg ttgcctcagg gcaagatgga 15480

ttgcttaatt ctatactcat cctctacagg gagttggcaa gattcaaaga caaccaaaga 15540

agtcaacaag ggatgttcca cgcttacccc gtattggtaa gtagcaggca acgagaactt 15600

atatctagga tcacccgcaa attctggggg cacattcttc tttactccgg gaacaaaaag 15660

ttgataaata agtttatcca gaatctcaag tccggctatc tgatactaga cttacaccag 15720

aatatcttcg ttaagaatct atccaagtca gagaaacaga ttattatgac gggggtttg 15780

aaacgtgagt gggtttttaa ggtaacagtc aaggagacca aagaatggta taagttagtc 15840

ggatacagtg ccctgattaa ggactaattg gttgaactcc ggaaccctaa tcctgcccta 15900

ggtggttagg cattatttgc aatatattaa agaaaacttt gaaaatacga agtttctatt 15960

cccagctttg tctggtggcc ggcatggtcc cagcctcctc gctggcgccg gctgggcaac 16020

attccgaggg gaccgtcccc tcggtaatgg cgaatgggac gcggccgatc cggctgctaa 16080

caaagcccga aaggaagctg agttggctgc tgccaccgct gagcaataac tagcataacc 16140

ccttggggcc tctaaacggg tcttgagggg ttttttgctg aaaggaggaa ctatatccgg 16200

atgcggccgc gggccctatg gtacccagct tttgttccct ttagtgaggg ttaattccga 16260

gcttggcgta atcatggtca tagctgtttc ctgtgtgaaa ttgttatccg ctcacaattc 16320

cacacaacat aggagccgga agcataaagt gtaaagcctg gggtgcctaa tgagtgaggt 16380

[0008] aactcacatt aattgcgttg cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc 16440

agctgcatta atgaatcggc caacgcgcgg ggagaggcgg tttgcgtatt gggcgctctt 16500

ccgcttcctc gctcactgac tcgctgcgct cggtcgttcg gctgcggcga gcggtatcag 16560

ctcactcaaa ggcggtaata cggttatcca cagaatcagg ggataacgca ggaaagaaca 16620

tgtgagcaaa aggccagcaa aaggccagga accgtaaaaa ggccgcgttg ctggcgtttt 16680

tccataggct cggcccccct gacgagcatc acaaaaatcg acgctcaagt cagaggtggc 16740

gaaacccgac aggactataa agataccagg cgttcccccc tggaagctcc ctcgtgcgct 16800

ctcctgttcc gaccctgccg cttaccggat acctgtccgc ctttctccct tcgggaagcg 16860

tggcgctttc tcaatgctca cgctgtaggt atctcagttc ggtgtaggtc gttcgctcca 16920

agctgggctg tgtgcacgaa ccccccgttc agcccgaccg ctgcgcctta tccggtaact 16980

atcgtcttga gtccaacccg gtaagacacg acttatcgcc actggcagca gccactggta 17040

acaggattag cagagcgagg tatgtaggcg gtgctacaga gttcttgaag tggtggccta 17100

actacggcta cactagaagg acagtatttg gtatctgcgc tctgctgaag ccagttacct 17160

tcggaaaaag agttggtagc tcttgatccg gcaaacaaac caccgctggt agcggtggtt 17220

tttttgtttg caagcagcag attacgcgca gaaaaaaagg atctcaagaa gatcctttga 17280

tcttttctac ggggtctgac gctcagtgga acgaaaactc acgttaaggg attttggtca 17340

tgagattatc aaaaaggatc ttcacctaga tccttttaaa ttaaaaatga agttttaaat 17400

caatctaaag tatatatgag taaacttggt ctgacagtta ccaatgctta atcagtgagg 17460

cacctatctc agcgatctgt ctatttcgtt catccatagt tgcctgactg ccccgtcgtgt 17520

agataactac gatacgggag ggcttaccat ctggccccag tgctgcaatg ataccgcgag 17580

acccacgctc accggctcca gatttatcag caataaacca gccagccgga agggccgagc 17640
gcagaagtgg tcctgcaact ttatccgcct ccatccagtc tattaattgt tgccgggaag 17700
ctagagtaag tagttcgcca gttaatagtt tgcgcaacgt tgttgccatt gctacaggca 17760
tcgtggtgtc acgctcgtcg tttggtatgg cttcattcag ctccggttcc caacgatcaa 17820
ggcgagttac atgatccccc atgttgtgaa aaaaagcggt tagctccttc ggtcctccga 17880
tcgttgtcag aagtaagttg gccgcagtgt tatcactcat gcttatggca gcactgcata 17940
attctcttac tgtcatgcca tccgtaagat gcttttctgt gactggtgag tactcaacca 18000
agtcattctg agaatagtgt atgcggcgac cgagttgctc ttgcccggcg tcaatacggg 18060
ataataccgc gccacatagc agaactttaa aagtgctcat cattggaaaa cgttcttcgg 18120
ggcgaaaact ctcaaggatc ttaccgctgt tgagatccag ttcgatgtaa cccactcgtg 18180
cacccaactg atcttcagca tcttttactt tcaccagcgt ttctgggtga gcaaaaacag 18240
gaaggcaaaa tgccgcaaaa aagggaataa gggcgacacg gaaatgttga atactcatac 18300
tcttcctttt tcaatattat tgaagcattt atcagggtta ttgtctcatg agcggataca 18360
tatttgaatg tatttagaaa aataaacaaa taggggttcc gcgcacattt ccccgaaaag 18420
tgccacctga aattgtaaac gttaatattt tgttaaaatt cgcgttaaat ttttgttaaa 18480
tcagctcatt ttttaaccaa taggccgaaa tcggcaaaat cccttataaa tcaaaagaat 18540
agaccgagat agggttgagt gttgttccag tttggaacaa gagtccacta ttaaagaacg 18600
tggactccaa cgtcaaaggg cgaaaaaccg tctatcaggg cgatggccca ctacgtgaac 18660
catcacccta atcaagtttt ttggggtcga ggtgccgtaa agcactaaat cggaacccta 18720
[0009] aagggagccc ccgatttaga gcttgacggg gaaagccggc gaacgtggcg agaaaggaag 18780
ggaagaaagc gaaaggagcg ggcgctaggg cgctggcaag tgtagcggtc acgctgcgcg 18840
taaccaccac acccgccgcg cttaatgcgc cgctacaggg cgcgtcccat tcgccattca 18900
ggctgcgcaa ctgttgggaa gggcgatcgg tgcgggcctc ttcgctatta cgccagccac 18960
cgcggtg 18967

<210> 2
<211> 1797
<212> DNA
<213> Chikungunya virus

<220>
<221> misc_feature
<223> soluble glycoprotein E2 strain 05.115

<220>
<221> CDS
<222> (1)..(1797)

<400> 2
ata ccc gtg cac atg aag tcc gac gct tcg aag ttc acc cat gag aaa 48
Ile Pro Val His Met Lys Ser Asp Ala Ser Lys Phe Thr His Glu Lys
1 5 10 15

ccg gag ggg tac tac aac tgg cac cac gga gca gta cag tac tca gga 96
Pro Glu Gly Tyr Tyr Asn Trp His His Gly Ala Val Gln Tyr Ser Gly
20 25 30

ggc cgg ttc acc atc cct aca ggt gct ggc aaa cca ggg gac agc ggc 144
Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly Lys Pro Gly Asp Ser Gly
35 40 45

aga ccg atc ttc gac aac aag gga cgc gtg gtg gcc ata gtc tta gga 192

```
Arg Pro Ile Phe Asp Asn Lys Gly Arg Val Val Ala Ile Val Leu Gly
     50                  55                  60

gga gct aat gaa gga gcc cgt aca gcc ctc tcg gtg gtg acc tgg aat       240
Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu Ser Val Val Thr Trp Asn
 65                  70                  75                  80

aaa gac att gtc act aaa atc acc ccc gag ggg gcc gaa gag tgg agt       288
Lys Asp Ile Val Thr Lys Ile Thr Pro Glu Gly Ala Glu Glu Trp Ser
                 85                  90                  95

ctt gcc atc cca gtt atg tgc ctg ttg gca aac acc acg ttc ccc tgc       336
Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr Thr Phe Pro Cys
            100                 105                 110

tcc cag ccc cct tgc acg ccc tgc tgc tac gaa aag gaa ccg gag gaa       384
Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr Glu Lys Glu Pro Glu Glu
        115                 120                 125

acc cta cgc atg ctt gag gac aac gtc atg aga cct ggg tac tat cag       432
Thr Leu Arg Met Leu Glu Asp Asn Val Met Arg Pro Gly Tyr Tyr Gln
    130                 135                 140

ctg cta caa gca tcc tta aca tgt tct ccc cac cgc cag cga cgc agc       480
Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro His Arg Gln Arg Arg Ser
145                 150                 155                 160

acc aag gac aac ttc aat gtc tat aaa gcc aca aga cca tac tta gct       528
Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala
                165                 170                 175

cac tgt ccc gac tgt gga gaa ggg cac tcg tgc cat agt ccc gta gca       576
His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala
            180                 185                 190

cta gaa cgc atc aga aat gaa gcg aca gac ggg acg ctg aaa atc cag       624
Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln
        195                 200                 205
```

[0010]

```
gtc tcc ttg caa atc gga ata aag acg gat gac agc cac gat tgg acc       672
Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr
    210                 215                 220

aag ctg cgt tat atg gac aac cac atg cca gca gac gca gag agg gcg       720
Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala
225                 230                 235                 240

ggg cta ttt gta aga aca tca gca ccg tgt acg att act gga aca atg       768
Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met
                245                 250                 255

gga cac ttc atc ctg gcc cga tgt cca aaa ggg gaa act ctg acg gtg       816
Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val
            260                 265                 270

gga ttc act gac agt agg aag att agt cac tca tgt acg cac cca ttt       864
Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe
        275                 280                 285

cac cac gac cct cct gtg ata ggt cgg gaa aaa ttc cat tcc cga ccg       912
His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro
    290                 295                 300

cag cac ggt aaa gag ata cct tgc agc acg tac gtg cag agc acc gcc       960
Gln His Gly Lys Glu Ile Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala
305                 310                 315                 320

gca act acc gag gag ata gag gta cac atg ccc cca gac acc cct gat      1008
Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp
                325                 330                 335

cgc aca tta atg tca caa cag tcc ggc aac gta aag atc aca gtc aat      1056
Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn
            340                 345                 350

ggc cag acg gtg cgg tac aag tgt aat tgc ggt ggc tca aat gaa gga      1104
Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly
```

```
        355                 360                 365

cta aca act aca gac aaa gtg att aat aac tgc aag gtt gat caa tgt        1152
Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys
    370                 375                 380

cat gcc gcg gtc acc aat cac aaa aag tgg cag tat aac tcc cct ctg        1200
His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu
385                 390                 395                 400

gtc ccg cgt aat gct gaa ctt ggg gac cga aaa gga aaa att cac atc        1248
Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile
                405                 410                 415

ccg ttt ccg ctg gca aat gta aca tgc agg gtg cct aaa gca agg aac        1296
Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn
            420                 425                 430

ccc acc gtg acg tac ggg aaa aac caa gtc atc atg cta ctg tat cct        1344
Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro
        435                 440                 445

gac cac cca aca ctc ctg tcc tac cgg aat atg gga gaa gaa cca aac        1392
Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn
    450                 455                 460

tat caa gaa gag tgg gtg atg cat aag aag gaa gtc gtg cta acc gtg        1440
Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val
465                 470                 475                 480

ccg act gaa ggg ctc gag gtc acg tgg ggc aac aac gag ccg tat aag        1488
Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys
                485                 490                 495

tat tgg ccg cag tta tct aca aac ggt aca gcc cat ggc cac ccg cat        1536
Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His
            500                 505                 510
```

[0011]

```
gag ata att ctg tat tat tat gag ctg tac ccc act atg act gta gta        1584
Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val Val
        515                 520                 525

gtt gtg tca gtg gcc acg ttc ata ctc ctg tcg atg gtg ggt atg gca        1632
Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met Ala
    530                 535                 540

gcg ggg atg tgc atg tgt gca cga cgc aga tgc atc aca ccg tat gaa        1680
Ala Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr Glu
545                 550                 555                 560

ctg aca cca gga gct acc gtc cct ttc ctg ctt agc cta ata tgc tgc        1728
Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys Cys
                565                 570                 575

atc aga aca gct aaa gcg gcc aca tac caa gag gct gcg ata tac ctg        1776
Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr Leu
            580                 585                 590

tgg aac gag cag caa cct tta                                            1797
Trp Asn Glu Gln Gln Pro Leu
        595


<210>  3
<211>  599
<212>  PRT
<213>  Chikungunya virus

<400>  3

Ile Pro Val His Met Lys Ser Asp Ala Ser Lys Phe Thr His Glu Lys
1               5                   10                  15


Pro Glu Gly Tyr Tyr Asn Trp His His Gly Ala Val Gln Tyr Ser Gly
            20                  25                  30
```

```
Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly Lys Pro Gly Asp Ser Gly
        35                  40                  45

Arg Pro Ile Phe Asp Asn Lys Gly Arg Val Val Ala Ile Val Leu Gly
    50                  55                  60

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

Lys Asp Ile Val Thr Lys Ile Thr Pro Glu Gly Ala Glu Glu Trp Ser
                85                  90                  95

Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr Thr Phe Pro Cys
            100                 105                 110

Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr Glu Lys Glu Pro Glu Glu
        115                 120                 125

Thr Leu Arg Met Leu Glu Asp Asn Val Met Arg Pro Gly Tyr Tyr Gln
    130                 135                 140

Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro His Arg Gln Arg Arg Ser
145                 150                 155                 160

Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala
                165                 170                 175
```

[0012]

```
His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala
            180                 185                 190

Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln
        195                 200                 205

Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr
    210                 215                 220

Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala
225                 230                 235                 240

Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met
                245                 250                 255

Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val
            260                 265                 270

Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe
        275                 280                 285

His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro
    290                 295                 300

Gln His Gly Lys Glu Ile Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala
305                 310                 315                 320

Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp
                325                 330                 335
```

```
Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn
            340                 345                 350

Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly
        355                 360                 365

Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys
    370                 375                 380

His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu
385                 390                 395                 400

Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile
                405                 410                 415

Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn
            420                 425                 430

Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro
        435                 440                 445

Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn
    450                 455                 460

Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val
465                 470                 475                 480

Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys
                485                 490                 495
```

[0013]

```
Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His
            500                 505                 510

Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val Val
        515                 520                 525

Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met Ala
    530                 535                 540

Ala Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr Glu
545                 550                 555                 560

Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys Cys
                565                 570                 575

Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr Leu
            580                 585                 590

Trp Asn Glu Gln Gln Pro Leu
        595
```

<210>　4
<211>　1797
<212>　DNA
<213>　Chikungunya virus

<220>
<221>　misc_feature
<223>　soluble glycoprotein E2 strain 06.21

```
        <220>
        <221>  CDS
        <222>  (1)..(1797)

        <400>  4
        ata ccc gtg cac atg aag tcc gac gct tcg aag ttc acc cat gag aaa       48
        Ile Pro Val His Met Lys Ser Asp Ala Ser Lys Phe Thr His Glu Lys
        1               5                   10                  15

        ccg gag ggg tac tac aac tgg cac cac gga gca gta cag tac tca gga       96
        Pro Glu Gly Tyr Tyr Asn Trp His His Gly Ala Val Gln Tyr Ser Gly
                    20                  25                  30

        ggc cgg ttc acc atc cct aca ggt gct ggc aaa cca ggg gac agc ggc      144
        Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly Lys Pro Gly Asp Ser Gly
                35                  40                  45

        aga ccg atc ttc gac aac aag gga cgc gtg gtg gcc ata gtc tta gga      192
        Arg Pro Ile Phe Asp Asn Lys Gly Arg Val Val Ala Ile Val Leu Gly
            50                  55                  60

        gga gct aat gaa gga gcc cgt aca gcc ctc tcg gtg gtg acc tgg aat      240
        Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu Ser Val Val Thr Trp Asn
        65                  70                  75                  80

        aaa gac att gtc act aaa atc acc ccc gag ggg gcc gaa gag tgg agt      288
        Lys Asp Ile Val Thr Lys Ile Thr Pro Glu Gly Ala Glu Glu Trp Ser
                        85                  90                  95

        ctt gcc atc cca gtt atg tgc ctg ttg gca aac acc acg ttc ccc tgc      336
        Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr Thr Phe Pro Cys
                    100                 105                 110

        tcc cag ccc cct tgc acg ccc tgc tgc tac gaa aag gaa ccg gag gaa      384
        Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr Glu Lys Glu Pro Glu Glu
                115                 120                 125

[0014]  acc cta cgc atg ctt gag gac aac gtc atg aga cct ggg tac tat cag      432
        Thr Leu Arg Met Leu Glu Asp Asn Val Met Arg Pro Gly Tyr Tyr Gln
            130                 135                 140

        ctg cta caa gca tcc tta aca tgt tct ccc cac cgc cag cga cgc agc      480
        Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro His Arg Gln Arg Arg Ser
        145                 150                 155                 160

        acc aag gac aac ttc aat gtc tat aaa gcc aca aga cca tac tta gct      528
        Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala
                        165                 170                 175

        cac tgt ccc gac tgt gga gaa ggg cac tcg tgc cat agt ccc gta gca      576
        His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala
                    180                 185                 190

        cta gaa cgc atc aga aat gaa gcg aca gac ggg acg ctg aaa atc cag      624
        Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln
                195                 200                 205

        gtc tcc ttg caa atc gga ata aag acg gat gac agc cac gat tgg acc      672
        Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr
            210                 215                 220

        aag ctg cgt tat atg gac aac cac atg cca gca gac gca gag agg gcg      720
        Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala
        225                 230                 235                 240

        ggg cta ttt gta aga aca tca gca ccg tgt acg att act gga aca atg      768
        Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met
                        245                 250                 255

        gga cac ttc atc ctg gcc cga tgt cca aaa ggg gaa act ctg acg gtg      816
        Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val
                    260                 265                 270

        gga ttc act gac agt agg aag att agt cac tca tgt acg cac cca ttt      864
        Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe
                275                 280                 285
```

```
       cac cac gac cct cct gtg ata ggt cgg gaa aaa ttc cat tcc cga ccg         912
       His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro
           290                 295                 300

       cag cac ggt aaa gag cta cct tgc agc acg tac gtg cag agc acc gcc         960
       Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala
       305                 310                 315                 320

       gca act acc gag gag ata gag gta cac atg ccc cca gac acc cct gat        1008
       Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp
                       325                 330                 335

       cgc aca tta atg tca caa cag tcc ggc aac gta aag atc aca gtc aat        1056
       Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn
                   340                 345                 350

       ggc cag acg gtg cgg tac aag tgt aat tgc ggt ggc tca aat gaa gga        1104
       Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly
               355                 360                 365

       cta aca act aca gac aaa gtg att aat aac tgc aag gtt gat caa tgt        1152
       Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys
           370                 375                 380

       cat gcc gcg gtc acc aat cac aaa aag tgg cag tat aac tcc cct ctg        1200
       His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu
       385                 390                 395                 400

       gtc ccg cgt aat gct gaa ctt ggg gac cga aaa gga aaa att cac atc        1248
       Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile
                       405                 410                 415

       ccg ttt ccg ctg gca aat gta aca tgc agg gtg cct aaa gca agg aac        1296
       Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn
                   420                 425                 430

       ccc acc gtg acg tac ggg aaa aac caa gtc atc atg cta ctg tat cct        1344
[0015] Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro
               435                 440                 445

       gac cac cca aca ctc ctg tcc tac cgg aat atg gga gaa gaa cca aac        1392
       Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn
           450                 455                 460

       tat caa gaa gag tgg gtg atg cat aag aag gaa gtc gtg cta acc gtg        1440
       Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val
       465                 470                 475                 480

       ccg act gaa ggg ctc gag gtc acg tgg ggc aac aac gag ccg tat aag        1488
       Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys
                       485                 490                 495

       tat tgg ccg cag tta tct aca aac ggt aca gcc cat ggc cac ccg cat        1536
       Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His
                   500                 505                 510

       gag ata att ctg tat tat tat gag ctg tac ccc act atg act gta gta        1584
       Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val Val
               515                 520                 525

       gtt gtg tca gtg gcc acg ttc ata ctc ctg tcg atg gtg ggt atg gca        1632
       Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met Ala
           530                 535                 540

       gcg ggg atg tgc atg cgt gca cga cgc aga tgc atc aca ccg tat gaa        1680
       Ala Gly Met Cys Met Arg Ala Arg Arg Arg Cys Ile Thr Pro Tyr Glu
       545                 550                 555                 560

       ctg aca cca gga gct acc gtc cct ttc ctg ctt agc cta ata tgc tgc        1728
       Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys Cys
                       565                 570                 575

       atc aga aca gct aaa gcg gcc aca tac caa gag gct gcg ata tac ctg        1776
       Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr Leu
                   580                 585                 590
```

```
tgg aac gag cag caa cct tta                                            1797
Trp Asn Glu Gln Gln Pro Leu
        595


<210>  5
<211>  599
<212>  PRT
<213>  Chikungunya virus

<400>  5

Ile Pro Val His Met Lys Ser Asp Ala Ser Lys Phe Thr His Glu Lys
1               5                   10                  15


Pro Glu Gly Tyr Tyr Asn Trp His His Gly Ala Val Gln Tyr Ser Gly
            20                  25                  30


Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly Lys Pro Gly Asp Ser Gly
        35                  40                  45


Arg Pro Ile Phe Asp Asn Lys Gly Arg Val Val Ala Ile Val Leu Gly
    50                  55                  60


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


Lys Asp Ile Val Thr Lys Ile Thr Pro Glu Gly Ala Glu Glu Trp Ser
                85                  90                  95


Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr Thr Phe Pro Cys
            100                 105                 110
```

[0016]

```
Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr Glu Lys Glu Pro Glu Glu
        115                 120                 125


Thr Leu Arg Met Leu Glu Asp Asn Val Met Arg Pro Gly Tyr Tyr Gln
    130                 135                 140


Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro His Arg Gln Arg Arg Ser
145                 150                 155                 160


Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala
                165                 170                 175


His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala
            180                 185                 190


Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln
        195                 200                 205


Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr
    210                 215                 220


Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala
225                 230                 235                 240


Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met
                245                 250                 255


Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val
```

260 265 270

Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe
275 280 285

His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro
290 295 300

Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala
305 310 315 320

Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp
325 330 335

Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn
340 345 350

Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly
355 360 365

Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys
370 375 380

His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu
385 390 395 400

Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile
405 410 415

[0017] Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn
420 425 430

Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro
435 440 445

Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn
450 455 460

Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val
465 470 475 480

Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys
485 490 495

Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His
500 505 510

Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val Val
515 520 525

Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met Ala
530 535 540

Ala Gly Met Cys Met Arg Ala Arg Arg Arg Cys Ile Thr Pro Tyr Glu
545 550 555 560

Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys Cys
565 570 575

Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr Leu
580 585 590

Trp Asn Glu Gln Gln Pro Leu
595

<210> 6
<211> 1797
<212> DNA
<213> Chikungunya virus

<220>
<221> misc_feature
<223> soluble glycoprotein E2 strain 06.27

<220>
<221> CDS
<222> (1)..(1797)

<400> 6
ata ccc gtg cac atg aag tcc gac gct tcg aag ttc acc cat gag aaa 48
Ile Pro Val His Met Lys Ser Asp Ala Ser Lys Phe Thr His Glu Lys
1 5 10 15

ccg gag ggg tac tac aac tgg cac cac gga gca gta cag tac tca gga 96
Pro Glu Gly Tyr Tyr Asn Trp His His Gly Ala Val Gln Tyr Ser Gly
20 25 30

ggc cgg ttc acc atc cct aca ggt gct ggc aaa cca ggg gac agc ggc 144
Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly Lys Pro Gly Asp Ser Gly
35 40 45

aga ccg atc ttc gac aac aag gga cgc gtg gtg gcc ata gtc tta gga 192
[0018] Arg Pro Ile Phe Asp Asn Lys Gly Arg Val Val Ala Ile Val Leu Gly
50 55 60

gga gct aat gaa gga gcc cgt aca gcc ctc tcg gtg gtg acc tgg aat 240
Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu Ser Val Val Thr Trp Asn
65 70 75 80

aaa gac att gtc act aaa atc acc ccc gag ggg gcc gaa gag tgg agt 288
Lys Asp Ile Val Thr Lys Ile Thr Pro Glu Gly Ala Glu Glu Trp Ser
85 90 95

ctt gcc atc cca gtt atg tgc ctg ttg gca aac acc acg ttc ccc tgc 336
Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr Thr Phe Pro Cys
100 105 110

tcc cag ccc cct tgc acg ccc tgc tgc tac gaa aag gaa ccg gag gaa 384
Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr Glu Lys Glu Pro Glu Glu
115 120 125

acc cta cgc atg ctt gag gac aac gtc atg aga cct ggg tac tat cag 432
Thr Leu Arg Met Leu Glu Asp Asn Val Met Arg Pro Gly Tyr Tyr Gln
130 135 140

ctg cta caa gca tcc tta aca tgt tct ccc cac cgc cag cga cgc agc 480
Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro His Arg Gln Arg Arg Ser
145 150 155 160

acc aag gac aac ttc aat gtc tat aaa gcc aca aga cca tac tta gct 528
Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala
165 170 175

cac tgt ccc gac tgt gga gaa ggg cac tcg tgc cat agt ccc gta gca 576
His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala
180 185 190

