



(12) 发明专利申请

(10) 申请公布号 CN 114728054 A

(43) 申请公布日 2022.07.08

(21) 申请号 202080071104.5

(74) 专利代理机构 中国贸促会专利商标事务所
有限公司 11038

(22) 申请日 2020.09.11

专利代理师 李程达

(30) 优先权数据

19197202.5 2019.09.13 EP

(51) Int.Cl.

A61K 39/12 (2006.01)

(85) PCT国际申请进入国家阶段日

A61P 31/14 (2006.01)

2022.04.11

(86) PCT国际申请的申请数据

PCT/EP2020/075541 2020.09.11

(87) PCT国际申请的公布数据

W02021/048402 EN 2021.03.18

(71) 申请人 勒芬天主教大学

地址 比利时勒芬

(72) 发明人 K·达迈尔 V·莱门斯 J·奈茨

L·桑切斯费利佩

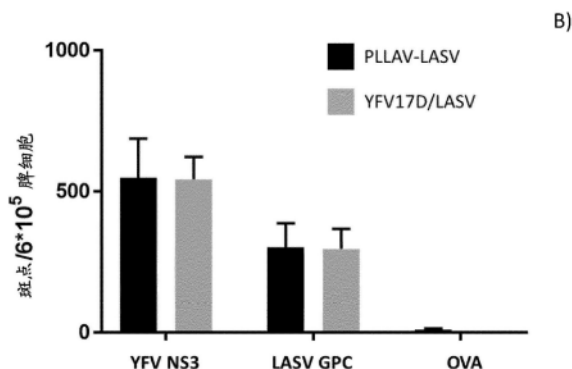
权利要求书1页 说明书15页
序列表15页 附图7页

(54) 发明名称

拉沙病毒疫苗

(57) 摘要

本发明涉及包含活的、感染性的、减毒的黄病毒序列的多核苷酸,其中编码沙粒病毒糖蛋白的至少一部分的核苷酸序列位于所述黄病毒的E基因和NS1基因之间的基因间区域,使得表达嵌合病毒,其特征在于所述黄病毒E蛋白的C末端和所述黄病毒NS1蛋白的信号肽的N末端的编码序列按以下顺序包含:-黄病毒NS1蛋白的另一信号肽,-缺乏N末端信号序列和GP2跨膜结构域的沙粒病毒糖蛋白,-黄病毒E蛋白的TM1和TM2结构域。



1. 多核苷酸,其包含活的、感染性的、减毒的黄病毒的序列,其中编码沙粒病毒糖蛋白的至少一部分的核苷酸序列位于所述黄病毒的E基因和NS1基因之间的基因间区域,使得表达嵌合病毒,其特征在于所述黄病毒E蛋白的C末端和所述黄病毒NS1蛋白的信号肽的N末端的编码序列按以下顺序包含:

- 黄病毒NS1蛋白的另一信号肽,
- 缺乏N末端信号序列和GP2跨膜结构域的沙粒病毒糖蛋白,
- 黄病毒E蛋白的TM1和TM2结构域。

2. 根据权利要求1所述的多核苷酸,其中所述活的、感染性的、减毒的黄病毒的序列是黄热病病毒,通常是YF17D株。

3. 根据权利要求1或2所述的多核苷酸,其中所述沙粒病毒是Mammarena病毒。

4. 根据权利要求1至3中任一项所述的多核苷酸,其中所述沙粒病毒是拉沙病毒。

5. 根据权利要求1至4中任一项所述的多核苷酸,其中编码的蛋白质序列包含SEQ ID NO:2或SEQ ID NO:4的序列。

6. 根据权利要求1至5中任一项所述的多核苷酸,其中所述糖蛋白包含R207C、G360C和E329P稳定突变。

7. 根据权利要求1至6中任一项所述的多核苷酸,其中所述糖蛋白包含R246A蛋白水解切割位点。

8. 根据权利要求1至7中任一项所述的多核苷酸,其中所述NS1蛋白的信号肽包含序列DQGCAINFG[SEQ ID NO:10]或由其组成。

9. 根据权利要求1至8中任一项所述的多核苷酸,其中所述黄病毒E蛋白的TM1结构域具有SEQ ID NO:14的序列,或其中所述黄病毒E蛋白的TM2结构域具有SEQ ID NO:15的序列,或其中所述嵌合病毒在NS1信号序列和GP1结构域连接处的序列包含SEQ ID NO:11的序列。

10. 根据权利要求1至9中任一项所述的多核苷酸,其中所述嵌合病毒在GP2结构域和TM1结构域的连接处的序列包含SEQ ID NO:12的序列,或其中所述嵌合病毒在TM2结构域和NS1蛋白的连接处的序列包含SEQ ID NO:13的序列。