cta gaa cgc atc aga aat gaa gcg aca gac ggg acg ctg aaa atc cag 624
Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln
195 200 205

```
       gtc tcc ttg caa atc gga ata aag acg gat gac agc cac gat tgg acc       672
       Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr
           210                 215                 220

       aag ctg cgt tat atg gac aac cac atg cca gca gac gca gag agg gcg       720
       Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala
       225                 230                 235                 240

       ggg cta ttt gta aga aca tca gca ccg tgt acg att act gga aca atg       768
       Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met
                       245                 250                 255

       gga cac ttc atc ctg gcc cga tgt cca aaa ggg gaa act ctg acg gtg       816
       Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val
                   260                 265                 270

       gga ttc act gac agt agg aag att agt cac tca tgt acg cac cca ttt       864
       Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe
               275                 280                 285

       cac cac gac cct cct gtg ata ggt cgg gaa aaa ttc cat tcc cga ccg       912
       His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro
           290                 295                 300

       cag cac ggt aaa gag cta cct tgc agc acg tac gtg cag agc acc gcc       960
       Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala
       305                 310                 315                 320

       gca act acc gag gag ata gag gta cac atg ccc cca gac acc cct gat      1008
       Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp
                       325                 330                 335

       cgc aca tta atg tca caa cag tcc ggc aac gta aag atc aca gtc aat      1056
       Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn
                   340                 345                 350

       ggc cag acg gtg cgg tac aag tgt aat tgc ggt ggc tca aat gaa gga      1104
[0019] Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly
               355                 360                 365

       cta aca act aca gac aaa gtg att aat aac tgc aag gtt gat caa tgt      1152
       Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys
           370                 375                 380

       cat gcc gcg gtc acc aat cac aaa aag tgg cag tat aac tcc cct ctg      1200
       His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu
       385                 390                 395                 400

       gtc ccg cgt aat gct gaa ctt ggg gac cga aaa gga aaa att cac atc      1248
       Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile
                       405                 410                 415

       ccg ttt ccg ctg gca aat gta aca tgc agg gtg cct aaa gca agg aac      1296
       Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn
                   420                 425                 430

       ccc acc gtg acg tac ggg aaa aac caa gtc atc atg cta ctg tat cct      1344
       Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro
               435                 440                 445

       gac cac cca aca ctc ctg tcc tac cgg aat atg gga gaa gaa cca aac      1392
       Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn
           450                 455                 460

       tat caa gaa gag tgg gtg atg cat aag aag gaa gtc gtg cta acc gtg      1440
       Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val
       465                 470                 475                 480

       ccg act gaa ggg ctc gag gtc acg tgg ggc aac aac gag ccg tat aag      1488
       Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys
                       485                 490                 495

       tat tgg ccg cag tta tct aca aac ggt aca gcc cat ggc cac ccg cat      1536
       Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His
                   500                 505                 510

       gag ata att ctg tat tat tat gag ctg tac ccc act atg act gta gta      1584
```

```
Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val Val
        515                 520                 525

gtt gtg tca gtg gcc acg ttc ata ctc ctg tcg atg gtg ggt atg gca            1632
Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met Ala
    530                 535                 540

gcg ggg atg tgc atg tgt gca cga cgc aga tgc atc aca ccg tat gaa            1680
Ala Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr Glu
545                 550                 555                 560

ctg aca cca gga gct acc gtc cct ttc ctg ctt agc cta ata tgc tgc            1728
Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys Cys
                565                 570                 575

atc aga aca gct aaa gcg gcc aca tac caa gag gct gcg ata tac ctg            1776
Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr Leu
            580                 585                 590

tgg aac gag cag caa cct tta                                                1797
Trp Asn Glu Gln Gln Pro Leu
        595


<210>  7
<211>  599
<212>  PRT
<213>  Chikungunya virus

<400>  7

Ile Pro Val His Met Lys Ser Asp Ala Ser Lys Phe Thr His Glu Lys
1               5                   10                  15


Pro Glu Gly Tyr Tyr Asn Trp His His Gly Ala Val Gln Tyr Ser Gly
            20                  25                  30

[0020]
Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly Lys Pro Gly Asp Ser Gly
        35                  40                  45


Arg Pro Ile Phe Asp Asn Lys Gly Arg Val Val Ala Ile Val Leu Gly
    50                  55                  60


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


Lys Asp Ile Val Thr Lys Ile Thr Pro Glu Gly Ala Glu Glu Trp Ser
                85                  90                  95


Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr Thr Phe Pro Cys
            100                 105                 110


Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr Glu Lys Glu Pro Glu Glu
        115                 120                 125


Thr Leu Arg Met Leu Glu Asp Asn Val Met Arg Pro Gly Tyr Tyr Gln
    130                 135                 140


Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro His Arg Gln Arg Arg Ser
145                 150                 155                 160


Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala
                165                 170                 175


His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala
            180                 185                 190
```

```
       Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln
               195                 200                 205

       Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr
           210                 215                 220

       Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala
       225                 230                 235                 240

       Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met
                       245                 250                 255

       Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val
                   260                 265                 270

       Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe
               275                 280                 285

       His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro
           290                 295                 300

       Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala
       305                 310                 315                 320

       Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp
                       325                 330                 335

[0021] Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn
                   340                 345                 350

       Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly
               355                 360                 365

       Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys
           370                 375                 380

       His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu
       385                 390                 395                 400

       Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile
                       405                 410                 415

       Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn
                   420                 425                 430

       Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro
               435                 440                 445

       Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn
           450                 455                 460

       Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val
       465                 470                 475                 480

       Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys
                       485                 490                 495
```

```
Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His
            500                 505                 510

Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val Val
        515                 520                 525

Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met Ala
    530                 535                 540

Ala Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr Glu
545                 550                 555                 560

Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys Cys
                565                 570                 575

Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr Leu
            580                 585                 590

Trp Asn Glu Gln Gln Pro Leu
        595


<210>  8
<211>  1797
<212>  DNA
<213>  Chikungunya virus


<220>
<221>  misc_feature
<223>  soluble glycoprotein E2 strain 06.49

[0022]
<220>
<221>  CDS
<222>  (1)..(1797)

<400>  8
ata ccc gtg cac atg aag tcc gac gct tcg aag ttc acc cat gag aaa          48
Ile Pro Val His Met Lys Ser Asp Ala Ser Lys Phe Thr His Glu Lys
1               5                   10                  15

ccg gag ggg tac tac aac tgg cac cac gga gca gta cag tac tca gga          96
Pro Glu Gly Tyr Tyr Asn Trp His His Gly Ala Val Gln Tyr Ser Gly
            20                  25                  30

ggc cgg ttc acc atc cct aca ggt gct ggc aaa cca ggg gac agc ggc         144
Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly Lys Pro Gly Asp Ser Gly
        35                  40                  45

aga ccg atc ttc gac aac aag gga cgc gtg gtg gcc ata gtc tta gga         192
Arg Pro Ile Phe Asp Asn Lys Gly Arg Val Val Ala Ile Val Leu Gly
    50                  55                  60

gga gct aat gaa gga gcc cgt aca gcc ctc tcg gtg gtg acc tgg aat         240
Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu Ser Val Val Thr Trp Asn
65                  70                  75                  80

aaa gac att gtc act aaa atc acc ccc gag ggg gcc gaa gag tgg agt         288
Lys Asp Ile Val Thr Lys Ile Thr Pro Glu Gly Ala Glu Glu Trp Ser
                85                  90                  95

ctt gcc atc cca gtt atg tgc ctg ttg gca aac acc acg ttc ccc tgc         336
Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr Thr Phe Pro Cys
            100                 105                 110

tcc cag ccc cct tgc acg ccc tgc tgc tac gaa aag gaa ccg gag gaa         384
Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr Glu Lys Glu Pro Glu Glu
        115                 120                 125

acc cta cgc atg ctt gag gac aac gtc atg aga cct ggg tac tat cag         432
```

[0023]

```
Thr Leu Arg Met Leu Glu Asp Asn Val Met Arg Pro Gly Tyr Tyr Gln
    130                 135                 140

ctg cta caa gca tcc tta aca tgt tct ccc cac cgc cag cga cgc agc          480
Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro His Arg Gln Arg Arg Ser
145                 150                 155                 160

acc aag gac aac ttc aat gtc tat aaa gcc aca aga cca tac tta gct          528
Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala
                165                 170                 175

cac tgt ccc gac tgt gga gaa ggg cac tcg tgc cat agt ccc gta gca          576
His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala
            180                 185                 190

cta gaa cgc atc aga aat gaa gcg aca gac ggg acg ctg aaa atc cag          624
Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln
        195                 200                 205

gtc tcc ttg caa atc gga ata aag acg gat gac agc cac gat tgg acc          672
Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr
    210                 215                 220

aag ctg cgt tat atg gac aac cac atg cca gca gac gca gag agg gcg          720
Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala
225                 230                 235                 240

ggg cta ttt gta aga aca tca gca ccg tgt acg att act gga aca atg          768
Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met
                245                 250                 255

gga cac ttc atc ctg gcc cga tgt cca aaa ggg gaa act ctg acg gtg          816
Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val
            260                 265                 270

gga ttc act gac agt agg aag att agt cac tca tgt acg cac cca ttt          864
Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe
        275                 280                 285

cac cac gac cct cct gtg ata ggt cgg gaa aaa ttc cat tcc cga ccg          912
His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro
    290                 295                 300

cag cac ggt aaa gag cta cct tgc agc acg tac gtg cag agc acc gcc          960
Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala
305                 310                 315                 320

gca act acc gag gag ata gag gta cac atg ccc cca gac acc cct gat         1008
Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp
                325                 330                 335

cgc aca tta atg tca caa cag tcc ggc aac gta aag atc aca gtc aat         1056
Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn
            340                 345                 350

ggc cag acg gtg cgg tac aag tgt aat tgc ggt ggc tca aat gaa gga         1104
Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly
        355                 360                 365

cta aca act aca gac aaa gtg att aat aac tgc aag gtt gat caa tgt         1152
Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys
    370                 375                 380

cat gcc gcg gtc acc aat cac aaa aag tgg cag tat aac tcc cct ctg         1200
His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu
385                 390                 395                 400

gtc ccg cgt aat gct gaa ctt ggg gac cga aaa gga aaa att cac atc         1248
Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile
                405                 410                 415

ccg ttt ccg ctg gca aat gta aca tgc agg gtg cct aaa gca agg aac         1296
Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn
            420                 425                 430

ccc acc gtg acg tac ggg aaa aac caa gtc atc atg cta ctg tat cct         1344
Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro
```

```
               435                 440                 445

       gac cac cca aca ctc ctg tcc tac cgg aat atg gga gaa gaa cca aac        1392
       Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn
           450                 455                 460

       tat caa gaa gag tgg gtg atg cat aag aag gaa gtc gtg cta acc gtg        1440
       Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val
       465                 470                 475                 480

       ccg act gaa ggg ctc gag gtc acg tgg ggc aac aac gag ccg tat aag        1488
       Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys
                       485                 490                 495

       tat tgg ccg cag tta tct aca aac ggt aca gcc cat ggc cac ccg cac        1536
       Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His
                   500                 505                 510

       gag ata att ctg tat tat tat gag ctg tac ccc act atg act gta gta        1584
       Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val Val
               515                 520                 525

       gtt gtg tca gtg gcc acg ttc ata ctc ctg tcg atg gtg ggt atg gca        1632
       Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met Ala
           530                 535                 540

       gcg ggg atg tgc atg tgt gca cga cgc aga tgc atc aca ccg tat gaa        1680
       Ala Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr Glu
       545                 550                 555                 560

       ctg aca cca gga gct acc gtc cct ttc ctg ctt agc cta ata tgc tgc        1728
       Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys Cys
                       565                 570                 575

       atc aga aca gct aaa gcg gcc aca tac caa gag gct gcg ata tac ctg        1776
       Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr Leu
                   580                 585                 590

[0024] tgg aac gag cag caa cct tta                                            1797
       Trp Asn Glu Gln Gln Pro Leu
               595


       <210>  9
       <211>  599
       <212>  PRT
       <213>  Chikungunya virus

       <400>  9

       Ile Pro Val His Met Lys Ser Asp Ala Ser Lys Phe Thr His Glu Lys
       1               5                   10                  15


       Pro Glu Gly Tyr Tyr Asn Trp His His Gly Ala Val Gln Tyr Ser Gly
                   20                  25                  30


       Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly Lys Pro Gly Asp Ser Gly
               35                  40                  45


       Arg Pro Ile Phe Asp Asn Lys Gly Arg Val Val Ala Ile Val Leu Gly
           50                  55                  60


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


       Lys Asp Ile Val Thr Lys Ile Thr Pro Glu Gly Ala Glu Glu Trp Ser
                       85                  90                  95


       Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr Thr Phe Pro Cys
                   100                 105                 110
```

Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr Glu Lys Glu Pro Glu Glu
115 120 125

Thr Leu Arg Met Leu Glu Asp Asn Val Met Arg Pro Gly Tyr Tyr Gln
130 135 140

Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro His Arg Gln Arg Arg Ser
145 150 155 160

Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala
165 170 175

His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala
180 185 190

Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln
195 200 205

Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr
210 215 220

Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala
225 230 235 240

Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met
245 250 255

[0025] Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val
260 265 270

Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe
275 280 285

His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro
290 295 300

Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala
305 310 315 320

Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp
325 330 335

Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn
340 345 350

Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly
355 360 365

Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys
370 375 380

His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu
385 390 395 400

Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile
405 410 415

Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn
420 425 430

Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro
435 440 445

Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn
450 455 460

Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val
465 470 475 480

Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys
485 490 495

Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His
500 505 510

Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val Val
515 520 525

Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met Ala
530 535 540

Ala Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr Glu
545 550 555 560

Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys Cys
565 570 575

[0026]

Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr Leu
580 585 590

Trp Asn Glu Gln Gln Pro Leu
595

<210> 10
<211> 1092
<212> DNA
<213> Chikungunya virus

<220>
<221> misc_feature
<223> ectodomain glycoprotein E2 strain 05.115

<220>
<221> CDS
<222> (1)..(1092)

<400> 10
gac aac ttc aat gtc tat aaa gcc aca aga cca tac tta gct cac tgt 48
Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala His Cys
1 5 10 15

ccc gac tgt gga gaa ggg cac tcg tgc cat agt ccc gta gca cta gaa 96
Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala Leu Glu
20 25 30

cgc atc aga aat gaa gcg aca gac ggg acg ctg aaa atc cag gtc tcc 144
Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln Val Ser
35 40 45

ttg caa atc gga ata aag acg gat gac agc cac gat tgg acc aag ctg 192
Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr Lys Leu

```
            50                  55                  60

        cgt tat atg gac aac cac atg cca gca gac gca gag agg gcg ggg cta       240
        Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala Gly Leu
        65                  70                  75                  80

        ttt gta aga aca tca gca ccg tgt acg att act gga aca atg gga cac       288
        Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met Gly His
                        85                  90                  95

        ttc atc ctg gcc cga tgt cca aaa ggg gaa act ctg acg gtg gga ttc       336
        Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val Gly Phe
                    100                 105                 110

        act gac agt agg aag att agt cac tca tgt acg cac cca ttt cac cac       384
        Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe His His
                115                 120                 125

        gac cct cct gtg ata ggt cgg gaa aaa ttc cat tcc cga ccg cag cac       432
        Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro Gln His
            130                 135                 140

        ggt aaa gag ata cct tgc agc acg tac gtg cag agc acc gcc gca act       480
        Gly Lys Glu Ile Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala Ala Thr
        145                 150                 155                 160

        acc gag gag ata gag gta cac atg ccc cca gac acc cct gat cgc aca       528
        Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp Arg Thr
                        165                 170                 175

        tta atg tca caa cag tcc ggc aac gta aag atc aca gtc aat ggc cag       576
        Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn Gly Gln
                    180                 185                 190

        acg gtg cgg tac aag tgt aat tgc ggt ggc tca aat gaa gga cta aca       624
        Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly Leu Thr
                195                 200                 205

[0027]  act aca gac aaa gtg att aat aac tgc aag gtt gat caa tgt cat gcc       672
        Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys His Ala
            210                 215                 220

        gcg gtc acc aat cac aaa aag tgg cag tat aac tcc cct ctg gtc ccg       720
        Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu Val Pro
        225                 230                 235                 240

        cgt aat gct gaa ctt ggg gac cga aaa gga aaa att cac atc ccg ttt       768
        Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile Pro Phe
                        245                 250                 255

        ccg ctg gca aat gta aca tgc agg gtg cct aaa gca agg aac ccc acc       816
        Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn Pro Thr
                    260                 265                 270

        gtg acg tac ggg aaa aac caa gtc atc atg cta ctg tat cct gac cac       864
        Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro Asp His
                275                 280                 285

        cca aca ctc ctg tcc tac cgg aat atg gga gaa gaa cca aac tat caa       912
        Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn Tyr Gln
            290                 295                 300

        gaa gag tgg gtg atg cat aag aag gaa gtc gtg cta acc gtg ccg act       960
        Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val Pro Thr
        305                 310                 315                 320

        gaa ggg ctc gag gtc acg tgg ggc aac aac gag ccg tat aag tat tgg      1008
        Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys Tyr Trp
                        325                 330                 335

        ccg cag tta tct aca aac ggt aca gcc cat ggc cac ccg cat gag ata      1056
        Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His Glu Ile
                    340                 345                 350

        att ctg tat tat tat gag ctg tac ccc act atg act                      1092
        Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr
                355                 360
```

[0028]

```
<210>  11
<211>  364
<212>  PRT
<213>  Chikungunya virus

<400>  11

Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala His Cys
1               5                   10                  15

Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala Leu Glu
            20                  25                  30

Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln Val Ser
        35                  40                  45

Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr Lys Leu
    50                  55                  60

Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala Gly Leu
65                  70                  75                  80

Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met Gly His
                85                  90                  95

Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val Gly Phe
            100                 105                 110

Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe His His
        115                 120                 125

Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro Gln His
    130                 135                 140

Gly Lys Glu Ile Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala Ala Thr
145                 150                 155                 160

Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp Arg Thr
                165                 170                 175

Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn Gly Gln
            180                 185                 190

Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly Leu Thr
        195                 200                 205

Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys His Ala
    210                 215                 220

Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu Val Pro
225                 230                 235                 240

Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile Pro Phe
                245                 250                 255

Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn Pro Thr
            260                 265                 270
```

```
Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro Asp His
        275                 280                 285

Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn Tyr Gln
    290                 295                 300

Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val Pro Thr
305                 310                 315                 320

Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys Tyr Trp
                325                 330                 335

Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His Glu Ile
            340                 345                 350

Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr
        355                 360


<210>  12
<211>  1092
<212>  DNA
<213>  Chikungunya virus


<220>
<221>  misc_feature
<223>  ectodomain glycoprotein E2 strain 06.21

<220>
<221>  CDS
<222>  (1)..(1092)
```

[0029]

```
<400>  12
gac aac ttc aat gtc tat aaa gcc aca aga cca tac tta gct cac tgt         48
Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala His Cys
1               5                   10                  15

ccc gac tgt gga gaa ggg cac tcg tgc cat agt ccc gta gca cta gaa         96
Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala Leu Glu
            20                  25                  30

cgc atc aga aat gaa gcg aca gac ggg acg ctg aaa atc cag gtc tcc        144
Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln Val Ser
        35                  40                  45

ttg caa atc gga ata aag acg gat gac agc cac gat tgg acc aag ctg        192
Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr Lys Leu
    50                  55                  60

cgt tat atg gac aac cac atg cca gca gac gca gag agg gcg ggg cta        240
Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala Gly Leu
65                  70                  75                  80

ttt gta aga aca tca gca ccg tgt acg att act gga aca atg gga cac        288
Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met Gly His
                85                  90                  95

ttc atc ctg gcc cga tgt cca aaa ggg gaa act ctg acg gtg gga ttc        336
Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val Gly Phe
            100                 105                 110

act gac agt agg aag att agt cac tca tgt acg cac cca ttt cac cac        384
Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe His His
        115                 120                 125

gac cct cct gtg ata ggt cgg gaa aaa ttc cat tcc cga ccg cag cac        432
Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro Gln His
    130                 135                 140

ggt aaa gag cta cct tgc agc acg tac gtg cag agc acc gcc gca act        480
Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala Ala Thr
```

```
145                 150                 155                 160

acc gag gag ata gag gta cac atg ccc cca gac acc cct gat cgc aca        528
Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp Arg Thr
                165                 170                 175

tta atg tca caa cag tcc ggc aac gta aag atc aca gtc aat ggc cag        576
Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn Gly Gln
            180                 185                 190

acg gtg cgg tac aag tgt aat tgc ggt ggc tca aat gaa gga cta aca        624
Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly Leu Thr
        195                 200                 205

act aca gac aaa gtg att aat aac tgc aag gtt gat caa tgt cat gcc        672
Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys His Ala
    210                 215                 220

gcg gtc acc aat cac aaa aag tgg cag tat aac tcc cct ctg gtc ccg        720
Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu Val Pro
225                 230                 235                 240

cgt aat gct gaa ctt ggg gac cga aaa gga aaa att cac atc ccg ttt        768
Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile Pro Phe
                245                 250                 255

ccg ctg gca aat gta aca tgc agg gtg cct aaa gca agg aac ccc acc        816
Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn Pro Thr
            260                 265                 270

gtg acg tac ggg aaa aac caa gtc atc atg cta ctg tat cct gac cac        864
Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro Asp His
        275                 280                 285

cca aca ctc ctg tcc tac cgg aat atg gga gaa gaa cca aac tat caa        912
Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn Tyr Gln
    290                 295                 300
```

[0030]

```
gaa gag tgg gtg atg cat aag aag gaa gtc gtg cta acc gtg ccg act        960
Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val Pro Thr
305                 310                 315                 320

gaa ggg ctc gag gtc acg tgg ggc aac aac gag ccg tat aag tat tgg       1008
Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys Tyr Trp
                325                 330                 335

ccg cag tta tct aca aac ggt aca gcc cat ggc cac ccg cat gag ata       1056
Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His Glu Ile
            340                 345                 350

att ctg tat tat tat gag ctg tac ccc act atg act                       1092
Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr
        355                 360


<210>  13
<211>  364
<212>  PRT
<213>  Chikungunya virus

<400>  13

Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala His Cys
1               5                   10                  15


Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala Leu Glu
            20                  25                  30


Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln Val Ser
        35                  40                  45


Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr Lys Leu
    50                  55                  60
```

```
       Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala Gly Leu
       65                  70                  75                  80

       Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met Gly His
                       85                  90                  95

       Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val Gly Phe
                   100                 105                 110

       Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe His His
               115                 120                 125

       Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro Gln His
           130                 135                 140

       Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala Ala Thr
       145                 150                 155                 160

       Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp Arg Thr
                       165                 170                 175

       Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn Gly Gln
                   180                 185                 190

       Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly Leu Thr
               195                 200                 205

[0031] Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys His Ala
           210                 215                 220

       Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu Val Pro
       225                 230                 235                 240

       Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile Pro Phe
                       245                 250                 255

       Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn Pro Thr
                   260                 265                 270

       Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro Asp His
               275                 280                 285

       Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn Tyr Gln
           290                 295                 300

       Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val Pro Thr
       305                 310                 315                 320

       Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys Tyr Trp
                       325                 330                 335

       Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His Glu Ile
                   340                 345                 350

       Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr
               355                 360
```

<210> 14
<211> 1092
<212> DNA
<213> Chikungunya virus

<220>
<221> misc_feature
<223> ectodomain glycoprotein E2 strain 06.27

<220>
<221> CDS
<222> (1)..(1092)

<400> 14

```
gac aac ttc aat gtc tat aaa gcc aca aga cca tac tta gct cac tgt       48
Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala His Cys
1               5                   10                  15

ccc gac tgt gga gaa ggg cac tcg tgc cat agt ccc gta gca cta gaa       96
Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala Leu Glu
            20                  25                  30

cgc atc aga aat gaa gcg aca gac ggg acg ctg aaa atc cag gtc tcc      144
Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln Val Ser
        35                  40                  45

ttg caa atc gga ata aag acg gat gac agc cac gat tgg acc aag ctg      192
Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr Lys Leu
    50                  55                  60

cgt tat atg gac aac cac atg cca gca gac gca gag agg gcg ggg cta      240
Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala Gly Leu
65                  70                  75                  80

ttt gta aga aca tca gca ccg tgt acg att act gga aca atg gga cac      288
Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met Gly His
                85                  90                  95
```

[0032]

```
ttc atc ctg gcc cga tgt cca aaa ggg gaa act ctg acg gtg gga ttc      336
Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val Gly Phe
            100                 105                 110

act gac agt agg aag att agt cac tca tgt acg cac cca ttt cac cac      384
Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe His His
        115                 120                 125

gac cct cct gtg ata ggt cgg gaa aaa ttc cat tcc cga ccg cag cac      432
Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro Gln His
    130                 135                 140

ggt aaa gag cta cct tgc agc acg tac gtg cag agc acc gcc gca act      480
Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala Ala Thr
145                 150                 155                 160

acc gag gag ata gag gta cac atg ccc cca gac acc cct gat cgc aca      528
Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp Arg Thr
                165                 170                 175

tta atg tca caa cag tcc ggc aac gta aag atc aca gtc aat ggc cag      576
Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn Gly Gln
            180                 185                 190

acg gtg cgg tac aag tgt aat tgc ggt ggc tca aat gaa gga cta aca      624
Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly Leu Thr
        195                 200                 205

act aca gac aaa gtg att aat aac tgc aag gtt gat caa tgt cat gcc      672
Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys His Ala
    210                 215                 220

gcg gtc acc aat cac aaa aag tgg cag tat aac tcc cct ctg gtc ccg      720
Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu Val Pro
225                 230                 235                 240

cgt aat gct gaa ctt ggg gac cga aaa gga aaa att cac atc ccg ttt      768
Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile Pro Phe
```

245 250 255

ccg ctg gca aat gta aca tgc agg gtg cct aaa gca agg aac ccc acc 816
Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn Pro Thr
260 265 270

gtg acg tac ggg aaa aac caa gtc atc atg cta ctg tat cct gac cac 864
Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro Asp His
275 280 285

cca aca ctc ctg tcc tac cgg aat atg gga gaa gaa cca aac tat caa 912
Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn Tyr Gln
290 295 300

gaa gag tgg gtg atg cat aag aag gaa gtc gtg cta acc gtg ccg act 960
Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val Pro Thr
305 310 315 320

gaa ggg ctc gag gtc acg tgg ggc aac aac gag ccg tat aag tat tgg 1008
Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys Tyr Trp
325 330 335

ccg cag tta tct aca aac ggt aca gcc cat ggc cac ccg cat gag ata 1056
Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His Glu Ile
340 345 350

att ctg tat tat tat gag ctg tac ccc act atg act 1092
Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr
355 360

<210> 15
<211> 364
<212> PRT
<213> Chikungunya virus

<400> 15

[0033] Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu Ala His Cys
1 5 10 15

Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val Ala Leu Glu
20 25 30

Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile Gln Val Ser
35 40 45

Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp Thr Lys Leu
50 55 60

Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg Ala Gly Leu
65 70 75 80

Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr Met Gly His
85 90 95

Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr Val Gly Phe
100 105 110

Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro Phe His His
115 120 125

Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg Pro Gln His
130 135 140

Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr Ala Ala Thr
145 150 155 160

Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro Asp Arg Thr
165 170 175

Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val Asn Gly Gln
180 185 190

Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu Gly Leu Thr
195 200 205

Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln Cys His Ala
210 215 220

Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro Leu Val Pro
225 230 235 240

Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly Lys Ile His Ile Pro Phe
245 250 255

Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg Asn Pro Thr
260 265 270

Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr Pro Asp His
275 280 285

Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro Asn Tyr Gln
290 295 300

Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr Val Pro Thr
[0034] 305 310 315 320

Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr Lys Tyr Trp
325 330 335

Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro His Glu Ile
340 345 350

Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr
355 360

<210> 16
<211> 2772
<212> DNA
<213> Artificial Sequence

<220>
<223> polyprotein E2-6K-E1 France/2010 CHIKV strains (GenBank: CCA61130.1)

<220>
<221> CDS
<222> (1)..(2772)

<400> 16
agc acc aag gac aac ttc aat gtc tat aaa gcc aca aga cca tac tta 48
Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu
1 5 10 15

gct cac tgt ccc gac tgt gga gaa ggg cac tcg tgc cat agt ccc gta 96
Ala His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val
20 25 30

gca cta gaa cgc atc aga aat gaa gcg aca gac ggg acg ctg aaa atc 144

```
Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile
        35                  40                  45

cag gtc tcc ttg caa atc gga ata aag acg gat gac agc cac gat tgg         192
Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp
    50                  55                  60

acc aag ctg cgt tat atg gac aac cac atg cca gca gac gca gag agg         240
Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg
65                  70                  75                  80

gcg ggg cta ttt gta aga aca tca gca ccg tgt acg att act gga aca         288
Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr
                85                  90                  95

atg gga cac ttc atc ctg gcc cga tgt cca aaa ggg gaa act ctg acg         336
Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr
            100                 105                 110

gtg gga ttc act gac agt agg aag att agt cat tca tgt acg cac cca         384
Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro
        115                 120                 125

ttt cac cac gac cct cct gtg ata ggt cgg gaa aaa ttc cat tcc cga         432
Phe His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg
    130                 135                 140

ccg cag cac ggt aaa gag cta cct tgc agc acg tac gtg cag agc acc         480
Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr
145                 150                 155                 160

gcc gca act acc gag gag ata gag gta cac atg ccc cca gac acc cct         528
Ala Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro
                165                 170                 175

gat cgc aca tta atg tca caa cag tcc ggc aac gta aag atc aca gtc         576
Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val
            180                 185                 190
```