11. 根据权利要求1至10中任一项所述的多核苷酸,其编码包含SEQ ID NO:11、SEQ ID NO:12和SEQ ID NO:13的序列;以及在SEQ ID NO:11、SEQ ID NO:12和SEQ ID NO:13之外与SEQ ID NO:2或SEQ ID NO:4至少95%、96%、97%、98%或99%同一的序列。

12. 一种活的、感染性的、减毒的嵌合黄病毒,其中至少一部分沙粒病毒糖蛋白位于所述黄病毒的E蛋白和NS1蛋白之间,使得所述E蛋白的C末端和所述NS1蛋白的信号肽的N末端的所述病毒按以下顺序包含:

- 黄病毒NS1蛋白的另一信号肽,
- 缺乏N末端信号序列和GP2跨膜结构域的沙粒病毒糖蛋白,
- 黄病毒E蛋白的TM1和TM2结构域。

13. 根据权利要求12所述的嵌合黄病毒,其中所述黄病毒是YFV。

14. 根据权利要求12或13所述的嵌合黄病毒,其中所述沙粒病毒是拉沙病毒。

15. 根据权利要求1至11中任一项所述的多核苷酸或根据权利要求12至14中任一项所述的嵌合病毒,其用作药物。

拉沙病毒疫苗

技术领域

[0001] 本发明涉及基于嵌合黄病毒的疫苗。本发明还涉及针对病毒例如拉沙病毒的疫苗。

背景技术

[0002] 目前,没有获得批准的针对拉沙病毒(LASV)的许可人类疫苗。已经产生了许多不同的候选疫苗,涉及多种平台技术。最先进的候选物是基于VSV的LASV(VSV-LASV-GPC)、Mopeia病毒(MOPV)/LASV重配病毒(克隆ML29)和称为INO-4500(pLASV-GPC)的DNA疫苗。基于VSV的LASV候选疫苗包含具有复制能力的VSV,其表达LASV的糖蛋白。ML29是拉沙病毒和Mopeia病毒之间的重配株,携带MOPV的L区段和LASV的S区段(核蛋白和糖蛋白)。INO-4500是一种编码来自Josiah f株的LASV-GPC基因的DNA疫苗,它来自Inovio公司(pLASV-GPC)。

[0003] 除了使用上述不同的方法外,黄热病病毒17D也已被用作拉沙病毒糖蛋白(GPC)或其亚基GP1和GP2的载体(Bredenbeek等人(2006) *Virology* 345,299-304和Jiang等人(2011) *Vaccine* 29,1248-1257)。在这些构建体中,GP基因(缺乏信号肽,SSP)(或GP1或GP2序列)被插入YF-E/NS1之间。这些构建体在C末端具有源自YF-E、WNV-E或人工设计序列的插入融合序列。这些构建体需要转染在细胞中,从中衍生的病毒被用作疫苗。

[0004] 唯一刚刚开始I期临床试验的疫苗是INO-4500(pLASV-GPC)。这种疫苗需要通过皮肤电穿孔提供多个高剂量,以实现全面保护并增强疫苗免疫反应。这种多剂量施用方案在LASV地方性流行并且已发生主要暴发的西非农村地区的实施将非常具有挑战性。

[0005] 关于其他候选物,ML29被EU归为风险组2,以及被US CDC归为风险组3,这是进一步开发该疫苗的障碍。VSV-LASV-GPC仍然需要冷链来保存它,这涉及到高成本,并且仍然没有关于其安全性的研究。涉及YF17D作为表达拉沙糖蛋白前体的载体的方法在NHP研究中并不成功(0%存活率,狨猴)。

[0006] 此外,该候选疫苗显示出遗传不稳定性问题,无法按照疫苗生产的要求扩大技术规模。

发明内容

[0007] 我们使用我们的PLLAV(质粒启动减毒活疫苗)技术和黄热病减毒活疫苗株(YFV-17D)作为载体,通过将LASV-GPC(在切割位点R246A中具有突变以保持GP1和GP2结合,以及额外的突变R207C、G360C和E329P)插入黄热病E/NS1基因间区域来设计转基因疫苗,如下所示:删除N末端(Nt)信号肽,将NS1的前9个氨基酸(27个核苷酸)添加到LASV-GPC的Nt,以允许正确释放LASV-GPC蛋白,删除跨膜结构域,将胞外结构域与WNV跨膜结构域1和2融合。由此产生的PLLAV-YFV17D-LASV-GPC启动了表达功能性LASV-GPC和YFV-17D蛋白的可行减毒活病毒。PLLAV-YFV17D-LASV-GPC构建体可以直接用作疫苗,这涉及该疫苗是热稳定的。该疫苗在单次注射后诱导针对LASV和YFV的免疫反应。已经生成了第二种类似的构建体,其中切割位点已恢复(R246A突变恢复为R246R)。(附加数据中的附加信息)