[0035]

```
aat ggc cag acg gtg cgg tat aag tgt aat tgc ggt ggc tca aat gaa         624
Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu
        195                 200                 205

gga cta aca act aca gac aaa gtg att aat aac tgc aag gtt gat caa         672
Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln
    210                 215                 220

tgt cat gcc gcg gtc acc aat cac aaa aag tgg cag tat aac tcc cct         720
Cys His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro
225                 230                 235                 240

ctg gtc ccg cgt aat gct gaa ctt ggg gac cga caa gga aaa att cac         768
Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Gln Gly Lys Ile His
                245                 250                 255

atc ccg ttt ccg ctg gca aat gta aca tgc agg gtg cct aaa gca agg         816
Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg
            260                 265                 270

aac ccc acc gtg acg tac ggg aaa aac caa gtc atc atg cta ctg tat         864
Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr
        275                 280                 285

cct gac cac cca aca ctc ctg tcc tac cgg aat atg gga gaa gaa cca         912
Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro
    290                 295                 300

aac tat caa gaa gag tgg gtg atg cat aag aag gaa gtc gtg cta acc         960
Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr
305                 310                 315                 320

gtg ccg act gaa ggg ctc gag gtc acg tgg ggc aac aac gag ccg tat        1008
Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr
                325                 330                 335

aag tat tgg ccg cag tta tct aca aac ggt aca gcc cat ggc cac ccg        1056
Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro
```

```
                   340                 345                 350

       cat gag ata att ctg tat tat tat gag ctg tac cct act atg act gta      1104
       His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val
               355                 360                 365

       gta gtt gta tca gtg gcc acg ttc ata ctc ctg tca atg gtg ggt atg      1152
       Val Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met
           370                 375                 380

       gca gcg ggg atg tgc atg tgt gca cga cgc aga tgc atc aca ccg tat      1200
       Ala Ala Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr
       385                 390                 395                 400

       gaa ctg aca cca gga gct acc gtc cct ttc ctg ctt agc cta ata tgc      1248
       Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys
                       405                 410                 415

       tgc atc aga aca gct aaa gcg gcc aca tac caa gag gct gcg ata tac      1296
       Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr
                   420                 425                 430

       ctg tgg aac gag cag caa cct ttg ttt tgg cta caa gcc ctt att ccg      1344
       Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu Gln Ala Leu Ile Pro
               435                 440                 445

       ctg gca gcc ctg att gtt cta tgc aac tgt ctg aga ctc tta cca tgc      1392
       Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu Arg Leu Leu Pro Cys
           450                 455                 460

       tgc tgt aaa acg ttg gct ttt tta gcc gta atg agc gtc ggt gcc cac      1440
       Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met Ser Val Gly Ala His
       465                 470                 475                 480

       act gtg agc gcg tac gaa cac gta aca gtg atc ccg aac acg gtg gga      1488
       Thr Val Ser Ala Tyr Glu His Val Thr Val Ile Pro Asn Thr Val Gly
                       485                 490                 495

[0036] gta ccg tat aag act cta gtc aat aga cct ggc tac agc ccc atg gta      1536
       Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly Tyr Ser Pro Met Val
                   500                 505                 510

       ttg gag atg gaa cta ctg tca gtc act ttg gag cca aca cta tcg ctt      1584
       Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu Pro Thr Leu Ser Leu
               515                 520                 525

       gat tac atc acg tgc gag tac aaa acc gtc atc ccg tct ccg tac gtg      1632
       Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile Pro Ser Pro Tyr Val
           530                 535                 540

       aag tgc tgc ggt aca gca gag tgc aag gac aaa aac cta cct gac tac      1680
       Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys Asn Leu Pro Asp Tyr
       545                 550                 555                 560

       agc tgt aag gtc ttc acc ggc gtc tac cca ttt atg tgg ggc ggc gcc      1728
       Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe Met Trp Gly Gly Ala
                       565                 570                 575

       tac tgc ttc tgc gac gct gaa aat acg cag ttg agc gaa gca cat gtg      1776
       Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu Ser Glu Ala His Val
                   580                 585                 590

       gag aag tcc gaa tca tgc aaa aca gaa ttt gca tca gcg tac agg gct      1824
       Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala Ser Ala Tyr Arg Ala
               595                 600                 605

       cat acc gca tct gca tca gct aag ctc cgc gtc ctt tac caa gga aat      1872
       His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val Leu Tyr Gln Gly Asn
           610                 615                 620

       aac atc act gta act gcc tat gca aac ggc gac cat gcc gtc aca gtt      1920
       Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp His Ala Val Thr Val
       625                 630                 635                 640

       aag gac gcc aaa ttc att gtg ggg cca atg tct tca gcc tgg aca cct      1968
       Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser Ser Ala Trp Thr Pro
                       645                 650                 655
```

```
       ttc gac aac aaa att gtg gtg tac aaa ggt gac gtc tat aac atg gac      2016
       Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp Val Tyr Asn Met Asp
                   660                 665                 670

       tac ccg ccc ttt ggc gca gga aga cca gga caa ttt ggc gat atc caa      2064
       Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln Phe Gly Asp Ile Gln
               675                 680                 685

       agt cgc aca cct gag agt aaa gac gtc tat gct aat aca caa ctg gta      2112
       Ser Arg Thr Pro Glu Ser Lys Asp Val Tyr Ala Asn Thr Gln Leu Val
           690                 695                 700

       ctg cag aga ccg gct gtg ggt acg gta cac gtg cca tac tct cag gca      2160
       Leu Gln Arg Pro Ala Val Gly Thr Val His Val Pro Tyr Ser Gln Ala
       705                 710                 715                 720

       cca tct ggc ttt aag tat tgg cta aaa gaa cgc ggg gcg tca ctg cag      2208
       Pro Ser Gly Phe Lys Tyr Trp Leu Lys Glu Arg Gly Ala Ser Leu Gln
                       725                 730                 735

       cac aca gca cca ttt ggc tgc caa ata gca aca aac ccg gta aga gcg      2256
       His Thr Ala Pro Phe Gly Cys Gln Ile Ala Thr Asn Pro Val Arg Ala
                   740                 745                 750

       gtg aac tgc gcc gta ggg aac atg ccc atc tcc atc gac ata ccg gaa      2304
       Val Asn Cys Ala Val Gly Asn Met Pro Ile Ser Ile Asp Ile Pro Glu
               755                 760                 765

       gcg gcc ttc act agg gtc gtc gac gcg ccc tct tta acg gac atg tcg      2352
       Ala Ala Phe Thr Arg Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser
           770                 775                 780

       tgc gag gta cca gcc tgc acc cat tcc tca gac ttt ggg ggc gtc gcc      2400
       Cys Glu Val Pro Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala
       785                 790                 795                 800

       att att aaa tat gca gcc agc aag aaa ggc aag tgt gcg gtg cat tcg      2448
[0037] Ile Ile Lys Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser
                       805                 810                 815

       atg act aac gcc gtc act att cgg gaa gct gag ata gaa gtt gaa ggg      2496
       Met Thr Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly
                   820                 825                 830

       aat tct cag ctg caa atc tct ttc tcg acg gcc tta gcc agc gcc gaa      2544
       Asn Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
               835                 840                 845

       ttc cgc gta caa gtc tgt tct aca caa gta cac tgt gca gct gag tgc      2592
       Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu Cys
           850                 855                 860

       cac ccc ccg aag gac cac ata gtc aac tac ccg gcg tca cat acc acc      2640
       His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His Thr Thr
       865                 870                 875                 880

       ctc ggg gtc cag gac atc tcc gct acg gcg atg tca tgg gtg cag aag      2688
       Leu Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp Val Gln Lys
                       885                 890                 895

       atc acg gga ggt gtg gga ctg gtt gtt gct gtt gcc gca ctg att cta      2736
       Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala Ala Leu Ile Leu
                   900                 905                 910

       atc gtg gtg cta tgc gtg tcg ttc agc agg cac taa                      2772
       Ile Val Val Leu Cys Val Ser Phe Ser Arg His
               915                 920
```

<210> 17
<211> 923
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 17

```
Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu
1               5                   10                  15

Ala His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val
            20                  25                  30

Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile
        35                  40                  45

Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp
    50                  55                  60

Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg
65                  70                  75                  80

Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr
                85                  90                  95

Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr
            100                 105                 110

Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys Thr His Pro
        115                 120                 125

Phe His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg
    130                 135                 140
```

[0038]

```
Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val Gln Ser Thr
145                 150                 155                 160

Ala Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro
                165                 170                 175

Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val
            180                 185                 190

Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly Ser Asn Glu
        195                 200                 205

Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln
    210                 215                 220

Cys His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro
225                 230                 235                 240

Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Gln Gly Lys Ile His
                245                 250                 255

Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg
            260                 265                 270

Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr
        275                 280                 285

Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro
    290                 295                 300
```

Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val Val Leu Thr
305 310 315 320

Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr
325 330 335

Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro
340 345 350

His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val
355 360 365

Val Val Val Ser Val Ala Thr Phe Ile Leu Leu Ser Met Val Gly Met
370 375 380

Ala Ala Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr
385 390 395 400

Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys
405 410 415

Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Ile Tyr
420 425 430

Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu Gln Ala Leu Ile Pro
435 440 445

[0039] Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu Arg Leu Leu Pro Cys
450 455 460

Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met Ser Val Gly Ala His
465 470 475 480

Thr Val Ser Ala Tyr Glu His Val Thr Val Ile Pro Asn Thr Val Gly
485 490 495

Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly Tyr Ser Pro Met Val
500 505 510

Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu Pro Thr Leu Ser Leu
515 520 525

Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile Pro Ser Pro Tyr Val
530 535 540

Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys Asn Leu Pro Asp Tyr
545 550 555 560

Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe Met Trp Gly Gly Ala
565 570 575

Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu Ser Glu Ala His Val
580 585 590

Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala Ser Ala Tyr Arg Ala
595 600 605

```
His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val Leu Tyr Gln Gly Asn
    610                 615                 620

Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp His Ala Val Thr Val
625                 630                 635                 640

Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser Ser Ala Trp Thr Pro
                645                 650                 655

Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp Val Tyr Asn Met Asp
            660                 665                 670

Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln Phe Gly Asp Ile Gln
        675                 680                 685

Ser Arg Thr Pro Glu Ser Lys Asp Val Tyr Ala Asn Thr Gln Leu Val
    690                 695                 700

Leu Gln Arg Pro Ala Val Gly Thr Val His Val Pro Tyr Ser Gln Ala
705                 710                 715                 720

Pro Ser Gly Phe Lys Tyr Trp Leu Lys Glu Arg Gly Ala Ser Leu Gln
                725                 730                 735

His Thr Ala Pro Phe Gly Cys Gln Ile Ala Thr Asn Pro Val Arg Ala
            740                 745                 750
```

[0040]

```
Val Asn Cys Ala Val Gly Asn Met Pro Ile Ser Ile Asp Ile Pro Glu
        755                 760                 765

Ala Ala Phe Thr Arg Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser
    770                 775                 780

Cys Glu Val Pro Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala
785                 790                 795                 800

Ile Ile Lys Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser
                805                 810                 815

Met Thr Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly
            820                 825                 830

Asn Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
        835                 840                 845

Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu Cys
    850                 855                 860

His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His Thr Thr
865                 870                 875                 880

Leu Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp Val Gln Lys
                885                 890                 895

Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala Ala Leu Ile Leu
            900                 905                 910
```

Ile Val Val Leu Cys Val Ser Phe Ser Arg His
915 920

<210> 18
<211> 2772
<212> DNA
<213> Artificial Sequence

<220>
<223> polyprotein E2-6K-E1 France/2010 CHIKV strains (GenBank: CCA61131.1)

<220>
<221> CDS
<222> (1)..(2772)

<400> 18
agt att aag gac cac ttc aat gtc tat aaa gcc aca aga ccg tac cta 48
Ser Ile Lys Asp His Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu
1 5 10 15

gct cac tgt ccc gac tgt gga gaa ggg cac tcg tgc cat agt ccc gta 96
Ala His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val
20 25 30

gcg cta gaa cgc atc aga aac gaa gcg aca gac ggg acg ttg aaa atc 144
Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile
35 40 45

cag gtt tcc ttg caa atc gga ata aag acg gat gat agc cat gat tgg 192
Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp
50 55 60

acc aag ctg cgt tat atg gac aat cac atg cca gca gac gca gag cgg 240
Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg
65 70 75 80

[0041] gcc ggg cta ttt gta aga acg tca gca ccg tgc acg att act gga aca 288
Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr
85 90 95

atg gga cac ttc att ctg gcc cga tgt ccg aaa gga gaa act ctg aca 336
Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr
100 105 110

gtg ggg ttc act gac ggt agg aag atc agt cac tca tgt acg cac cca 384
Val Gly Phe Thr Asp Gly Arg Lys Ile Ser His Ser Cys Thr His Pro
115 120 125

ttt cac cat gac cct cct gtg ata ggc cgg gaa aaa ttc cat tcc cga 432
Phe His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg
130 135 140

ccg cag cac ggt agg gaa cta cct tgc agc acg tac gcg cag agc acc 480
Pro Gln His Gly Arg Glu Leu Pro Cys Ser Thr Tyr Ala Gln Ser Thr
145 150 155 160

gct gca act gcc gag gag ata gag gta cat atg ccc cca gac act cca 528
Ala Ala Thr Ala Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro
165 170 175

gat cgc aca tta atg tca caa cag tcc ggc aat gta aag atc aca gtc 576
Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val
180 185 190

aat agt cag acg gtg cgg tac aag tgc aat tgt ggt gac tca agt gaa 624
Asn Ser Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Asp Ser Ser Glu
195 200 205

gga tta acc act aca gat aaa gtg att aat aac tgc aag gtt gat caa 672
Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln
210 215 220

tgc cat gcc gcg gtc acc aat cac aaa aaa tgg cag tat aat tcc cct 720
Cys His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro

```
225                 230                 235                 240

ctg gtc ccg cgc aat gct gaa ttc ggg gac cgg aaa gga aaa gtt cac         768
Leu Val Pro Arg Asn Ala Glu Phe Gly Asp Arg Lys Gly Lys Val His
                245                 250                 255

att cca ttt cct ctg gca aat gtg aca tgc agg gtg cct aaa gca aga         816
Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg
            260                 265                 270

aac ccc acc gtg acg tac gga aaa aac caa gtc atc atg ttg ctg tat         864
Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr
        275                 280                 285

ccg gac cac cca acg ctc ctg tcc tac agg aat atg gga gaa gaa cca         912
Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro
    290                 295                 300

aac tat caa gaa gag tgg gtg acg cat aag aag gag atc agg tta acc         960
Asn Tyr Gln Glu Glu Trp Val Thr His Lys Lys Glu Ile Arg Leu Thr
305                 310                 315                 320

gta ccg act gaa ggg ctc gag gtc acg tgg ggt aac aat gag ccg tac        1008
Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr
                325                 330                 335

aag tat tgg ccg cag tta tcc aca aac ggt aca gcc cac ggc cac ccg        1056
Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro
            340                 345                 350

cat gag ata att ttg tat tat tat gag ctg tac cca act atg act gtg        1104
His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val
        355                 360                 365

gta gtt ttg tca gtg gcc tcg ttc ata ctc ctg tcg atg gtg ggt gtg        1152
Val Val Leu Ser Val Ala Ser Phe Ile Leu Leu Ser Met Val Gly Val
    370                 375                 380
```

[0042]

```
gca gtg ggg atg tgc atg tgt gca cga cgc aga tgc att aca ccg tac        1200
Ala Val Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr
385                 390                 395                 400

gaa ctg aca cca gga gct acc gtc cct ttc ctg ctt agc cta ata tgc        1248
Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys
                405                 410                 415

tgc att aga aca gct aaa gcg gcc aca tac cag gag gcc gcg gta tac        1296
Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Val Tyr
            420                 425                 430

ctg tgg aac gag cag cag cct tta ttt tgg ctg caa gcc ctt att ccg        1344
Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu Gln Ala Leu Ile Pro
        435                 440                 445

ctg gca gcc ctg att gtc cta tgt aac tgt ctg aga ctc tta cca tgc        1392
Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu Arg Leu Leu Pro Cys
    450                 455                 460

tgt tgt aaa atg ttg act ttt tta gcc gta ctg agc gtc ggt gcc cac        1440
Cys Cys Lys Met Leu Thr Phe Leu Ala Val Leu Ser Val Gly Ala His
465                 470                 475                 480

act gtg agc gcg tac gaa cac gta aca gtg atc ccg aac acg gtg gga        1488
Thr Val Ser Ala Tyr Glu His Val Thr Val Ile Pro Asn Thr Val Gly
                485                 490                 495

gta ccg tat aag act cta gtc aac aga ccg ggc tac agc ccc atg gta        1536
Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly Tyr Ser Pro Met Val
            500                 505                 510

ctg gag atg gag ctt ctg tct gtc acc ttg gaa cca acg cta tcg ctt        1584
Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu Pro Thr Leu Ser Leu
        515                 520                 525

gat tac atc acg tgc gag tat aaa acc gtt atc ccg tct ccg tac gtg        1632
Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile Pro Ser Pro Tyr Val
    530                 535                 540
```

```
aaa tgc tgc ggt aca gca gag tgt aag gac aag agc cta cct gat tac          1680
Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys Ser Leu Pro Asp Tyr
545                 550                 555                 560

agc tgt aag gtc ttc acc ggc gtc tac cca ttc atg tgg ggc ggc gcc          1728
Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe Met Trp Gly Gly Ala
                565                 570                 575

tac tgc ttc tgc gac acc gaa aat acg caa ttg agc gaa gca cat gtg          1776
Tyr Cys Phe Cys Asp Thr Glu Asn Thr Gln Leu Ser Glu Ala His Val
            580                 585                 590

gag aag tcc gaa tca tgc aaa aca gaa ttt gca tca gca tac agg gct          1824
Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala Ser Ala Tyr Arg Ala
        595                 600                 605

cat acc gca tcc gca tca gct aag ctc cgc gtc ctt tac caa gga aat          1872
His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val Leu Tyr Gln Gly Asn
    610                 615                 620

aat atc act gta gct gct tat gca aac ggc gac cat gcc gtc aca gtt          1920
Asn Ile Thr Val Ala Ala Tyr Ala Asn Gly Asp His Ala Val Thr Val
625                 630                 635                 640

aag gac gct aaa ttc ata gtg ggg cca atg tct tca gcc tgg aca cct          1968
Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser Ser Ala Trp Thr Pro
                645                 650                 655

ttc gac aat aaa atc gtg gtg tac aaa ggc gac gtc tac aac atg gac          2016
Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp Val Tyr Asn Met Asp
            660                 665                 670

tac ccg ccc ttc ggc gca gga aga cca gga caa ttt ggc gac atc caa          2064
Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln Phe Gly Asp Ile Gln
        675                 680                 685
```

[0043]

```
agt cgc acg cct gag agc gaa gac gtc tat gct aac aca caa ctg gta          2112
Ser Arg Thr Pro Glu Ser Glu Asp Val Tyr Ala Asn Thr Gln Leu Val
    690                 695                 700

ctg cag aga ccg tcc gcg ggt acg gtg cac gtg ccg tac tct cag gca          2160
Leu Gln Arg Pro Ser Ala Gly Thr Val His Val Pro Tyr Ser Gln Ala
705                 710                 715                 720

cca tct ggc ttc aag tat tgg cta aaa gaa cga ggg gcg tcg ctg cag          2208
Pro Ser Gly Phe Lys Tyr Trp Leu Lys Glu Arg Gly Ala Ser Leu Gln
                725                 730                 735

cac aca gca cca ttt ggc tgt caa ata gca aca aac ccg gta aga gcg          2256
His Thr Ala Pro Phe Gly Cys Gln Ile Ala Thr Asn Pro Val Arg Ala
            740                 745                 750

atg aac tgc gcc gta ggg aac atg cct atc tcc atc gac ata ccg gac          2304
Met Asn Cys Ala Val Gly Asn Met Pro Ile Ser Ile Asp Ile Pro Asp
        755                 760                 765

gcg gcc ttt acc agg gtc gtc gac gcg cca tct tta acg gac atg tcg          2352
Ala Ala Phe Thr Arg Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser
    770                 775                 780

tgt gag gta tca gcc tgc acc cat tcc tca gac ttt ggg ggc gtt gcc          2400
Cys Glu Val Ser Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala
785                 790                 795                 800

atc att aaa tat gca gcc agt aag aaa ggc aag tgt gca gtg cac tcg          2448
Ile Ile Lys Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser
                805                 810                 815

atg act aac gcc gtc act att cgg gaa gct gaa ata gaa gta gaa ggg          2496
Met Thr Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly
            820                 825                 830

aac tct cag ttg caa atc tct ttt tcg acg gcc cta gcc agc gcc gaa          2544
Asn Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
        835                 840                 845
```

```
ttc cgc gta caa gtc tgt tct aca caa gta cac tgt gca gcc gag tgc      2592
Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu Cys
    850                 855                 860

cat cca ccg aaa gac cat ata gtc aat tac ccg gcg tca cac acc acc      2640
His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His Thr Thr
865                 870                 875                 880

ccc ggg gtc caa gac att tcc gct acg gcg atg tca tgg gtg cag aag      2688
Pro Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp Val Gln Lys
                885                 890                 895

atc acg gga ggt gtg gga ctg gtt gtc gct gtt gca gca ctg atc cta      2736
Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala Ala Leu Ile Leu
            900                 905                 910

atc gtg gtg cta tgc gtg tcg ttt agc agg cac taa                      2772
Ile Val Val Leu Cys Val Ser Phe Ser Arg His
        915                 920


<210>  19
<211>  923
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic Construct

<400>  19

Ser Ile Lys Asp His Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu
1               5                   10                  15

Ala His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His Ser Pro Val
            20                  25                  30
```

[0044]

```
Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr Leu Lys Ile
        35                  40                  45

Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser His Asp Trp
    50                  55                  60

Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp Ala Glu Arg
65                  70                  75                  80

Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile Thr Gly Thr
                85                  90                  95

Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu Thr Leu Thr
            100                 105                 110

Val Gly Phe Thr Asp Gly Arg Lys Ile Ser His Ser Cys Thr His Pro
        115                 120                 125

Phe His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe His Ser Arg
    130                 135                 140

Pro Gln His Gly Arg Glu Leu Pro Cys Ser Thr Tyr Ala Gln Ser Thr
145                 150                 155                 160

Ala Ala Thr Ala Glu Glu Ile Glu Val His Met Pro Pro Asp Thr Pro
                165                 170                 175

Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys Ile Thr Val
            180                 185                 190
```

```
Asn Ser Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Asp Ser Ser Glu
        195                 200                 205

Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys Val Asp Gln
    210                 215                 220

Cys His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr Asn Ser Pro
225                 230                 235                 240

Leu Val Pro Arg Asn Ala Glu Phe Gly Asp Arg Lys Gly Lys Val His
                245                 250                 255

Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro Lys Ala Arg
            260                 265                 270

Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met Leu Leu Tyr
        275                 280                 285

Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly Glu Glu Pro
    290                 295                 300

Asn Tyr Gln Glu Glu Trp Val Thr His Lys Lys Glu Ile Arg Leu Thr
305                 310                 315                 320

Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn Glu Pro Tyr
                325                 330                 335
```

[0045]

```
Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His Gly His Pro
            340                 345                 350

His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr Met Thr Val
        355                 360                 365

Val Val Leu Ser Val Ala Ser Phe Ile Leu Leu Ser Met Val Gly Val
    370                 375                 380

Ala Val Gly Met Cys Met Cys Ala Arg Arg Arg Cys Ile Thr Pro Tyr
385                 390                 395                 400

Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu Leu Ser Leu Ile Cys
                405                 410                 415

Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln Glu Ala Ala Val Tyr
            420                 425                 430

Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu Gln Ala Leu Ile Pro
        435                 440                 445

Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu Arg Leu Leu Pro Cys
    450                 455                 460

Cys Cys Lys Met Leu Thr Phe Leu Ala Val Leu Ser Val Gly Ala His
465                 470                 475                 480

Thr Val Ser Ala Tyr Glu His Val Thr Val Ile Pro Asn Thr Val Gly
                485                 490                 495
```

Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly Tyr Ser Pro Met Val
500 505 510

Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu Pro Thr Leu Ser Leu
515 520 525

Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile Pro Ser Pro Tyr Val
530 535 540

Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys Ser Leu Pro Asp Tyr
545 550 555 560

Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe Met Trp Gly Gly Ala
565 570 575

Tyr Cys Phe Cys Asp Thr Glu Asn Thr Gln Leu Ser Glu Ala His Val
580 585 590

Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala Ser Ala Tyr Arg Ala
595 600 605

His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val Leu Tyr Gln Gly Asn
610 615 620

Asn Ile Thr Val Ala Ala Tyr Ala Asn Gly Asp His Ala Val Thr Val
625 630 635 640

[0046] Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser Ser Ala Trp Thr Pro
645 650 655

Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp Val Tyr Asn Met Asp
660 665 670

Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln Phe Gly Asp Ile Gln
675 680 685

Ser Arg Thr Pro Glu Ser Glu Asp Val Tyr Ala Asn Thr Gln Leu Val
690 695 700

Leu Gln Arg Pro Ser Ala Gly Thr Val His Val Pro Tyr Ser Gln Ala
705 710 715 720

Pro Ser Gly Phe Lys Tyr Trp Leu Lys Glu Arg Gly Ala Ser Leu Gln
725 730 735

His Thr Ala Pro Phe Gly Cys Gln Ile Ala Thr Asn Pro Val Arg Ala
740 745 750

Met Asn Cys Ala Val Gly Asn Met Pro Ile Ser Ile Asp Ile Pro Asp
755 760 765

Ala Ala Phe Thr Arg Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser
770 775 780

Cys Glu Val Ser Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala
785 790 795 800

Ile Ile Lys Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser
805 810 815

Met Thr Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly
820 825 830

Asn Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
835 840 845

Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu Cys
850 855 860

His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His Thr Thr
865 870 875 880

Pro Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp Val Gln Lys
885 890 895

Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala Ala Leu Ile Leu
900 905 910

Ile Val Val Leu Cys Val Ser Phe Ser Arg His
915 920

<210> 20
<211> 3744
<212> DNA
<213> Artificial Sequence

[0047] <220>
<223> proteins C-E3-E2-6K-E1 strain S27 (GenBank: AF339485.1)

<220>
<221> CDS
<222> (1)..(3744)

<400> 20
atg gag ttc atc cca acc caa act ttt tac aac agg agg tac cag cct 48
Met Glu Phe Ile Pro Thr Gln Thr Phe Tyr Asn Arg Arg Tyr Gln Pro
1 5 10 15

cga ccc tgg act ccg cgc cct act atc caa gtc atc agg ccc aga ccg 96
Arg Pro Trp Thr Pro Arg Pro Thr Ile Gln Val Ile Arg Pro Arg Pro
20 25 30

cgc ccg cag agg caa gct ggg caa ctt gcc cag ctg atc tca gca gtt 144
Arg Pro Gln Arg Gln Ala Gly Gln Leu Ala Gln Leu Ile Ser Ala Val
35 40 45

aat aaa ctg aca atg cgc gcg gta ccc caa cag aag cca cgc agg aat 192
Asn Lys Leu Thr Met Arg Ala Val Pro Gln Gln Lys Pro Arg Arg Asn
50 55 60

cgg aag aat aag aag caa aag caa aag caa cag gcg cca caa aac aac 240
Arg Lys Asn Lys Lys Gln Lys Gln Lys Gln Gln Ala Pro Gln Asn Asn
65 70 75 80

aca aac caa aag aag cag cca cct aaa aag aaa cca gct caa aag aaa 288
Thr Asn Gln Lys Lys Gln Pro Pro Lys Lys Lys Pro Ala Gln Lys Lys
85 90 95

aag aag ccg ggc cgc aga gag agg atg tgc atg aaa atc gaa aat gac 336
Lys Lys Pro Gly Arg Arg Glu Arg Met Cys Met Lys Ile Glu Asn Asp
100 105 110

tgt att ttc gaa gtc aag cac gaa ggt aag gta aca ggt tac gcg tgc 384
Cys Ile Phe Glu Val Lys His Glu Gly Lys Val Thr Gly Tyr Ala Cys
115 120 125

```
ttg gtg ggg gac aaa gta atg aaa cca gca cac gta aag ggg acc atc           432
Leu Val Gly Asp Lys Val Met Lys Pro Ala His Val Lys Gly Thr Ile
    130                 135                 140

gat aac gcg gac ctg gcc aaa ttg gcc ttt aag cgg tca tct aag tac           480
Asp Asn Ala Asp Leu Ala Lys Leu Ala Phe Lys Arg Ser Ser Lys Tyr
145                 150                 155                 160

gac ctt gaa tgc gcg cag ata ccc gtg cac atg aag tcc gac gct tcg           528
Asp Leu Glu Cys Ala Gln Ile Pro Val His Met Lys Ser Asp Ala Ser
                165                 170                 175

aag ttc acc cat gag aaa ccg gag ggg tac tac aac tgg cac cac gga           576
Lys Phe Thr His Glu Lys Pro Glu Gly Tyr Tyr Asn Trp His His Gly
            180                 185                 190

gca gta cag tac tca gga ggc cgg ttc acc atc cct aca ggt gcg ggc           624
Ala Val Gln Tyr Ser Gly Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly
        195                 200                 205

aaa cca ggg gac agc ggt aga ccg atc ttc gac aac aag gga cgc gtg           672
Lys Pro Gly Asp Ser Gly Arg Pro Ile Phe Asp Asn Lys Gly Arg Val
    210                 215                 220

gtg gcc ata gtc tta gga gga gct aat gaa gga gcc cgt aca gcc ctc           720
Val Ala Ile Val Leu Gly Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu
225                 230                 235                 240

tca gtg gtg acc tgg aat aaa gac att gtc act aaa atc acc cct gag           768
Ser Val Val Thr Trp Asn Lys Asp Ile Val Thr Lys Ile Thr Pro Glu
                245                 250                 255

gga gcc gaa gag tgg agt ctt gcc atc cca gtt atg tgc ctg ttg gca           816
Gly Ala Glu Glu Trp Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala
            260                 265                 270
```

[0048]

```
aat acc acg ttc ccc tgc tcc cag ccc cct tgc ata ccc tgc tgc tac           864
Asn Thr Thr Phe Pro Cys Ser Gln Pro Pro Cys Ile Pro Cys Cys Tyr
        275                 280                 285

gaa aag gaa ccg gag gaa acc cta cgc atg ctt gag gac aac gtc atg           912
Glu Lys Glu Pro Glu Glu Thr Leu Arg Met Leu Glu Asp Asn Val Met
    290                 295                 300

aga cct ggg tac tat cag ctg cta caa gca tca tta aca tgt tct ccc           960
Arg Pro Gly Tyr Tyr Gln Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro
305                 310                 315                 320

cac cgc cag cga cgc agc acc aag gac aac ttc aat gtc tat aaa gcc          1008
His Arg Gln Arg Arg Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala
                325                 330                 335

aca aga cca tac tta gct cac tgt ccc gac tgt gga gaa ggg cac tcg          1056
Thr Arg Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser
            340                 345                 350

tgc cat agt ccc gta gca cta gaa cgc atc aga aat gaa gcg aca gac          1104
Cys His Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp
        355                 360                 365

ggg acg ctg aaa atc cag gtc tcc ttg caa att gga ata ggg acg gat          1152
Gly Thr Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Gly Thr Asp
    370                 375                 380

gat agc cat gat tgg acc aag ctg cgt tac atg gac aat cac ata cca          1200
Asp Ser His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Ile Pro
385                 390                 395                 400

gca gac gca ggg agg gcc ggg cta ttt gta aga aca tca gca cca tgc          1248
Ala Asp Ala Gly Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys
                405                 410                 415

acg att act gga aca atg gga cac ttc atc ctg gcc cga tgt ccg aaa          1296
Thr Ile Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys
            420                 425                 430
```

[0049]

```
gga gaa act ctg acg gtg gga ttc act gac agt agg aag att agt cac        1344
Gly Glu Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His
        435                 440                 445

tca tgt acg cac cca ttt cac cac gac cct cct gtg ata ggc cgg gaa        1392
Ser Cys Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu
    450                 455                 460

aaa ttc cat tcc cga ccg cag cac ggt aaa gag cta cct tgc agc acg        1440
Lys Phe His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr
465                 470                 475                 480

tac gtg cag agc aac gcc gca act gcc gag gag ata gag gta cac atg        1488
Tyr Val Gln Ser Asn Ala Ala Thr Ala Glu Glu Ile Glu Val His Met
                485                 490                 495

ccc cca gac acc cct gat cgc aca ttg ctg tca caa cag tcc ggc aac        1536
Pro Pro Asp Thr Pro Asp Arg Thr Leu Leu Ser Gln Gln Ser Gly Asn
            500                 505                 510

gta aag atc aca gtc aac ggc cgg acg gtg cgg tat aag tgt aat tgc        1584
Val Lys Ile Thr Val Asn Gly Arg Thr Val Arg Tyr Lys Cys Asn Cys
        515                 520                 525

ggt ggc tca aat gaa gga cta ata act aca gat aaa gtg att aat aac        1632
Gly Gly Ser Asn Glu Gly Leu Ile Thr Thr Asp Lys Val Ile Asn Asn
    530                 535                 540

tgc aag gtt gat caa tgt cat gcc gcg gtc acc aat cac aaa aag tgg        1680
Cys Lys Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp
545                 550                 555                 560

cag tat aac tcc cct ctg gtc ccg cgt aac gct gaa ctc ggg gac cga        1728
Gln Tyr Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg
                565                 570                 575

aaa gga aaa att cac atc ccg ttt ccg ctg gca aat gta aca tgc atg        1776
Lys Gly Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Met
            580                 585                 590

gtg cct aaa gca agg aac ccc acc gtg acg tac ggg aaa aac caa gtc        1824
Val Pro Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val
        595                 600                 605

atc atg cta ctg tat cct gac cac cca aca ctc ctg tcc tac cgg agt        1872
Ile Met Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Ser
    610                 615                 620

atg gga gaa gaa cca aac tat caa gaa gag tgg gtg acg cac aag aag        1920
Met Gly Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Thr His Lys Lys
625                 630                 635                 640

gag gtc gtg cta acc gtg ccg act gaa ggg ctc gag gtt acg tgg ggc        1968
Glu Val Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly
                645                 650                 655

aac aac gag ccg tat aag tat tgg ccg cag tta tct gca aac ggt aca        2016
Asn Asn Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Ala Asn Gly Thr
            660                 665                 670

gcc cac ggc cac ccg cat gag ata atc ttg tac tat tat gag ctg tac        2064
Ala His Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr
        675                 680                 685

cct act atg act gta gta gtt gtg tca gtg gcc tcg ttc ata ctc ctg        2112
Pro Thr Met Thr Val Val Val Val Ser Val Ala Ser Phe Ile Leu Leu
    690                 695                 700

tcg atg gtg ggt atg gca gtg ggg atg tgc atg tgt gca cga cgc aga        2160
Ser Met Val Gly Met Ala Val Gly Met Cys Met Cys Ala Arg Arg Arg
705                 710                 715                 720

tgc atc aca cca tac gaa ctg aca cca gga gct acc gtc cct ttc ctg        2208
Cys Ile Thr Pro Tyr Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu
                725                 730                 735

ctt agc cta ata tgc tgc atc aga aca gct aaa gcg gcc aca tac caa        2256
```

Leu Ser Leu Ile Cys Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln
740 745 750

gag gct gcg gta tac ctg tgg aac gag cag caa cct ttg ttt tgg cta 2304
Glu Ala Ala Val Tyr Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu
755 760 765

caa gcc ctt att ccg ctg gca gcc ctg att gtc cta tgc aac tgt ctg 2352
Gln Ala Leu Ile Pro Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu
770 775 780

aga ctc tta cca tgc tgt tgt aaa acg ttg gct ttt tta gcc gta atg 2400
Arg Leu Leu Pro Cys Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met
785 790 795 800

agc atc ggt gcc cac act gtg agc gcg tac gaa cac gta aca gtg atc 2448
Ser Ile Gly Ala His Thr Val Ser Ala Tyr Glu His Val Thr Val Ile
805 810 815

ccg aac acg gtg gga gta ccg tat aag act cta gtc aac aga ccg ggc 2496
Pro Asn Thr Val Gly Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly
820 825 830

tac agc ccc atg gta ctg gag atg gag cta ctg tca gtc act ttg gag 2544
Tyr Ser Pro Met Val Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu
835 840 845

cca acg cta tcg ctt gat tac atc acg tgc gaa tac aaa acc gtc atc 2592
Pro Thr Leu Ser Leu Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile
850 855 860

ccg tct ccg tac gtg aaa tgc tgc ggt aca gca gag tgc aag gac aaa 2640
Pro Ser Pro Tyr Val Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys
865 870 875 880

aac cta cct gac tac agc tgt aag gtc ttc acc ggc gtc tac cca ttt 2688
Asn Leu Pro Asp Tyr Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe
885 890 895

[0050]

atg tgg ggc ggc gcc tac tgc ttc tgc gac gct gaa aac acg caa ttg 2736
Met Trp Gly Gly Ala Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu
900 905 910

agc gaa gca cat gtg gag aag tcc gaa tca tgc aaa aca gaa ttt gca 2784
Ser Glu Ala His Val Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala
915 920 925

tca gca tac agg gct cat acc gca tcc gca tca gct aag ctc cgc gtc 2832
Ser Ala Tyr Arg Ala His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val
930 935 940

ctt tac caa gga aat aac atc act gta act gcc tat gca aac ggc gac 2880
Leu Tyr Gln Gly Asn Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp
945 950 955 960

cat gcc gtc aca gtt aag gac gcc aaa ttc att gtg ggg cca atg tct 2928
His Ala Val Thr Val Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser
965 970 975

tca gcc tgg aca cct ttc gac aac aaa atc gtg gtg tac aaa ggt gac 2976
Ser Ala Trp Thr Pro Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp
980 985 990

gtt tac aac atg gac tac ccg ccc ttt ggc gca gga aga cca gga caa 3024
Val Tyr Asn Met Asp Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln
995 1000 1005

ttt ggc gat atc caa agt cgc acg cct gag agc aaa gac gtc tat 3069
Phe Gly Asp Ile Gln Ser Arg Thr Pro Glu Ser Lys Asp Val Tyr
1010 1015 1020

gct aac aca caa ctg gta ctg cag aga ccg gct gcg ggt acg gta 3114
Ala Asn Thr Gln Leu Val Leu Gln Arg Pro Ala Ala Gly Thr Val
1025 1030 1035

cac gtg cca tac tct cag gca cca tct ggc ttt aag tat tgg tta 3159
His Val Pro Tyr Ser Gln Ala Pro Ser Gly Phe Lys Tyr Trp Leu

```
        1040                1045                1050

aaa gaa  cga ggg gcg tcg cta  cag cac aca gca cca  ttt ggc tgc      3204
Lys Glu  Arg Gly Ala Ser Leu  Gln His Thr Ala Pro  Phe Gly Cys
    1055                1060                1065

caa ata  gca aca aac ccg gta  aga gcg atg aac tgc  gcc gta ggg      3249
Gln Ile  Ala Thr Asn Pro Val  Arg Ala Met Asn Cys  Ala Val Gly
    1070                1075                1080

aac atg  ccc atc tcc atc gac  ata ccg gat gcg gcc  ttc act agg      3294
Asn Met  Pro Ile Ser Ile Asp  Ile Pro Asp Ala Ala  Phe Thr Arg
    1085                1090                1095

gtc gtc  gac gcg ccc tct tta  acg gac atg tca tgc  gag gta cca      3339
Val Val  Asp Ala Pro Ser Leu  Thr Asp Met Ser Cys  Glu Val Pro
    1100                1105                1110

gcc tgc  acc cat tcc tca gac  ttt ggg ggc gtc gcc  att att aaa      3384
Ala Cys  Thr His Ser Ser Asp  Phe Gly Gly Val Ala  Ile Ile Lys
    1115                1120                1125

tat gca  gtc agc aag aaa ggc  aag tgt gcg gtg cat  tcg atg acc      3429
Tyr Ala  Val Ser Lys Lys Gly  Lys Cys Ala Val His  Ser Met Thr
    1130                1135                1140

aac gcc  gtc act atc cgg gaa  gct gag ata gaa gtt  gaa ggg aat      3474
Asn Ala  Val Thr Ile Arg Glu  Ala Glu Ile Glu Val  Glu Gly Asn
    1145                1150                1155

tct cag  ctg caa atc tct ttc  tcg acg gcc ttg gcc  agc gcc gaa      3519
Ser Gln  Leu Gln Ile Ser Phe  Ser Thr Ala Leu Ala  Ser Ala Glu
    1160                1165                1170

ttc cgc  gta caa gtc tgt tct  aca caa gta cac tgt  gca gcc gag      3564
Phe Arg  Val Gln Val Cys Ser  Thr Gln Val His Cys  Ala Ala Glu
    1175                1180                1185
```

[0051]

```
tgc cac  cct ccg aag gac cac  ata gtc aac tac ccg  gcg tca cat      3609
Cys His  Pro Pro Lys Asp His  Ile Val Asn Tyr Pro  Ala Ser His
    1190                1195                1200

acc acc  ctc ggg gtc cag gac  att tcc gct acg gcg  atg tca tgg      3654
Thr Thr  Leu Gly Val Gln Asp  Ile Ser Ala Thr Ala  Met Ser Trp
    1205                1210                1215

gtg cag  aag atc acg gga ggt  gtg gga ctg gtt gtc  gct gtt gca      3699
Val Gln  Lys Ile Thr Gly Gly  Val Gly Leu Val Val  Ala Val Ala
    1220                1225                1230

gca ctg  att cta atc gtg gtg  cta tgc gtg tcg ttc  agc agg cac      3744
Ala Leu  Ile Leu Ile Val Val  Leu Cys Val Ser Phe  Ser Arg His
    1235                1240                1245


<210>  21
<211>  1248
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic Construct

<400>  21

Met Glu Phe Ile Pro Thr Gln Thr Phe Tyr Asn Arg Arg Tyr Gln Pro
1               5                   10                  15


Arg Pro Trp Thr Pro Arg Pro Thr Ile Gln Val Ile Arg Pro Arg Pro
            20                  25                  30


Arg Pro Gln Arg Gln Ala Gly Gln Leu Ala Gln Leu Ile Ser Ala Val
        35                  40                  45
```

```
Asn Lys Leu Thr Met Arg Ala Val Pro Gln Gln Lys Pro Arg Arg Asn
    50                  55                  60

Arg Lys Asn Lys Lys Gln Lys Gln Lys Gln Gln Ala Pro Gln Asn Asn
65                  70                  75                  80

Thr Asn Gln Lys Lys Gln Pro Pro Lys Lys Lys Pro Ala Gln Lys Lys
                85                  90                  95

Lys Lys Pro Gly Arg Arg Glu Arg Met Cys Met Lys Ile Glu Asn Asp
            100                 105                 110

Cys Ile Phe Glu Val Lys His Glu Gly Lys Val Thr Gly Tyr Ala Cys
        115                 120                 125

Leu Val Gly Asp Lys Val Met Lys Pro Ala His Val Lys Gly Thr Ile
    130                 135                 140

Asp Asn Ala Asp Leu Ala Lys Leu Ala Phe Lys Arg Ser Ser Lys Tyr
145                 150                 155                 160

Asp Leu Glu Cys Ala Gln Ile Pro Val His Met Lys Ser Asp Ala Ser
                165                 170                 175

Lys Phe Thr His Glu Lys Pro Glu Gly Tyr Tyr Asn Trp His His Gly
            180                 185                 190

Ala Val Gln Tyr Ser Gly Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly
        195                 200                 205
```

[0052]

```
Lys Pro Gly Asp Ser Gly Arg Pro Ile Phe Asp Asn Lys Gly Arg Val
    210                 215                 220

Val Ala Ile Val Leu Gly Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu
225                 230                 235                 240

Ser Val Val Thr Trp Asn Lys Asp Ile Val Thr Lys Ile Thr Pro Glu
                245                 250                 255

Gly Ala Glu Glu Trp Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala
            260                 265                 270

Asn Thr Thr Phe Pro Cys Ser Gln Pro Pro Cys Ile Pro Cys Cys Tyr
        275                 280                 285

Glu Lys Glu Pro Glu Glu Thr Leu Arg Met Leu Glu Asp Asn Val Met
    290                 295                 300

Arg Pro Gly Tyr Tyr Gln Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro
305                 310                 315                 320

His Arg Gln Arg Arg Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala
                325                 330                 335

Thr Arg Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser
            340                 345                 350

Cys His Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp
```

355 360 365

Gly Thr Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Gly Thr Asp
370 375 380

Asp Ser His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Ile Pro
385 390 395 400

Ala Asp Ala Gly Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys
405 410 415

Thr Ile Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys
420 425 430

Gly Glu Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His
435 440 445

Ser Cys Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu
450 455 460

Lys Phe His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr
465 470 475 480

Tyr Val Gln Ser Asn Ala Ala Thr Ala Glu Glu Ile Glu Val His Met
485 490 495

Pro Pro Asp Thr Pro Asp Arg Thr Leu Leu Ser Gln Gln Ser Gly Asn
500 505 510

[0053]

Val Lys Ile Thr Val Asn Gly Arg Thr Val Arg Tyr Lys Cys Asn Cys
515 520 525

Gly Gly Ser Asn Glu Gly Leu Ile Thr Thr Asp Lys Val Ile Asn Asn
530 535 540

Cys Lys Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp
545 550 555 560

Gln Tyr Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg
565 570 575

Lys Gly Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Met
580 585 590

Val Pro Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val
595 600 605

Ile Met Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Ser
610 615 620

Met Gly Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Thr His Lys Lys
625 630 635 640

Glu Val Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly
645 650 655

Asn Asn Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Ala Asn Gly Thr
660 665 670

Ala His Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr
675 680 685

Pro Thr Met Thr Val Val Val Val Ser Val Ala Ser Phe Ile Leu Leu
690 695 700

Ser Met Val Gly Met Ala Val Gly Met Cys Met Cys Ala Arg Arg Arg
705 710 715 720

Cys Ile Thr Pro Tyr Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu
725 730 735

Leu Ser Leu Ile Cys Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln
740 745 750

Glu Ala Ala Val Tyr Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu
755 760 765

Gln Ala Leu Ile Pro Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu
770 775 780

Arg Leu Leu Pro Cys Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met
785 790 795 800

Ser Ile Gly Ala His Thr Val Ser Ala Tyr Glu His Val Thr Val Ile
805 810 815

[0054] Pro Asn Thr Val Gly Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly
820 825 830

Tyr Ser Pro Met Val Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu
835 840 845

Pro Thr Leu Ser Leu Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile
850 855 860

Pro Ser Pro Tyr Val Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys
865 870 875 880

Asn Leu Pro Asp Tyr Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe
885 890 895

Met Trp Gly Gly Ala Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu
900 905 910

Ser Glu Ala His Val Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala
915 920 925

Ser Ala Tyr Arg Ala His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val
930 935 940

Leu Tyr Gln Gly Asn Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp
945 950 955 960

His Ala Val Thr Val Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser
965 970 975

Ser Ala Trp Thr Pro Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp
980 985 990

Val Tyr Asn Met Asp Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln
995 1000 1005

Phe Gly Asp Ile Gln Ser Arg Thr Pro Glu Ser Lys Asp Val Tyr
1010 1015 1020

Ala Asn Thr Gln Leu Val Leu Gln Arg Pro Ala Ala Gly Thr Val
1025 1030 1035

His Val Pro Tyr Ser Gln Ala Pro Ser Gly Phe Lys Tyr Trp Leu
1040 1045 1050

Lys Glu Arg Gly Ala Ser Leu Gln His Thr Ala Pro Phe Gly Cys
1055 1060 1065

Gln Ile Ala Thr Asn Pro Val Arg Ala Met Asn Cys Ala Val Gly
1070 1075 1080

Asn Met Pro Ile Ser Ile Asp Ile Pro Asp Ala Ala Phe Thr Arg
1085 1090 1095

Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser Cys Glu Val Pro
1100 1105 1110

[0055] Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala Ile Ile Lys
1115 1120 1125

Tyr Ala Val Ser Lys Lys Gly Lys Cys Ala Val His Ser Met Thr
1130 1135 1140

Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly Asn
1145 1150 1155

Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
1160 1165 1170

Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu
1175 1180 1185

Cys His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His
1190 1195 1200

Thr Thr Leu Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp
1205 1210 1215

Val Gln Lys Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala
1220 1225 1230

Ala Leu Ile Leu Ile Val Val Leu Cys Val Ser Phe Ser Arg His
1235 1240 1245

<210> 22
<211> 1248
<212> PRT
<213> Artificial Sequence

<220>
<223> proteins C-E3-E2-6K-E1 strain 05.61

<400> 22

Met Glu Phe Ile Pro Thr Gln Thr Phe Tyr Asn Arg Arg Tyr Gln Pro
1 5 10 15

Arg Pro Trp Thr Pro Arg Pro Thr Ile Gln Val Ile Arg Pro Arg Pro
20 25 30

Arg Pro Gln Arg Gln Ala Gly Gln Leu Ala Gln Leu Ile Ser Ala Val
35 40 45

Asn Lys Leu Thr Met Arg Ala Val Pro Gln Gln Lys Pro Arg Arg Asn
50 55 60

Arg Lys Asn Lys Lys Gln Lys Gln Lys Gln Gln Ala Pro Gln Asn Asn
65 70 75 80

Thr Asn Gln Lys Lys Gln Pro Pro Lys Lys Lys Pro Ala Gln Lys Lys
85 90 95

Lys Lys Pro Gly Arg Arg Glu Arg Met Cys Met Lys Ile Glu Asn Asp
100 105 110

Cys Ile Phe Glu Val Lys His Glu Gly Lys Val Thr Gly Tyr Ala Cys
115 120 125

[0056] Leu Val Gly Asp Lys Val Met Lys Pro Ala His Val Lys Gly Thr Ile
130 135 140

Asp Asn Ala Asp Leu Ala Lys Leu Ala Phe Lys Arg Ser Ser Lys Tyr
145 150 155 160

Asp Leu Glu Cys Ala Gln Ile Pro Val His Met Lys Ser Asp Ala Ser
165 170 175

Lys Phe Thr His Glu Lys Pro Glu Gly Tyr Tyr Asn Trp His His Gly
180 185 190

Ala Val Gln Tyr Ser Gly Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly
195 200 205

Lys Pro Gly Asp Ser Gly Arg Pro Ile Phe Asp Asn Lys Gly Arg Val
210 215 220

Val Ala Ile Val Leu Gly Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu
225 230 235 240

Ser Val Val Thr Trp Asn Lys Asp Ile Val Thr Lys Ile Thr Pro Glu
245 250 255

Gly Ala Glu Glu Trp Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala
260 265 270

Asn Thr Thr Phe Pro Cys Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr
275 280 285

Glu Lys Glu Pro Glu Glu Thr Leu Arg Met Leu Glu Asp Asn Val Met
290 295 300

Arg Pro Gly Tyr Tyr Gln Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro
305 310 315 320

His Arg Gln Arg Arg Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala
325 330 335

Thr Arg Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser
340 345 350

Cys His Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp
355 360 365

Gly Thr Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp
370 375 380

Asp Ser His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro
385 390 395 400

Ala Asp Ala Glu Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys
405 410 415

Thr Ile Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys
420 425 430

[0057] Gly Glu Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His
435 440 445

Ser Cys Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu
450 455 460

Lys Phe His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr
465 470 475 480

Tyr Val Gln Ser Thr Ala Ala Thr Thr Glu Glu Ile Glu Val His Met
485 490 495

Pro Pro Asp Thr Pro Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn
500 505 510

Val Lys Ile Thr Val Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys
515 520 525

Gly Gly Ser Asn Glu Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn
530 535 540

Cys Lys Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp
545 550 555 560

Gln Tyr Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg
565 570 575

Lys Gly Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg
580 585 590

Val Pro Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val
595 600 605

Ile Met Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn
610 615 620

Met Gly Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys
625 630 635 640

Glu Val Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly
645 650 655

Asn Asn Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr
660 665 670

Ala His Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr
675 680 685

Pro Thr Met Thr Val Val Val Val Ser Val Ala Thr Phe Ile Leu Leu
690 695 700

Ser Met Val Gly Met Ala Ala Gly Met Cys Met Cys Ala Arg Arg Arg
705 710 715 720

Cys Ile Thr Pro Tyr Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu
725 730 735

Leu Ser Leu Ile Cys Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln
740 745 750

[0058]

Glu Ala Ala Ile Tyr Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu
755 760 765

Gln Ala Leu Ile Pro Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu
770 775 780

Arg Leu Leu Pro Cys Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met
785 790 795 800

Ser Val Gly Ala His Thr Val Ser Ala Tyr Glu His Val Thr Val Ile
805 810 815

Pro Asn Thr Val Gly Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly
820 825 830

Tyr Ser Pro Met Val Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu
835 840 845

Pro Thr Leu Ser Leu Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile
850 855 860

Pro Ser Pro Tyr Val Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys
865 870 875 880

Asn Leu Pro Asp Tyr Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe
885 890 895

Met Trp Gly Gly Ala Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu

900 905 910

Ser Glu Ala His Val Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala
915 920 925

Ser Ala Tyr Arg Ala His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val
930 935 940

Leu Tyr Gln Gly Asn Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp
945 950 955 960

His Ala Val Thr Val Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser
965 970 975

Ser Ala Trp Thr Pro Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp
980 985 990

Val Tyr Asn Met Asp Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln
995 1000 1005

Phe Gly Asp Ile Gln Ser Arg Thr Pro Glu Ser Lys Asp Val Tyr
1010 1015 1020

Ala Asn Thr Gln Leu Val Leu Gln Arg Pro Ala Ala Gly Thr Val
1025 1030 1035

His Val Pro Tyr Ser Gln Ala Pro Ser Gly Phe Lys Tyr Trp Leu
1040 1045 1050

[0059] Lys Glu Arg Gly Ala Ser Leu Gln His Thr Ala Pro Phe Gly Cys
1055 1060 1065

Gln Ile Ala Thr Asn Pro Val Arg Ala Val Asn Cys Ala Val Gly
1070 1075 1080

Asn Met Pro Ile Ser Ile Asp Ile Pro Glu Ala Ala Phe Thr Arg
1085 1090 1095

Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser Cys Glu Val Pro
1100 1105 1110

Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala Ile Ile Lys
1115 1120 1125

Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser Met Thr
1130 1135 1140

Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly Asn
1145 1150 1155

Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
1160 1165 1170

Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu
1175 1180 1185

Cys His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His
1190 1195 1200

Thr Thr Leu Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp
1205 1210 1215

Val Gln Lys Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala
1220 1225 1230