[0008] PLLAV-YFV17D-LASV-GPC是一种诱导YFV和拉沙病毒特异性免疫的双重疫苗。PLLAV-YFV17D-LASV-GPC也可用作生产组织培养衍生的减毒活疫苗的稳定种子,不仅在PLLAV模式中,而且出乎意料的是,重组YFV17D-LASV-GPC病毒似乎在基因上超过Bredenbeek等人和Jiang等人(上文引用)在现有技术中公开的。

[0009] 本发明进一步概括为以下陈述:

[0010] 1. 多核苷酸,其包含活的、感染性的、减毒的黄病毒的序列,其中编码沙粒病毒糖蛋白的至少一部分的核苷酸序列位于所述黄病毒的E基因和NS1基因之间的基因间区域,使得表达嵌合病毒,其特征在于所述黄病毒E蛋白的C末端和所述黄病毒NS1蛋白的信号肽的N末端的编码序列按以下顺序包含:

[0011] -黄病毒NS1蛋白的另一信号肽,

[0012] -缺乏N末端信号序列和GP2跨膜结构域的沙粒病毒糖蛋白,

[0013] -黄病毒E蛋白的TM1和TM2结构域。

[0014] 2. 根据权利要求1所述的多核苷酸,其中所述活的、感染性的、减毒的黄病毒的序列是黄热病病毒,通常是YF17D株。

[0015] 3. 根据权利要求1所述的多核苷酸,其中所述活的、感染性的、减毒的黄病毒主链是两种不同黄病毒的嵌合主链。

[0016] 4. 根据权利要求1至3中任一项所述的多核苷酸,其中所述沙粒病毒是Mammarena病毒。

[0017] 5. 根据权利要求1至4中任一项所述的多核苷酸,其中所述沙粒病毒是拉沙病毒。

[0018] 6. 根据权利要求1至3中任一项所述的多核苷酸,其中所述拉沙病毒株是Josiah株。

[0019] 7. 根据权利要求1至6中任一项所述的多核苷酸,其中所述糖蛋白包含R207C、G360C和E329P稳定突变。

[0020] 8. 根据权利要求1至7中任一项所述的多核苷酸,其中所述糖蛋白包含R246A蛋白水解切割位点。

[0021] 9. 根据权利要求1至8中任一项所述的多核苷酸,其中G蛋白的核苷酸序列被密码子优化以改善在哺乳动物细胞中的表达。

[0022] 10. 根据权利要求1至9中任一项所述的多核苷酸,其中所述NS1蛋白的信号肽包含序列DQGCAINFG[SEQ ID NO:10]或由其组成。

[0023] 11. 根据权利要求1至10中任一项所述的多核苷酸,其中所述黄病毒E蛋白的TM1和TM2结构域来自西尼罗河病毒。

[0024] 12. 根据权利要求1至11中任一项所述的多核苷酸,其中所述黄病毒E蛋白的TM1结构域具有SEQ ID NO:14的序列。

[0025] 13. 根据权利要求1至12中任一项所述的多核苷酸,其中所述黄病毒E蛋白的TM2结构域具有SEQ ID NO:15的序列。

[0026] 14. 根据权利要求1至13中任一项所述的多核苷酸,其中所述嵌合病毒在NS1信号序列和GP1结构域连接处的序列包含SEQ ID NO:11的序列。

[0027] 15. 根据权利要求1至14中任一项所述的多核苷酸,其中所述嵌合病毒在GP2结构域和TM1结构域的连接处的序列包含SEQ ID NO:12的序列。

[0028] 16. 根据权利要求1至14中任一项所述的多核苷酸,其中所述嵌合病毒在TM2结构域和NS1蛋白的连接处的序列包含SEQ ID NO:13的序列。

[0029] 在优选的实施方案中,连接黄病毒NS1信号序列、拉沙病毒G蛋白、TM2蛋白和第二NS1信号序列的连接序列为编码的蛋白质提供指纹。因此,编码序列的实施方案可以由含有SEQ ID NO:2或SEQ ID NO:4的序列的序列来定义,包括含有SEQ ID NO:11、SEQ ID NO:12和SEQ ID NO:13的序列;并且其中在SEQ ID NO:11、SEQ ID NO:12和SEQ ID NO:13之外,许多氨基酸可能不同于SEQ ID NO:2或SEQ ID NO:4,例如与SEQ ID NO:2或SEQ ID NO:4相比,差异高达20、高达10或高达5,或例如与SEQ ID NO:2或SEQ ID NO:4具有至少95%、96%、97%、98%或99%的序列同一性。