Ala Leu Ile Leu Ile Val Val Leu Cys Val Ser Phe Ser Arg His
1235 1240 1245

<210> 23
<211> 1248
<212> PRT
<213> Artificial Sequence

<220>
<223> proteins C-E3-E2-6K-E1 strain 05.209

<400> 23

Met Glu Phe Ile Pro Thr Gln Thr Phe Tyr Asn Arg Arg Tyr Gln Pro
1 5 10 15

Arg Pro Trp Thr Pro Arg Pro Thr Ile Gln Val Ile Arg Pro Arg Pro
20 25 30

Arg Pro Gln Arg Gln Ala Gly Gln Leu Ala Gln Leu Ile Ser Ala Val
35 40 45

Asn Lys Leu Thr Met Arg Ala Val Pro Gln Gln Lys Pro Arg Arg Asn
50 55 60

[0060] Arg Lys Asn Lys Lys Gln Lys Gln Lys Gln Gln Ala Pro Gln Asn Asn
65 70 75 80

Thr Asn Gln Lys Lys Gln Pro Pro Lys Lys Lys Pro Ala Gln Lys Lys
85 90 95

Lys Lys Pro Gly Arg Arg Glu Arg Met Cys Met Lys Ile Glu Asn Asp
100 105 110

Cys Ile Phe Glu Val Lys His Glu Gly Lys Val Thr Gly Tyr Ala Cys
115 120 125

Leu Val Gly Asp Lys Val Met Lys Pro Ala His Val Lys Gly Thr Ile
130 135 140

Asp Asn Ala Asp Leu Ala Lys Leu Ala Phe Lys Arg Ser Ser Lys Tyr
145 150 155 160

Asp Leu Glu Cys Ala Gln Ile Pro Val His Met Lys Ser Asp Ala Ser
165 170 175

Lys Phe Thr His Glu Lys Pro Glu Gly Tyr Tyr Asn Trp His His Gly
180 185 190

Ala Val Gln Tyr Ser Gly Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly
195 200 205

Lys Pro Gly Asp Ser Gly Arg Pro Ile Phe Asp Asn Lys Gly Arg Val

210 215 220

Val Ala Ile Val Leu Gly Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu
225 230 235 240

Ser Val Val Thr Trp Asn Lys Asp Ile Val Thr Lys Ile Thr Pro Glu
245 250 255

Gly Ala Glu Glu Trp Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala
260 265 270

Asn Thr Thr Phe Pro Cys Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr
275 280 285

Glu Lys Glu Pro Glu Glu Thr Leu Arg Met Leu Glu Asp Asn Val Met
290 295 300

Arg Pro Gly Tyr Tyr Gln Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro
305 310 315 320

His Arg Gln Arg Arg Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala
325 330 335

Thr Arg Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser
340 345 350

Cys His Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp
355 360 365

[0061] Gly Thr Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp
370 375 380

Asp Ser His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro
385 390 395 400

Ala Asp Ala Glu Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys
405 410 415

Thr Ile Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys
420 425 430

Gly Glu Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His
435 440 445

Ser Cys Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu
450 455 460

Lys Phe His Ser Arg Pro Arg His Gly Lys Glu Leu Pro Cys Ser Thr
465 470 475 480

Tyr Val Gln Ser Thr Ala Ala Thr Thr Glu Glu Ile Glu Val His Met
485 490 495

Pro Pro Asp Thr Pro Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn
500 505 510

Val Lys Ile Thr Val Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys
515 520 525

Gly Gly Ser Asn Glu Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn
530 535 540

Cys Lys Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp
545 550 555 560

Gln Tyr Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg
565 570 575

Lys Gly Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg
580 585 590

Val Pro Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val
595 600 605

Ile Met Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn
610 615 620

Met Gly Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys
625 630 635 640

Glu Val Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly
645 650 655

Asn Asn Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr
660 665 670

[0062] Ala His Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr
675 680 685

Pro Thr Met Thr Val Val Val Val Ser Val Ala Thr Phe Ile Leu Leu
690 695 700

Ser Met Val Gly Met Ala Ala Gly Met Cys Met Cys Ala Arg Arg Arg
705 710 715 720

Cys Ile Thr Pro Tyr Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu
725 730 735

Leu Ser Leu Ile Cys Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln
740 745 750

Glu Ala Ala Ile Tyr Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu
755 760 765

Gln Ala Leu Ile Pro Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu
770 775 780

Arg Leu Leu Pro Cys Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met
785 790 795 800

Ser Val Gly Ala His Thr Val Ser Ala Tyr Glu His Val Thr Val Ile
805 810 815

Pro Asn Thr Val Gly Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly
820 825 830

Tyr Ser Pro Met Val Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu
835 840 845

Pro Thr Leu Ser Leu Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile
850 855 860

Pro Ser Pro Tyr Val Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys
865 870 875 880

Asn Leu Pro Asp Tyr Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe
885 890 895

Met Trp Gly Gly Ala Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu
900 905 910

Ser Glu Ala His Val Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala
915 920 925

Ser Ala Tyr Arg Ala His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val
930 935 940

Leu Tyr Gln Gly Asn Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp
945 950 955 960

His Ala Val Thr Val Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser
965 970 975

[0063] Ser Ala Trp Thr Pro Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp
980 985 990

Val Tyr Asn Met Asp Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln
995 1000 1005

Phe Gly Asp Ile Gln Ser Arg Thr Pro Glu Ser Lys Asp Val Tyr
1010 1015 1020

Ala Asn Thr Gln Leu Val Leu Gln Arg Pro Ala Ala Gly Thr Val
1025 1030 1035

His Val Pro Tyr Ser Gln Ala Pro Ser Gly Phe Lys Tyr Trp Leu
1040 1045 1050

Lys Glu Arg Gly Ala Ser Leu Gln His Thr Ala Pro Phe Gly Cys
1055 1060 1065

Gln Ile Ala Thr Asn Pro Val Arg Ala Val Asn Cys Ala Val Gly
1070 1075 1080

Asn Met Pro Ile Ser Ile Asp Ile Pro Glu Ala Ala Phe Thr Arg
1085 1090 1095

Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser Cys Glu Val Pro
1100 1105 1110

Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala Ile Ile Lys
1115 1120 1125

Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser Met Thr
1130 1135 1140

Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly Asn
1145 1150 1155

Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
1160 1165 1170

Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu
1175 1180 1185

Cys His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His
1190 1195 1200

Thr Thr Leu Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp
1205 1210 1215

Val Gln Lys Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala
1220 1225 1230

Ala Leu Ile Leu Ile Val Val Leu Cys Val Ser Phe Ser Arg His
1235 1240 1245

<210> 24
<211> 1248
<212> PRT
<213> Artificial Sequence

[0064] <220>
<223> proteins C-E3-E2-6K-E1 strain 05.115

<400> 24

Met Glu Phe Ile Pro Thr Gln Thr Phe Tyr Asn Arg Arg Tyr Gln Pro
1 5 10 15

Arg Pro Trp Thr Pro Arg Pro Thr Ile Gln Val Ile Arg Pro Arg Pro
20 25 30

Arg Pro Gln Arg Gln Ala Gly Gln Leu Ala Gln Leu Ile Ser Ala Val
35 40 45

Asn Lys Leu Thr Met Arg Ala Val Pro Gln Gln Lys Pro Arg Arg Asn
50 55 60

Arg Lys Asn Lys Lys Gln Lys Gln Lys Gln Gln Ala Pro Gln Asn Asn
65 70 75 80

Thr Asn Gln Lys Lys Gln Pro Pro Lys Lys Lys Pro Ala Gln Lys Lys
85 90 95

Lys Lys Pro Gly Arg Arg Glu Arg Met Cys Met Lys Ile Glu Asn Asp
100 105 110

Cys Ile Phe Glu Val Lys His Glu Gly Lys Val Thr Gly Tyr Ala Cys
115 120 125

Leu Val Gly Asp Lys Val Met Lys Pro Ala His Val Lys Gly Thr Ile
130 135 140

Asp Asn Ala Asp Leu Ala Lys Leu Ala Phe Lys Arg Ser Ser Lys Tyr
145 150 155 160

Asp Leu Glu Cys Ala Gln Ile Pro Val His Met Lys Ser Asp Ala Ser
165 170 175

Lys Phe Thr His Glu Lys Pro Glu Gly Tyr Tyr Asn Trp His His Gly
180 185 190

Ala Val Gln Tyr Ser Gly Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly
195 200 205

Lys Pro Gly Asp Ser Gly Arg Pro Ile Phe Asp Asn Lys Gly Arg Val
210 215 220

Val Ala Ile Val Leu Gly Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu
225 230 235 240

Ser Val Val Thr Trp Asn Lys Asp Ile Val Thr Lys Ile Thr Pro Glu
245 250 255

Gly Ala Glu Glu Trp Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala
260 265 270

Asn Thr Thr Phe Pro Cys Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr
275 280 285

[0065] Glu Lys Glu Pro Glu Glu Thr Leu Arg Met Leu Glu Asp Asn Val Met
290 295 300

Arg Pro Gly Tyr Tyr Gln Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro
305 310 315 320

His Arg Gln Arg Arg Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala
325 330 335

Thr Arg Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser
340 345 350

Cys His Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp
355 360 365

Gly Thr Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp
370 375 380

Asp Ser His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro
385 390 395 400

Ala Asp Ala Glu Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys
405 410 415

Thr Ile Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys
420 425 430

Gly Glu Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His
435 440 445

```
Ser Cys Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu
    450                 455                 460

Lys Phe His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr
465                 470                 475                 480

Tyr Val Gln Ser Thr Ala Ala Thr Thr Glu Glu Ile Glu Val His Met
                485                 490                 495

Pro Pro Asp Thr Pro Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn
            500                 505                 510

Val Lys Ile Thr Val Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys
        515                 520                 525

Gly Gly Ser Asn Glu Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn
    530                 535                 540

Cys Lys Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp
545                 550                 555                 560

Gln Tyr Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg
                565                 570                 575

Lys Gly Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg
            580                 585                 590

Val Pro Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val
        595                 600                 605
```

[0066]

```
Ile Met Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn
    610                 615                 620

Met Gly Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys
625                 630                 635                 640

Glu Val Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly
                645                 650                 655

Asn Asn Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr
            660                 665                 670

Ala His Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr
        675                 680                 685

Pro Thr Met Thr Val Val Val Val Ser Val Ala Thr Phe Ile Leu Leu
    690                 695                 700

Ser Met Val Gly Met Ala Ala Gly Met Cys Met Cys Ala Arg Arg Arg
705                 710                 715                 720

Cys Ile Thr Pro Tyr Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu
                725                 730                 735

Leu Ser Leu Ile Cys Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln
            740                 745                 750

Glu Ala Ala Ile Tyr Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu
```

755 760 765

Gln Ala Leu Ile Pro Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu
770 775 780

Arg Leu Leu Pro Cys Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met
785 790 795 800

Ser Val Gly Ala His Thr Val Ser Ala Tyr Glu His Val Thr Val Ile
805 810 815

Pro Asn Thr Val Gly Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly
820 825 830

Tyr Ser Pro Met Val Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu
835 840 845

Pro Thr Leu Ser Leu Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile
850 855 860

Pro Ser Pro Tyr Val Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys
865 870 875 880

Asn Leu Pro Asp Tyr Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe
885 890 895

Met Trp Gly Gly Ala Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu
900 905 910

[0067] Ser Glu Ala His Val Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala
915 920 925

Ser Ala Tyr Arg Ala His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val
930 935 940

Leu Tyr Gln Gly Asn Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp
945 950 955 960

His Ala Val Thr Val Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser
965 970 975

Ser Ala Trp Thr Pro Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp
980 985 990

Val Tyr Asn Met Asp Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln
995 1000 1005

Phe Gly Asp Ile Gln Ser Arg Thr Pro Glu Ser Lys Asp Val Tyr
1010 1015 1020

Ala Asn Thr Gln Leu Val Leu Gln Arg Pro Ala Ala Gly Thr Val
1025 1030 1035

His Val Pro Tyr Ser Gln Ala Pro Ser Gly Phe Lys Tyr Trp Leu
1040 1045 1050

Lys Glu Arg Gly Ala Ser Leu Gln His Thr Ala Pro Phe Gly Cys
1055 1060 1065

Gln Ile Ala Thr Asn Pro Val Arg Ala Val Asn Cys Ala Val Gly
1070 1075 1080

Asn Met Pro Ile Ser Ile Asp Ile Pro Glu Ala Ala Phe Thr Arg
1085 1090 1095

Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser Cys Glu Val Pro
1100 1105 1110

Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala Ile Ile Lys
1115 1120 1125

Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser Met Thr
1130 1135 1140

Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly Asn
1145 1150 1155

Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
1160 1165 1170

Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu
1175 1180 1185

Cys His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His
1190 1195 1200

[0068] Thr Thr Leu Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp
1205 1210 1215

Val Gln Lys Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala
1220 1225 1230

Ala Leu Ile Leu Ile Val Val Leu Cys Val Ser Phe Ser Arg His
1235 1240 1245

<210> 25
<211> 1248
<212> PRT
<213> Artificial Sequence

<220>
<223> proteins C-E3-E2-6K-E1 strain 06.27

<400> 25

Met Glu Phe Ile Pro Thr Gln Thr Phe Tyr Asn Arg Arg Tyr Gln Pro
1 5 10 15

Arg Pro Trp Thr Pro Arg Pro Thr Ile Gln Val Ile Arg Pro Arg Pro
20 25 30

Arg Pro Gln Arg Gln Ala Gly Gln Leu Ala Gln Leu Ile Ser Ala Val
35 40 45

Asn Lys Leu Thr Met Arg Ala Val Pro Gln Gln Lys Pro Arg Arg Asn
50 55 60

Arg Lys Asn Lys Lys Gln Lys Gln Lys Gln Gln Ala Pro Gln Asn Asn

65 70 75 80

Thr Asn Gln Lys Lys Gln Pro Pro Lys Lys Lys Pro Ala Gln Lys Lys
85 90 95

Lys Lys Pro Gly Arg Arg Glu Arg Met Cys Met Lys Ile Glu Asn Asp
100 105 110

Cys Ile Phe Glu Val Lys His Glu Gly Lys Val Thr Gly Tyr Ala Cys
115 120 125

Leu Val Gly Asp Lys Val Met Lys Pro Ala His Val Lys Gly Thr Ile
130 135 140

Asp Asn Ala Asp Leu Ala Lys Leu Ala Phe Lys Arg Ser Ser Lys Tyr
145 150 155 160

Asp Leu Glu Cys Ala Gln Ile Pro Val His Met Lys Ser Asp Ala Ser
165 170 175

Lys Phe Thr His Glu Lys Pro Glu Gly Tyr Tyr Asn Trp His His Gly
180 185 190

Ala Val Gln Tyr Ser Gly Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly
195 200 205

Lys Pro Gly Asp Ser Gly Arg Pro Ile Phe Asp Asn Lys Gly Arg Val
210 215 220

[0069] Val Ala Ile Val Leu Gly Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu
225 230 235 240

Ser Val Val Thr Trp Asn Lys Asp Ile Val Thr Lys Ile Thr Pro Glu
245 250 255

Gly Ala Glu Glu Trp Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala
260 265 270

Asn Thr Thr Phe Pro Cys Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr
275 280 285

Glu Lys Glu Pro Glu Glu Thr Leu Arg Met Leu Glu Asp Asn Val Met
290 295 300

Arg Pro Gly Tyr Tyr Gln Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro
305 310 315 320

His Arg Gln Arg Arg Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala
325 330 335

Thr Arg Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser
340 345 350

Cys His Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp
355 360 365

Gly Thr Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp
370 375 380

Asp Ser His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro
385   390   395   400

Ala Asp Ala Glu Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys
                405   410   415

Thr Ile Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys
            420   425   430

Gly Glu Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His
        435   440   445

Ser Cys Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu
    450   455   460

Lys Phe His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr
465   470   475   480

Tyr Val Gln Ser Thr Ala Ala Thr Thr Glu Glu Ile Glu Val His Met
                485   490   495

Pro Pro Asp Thr Pro Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn
            500   505   510

Val Lys Ile Thr Val Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys
        515   520   525

[0070] Gly Gly Ser Asn Glu Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn
    530   535   540

Cys Lys Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp
545   550   555   560

Gln Tyr Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg
                565   570   575

Lys Gly Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg
            580   585   590

Val Pro Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val
        595   600   605

Ile Met Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn
    610   615   620

Met Gly Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys
625   630   635   640

Glu Val Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly
                645   650   655

Asn Asn Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr
            660   665   670

Ala His Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr
        675   680   685

Pro Thr Met Thr Val Val Val Val Ser Val Ala Thr Phe Ile Leu Leu
690 695 700

Ser Met Val Gly Met Ala Ala Gly Met Cys Met Cys Ala Arg Arg Arg
705 710 715 720

Cys Ile Thr Pro Tyr Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu
725 730 735

Leu Ser Leu Ile Cys Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln
740 745 750

Glu Ala Ala Ile Tyr Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu
755 760 765

Gln Ala Leu Ile Pro Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu
770 775 780

Arg Leu Leu Pro Cys Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met
785 790 795 800

Ser Val Gly Ala His Thr Val Ser Ala Tyr Glu His Val Thr Val Ile
805 810 815

Pro Asn Thr Val Gly Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly
820 825 830

[0071] Tyr Ser Pro Met Val Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu
835 840 845

Pro Thr Leu Ser Leu Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile
850 855 860

Pro Ser Pro Tyr Val Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys
865 870 875 880

Asn Leu Pro Asp Tyr Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe
885 890 895

Met Trp Gly Gly Ala Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu
900 905 910

Ser Glu Ala His Val Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala
915 920 925

Ser Ala Tyr Arg Ala His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val
930 935 940

Leu Tyr Gln Gly Asn Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp
945 950 955 960

His Ala Val Thr Val Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser
965 970 975

Ser Ala Trp Thr Pro Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp
980 985 990

Val Tyr Asn Met Asp Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln
995 1000 1005

Phe Gly Asp Ile Gln Ser Arg Thr Pro Glu Ser Lys Asp Val Tyr
1010 1015 1020

Ala Asn Thr Gln Leu Val Leu Gln Arg Pro Ala Val Gly Thr Val
1025 1030 1035

His Val Pro Tyr Ser Gln Ala Pro Ser Gly Phe Lys Tyr Trp Leu
1040 1045 1050

Lys Glu Arg Gly Ala Ser Leu Gln His Thr Ala Pro Phe Gly Cys
1055 1060 1065

Gln Ile Ala Thr Asn Pro Val Arg Ala Val Asn Cys Ala Val Gly
1070 1075 1080

Asn Met Pro Ile Ser Ile Asp Ile Pro Glu Ala Ala Phe Thr Arg
1085 1090 1095

Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser Cys Glu Val Pro
1100 1105 1110

Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala Ile Ile Lys
1115 1120 1125

Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser Met Thr
1130 1135 1140

[0072]

Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly Asn
1145 1150 1155

Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
1160 1165 1170

Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu
1175 1180 1185

Cys His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His
1190 1195 1200

Thr Thr Leu Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp
1205 1210 1215

Val Gln Lys Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala
1220 1225 1230

Ala Leu Ile Leu Ile Val Val Leu Cys Val Ser Phe Ser Arg His
1235 1240 1245

<210> 26
<211> 1248
<212> PRT
<213> Artificial Sequence

<220>
<223> proteins C-E3-E2-6K-E1 strain 06.21

<400> 26

Met Glu Phe Ile Pro Thr Gln Thr Phe Tyr Asn Arg Arg Tyr Gln Pro
1 5 10 15

Arg Pro Trp Thr Pro Arg Pro Thr Ile Gln Val Ile Arg Pro Arg Pro
20 25 30

Arg Pro Gln Arg Gln Ala Gly Gln Leu Ala Gln Leu Ile Ser Ala Val
35 40 45

Asn Lys Leu Thr Met Arg Ala Val Pro Gln Gln Lys Pro Arg Arg Asn
50 55 60

Arg Lys Asn Lys Lys Gln Lys Gln Lys Gln Gln Ala Pro Gln Asn Asn
65 70 75 80

Thr Asn Gln Lys Lys Gln Pro Pro Lys Lys Lys Pro Ala Gln Lys Lys
85 90 95

Lys Lys Pro Gly Arg Arg Glu Arg Met Cys Met Lys Ile Glu Asn Asp
100 105 110

Cys Ile Phe Glu Val Lys His Glu Gly Lys Val Thr Gly Tyr Ala Cys
115 120 125

Leu Val Gly Asp Lys Val Met Lys Pro Ala His Val Lys Gly Thr Ile
130 135 140

[0073] Asp Asn Ala Asp Leu Ala Lys Leu Ala Phe Lys Arg Ser Ser Lys Tyr
145 150 155 160

Asp Leu Glu Cys Ala Gln Ile Pro Val His Met Lys Ser Asp Ala Ser
165 170 175

Lys Phe Thr His Glu Lys Pro Glu Gly Tyr Tyr Asn Trp His His Gly
180 185 190

Ala Val Gln Tyr Ser Gly Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly
195 200 205

Lys Pro Gly Asp Ser Gly Arg Pro Ile Phe Asp Asn Lys Gly Arg Val
210 215 220

Val Ala Ile Val Leu Gly Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu
225 230 235 240

Ser Val Val Thr Trp Asn Lys Asp Ile Val Thr Lys Ile Thr Pro Glu
245 250 255

Gly Ala Glu Glu Trp Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala
260 265 270

Asn Thr Thr Phe Pro Cys Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr
275 280 285

Glu Lys Glu Pro Glu Glu Thr Leu Arg Met Leu Glu Asp Asn Val Met
290 295 300

Arg Pro Gly Tyr Tyr Gln Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro
305 310 315 320

His Arg Gln Arg Arg Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala
325 330 335

Thr Arg Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser
340 345 350

Cys His Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp
355 360 365

Gly Thr Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp
370 375 380

Asp Ser His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro
385 390 395 400

Ala Asp Ala Glu Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys
405 410 415

Thr Ile Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys
420 425 430

Gly Glu Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His
435 440 445

Ser Cys Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu
450 455 460

[0074]

Lys Phe His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr
465 470 475 480

Tyr Val Gln Ser Thr Ala Ala Thr Thr Glu Glu Ile Glu Val His Met
485 490 495

Pro Pro Asp Thr Pro Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn
500 505 510

Val Lys Ile Thr Val Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys
515 520 525

Gly Gly Ser Asn Glu Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn
530 535 540

Cys Lys Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp
545 550 555 560

Gln Tyr Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg
565 570 575

Lys Gly Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg
580 585 590

Val Pro Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val
595 600 605

Ile Met Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn

610 615 620

Met Gly Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys
625 630 635 640

Glu Val Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly
645 650 655

Asn Asn Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr
660 665 670

Ala His Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr
675 680 685

Pro Thr Met Thr Val Val Val Val Ser Val Ala Thr Phe Ile Leu Leu
690 695 700

Ser Met Val Gly Met Ala Ala Gly Met Cys Met Cys Ala Arg Arg Arg
705 710 715 720

Cys Ile Thr Pro Tyr Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu
725 730 735

Leu Ser Leu Ile Cys Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln
740 745 750

Glu Ala Ala Ile Tyr Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu
755 760 765

[0075]

Gln Ala Leu Ile Pro Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu
770 775 780

Arg Leu Leu Pro Cys Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met
785 790 795 800

Ser Val Gly Ala His Thr Val Ser Ala Tyr Glu His Val Thr Val Ile
805 810 815

Pro Asn Thr Val Gly Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly
820 825 830

Tyr Ser Pro Met Val Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu
835 840 845

Pro Thr Leu Ser Leu Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile
850 855 860

Pro Ser Pro Tyr Val Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys
865 870 875 880

Asn Leu Pro Asp Tyr Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe
885 890 895

Met Trp Gly Gly Ala Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu
900 905 910

Ser Glu Ala His Val Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala
915 920 925

```
Ser Ala Tyr Arg Ala His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val
    930                 935                 940

Leu Tyr Gln Gly Asn Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp
945                 950                 955                 960

His Ala Val Thr Val Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser
                965                 970                 975

Ser Ala Trp Thr Pro Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp
            980                 985                 990

Val Tyr Asn Met Asp Tyr Pro Pro  Phe Gly Ala Gly Arg  Pro Gly Gln
        995                 1000                 1005

Phe Gly  Asp Ile Gln Ser Arg  Thr Pro Glu Ser Lys  Asp Val Tyr
    1010                 1015                 1020

Ala Asn  Thr Gln Leu Val Leu  Gln Arg Pro Ala Val  Gly Thr Val
    1025                 1030                 1035

His Val  Pro Tyr Ser Gln Ala  Pro Ser Gly Phe Lys  Tyr Trp Leu
    1040                 1045                 1050

Lys Glu  Arg Gly Ala Ser Leu  Gln His Thr Ala Pro  Phe Gly Cys
    1055                 1060                 1065
```

[0076]

```
Gln Ile  Ala Thr Asn Pro Val  Arg Ala Val Asn Cys  Ala Val Gly
    1070                 1075                 1080

Asn Met  Pro Ile Ser Ile Asp  Ile Pro Glu Ala Ala  Phe Thr Arg
    1085                 1090                 1095

Val Val  Asp Ala Pro Ser Leu  Thr Asp Met Ser Cys  Glu Val Pro
    1100                 1105                 1110

Ala Cys  Thr His Ser Ser Asp  Phe Gly Gly Val Ala  Ile Ile Lys
    1115                 1120                 1125

Tyr Ala  Ala Ser Lys Lys Gly  Lys Cys Ala Val His  Ser Met Thr
    1130                 1135                 1140

Asn Ala  Val Thr Ile Arg Glu  Ala Glu Ile Glu Val  Glu Gly Asn
    1145                 1150                 1155

Ser Gln  Leu Gln Ile Ser Phe  Ser Thr Ala Leu Ala  Ser Ala Glu
    1160                 1165                 1170

Phe Arg  Val Gln Val Cys Ser  Thr Gln Val His Cys  Ala Ala Glu
    1175                 1180                 1185

Cys His  Pro Pro Lys Asp His  Ile Val Asn Tyr Pro  Ala Ser His
    1190                 1195                 1200

Thr Thr  Leu Gly Val Gln Asp  Ile Ser Ala Thr Ala  Met Ser Trp
    1205                 1210                 1215
```

Val Gln Lys Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala
1220 1225 1230

Ala Leu Ile Leu Ile Val Val Leu Cys Val Ser Phe Ser Arg His
1235 1240 1245

<210> 27
<211> 3747
<212> DNA
<213> Artificial Sequence

<220>
<223> proteins C-E3-E2-6K-E1 strain 06.49 (GenBank: AM258994.1)

<220>
<221> CDS
<222> (1)..(3747)

<400> 27
atg gag ttc atc cca acc caa act ttt tac aat agg agg tac cag cct 48
Met Glu Phe Ile Pro Thr Gln Thr Phe Tyr Asn Arg Arg Tyr Gln Pro
1 5 10 15

cga ccc tgg act ccg cgc cct act atc caa gtc atc agg ccc aga ccg 96
Arg Pro Trp Thr Pro Arg Pro Thr Ile Gln Val Ile Arg Pro Arg Pro
20 25 30

cgc cct cag agg caa gct ggg caa ctt gcc cag ctg atc tca gca gtt 144
Arg Pro Gln Arg Gln Ala Gly Gln Leu Ala Gln Leu Ile Ser Ala Val
35 40 45

aat aaa ctg aca atg cgc gcg gta ccc caa cag aag cca cgc agg aat 192
Asn Lys Leu Thr Met Arg Ala Val Pro Gln Gln Lys Pro Arg Arg Asn
50 55 60

[0077] cgg aag aat aag aag caa aag caa aaa caa cag gcg cca caa aac aac 240
Arg Lys Asn Lys Lys Gln Lys Gln Lys Gln Gln Ala Pro Gln Asn Asn
65 70 75 80

aca aat caa aag aag cag cca cct aaa aag aaa ccg gct caa aag aaa 288
Thr Asn Gln Lys Lys Gln Pro Pro Lys Lys Lys Pro Ala Gln Lys Lys
85 90 95

aag aag ccg ggc cgc aga gag agg atg tgc atg aaa atc gaa aat gat 336
Lys Lys Pro Gly Arg Arg Glu Arg Met Cys Met Lys Ile Glu Asn Asp
100 105 110

tgt att ttc gaa gtc aag cac gaa ggt aag gta aca ggt tac gcg tgc 384
Cys Ile Phe Glu Val Lys His Glu Gly Lys Val Thr Gly Tyr Ala Cys
115 120 125

ctg gtg ggg gac aaa gta atg aaa cca gca cac gta aag ggg acc atc 432
Leu Val Gly Asp Lys Val Met Lys Pro Ala His Val Lys Gly Thr Ile
130 135 140

gat aac gcg gac ctg gcc aaa ctg gcc ttt aag cgg tca tct aag tat 480
Asp Asn Ala Asp Leu Ala Lys Leu Ala Phe Lys Arg Ser Ser Lys Tyr
145 150 155 160

gac ctt gaa tgc gcg cag ata ccc gtg cac atg aag tcc gac gct tcg 528
Asp Leu Glu Cys Ala Gln Ile Pro Val His Met Lys Ser Asp Ala Ser
165 170 175

aag ttc acc cat gag aaa ccg gag ggg tac tac aac tgg cac cac gga 576
Lys Phe Thr His Glu Lys Pro Glu Gly Tyr Tyr Asn Trp His His Gly
180 185 190

gca gta cag tac tca gga ggc cgg ttc acc atc cct aca ggt gct ggc 624
Ala Val Gln Tyr Ser Gly Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly
195 200 205

aaa cca ggg gac agc ggc aga ccg atc ttc gac aac aag gga cgc gtg 672
Lys Pro Gly Asp Ser Gly Arg Pro Ile Phe Asp Asn Lys Gly Arg Val

```
            210                 215                 220

        gtg gcc ata gtc tta gga gga gct aat gaa gga gcc cgt aca gcc ctc        720
        Val Ala Ile Val Leu Gly Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu
        225                 230                 235                 240

        tcg gtg gtg acc tgg aat aaa gac att gtc act aaa atc acc ccc gag        768
        Ser Val Val Thr Trp Asn Lys Asp Ile Val Thr Lys Ile Thr Pro Glu
                        245                 250                 255

        ggg gcc gaa gag tgg agt ctt gcc atc cca gtt atg tgc ctg ttg gca        816
        Gly Ala Glu Glu Trp Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala
                    260                 265                 270

        aac acc acg ttc ccc tgc tcc cag ccc cct tgc acg ccc tgc tgc tac        864
        Asn Thr Thr Phe Pro Cys Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr
                275                 280                 285

        gaa aag gaa ccg gag gaa acc cta cgc atg ctt gag gac aac gtc atg        912
        Glu Lys Glu Pro Glu Glu Thr Leu Arg Met Leu Glu Asp Asn Val Met
            290                 295                 300

        aga cct ggg tac tat cag ctg cta caa gca tcc tta aca tgt tct ccc        960
        Arg Pro Gly Tyr Tyr Gln Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro
        305                 310                 315                 320

        cac cgc cag cga cgc agc acc aag gac aac ttc aat gtc tat aaa gcc       1008
        His Arg Gln Arg Arg Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala
                        325                 330                 335

        aca aga cca tac tta gct cac tgt ccc gac tgt gga gaa ggg cac tcg       1056
        Thr Arg Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser
                    340                 345                 350

        tgc cat agt ccc gta gca cta gaa cgc atc aga aat gaa gcg aca gac       1104
        Cys His Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp
                355                 360                 365

[0078]  ggg acg ctg aaa atc cag gtc tcc ttg caa atc gga ata aag acg gat       1152
        Gly Thr Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp
            370                 375                 380

        gac agc cac gat tgg acc aag ctg cgt tat atg gac aac cac atg cca       1200
        Asp Ser His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro
        385                 390                 395                 400

        gca gac gca gag agg gcg ggg cta ttt gta aga aca tca gca ccg tgt       1248
        Ala Asp Ala Glu Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys
                        405                 410                 415

        acg att act gga aca atg gga cac ttc atc ctg gcc cga tgt cca aaa       1296
        Thr Ile Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys
                    420                 425                 430

        ggg gaa act ctg acg gtg gga ttc act gac agt agg aag att agt cac       1344
        Gly Glu Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His
                435                 440                 445

        tca tgt acg cac cca ttt cac cac gac cct cct gtg ata ggt cgg gaa       1392
        Ser Cys Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu
            450                 455                 460

        aaa ttc cat tcc cga ccg cag cac ggt aaa gag cta cct tgc agc acg       1440
        Lys Phe His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr
        465                 470                 475                 480

        tac gtg cag agc acc gcc gca act acc gag gag ata gag gta cac atg       1488
        Tyr Val Gln Ser Thr Ala Ala Thr Thr Glu Glu Ile Glu Val His Met
                        485                 490                 495

        ccc cca gac acc cct gat cgc aca tta atg tca caa cag tcc ggc aac       1536
        Pro Pro Asp Thr Pro Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn
                    500                 505                 510

        gta aag atc aca gtc aat ggc cag acg gtg cgg tac aag tgt aat tgc       1584
        Val Lys Ile Thr Val Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys
                515                 520                 525
```

```
ggt ggc tca aat gaa gga cta aca act aca gac aaa gtg att aat aac      1632
Gly Gly Ser Asn Glu Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn
    530                 535                 540

tgc aag gtt gat caa tgt cat gcc gcg gtc acc aat cac aaa aag tgg      1680
Cys Lys Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp
545                 550                 555                 560

cag tat aac tcc cct ctg gtc ccg cgt aat gct gaa ctt ggg gac cga      1728
Gln Tyr Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg
                565                 570                 575

aaa gga aaa att cac atc ccg ttt ccg ctg gca aat gta aca tgc agg      1776
Lys Gly Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg
            580                 585                 590

gtg cct aaa gca agg aac ccc acc gtg acg tac ggg aaa aac caa gtc      1824
Val Pro Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val
        595                 600                 605

atc atg cta ctg tat cct gac cac cca aca ctc ctg tcc tac cgg aat      1872
Ile Met Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn
    610                 615                 620

atg gga gaa gaa cca aac tat caa gaa gag tgg gtg atg cat aag aag      1920
Met Gly Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys
625                 630                 635                 640

gaa gtc gtg cta acc gtg ccg act gaa ggg ctc gag gtc acg tgg ggc      1968
Glu Val Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly
                645                 650                 655

aac aac gag ccg tat aag tat tgg ccg cag tta tct aca aac ggt aca      2016
Asn Asn Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr
            660                 665                 670
```

[0079]

```
gcc cat ggc cac ccg cac gag ata att ctg tat tat tat gag ctg tac      2064
Ala His Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr
        675                 680                 685

ccc act atg act gta gta gtt gtg tca gtg gcc acg ttc ata ctc ctg      2112
Pro Thr Met Thr Val Val Val Val Ser Val Ala Thr Phe Ile Leu Leu
    690                 695                 700

tcg atg gtg ggt atg gca gcg ggg atg tgc atg tgt gca cga cgc aga      2160
Ser Met Val Gly Met Ala Ala Gly Met Cys Met Cys Ala Arg Arg Arg
705                 710                 715                 720

tgc atc aca ccg tat gaa ctg aca cca gga gct acc gtc cct ttc ctg      2208
Cys Ile Thr Pro Tyr Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu
                725                 730                 735

ctt agc cta ata tgc tgc atc aga aca gct aaa gcg gcc aca tac caa      2256
Leu Ser Leu Ile Cys Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln
            740                 745                 750

gag gct gcg ata tac ctg tgg aac gag cag caa cct ttg ttt tgg cta      2304
Glu Ala Ala Ile Tyr Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu
        755                 760                 765

caa gcc ctt att ccg ctg gca gcc ctg att gtt cta tgc aac tgt ctg      2352
Gln Ala Leu Ile Pro Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu
    770                 775                 780

aga ctc tta cca tgc tgc tgt aaa acg ttg gct ttt tta gcc gta atg      2400
Arg Leu Leu Pro Cys Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met
785                 790                 795                 800

agc gtc ggt gcc cac act gtg agc gcg tac gaa cac gta aca gtg atc      2448
Ser Val Gly Ala His Thr Val Ser Ala Tyr Glu His Val Thr Val Ile
                805                 810                 815

ccg aac acg gtg gga gta ccg tat aag act cta gtc aat aga cct ggc      2496
Pro Asn Thr Val Gly Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly
            820                 825                 830
```

```
tac agc ccc atg gta ttg gag atg gaa cta ctg tca gtc act ttg gag         2544
Tyr Ser Pro Met Val Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu
        835                 840                 845

cca aca cta tcg ctt gat tac atc acg tgc gag tac aaa acc gtc atc         2592
Pro Thr Leu Ser Leu Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile
850                     855                 860

ccg tct ccg tac gtg aag tgc tgc ggt aca gca gag tgc aag gac aaa         2640
Pro Ser Pro Tyr Val Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys
865                 870                 875                 880

aac cta cct gac tac agc tgt aag gtc ttc acc ggc gtc tac cca ttt         2688
Asn Leu Pro Asp Tyr Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe
                885                 890                 895

atg tgg ggc ggc gcc tac tgc ttc tgc gac gct gaa aac acg cag ttg         2736
Met Trp Gly Gly Ala Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu
            900                 905                 910

agc gaa gca cac gtg gag aag tcc gaa tca tgc aaa aca gaa ttt gca         2784
Ser Glu Ala His Val Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala
        915                 920                 925

tca gca tac agg gct cat acc gca tct gca tca gct aag ctc cgc gtc         2832
Ser Ala Tyr Arg Ala His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val
    930                 935                 940

ctt tac caa gga aat aac atc act gta act gcc tat gca aac ggc gac         2880
Leu Tyr Gln Gly Asn Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp
945                 950                 955                 960

cat gcc gtc aca gtt aag gac gcc aaa ttc att gtg ggg cca atg tct         2928
His Ala Val Thr Val Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser
                965                 970                 975
```

[0080]

```
tca gcc tgg aca cct ttc gac aac aaa att gtg gtg tac aaa ggt gac         2976
Ser Ala Trp Thr Pro Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp
            980                 985                 990

gtc tat aac atg gac tac ccg ccc  ttt ggc gca gga aga  cca gga caa       3024
Val Tyr Asn Met Asp Tyr Pro Pro  Phe Gly Ala Gly Arg  Pro Gly Gln
        995                 1000                 1005

ttt ggc  gat atc caa agt cgc  aca cct gag agt aaa  gac gtc tat         3069
Phe Gly  Asp Ile Gln Ser Arg  Thr Pro Glu Ser Lys  Asp Val Tyr
    1010                 1015                 1020

gct aat  aca caa ctg gta ctg  cag aga ccg gct gtg  ggt acg gta         3114
Ala Asn  Thr Gln Leu Val Leu  Gln Arg Pro Ala Val  Gly Thr Val
    1025                 1030                 1035

cac gtg  cca tac tct cag gca  cca tct ggc ttt aag  tat tgg cta         3159
His Val  Pro Tyr Ser Gln Ala  Pro Ser Gly Phe Lys  Tyr Trp Leu
    1040                 1045                 1050

aaa gaa  cgc ggg gcg tcg ctg  cag cac aca gca cca  ttt ggc tgc         3204
Lys Glu  Arg Gly Ala Ser Leu  Gln His Thr Ala Pro  Phe Gly Cys
    1055                 1060                 1065

caa ata  gca aca aac ccg gta  aga gcg gtg aac tgc  gcc gta ggg         3249
Gln Ile  Ala Thr Asn Pro Val  Arg Ala Val Asn Cys  Ala Val Gly
    1070                 1075                 1080

aac atg  ccc atc tcc atc gac  ata ccg gaa gcg gcc  ttc act agg         3294
Asn Met  Pro Ile Ser Ile Asp  Ile Pro Glu Ala Ala  Phe Thr Arg
    1085                 1090                 1095

gtc gtc  gac gcg ccc tct tta  acg gac atg tcg tgc  gag gta cca         3339
Val Val  Asp Ala Pro Ser Leu  Thr Asp Met Ser Cys  Glu Val Pro
    1100                 1105                 1110

gcc tgc  acc cat tcc tca gac  ttt ggg ggc gtc gcc  att att aaa         3384
Ala Cys  Thr His Ser Ser Asp  Phe Gly Gly Val Ala  Ile Ile Lys
    1115                 1120                 1125

tat gca  gcc agc aag aaa ggc  aag tgt gcg gtg cat  tcg atg act         3429
```

Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser Met Thr
1130 1135 1140

aac gcc gtc act att cgg gaa gct gag ata gaa gtt gaa ggg aat 3474
Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly Asn
1145 1150 1155

tct cag ctg caa atc tct ttc tcg acg gcc tta gcc agc gcc gaa 3519
Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
1160 1165 1170

ttc cgc gta caa gtc tgt tct aca caa gta cac tgt gca gcc gag 3564
Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu
1175 1180 1185

tgc cac ccc ccg aag gac cac ata gtc aac tac ccg gcg tca cat 3609
Cys His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His
1190 1195 1200

acc acc ctc ggg gtc cag gac atc tcc gct acg gcg atg tca tgg 3654
Thr Thr Leu Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp
1205 1210 1215

gtg cag aag atc acg gga ggt gtg gga ctg gtt gtt gct gtt gcc 3699
Val Gln Lys Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala
1220 1225 1230

gca ctg att cta atc gtg gtg cta tgc gtg tcg ttc agc agg cac 3744
Ala Leu Ile Leu Ile Val Val Leu Cys Val Ser Phe Ser Arg His
1235 1240 1245

taa 3747

[0081]

<210> 28
<211> 1248
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 28

Met Glu Phe Ile Pro Thr Gln Thr Phe Tyr Asn Arg Arg Tyr Gln Pro
1 5 10 15

Arg Pro Trp Thr Pro Arg Pro Thr Ile Gln Val Ile Arg Pro Arg Pro
20 25 30

Arg Pro Gln Arg Gln Ala Gly Gln Leu Ala Gln Leu Ile Ser Ala Val
35 40 45

Asn Lys Leu Thr Met Arg Ala Val Pro Gln Gln Lys Pro Arg Arg Asn
50 55 60

Arg Lys Asn Lys Lys Gln Lys Gln Lys Gln Gln Ala Pro Gln Asn Asn
65 70 75 80

Thr Asn Gln Lys Lys Gln Pro Pro Lys Lys Lys Pro Ala Gln Lys Lys
85 90 95

Lys Lys Pro Gly Arg Arg Glu Arg Met Cys Met Lys Ile Glu Asn Asp
100 105 110

Cys Ile Phe Glu Val Lys His Glu Gly Lys Val Thr Gly Tyr Ala Cys
115 120 125

Leu Val Gly Asp Lys Val Met Lys Pro Ala His Val Lys Gly Thr Ile

序 列 表

```
    130                 135                 140

Asp Asn Ala Asp Leu Ala Lys Leu Ala Phe Lys Arg Ser Ser Lys Tyr
145                 150                 155                 160

Asp Leu Glu Cys Ala Gln Ile Pro Val His Met Lys Ser Asp Ala Ser
                165                 170                 175

Lys Phe Thr His Glu Lys Pro Glu Gly Tyr Tyr Asn Trp His His Gly
            180                 185                 190

Ala Val Gln Tyr Ser Gly Gly Arg Phe Thr Ile Pro Thr Gly Ala Gly
        195                 200                 205

Lys Pro Gly Asp Ser Gly Arg Pro Ile Phe Asp Asn Lys Gly Arg Val
    210                 215                 220

Val Ala Ile Val Leu Gly Gly Ala Asn Glu Gly Ala Arg Thr Ala Leu
225                 230                 235                 240

Ser Val Val Thr Trp Asn Lys Asp Ile Val Thr Lys Ile Thr Pro Glu
                245                 250                 255

Gly Ala Glu Glu Trp Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala
            260                 265                 270

Asn Thr Thr Phe Pro Cys Ser Gln Pro Pro Cys Thr Pro Cys Cys Tyr
        275                 280                 285

[0082]
Glu Lys Glu Pro Glu Glu Thr Leu Arg Met Leu Glu Asp Asn Val Met
    290                 295                 300

Arg Pro Gly Tyr Tyr Gln Leu Leu Gln Ala Ser Leu Thr Cys Ser Pro
305                 310                 315                 320

His Arg Gln Arg Arg Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala
                325                 330                 335

Thr Arg Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser
            340                 345                 350

Cys His Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp
        355                 360                 365

Gly Thr Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp
    370                 375                 380

Asp Ser His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro
385                 390                 395                 400

Ala Asp Ala Glu Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys
                405                 410                 415

Thr Ile Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys
            420                 425                 430

Gly Glu Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His
        435                 440                 445
```

```
Ser Cys Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu
    450                 455                 460

Lys Phe His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr
465                 470                 475                 480

Tyr Val Gln Ser Thr Ala Ala Thr Thr Glu Glu Ile Glu Val His Met
                485                 490                 495

Pro Pro Asp Thr Pro Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn
            500                 505                 510

Val Lys Ile Thr Val Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys
        515                 520                 525

Gly Gly Ser Asn Glu Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn
    530                 535                 540

Cys Lys Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp
545                 550                 555                 560

Gln Tyr Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg
                565                 570                 575

Lys Gly Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg
            580                 585                 590
```

[0083]

```
Val Pro Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val
        595                 600                 605

Ile Met Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn
    610                 615                 620

Met Gly Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys
625                 630                 635                 640

Glu Val Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly
                645                 650                 655

Asn Asn Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr
            660                 665                 670

Ala His Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr
        675                 680                 685

Pro Thr Met Thr Val Val Val Val Ser Val Ala Thr Phe Ile Leu Leu
    690                 695                 700

Ser Met Val Gly Met Ala Ala Gly Met Cys Met Cys Ala Arg Arg Arg
705                 710                 715                 720

Cys Ile Thr Pro Tyr Glu Leu Thr Pro Gly Ala Thr Val Pro Phe Leu
                725                 730                 735

Leu Ser Leu Ile Cys Cys Ile Arg Thr Ala Lys Ala Ala Thr Tyr Gln
            740                 745                 750
```

Glu Ala Ala Ile Tyr Leu Trp Asn Glu Gln Gln Pro Leu Phe Trp Leu
755 760 765

Gln Ala Leu Ile Pro Leu Ala Ala Leu Ile Val Leu Cys Asn Cys Leu
770 775 780

Arg Leu Leu Pro Cys Cys Cys Lys Thr Leu Ala Phe Leu Ala Val Met
785 790 795 800

Ser Val Gly Ala His Thr Val Ser Ala Tyr Glu His Val Thr Val Ile
805 810 815

Pro Asn Thr Val Gly Val Pro Tyr Lys Thr Leu Val Asn Arg Pro Gly
820 825 830

Tyr Ser Pro Met Val Leu Glu Met Glu Leu Leu Ser Val Thr Leu Glu
835 840 845

Pro Thr Leu Ser Leu Asp Tyr Ile Thr Cys Glu Tyr Lys Thr Val Ile
850 855 860

Pro Ser Pro Tyr Val Lys Cys Cys Gly Thr Ala Glu Cys Lys Asp Lys
865 870 875 880

Asn Leu Pro Asp Tyr Ser Cys Lys Val Phe Thr Gly Val Tyr Pro Phe
885 890 895

[0084] Met Trp Gly Gly Ala Tyr Cys Phe Cys Asp Ala Glu Asn Thr Gln Leu
900 905 910

Ser Glu Ala His Val Glu Lys Ser Glu Ser Cys Lys Thr Glu Phe Ala
915 920 925

Ser Ala Tyr Arg Ala His Thr Ala Ser Ala Ser Ala Lys Leu Arg Val
930 935 940

Leu Tyr Gln Gly Asn Asn Ile Thr Val Thr Ala Tyr Ala Asn Gly Asp
945 950 955 960

His Ala Val Thr Val Lys Asp Ala Lys Phe Ile Val Gly Pro Met Ser
965 970 975

Ser Ala Trp Thr Pro Phe Asp Asn Lys Ile Val Val Tyr Lys Gly Asp
980 985 990

Val Tyr Asn Met Asp Tyr Pro Pro Phe Gly Ala Gly Arg Pro Gly Gln
995 1000 1005

Phe Gly Asp Ile Gln Ser Arg Thr Pro Glu Ser Lys Asp Val Tyr
1010 1015 1020

Ala Asn Thr Gln Leu Val Leu Gln Arg Pro Ala Val Gly Thr Val
1025 1030 1035

His Val Pro Tyr Ser Gln Ala Pro Ser Gly Phe Lys Tyr Trp Leu
1040 1045 1050

Lys Glu Arg Gly Ala Ser Leu Gln His Thr Ala Pro Phe Gly Cys
1055 1060 1065

Gln Ile Ala Thr Asn Pro Val Arg Ala Val Asn Cys Ala Val Gly
1070 1075 1080

Asn Met Pro Ile Ser Ile Asp Ile Pro Glu Ala Ala Phe Thr Arg
1085 1090 1095

Val Val Asp Ala Pro Ser Leu Thr Asp Met Ser Cys Glu Val Pro
1100 1105 1110

Ala Cys Thr His Ser Ser Asp Phe Gly Gly Val Ala Ile Ile Lys
1115 1120 1125

Tyr Ala Ala Ser Lys Lys Gly Lys Cys Ala Val His Ser Met Thr
1130 1135 1140

Asn Ala Val Thr Ile Arg Glu Ala Glu Ile Glu Val Glu Gly Asn
1145 1150 1155

Ser Gln Leu Gln Ile Ser Phe Ser Thr Ala Leu Ala Ser Ala Glu
1160 1165 1170

Phe Arg Val Gln Val Cys Ser Thr Gln Val His Cys Ala Ala Glu
1175 1180 1185

Cys His Pro Pro Lys Asp His Ile Val Asn Tyr Pro Ala Ser His
1190 1195 1200

[0085]

Thr Thr Leu Gly Val Gln Asp Ile Ser Ala Thr Ala Met Ser Trp
1205 1210 1215

Val Gln Lys Ile Thr Gly Gly Val Gly Leu Val Val Ala Val Ala
1220 1225 1230

Ala Leu Ile Leu Ile Val Val Leu Cys Val Ser Phe Ser Arg His
1235 1240 1245

<210> 29
<211> 1170
<212> DNA
<213> Artificial Sequence

<220>
<223> soluble glycoprotein E2 without stem region strain 06.49

<220>
<221> CDS
<222> (1)..(1170)

<400> 29
cgt acg atg agt ctt gcc atc cca gtt atg tgc ctg ttg gca aac acc 48
Arg Thr Met Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr
1 5 10 15

acg ttt cct tgc tcc cag gac aac ttc aat gtc tat aaa gcc aca aga 96
Thr Phe Pro Cys Ser Gln Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg
20 25 30

cca tac tta gct cac tgt ccc gac tgt gga gaa ggg cac tcg tgc cat 144
Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His
35 40 45

```
agt ccc gta gca cta gaa cgc atc aga aat gaa gcg aca gac ggg acg         192
Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr
    50                  55                  60

ctg aaa atc cag gtc tcc ttg caa atc gga ata aag acg gat gac agc         240
Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser
65                  70                  75                  80

cac gat tgg acc aag ctg cgt tat atg gac aac cac atg cca gca gac         288
His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp
                85                  90                  95

gca gag agg gcg ggg cta ttt gta aga aca tca gca ccg tgt acg att         336
Ala Glu Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile
            100                 105                 110

act gga aca atg gga cac ttc atc ctg gcc cga tgt ccc aaa ggc gaa         384
Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu
        115                 120                 125

act ctg acg gtg gga ttc act gac agt agg aag att agt cac tca tgt         432
Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys
    130                 135                 140

acg cac cca ttt cac cac gac cct cct gtg ata ggt cgg gag aaa ttc         480
Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe
145                 150                 155                 160

cat tcc cga ccg cag cac ggt aaa gag cta cct tgc agc acc tac gtg         528
His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val
                165                 170                 175

cag agc acc gcc gca act acc gag gag ata gag gta cac atg ccc cca         576
Gln Ser Thr Ala Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro
            180                 185                 190

[0086]
gac acc cct gat cgc aca tta atg tca caa cag tcc ggc aac gta aag         624
Asp Thr Pro Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys
        195                 200                 205

atc aca gtc aat ggc cag acg gtg cgg tac aag tgt aat tgc ggt ggc         672
Ile Thr Val Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly
    210                 215                 220

tca aat gaa gga cta aca act aca gac aaa gtg att aat aac tgc aag         720
Ser Asn Glu Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys
225                 230                 235                 240

gtt gat caa tgt cat gcc gcg gtc acc aat cac aaa aag tgg cag tat         768
Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr
                245                 250                 255

aac tcc cct ctg gtc ccg cgt aat gct gaa ctt ggg gac agg aaa gga         816
Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly
            260                 265                 270

aaa att cac atc ccg ttt ccg ctg gca aat gta aca tgc agg gtg cct         864
Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro
        275                 280                 285

aaa gca agg aac ccc acc gtg acg tat ggc aaa aac caa gtc atc atg         912
Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met
    290                 295                 300

cta ctg tat cct gac cac cca aca ctc ctg tcc tac cgg aat atg gga         960
Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly
305                 310                 315                 320

gaa gaa cca aac tat caa gaa gag tgg gtg atg cat aag aag gaa gtc        1008
Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val
                325                 330                 335

gtg cta acc gtg ccg act gaa ggg ctc gag gtc acg tgg ggc aac aac        1056
Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn
            340                 345                 350
```

```
gag ccg tat aag tat tgg ccg cag tta tct aca aac ggt aca gcc cat      1104
Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His
        355                 360                 365

ggc cac ccg cac gag ata att ctg tat tat tat gag ctg tac ccc act      1152
Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr
    370                 375                 380

atg act taa tag gcg cgc                                              1170
Met Thr         Ala Arg
385


<210>  30
<211>  386
<212>  PRT
<213>  Artificial Sequence

<220>
<223>  Synthetic Construct

<400>  30

Arg Thr Met Ser Leu Ala Ile Pro Val Met Cys Leu Leu Ala Asn Thr
1               5                   10                  15


Thr Phe Pro Cys Ser Gln Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg
            20                  25                  30


Pro Tyr Leu Ala His Cys Pro Asp Cys Gly Glu Gly His Ser Cys His
        35                  40                  45


Ser Pro Val Ala Leu Glu Arg Ile Arg Asn Glu Ala Thr Asp Gly Thr
    50                  55                  60
```