[0030] 17. 根据权利要求1至16中任一项所述的多核苷酸,其为细菌人工染色体。

[0031] 18. 根据权利要求1至17中任一项所述的多核苷酸,其用作药物。

[0032] 19. 根据权利要求18所述的用作药物的多核苷酸,其中所述药物是疫苗。

[0033] 20. 根据权利要求1至17中任一项所述的多核苷酸序列,其用于针对沙粒病毒感染的疫苗接种。

[0034] 21. 一种活的、感染性的、减毒的嵌合黄病毒,其中至少一部分沙粒病毒糖蛋白位于所述黄病毒的E蛋白和NS1蛋白之间,使得所述E蛋白的C末端和所述NS1蛋白的信号肽的N末端的所述病毒按以下顺序包含:

[0035] -黄病毒NS1蛋白的另一信号肽,

[0036] -缺乏N末端信号序列和GP2跨膜结构域的沙粒病毒糖蛋白,

[0037] -黄病毒E蛋白的TM1和TM2结构域。

[0038] 22. 根据权利要求21所述的嵌合黄病毒,其中所述黄病毒是YFV。

[0039] 23. 根据权利要求21或22所述的嵌合黄病毒,其中所述沙粒病毒是拉沙病毒。

[0040] 24. 根据权利要求21至23中任一项所述的嵌合病毒,其用作药物。

[0041] 25. 根据权利要求21至24中任一项所述的嵌合病毒,其用于预防沙粒病毒感染。

[0042] 26. 由根据权利要求21至23中任一项的核苷酸编码的嵌合病毒,其用于预防沙粒病毒感染和用于预防黄病毒。

[0043] 27. 一种制备针对沙粒病毒感染的疫苗的方法,其包括以下步骤:

[0044] 提供BAC,其包括:

[0045] 诱导型细菌ori序列,其用于将所述BAC扩增至每个细菌细胞超过10个拷贝,和

[0046] 病毒表达盒,其包含根据权利要求1至16中任一项所述的沙粒病毒-黄病毒嵌合病毒的cDNA,并且包含用于在哺乳动物细胞中转录所述病毒cDNA和用于将转录的RNA加工成感染性RNA病毒的顺式调节元件,

[0047] -用步骤a)的BAC转染哺乳动物细胞并传代被感染的细胞,

[0048] -验证步骤b)的转染细胞的复制病毒的毒力和产生抗体并诱导针对所述沙粒病毒感染的保护的能力,

[0049] -将步骤c)中验证的病毒克隆到载体中,以及

[0050] -将载体配制成疫苗制剂。

[0051] 28. 根据权利要求27所述的方法,其中所述黄病毒是黄热病病毒。

[0052] 29. 根据权利要求27或28所述的方法,其中所述沙粒病毒是拉沙病毒。

[0053] 30. 根据权利要求27至29中任一项所述的方法,其中所述载体是BAC,其包含用于将所述BAC扩增至每个细菌细胞超过10个拷贝的诱导型细菌ori序列。

[0054] 详细说明

附图说明

[0055] 图1:1) PLLAV-YFV17D-LASV-GPC和2) PLLAV-YFV17D-LASV-GPCcs的示意图。

[0056] 图2:A) 与YFV17D相比,YFV17D-LASV-GPC的噬斑表型。B) 病毒稳定性:在YFV17D-LASV-GPC病毒连续传代(BHK-21J和VeroE6)期间收获的病毒样品的RT-PCR分析。C+,对照阳性PLLAV-YFV17D-LASV-GPC;-RT:无逆转录酶的RT-PCR反应;RNA:与病毒RNA的RT-PCR反应。

[0057] 图3:疫苗接种时间表示意图。用PLLAV-YFV17D-LASV-GPC (25ug, i. p.) 或YFV17D-LASV-GPC (375PFU) 接种AG129小鼠。

[0058] 图4:接种疫苗的AG129小鼠的细胞免疫分析。A) 用指定抗原刺激脾细胞48小时后的代表性IFN- γ ELISPOT孔。B) 用指定抗原刺激48小时后,IFN- γ ELISPOT中每60万个脾细胞的斑点。对于每只小鼠,一式两份地分析样品,并通过减去对照孔(卵白蛋白刺激)中的斑点数来标准化值。