[0087]

```
Leu Lys Ile Gln Val Ser Leu Gln Ile Gly Ile Lys Thr Asp Asp Ser
65                  70                  75                  80


His Asp Trp Thr Lys Leu Arg Tyr Met Asp Asn His Met Pro Ala Asp
                85                  90                  95


Ala Glu Arg Ala Gly Leu Phe Val Arg Thr Ser Ala Pro Cys Thr Ile
            100                 105                 110


Thr Gly Thr Met Gly His Phe Ile Leu Ala Arg Cys Pro Lys Gly Glu
        115                 120                 125


Thr Leu Thr Val Gly Phe Thr Asp Ser Arg Lys Ile Ser His Ser Cys
    130                 135                 140


Thr His Pro Phe His His Asp Pro Pro Val Ile Gly Arg Glu Lys Phe
145                 150                 155                 160


His Ser Arg Pro Gln His Gly Lys Glu Leu Pro Cys Ser Thr Tyr Val
                165                 170                 175


Gln Ser Thr Ala Ala Thr Thr Glu Glu Ile Glu Val His Met Pro Pro
            180                 185                 190


Asp Thr Pro Asp Arg Thr Leu Met Ser Gln Gln Ser Gly Asn Val Lys
        195                 200                 205


Ile Thr Val Asn Gly Gln Thr Val Arg Tyr Lys Cys Asn Cys Gly Gly
    210                 215                 220
```

```
        Ser Asn Glu Gly Leu Thr Thr Thr Asp Lys Val Ile Asn Asn Cys Lys
        225                 230                 235                 240

        Val Asp Gln Cys His Ala Ala Val Thr Asn His Lys Lys Trp Gln Tyr
                        245                 250                 255

        Asn Ser Pro Leu Val Pro Arg Asn Ala Glu Leu Gly Asp Arg Lys Gly
                    260                 265                 270

        Lys Ile His Ile Pro Phe Pro Leu Ala Asn Val Thr Cys Arg Val Pro
                275                 280                 285

        Lys Ala Arg Asn Pro Thr Val Thr Tyr Gly Lys Asn Gln Val Ile Met
            290                 295                 300

        Leu Leu Tyr Pro Asp His Pro Thr Leu Leu Ser Tyr Arg Asn Met Gly
        305                 310                 315                 320

        Glu Glu Pro Asn Tyr Gln Glu Glu Trp Val Met His Lys Lys Glu Val
                        325                 330                 335

        Val Leu Thr Val Pro Thr Glu Gly Leu Glu Val Thr Trp Gly Asn Asn
                    340                 345                 350

        Glu Pro Tyr Lys Tyr Trp Pro Gln Leu Ser Thr Asn Gly Thr Ala His
                355                 360                 365

[0088]  Gly His Pro His Glu Ile Ile Leu Tyr Tyr Tyr Glu Leu Tyr Pro Thr
            370                 375                 380

        Met Thr
        385

        <210>  31
        <211>  3762
        <212>  DNA
        <213>  Artificial Sequence

        <220>
        <223>  proteins C-E3-E2-6K-E1 strain 06.49

        <400>  31
        cgtacgatgg agttcatccc aacccaaact ttttacaata ggaggtacca gcctcgaccc        60

        tggactccgc gccctactat ccaagtcatc aggcccagac cgcgccctca gaggcaagct       120

        gggcaacttg cccagctgat ctcagcagtt aataaactga caatgagagc ggtaccccaa       180

        cagaagccac gcaggaatcg gaagaataag aagcaaaagc aaaaacaaca ggcgccacaa       240

        aacaacacaa atcaaaagaa gcagccacct aaaaagaaac cggctcaaaa gaaaaagaag       300

        ccgggccgca gagagaggat gtgcatgaaa atcgaaaatg attgtatttt cgaagtcaag       360

        cacgaaggta aggtaacagg ttacgcgtgc ctggtggggg acaaagtaat gaaaccagca       420

        cacgtcaagg gcaccatcga taacgcggac ctggccaaac tggcctttaa gcggtcatct       480

        aagtatgacc ttgaatgcgc tcagataccc gtgcacatga agtccgacgc ttcgaagttc       540

        acccatgaga aaccggaggg gtactacaac tggcaccacg gagcagtaca gtactcagga       600

        ggccggttca ccatccctac aggtgctggc aaaccagggg acagcggcag accgatcttc       660
```

gacaacaagg gacgcgtggt ggccatagtc ttaggaggag ctaatgaagg agcccgtaca 720
gccctctcgg tggtgacctg gaataaagac attgtcacta aaatcacccc cgaggggggcc 780
gaagagtgga gtcttgccat cccagttatg tgcctgttgg caaacaccac gttcccatgc 840
tcccagccac cttgcacgcc ctgctgctac gagaaggaac cggaggaaac cctacgcatg 900
cttgaggaca acgtcatgag acctgggtac tatcagctgc tacaagcatc cttaacatgt 960
tctccccacc gccagcgacg cagcaccaag gacaacttca atgtctataa agccacaaga 1020
ccatacttag ctcactgtcc cgactgtgga gaagggcact cgtgccatag tcccgtagca 1080
ctagaacgca tcagaaatga agcgacagac gggacgctga aaatccaggt ctccttgcaa 1140
atcggaataa agacggatga cagccacgat tggaccaagc tgcgttatat ggacaaccac 1200
atgccagcag acgcagagag ggcggggcta tttgtaagaa catcagcacc gtgtacgatt 1260
actggaacaa tgggacactt catcctggcc cgatgtccta aaggcgaaac tctgacggtg 1320
ggattcactg acagtaggaa gattagtcac tcatgtacgc acccatttca ccacgaccct 1380
cctgtgatag gtcgggagaa attccattcc cgaccgcagc acggtaaaga gctaccttgc 1440
agcacctacg tgcagagcac cgccgcaact accgaggaga tagaggtaca catgcccccca 1500
gacacccctg atcgcacatt aatgtcacaa cagtccggca acgtaaagat cacagtcaat 1560
ggccagacgg tgcggtacaa gtgtaattgc ggtggctcaa atgaaggact aacaactaca 1620
gacaaagtga ttaataactg caaggttgat caatgtcatg ccgcggtcac caatcacaaa 1680
aagtggcagt ataactcccc tctggtcccg cgtaatgctg aacttgggga ccggaaagga 1740
aaaattcaca tcccgtttcc gctggcaaat gtaacatgca gggtgcctaa agcaaggaac 1800

[0089] cccaccgtga cgtatggcaa aaaccaagtc atcatgctac tgtatcctga ccacccaaca 1860
ctcctgtcct accggaatat gggagaagaa ccaaactatc aagaagagtg ggtgatgcat 1920
aagaaggaag tcgtgctaac cgtgccgact gaagggctcg aggtcacgtg ggcaacaac 1980
gagccgtata agtattggcc gcagttatct acaaacggta cagcccatgg ccacccgcac 2040
gagataattc tgtattatta tgagctgtac cccactatga ctgtagtagt tgtgtcagtg 2100
gccacgttca tactcctgtc gatggtgggt atggcagcgg ggatgtgcat gtgtgcacga 2160
cgcagatgca tcacaccgta tgaactgaca ccaggagcta ccgtcccatt cctgcttagc 2220
ctaatatgct gcatcagaac agctaaagcg gccacatacc aagaggctgc gatatacctg 2280
tggaacgagc agcaaccttt gttttggcta caagccctta ttccgctggc agccctgatt 2340
gttctatgca actgtctgag actcttacca tgctgctgta aaacgttggc ttttttagcc 2400
gtaatgagcg tcggtgccca cactgtgagc gcttacgaac acgtaacagt gatcccgaac 2460
acggtgggag taccgtataa gactctagtc aatagacctg gctacagccc catggtattg 2520
gagatggaac tactgtcagt cactttggag ccaacactat cgcttgatta catcacgtgc 2580
gagtacaaaa ccgtcatccc gtctccttac gtgaagtgct gcggtacagc agagtgcaag 2640
gacaaaaacc tacctgacta cagctgtaag gtcttcaccg gcgtctaccc atttatgtgg 2700
ggcggcgcct actgcttctg cgacgctgaa aacacgcagt tgagcgaagc acacgtggag 2760
aagtccgaat catgcaaaac agaatttgca tcagcataca gggctcatac cgcatctgca 2820
tcagctaagc tccgcgtcct ttaccaagga aataacatca ctgtaactgc ctatgcaaac 2880
ggcgaccatg ccgtcacagt taaggacgcc aaattcattg tggggccaat gtcttcagcc 2940
tggacacctt tcgacaacaa aattgtggtg tacaaaggtg acgtctataa catggactac 3000

ccgcccttcg gcgcaggaag accaggacaa tttggcgata tccaaagtcg cacacctgag 3060
agtaaagacg tctatgctaa tacacaactg gtactgcaga gaccggctgt gggtacggta 3120
cacgtgccat actctcaggc accatctggc tttaagtatt ggctaaaaga acgcggggcg 3180
tcgctgcagc acacagcacc atttggctgc caaatagcaa caaacccggt aagagcggtg 3240
aactgcgccg tagggaacat gcccatctcc atcgacatac cggaagcggc cttcactagg 3300
gtcgtcgacg cgccctcttt aacggacatg tcgtgcgagg taccagcctg cacccattcc 3360
tcagactttg ggggcgtcgc cattattaaa tatgcagcca gcaagaaagg caagtgtgcg 3420
gtgcattcga tgactaacgc cgtcactatt cgggaagctg agatagaagt tgaagggaat 3480
tctcagctgc aaatctcttt ctcgacggcc ttagccagcg ccgaattccg cgtacaagtc 3540
tgttctacac aagtacactg tgcagccgag tgccaccccc cgaaggacca catagtcaac 3600
tacccggcgt cacataccac cctcggggtc caggacatct ccgctacggc gatgtcatgg 3660
gtgcagaaga tcacgggagg tgtgggactg gttgttgctg ttgccgcact gattctaatc 3720
gtggtgctat gcgtgtcgtt cagcaggcac taataggcgc gc 3762

<210> 32
<211> 54
<212> DNA
<213> Artificial Sequence

[0090] <220>
<223> Nucleotide sequence of E2 EP3

<220>
<221> CDS
<222> (1)..(54)

<400> 32
```
agc acc aag gac aac ttc aat gtc tat aaa gcc aca aga cca tac tta        48
Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu
1               5                   10                  15

gct cac                                                                 54
Ala His
```

<210> 33
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Synthetic Construct

<400> 33

```
Ser Thr Lys Asp Asn Phe Asn Val Tyr Lys Ala Thr Arg Pro Tyr Leu
1               5                   10                  15

Ala His
```

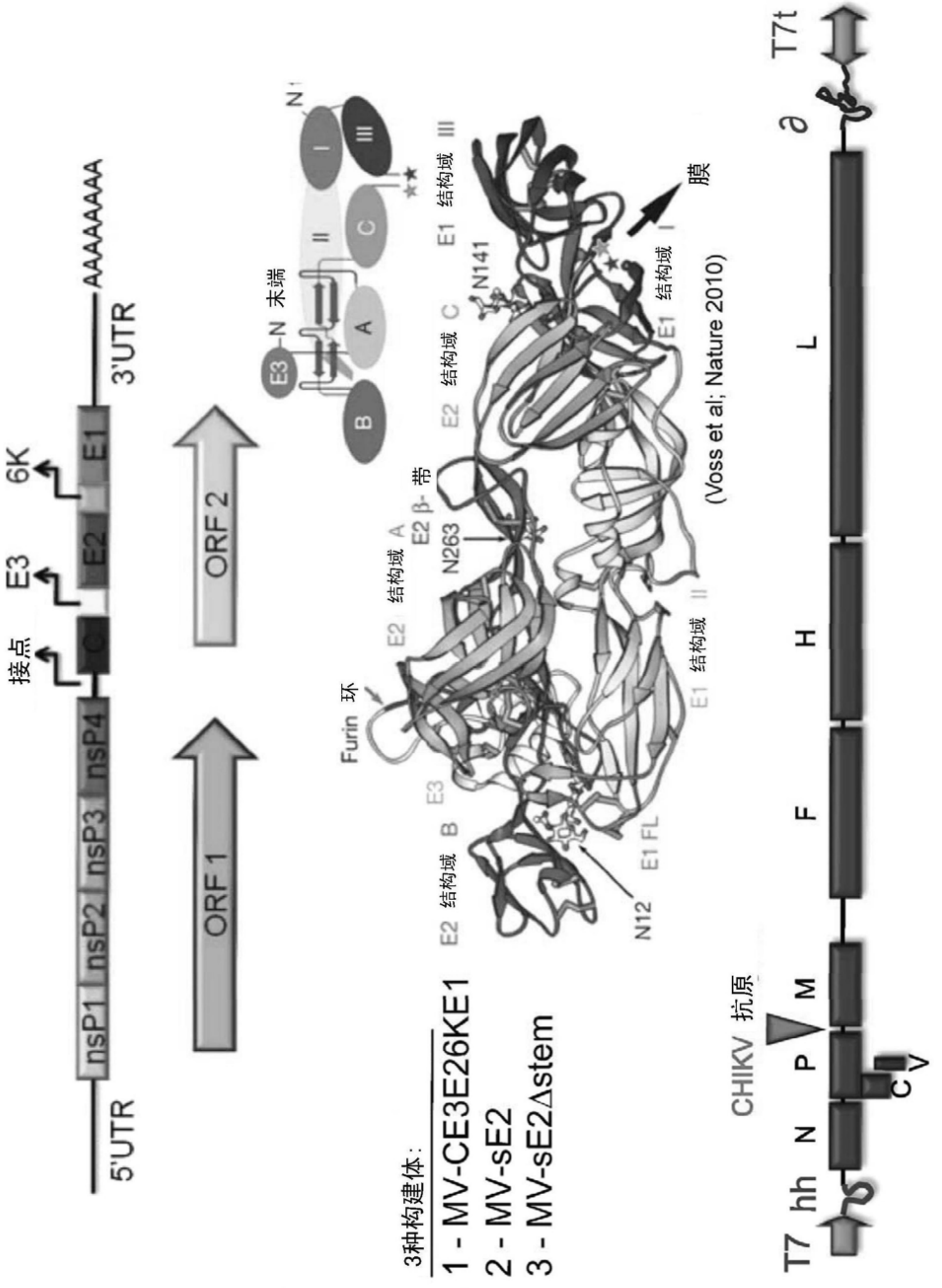

5'UTR
nsP1
nsP2
nsP3
nsP4
接点
C
E3
E2
6K
E1
3'UTR
AAAAAAAA
ORF 1
ORF 2
E3
N
N 末端
B
A
C
I
II
III
E2 结构域 B
E3
Furin 环
E2 结构域 A
N263
E2 β- 带
E2 结构域 C
N141
E1 结构域 III
E1 FL
N12
E1 结构域 II
E1 结构域 I
膜
(Voss et al; Nature 2010)
3种构建体：
1 - MV-CE3E26KE1
2 - MV-sE2
3 - MV-sE2Δstem
T7
hh
N
P
C
V
CHIKV 抗原
M
F
H
L
δ
T7t

图1A

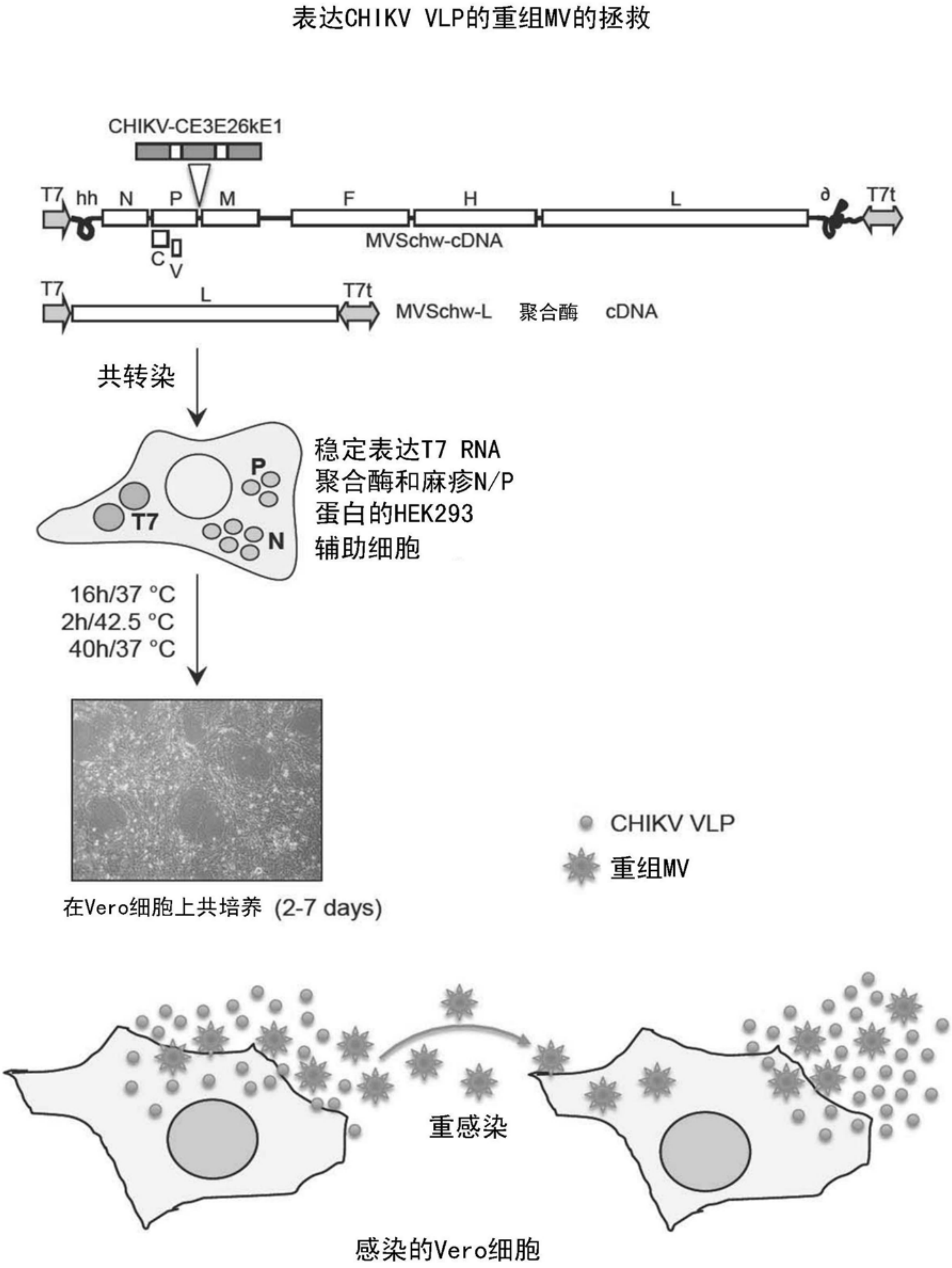

表达CHIKV VLP的重组MV的拯救
CHIKV-CE3E26kE1
T7
hh
N
P
M
F
H
L
∂
T7t
C
V
MVSchw-cDNA
T7
L
T7t
MVSchw-L 聚合酶 cDNA
共转染
P
T7
N
稳定表达T7 RNA
聚合酶和麻疹N/P
蛋白的HEK293
辅助细胞
16h/37 °C
2h/42.5 °C
40h/37 °C
在Vero细胞上共培养 (2-7 days)
CHIKV VLP
重组MV
重感染
感染的Vero细胞

图1B

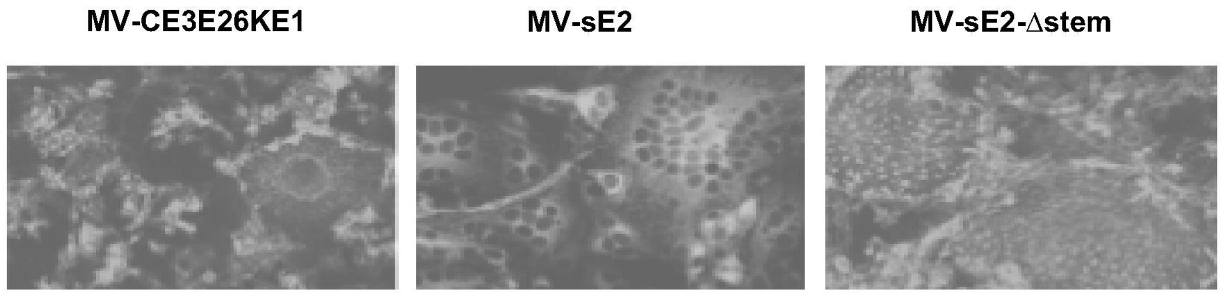

MV-CE3E26KE1
MV-sE2
MV-sE2-Δstem

图2

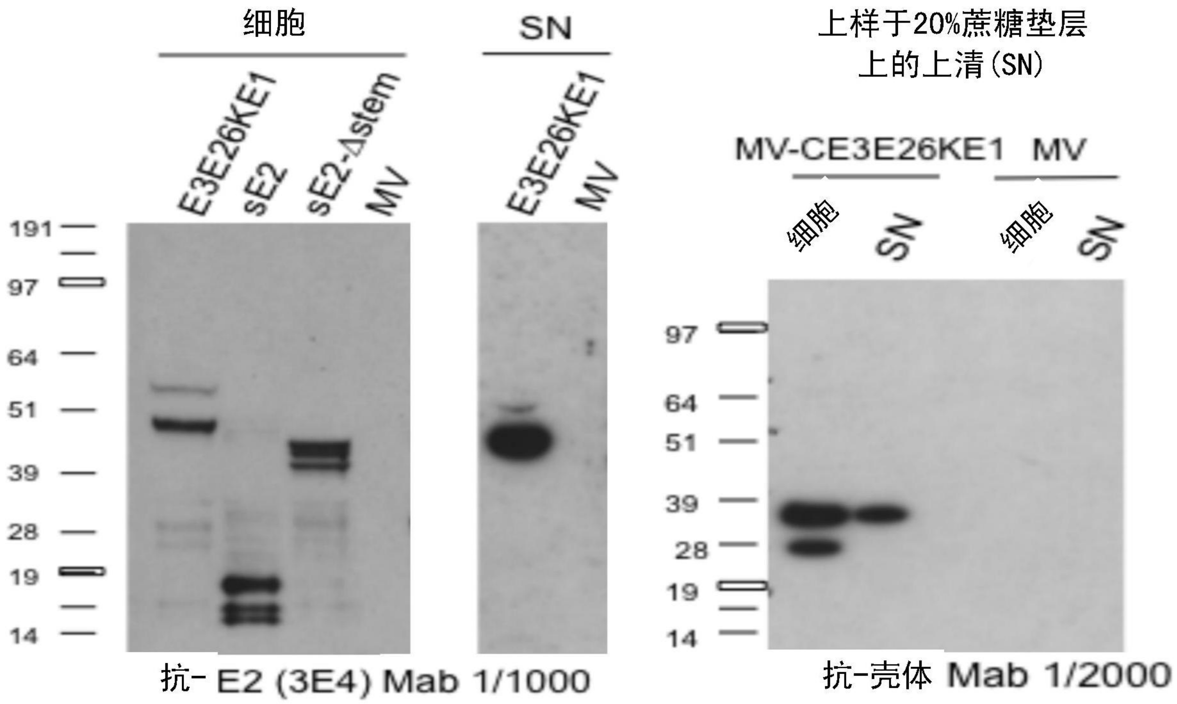

细胞
SN
E3E26KE1
sE2
sE2-Δstem
MV
E3E26KE1
MV
191
97
64
51
39
28
19
14
抗- E2 (3E4) Mab 1/1000
上样于20%蔗糖垫层
上的上清(SN)
MV-CE3E26KE1
MV
细胞
SN
细胞
SN
97
64
51
39
28
19
14
抗-壳体 Mab 1/2000

图3

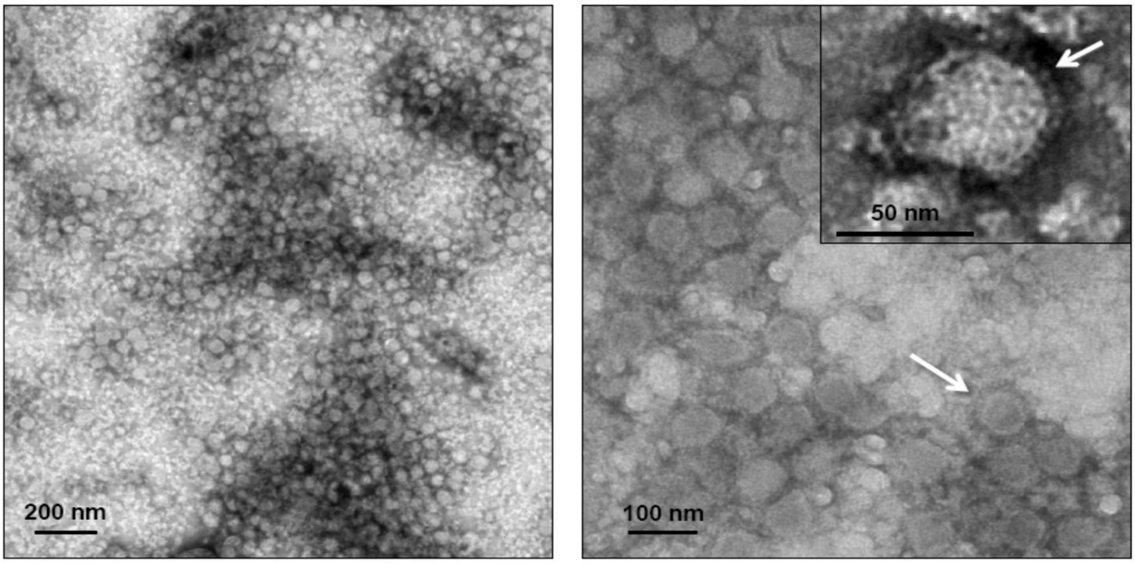


图4

SLAIPVMCLLANTTFPCSQDNFNVYKATRPYLAHCPDCGEGHSCHSPVAL
ERIRNEATDGTLKIQVSLQIGIKTDDSHDWTKLRYMDNHMPADAERAGLF
VRTSAPCTITGTMGHFILARCPKGETLTVGFTDSRKISHSCTHPFHHDPP
VIGRE*

图5

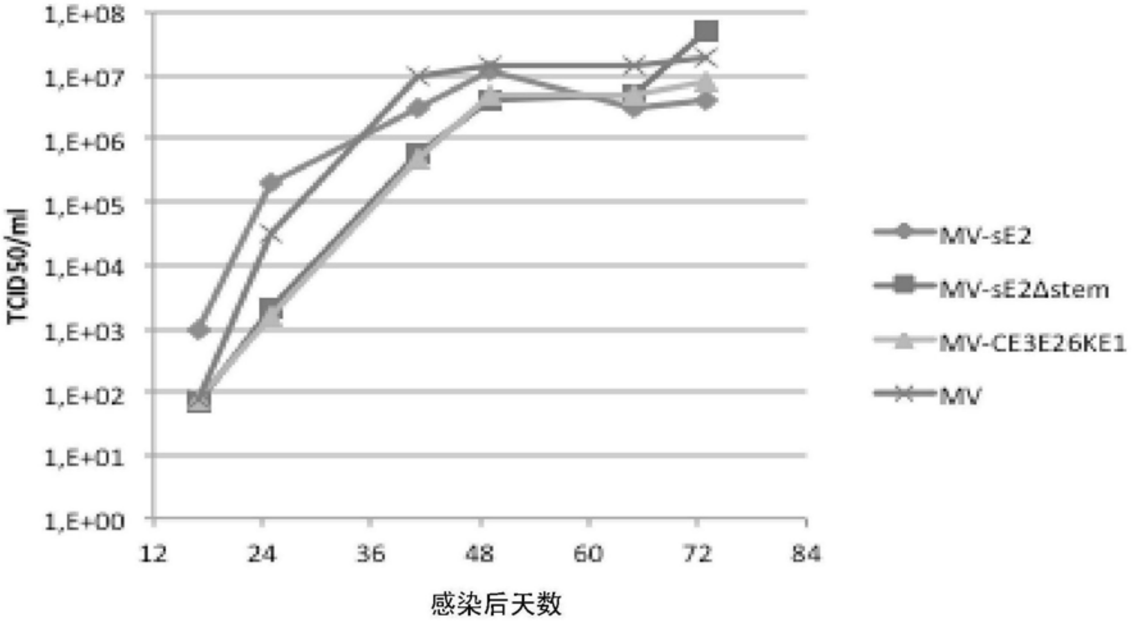


图6

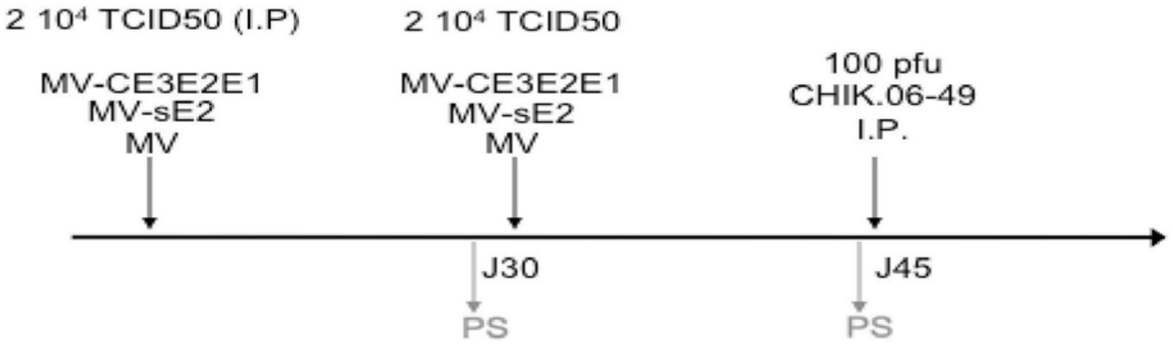

2 10⁴ TCID50 (I.P)
MV-CE3E2E1
MV-sE2
MV
2 10⁴ TCID50
MV-CE3E2E1
MV-sE2
MV
100 pfu
CHIK.06-49
I.P.
J30
PS
J45
PS

图7

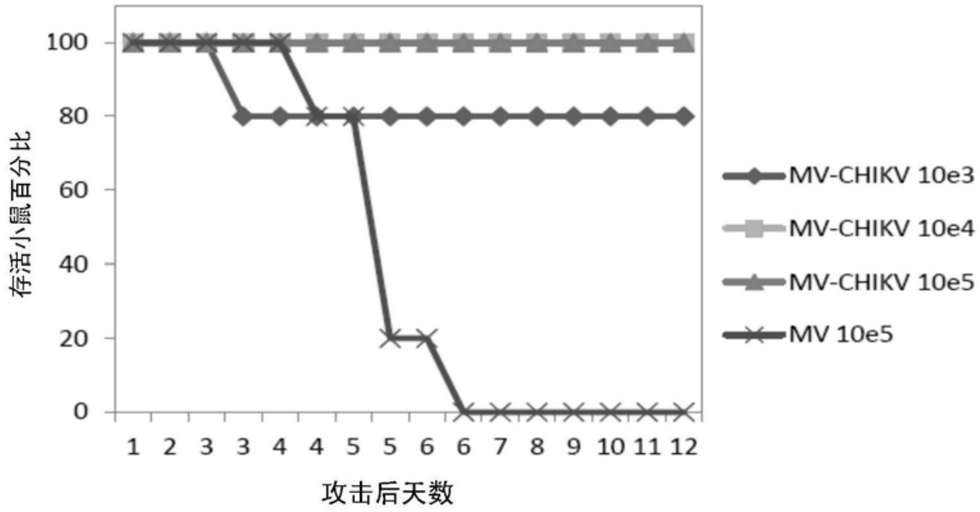

100
80
60
40
20
0
存活小鼠百分比
1 2 3 3 4 4 5 5 6 6 7 8 9 10 11 12
攻击后天数
MV-CHIKV 10e3
MV-CHIKV 10e4
MV-CHIKV 10e5
MV 10e5

图8

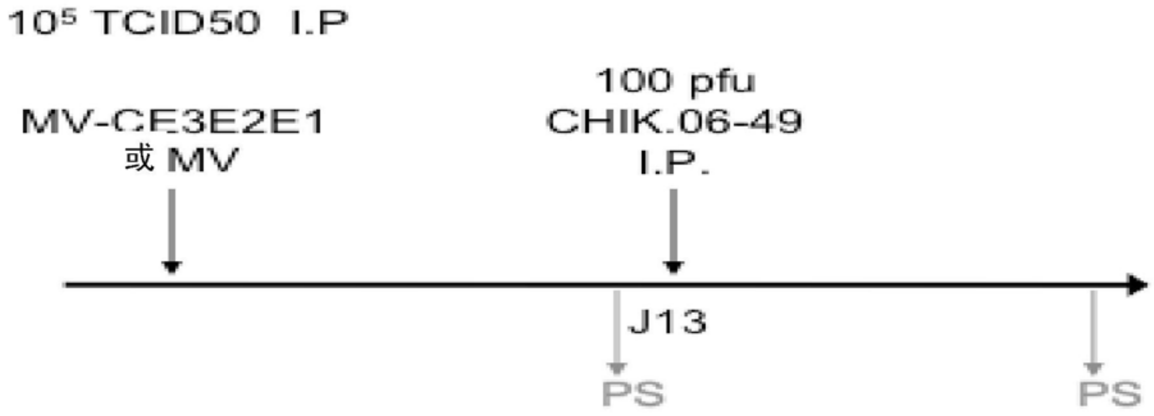

10⁵ TCID50 I.P
MV-CE3E2E1
或 MV
100 pfu
CHIK.06-49
I.P.
J13
PS
PS

图9

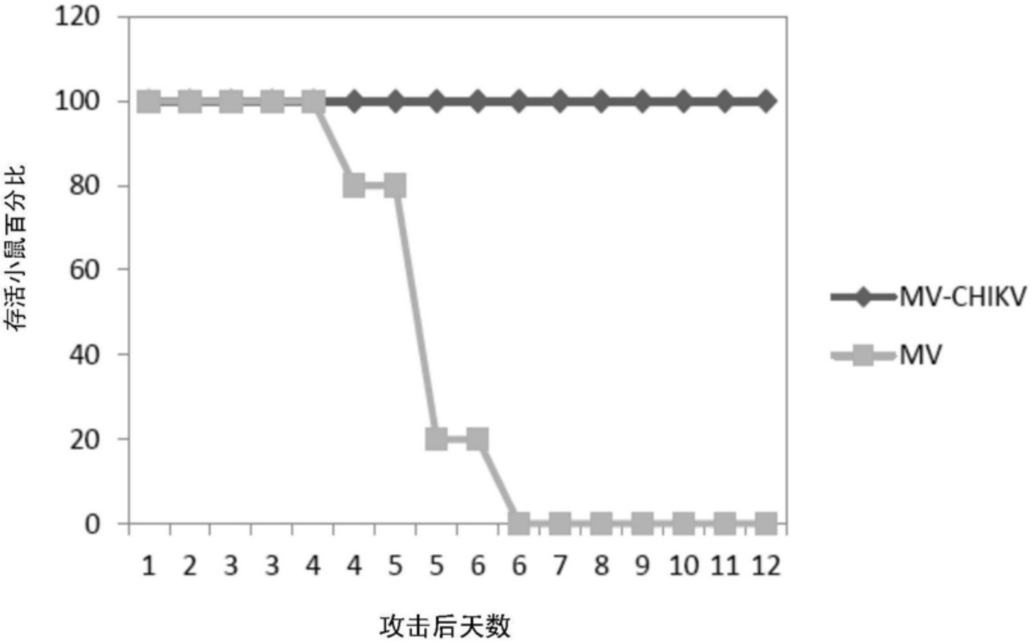
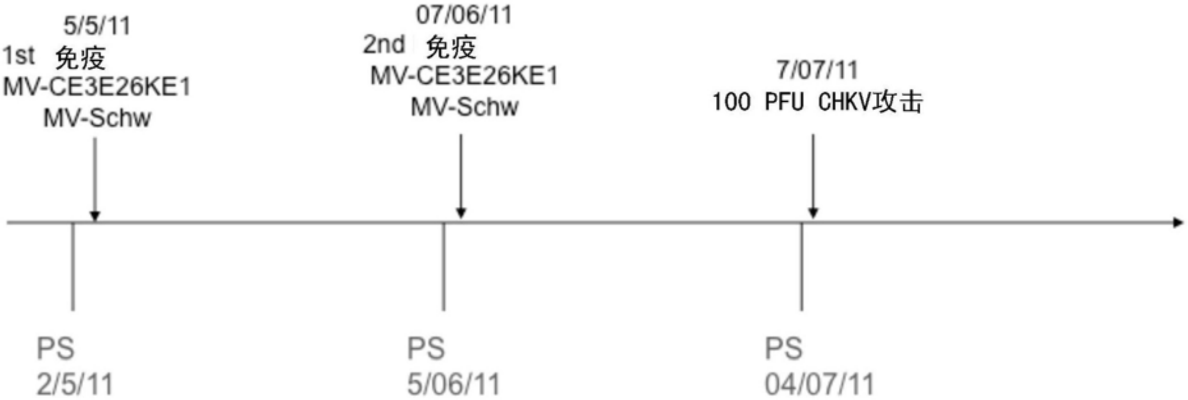


图10



MV-CE3E26KE1：
组1：6只小鼠，10^3 TCID/小鼠
组2：6只小鼠，10^4 TCID/小鼠
组3：6只小鼠，10^5 TCID/小鼠

MV-Schw：
组4：6只小鼠，10^5 TCID/小鼠

图11

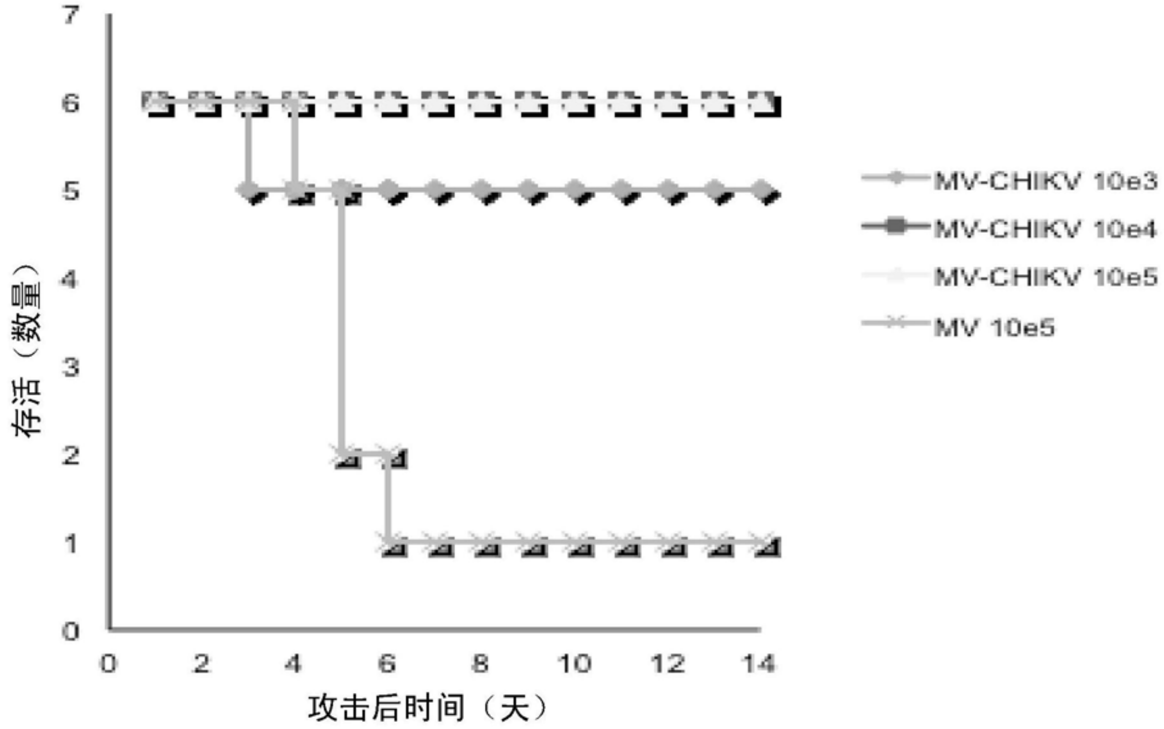

存活（数量）
0
1
2
3
4
5
6
7
0
2
4
6
8
10
12
14
攻击后时间（天）
MV-CHIKV 10e3
MV-CHIKV 10e4
MV-CHIKV 10e5
MV 10e5

图12

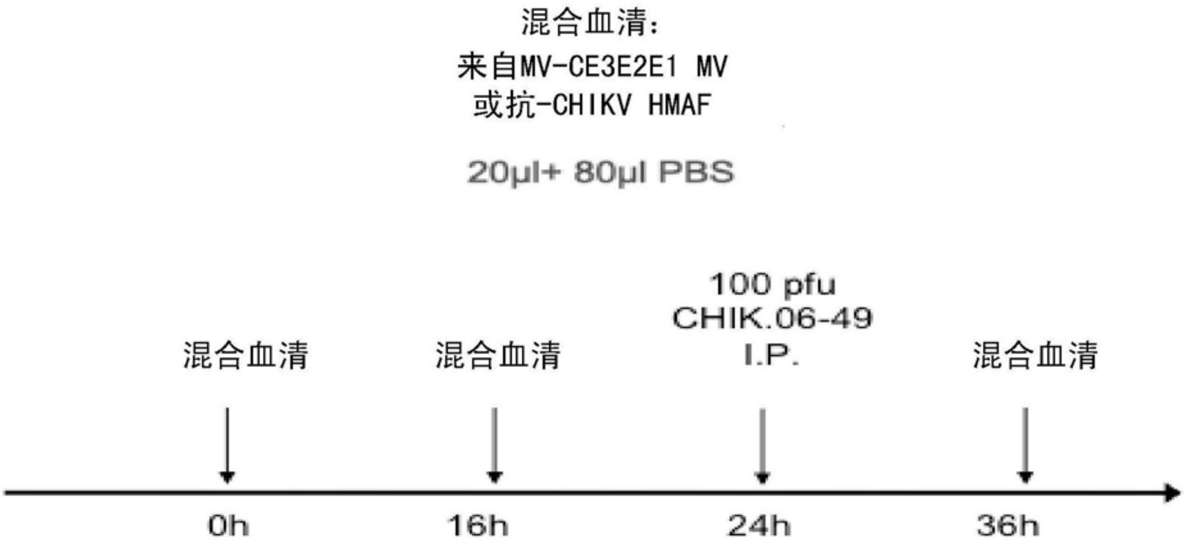

混合血清：
来自MV-CE3E2E1 MV
或抗-CHIKV HMAF
20μl+ 80μl PBS
混合血清
混合血清
100 pfu
CHIK.06-49
I.P.
混合血清
0h
16h
24h
36h

图13

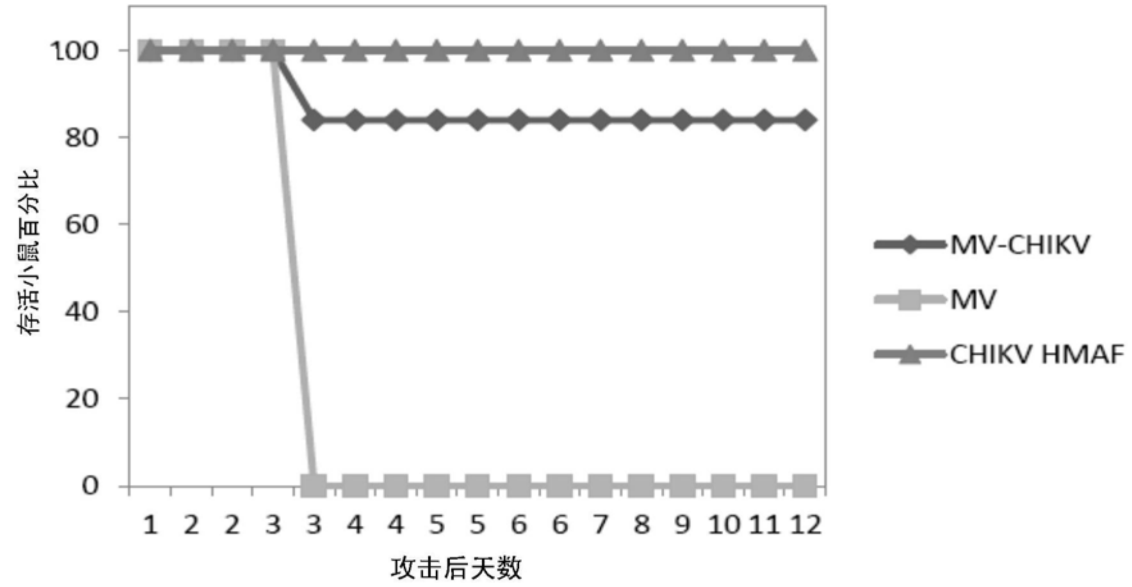

100
80
60
40
20
0
存活小鼠百分比
1 2 2 3 3 4 4 5 5 6 6 7 8 9 10 11 12
攻击后天数
MV-CHIKV
MV
CHIKV HMAF

图14

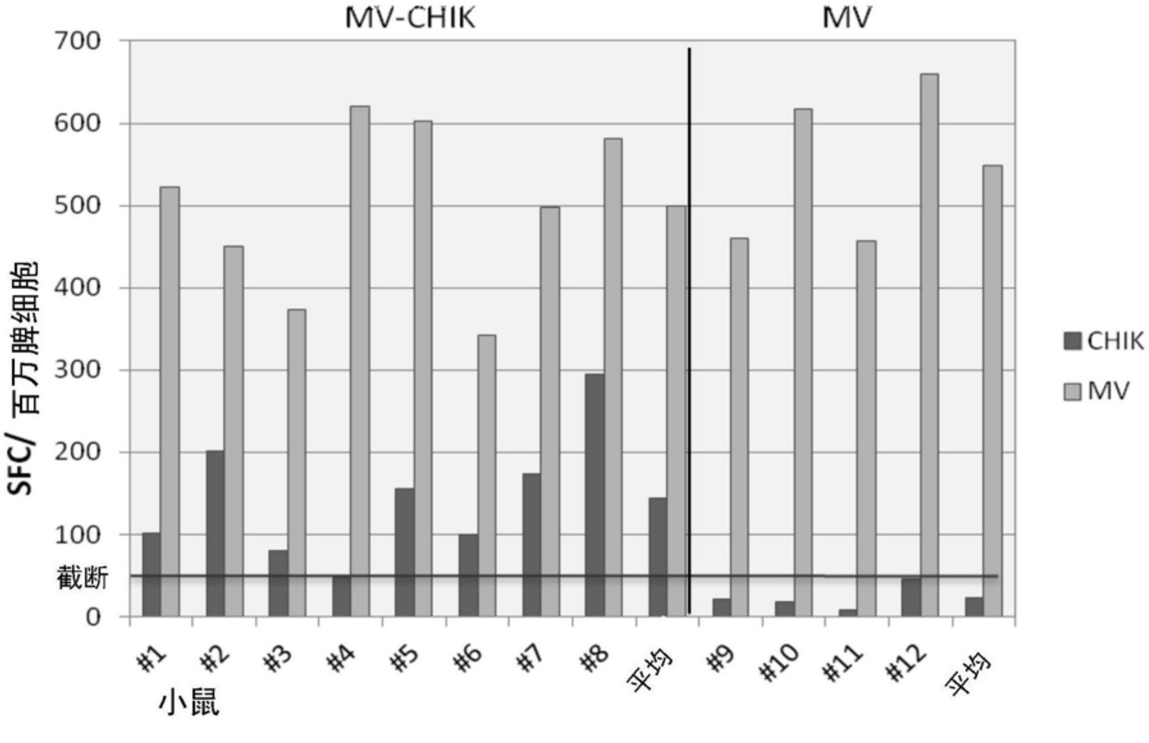

MV-CHIK
MV
700
600
500
400
300
200
100
截断
0
SFC/百万脾细胞
CHIK
MV
#1
#2
#3
#4
#5
#6
#7
#8
平均
#9
#10
#11
#12
平均
小鼠

图15

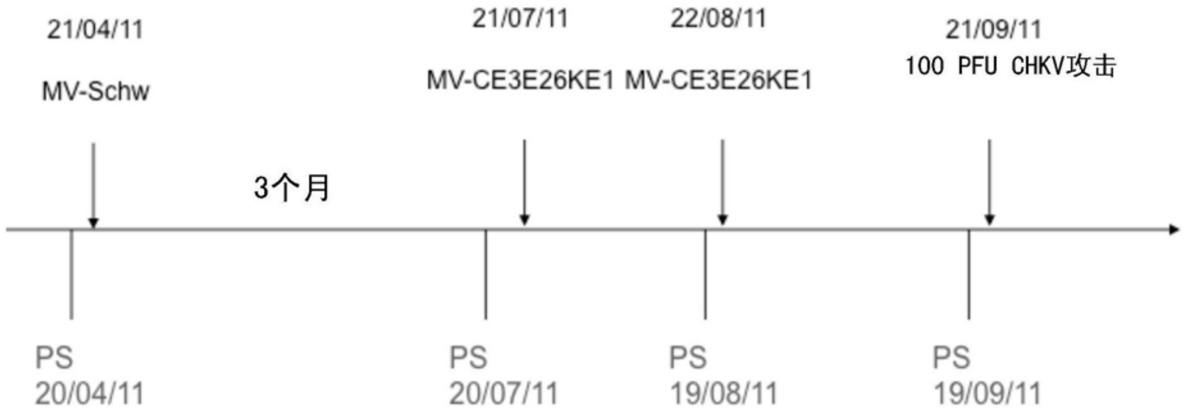

21/04/11
MV-Schw
21/07/11
MV-CE3E26KE1
22/08/11
MV-CE3E26KE1
21/09/11
100 PFU CHKV攻击
3个月
PS
20/04/11
PS
20/07/11
PS
19/08/11
PS
19/09/11

图16

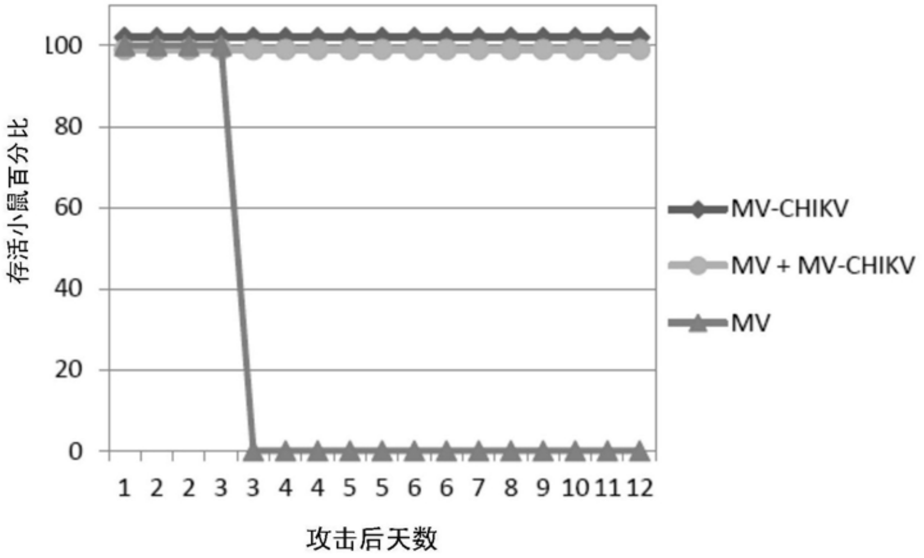

100
80
60
40
20
0
存活小鼠百分比
1 2 2 3 3 4 4 5 5 6 6 7 8 9 10 11 12
攻击后天数
MV-CHIKV
MV + MV-CHIKV
MV

图17

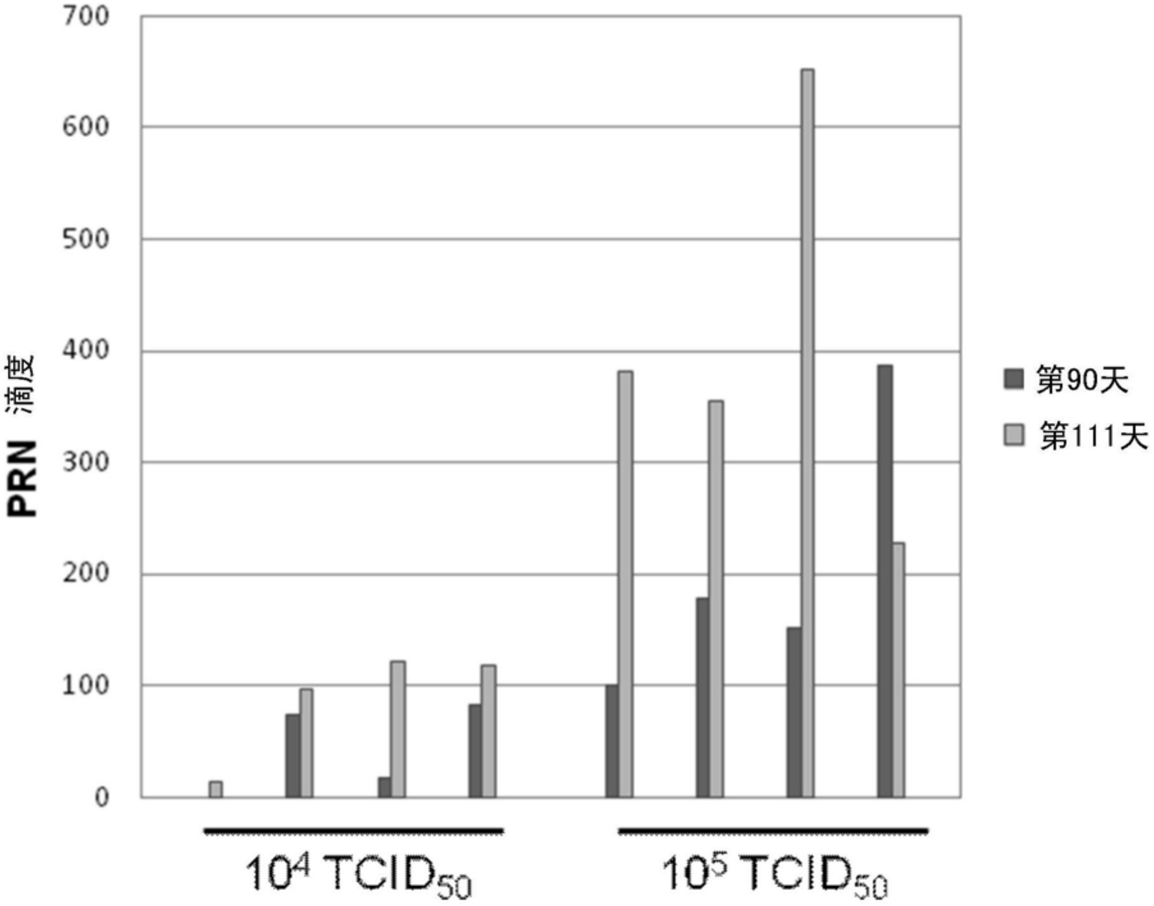

700
600
500
400
300
200
100
0
PRN 滴度
第90天
第111天
10^4 TCID$_{50}$
10^5 TCID$_{50}$

图18