[0059] 图5:A) 与YFV17D相比,YFV17D-LASV-GPCcs的噬斑表型。B) 通过免疫荧光检测用PLLAV-YFV17D-LASV-GPCcs转染的细胞的上清液感染的BHK21J细胞的LASV-GPC和YFV抗原的共表达。在感染后48h固定细胞并针对LASV-GPC(红色)和YFV(绿色)进行染色。

[0060] 图6:A) 疫苗接种时间表示意图。AG129小鼠皮下(SC)接种YFV17D-LASV-GPCcs (250PFU)。B) 接种疫苗的AG129小鼠的细胞免疫分析。用指定抗原刺激脾细胞48小时后的代表性IFN- γ ELISPOT孔。用指定抗原刺激48小时后,IFN- γ ELISPOT中每六十万个脾细胞的斑点。对于每只小鼠,一式两份分析样品,并通过减去对照孔(卵白蛋白肽刺激)中的斑点数来标准化值。

[0061] 本发明以黄热病病毒为例,但也适用于使用黄病毒物种的其他病毒主链,例如但不限于日本脑炎、登革热、墨累山谷脑炎(Murray Valley Encephalitis,MVE)、圣路易斯脑炎(St. Louis Encephalitis,SLE)、西尼罗河(WN)、蜱传脑炎(TBE)、俄罗斯春夏脑炎(RSSE)、库京病毒、波瓦桑病毒、科萨努尔森林病病毒、寨卡病毒、乌苏图病毒(Usutu virus)、韦塞尔斯布朗和鄂木斯克出血热病毒。

[0062] 本发明还适用于黄病毒科,该科包括黄病毒属,但也包括Pegivirus属、丙型肝炎病毒属和瘟病毒属。

[0063] 丙型肝炎病毒属包括例如丙型肝炎病毒(hepatitis C virus)和乙型肝炎病毒(GB病毒B)

[0064] Pegivirus属包括例如Pegivirus A(GB病毒A)、Pegivirus C(GB病毒C)和Pegivirus B(GB病毒D)。

[0065] 瘟病毒属包括例如牛病毒腹泻病毒1和经典猪瘟病毒(先前是猪霍乱病毒)。

[0066] 用作主链的黄病毒本身可以由由不同黄病毒部分组成的嵌合病毒组成。

[0067] 例如,C和NS1-5区域来自黄热病,而prME区域来自日本脑炎或寨卡病毒。

[0068] 本发明以拉沙病毒的G蛋白为例,但也适用于其他沙粒病毒的G蛋白。

[0069] 本发明涉及核苷酸序列和编码的蛋白质,其中在黄病毒的RNA或拷贝DNA(cDNA)中

插入了沙粒病毒的糖蛋白。

[0070] 在Burr等人.(2012) *Viruses* 4,2162-2181和在Nurnberg&Yorke (2012) *Viruses* 4,83-101中讨论了沙粒病毒的糖蛋白。沙粒病毒由两个RNA基因组片段和四种蛋白质组成,聚合酶L、包膜糖蛋白GP(在本发明中也称为G蛋白或GPC)、基质蛋白Z和核蛋白NP。

[0071] 在沙粒病毒生命周期中,GP前体(GPC)的生物合成和成熟由细胞信号肽酶和细胞酶枯草杆菌蛋白酶Kexin同工酶-1(SKI-1)/位点-1蛋白酶(S1P)执行,产生由GP1/GP2和稳定信号肽(SSP)形成的三方成熟GP复合物。

[0072] 根据血清学、遗传和地理数据,Mammarenavirus沙粒病毒分为两个主要亚组:Old World(OW)和New World(NW)复合体。Old World谱系包括原型LCMV和非洲大陆特有的其他病毒,包括拉沙(LASV)、Mopeia(MOPV)、Ippy和Mobala(MOBV)病毒。

[0073] 较大的New World复合体进一步分为三个进化枝,A、B和C。在人类疾病方面进化枝B最重要,因为它包含在南美洲引起出血热(HF)的主要病毒,即Junin(JUNV)、Machupo(MACV)、Guanarito(GTOV)和Sabia(SABV)病毒,还有其他非致病性病毒,如Tacaribe(TCRV)和Amapari病毒(AMPV)。

[0074] 本发明设想使用基于任何上述组、亚组或物种的G蛋白的嵌合构建体。优选实施方案是基于插入黄病毒RNA或cDNA的LASV的G蛋白的构建体。

[0075] 本发明设想使用基于Reptarenavirusse或Hartmanivirussese的G蛋白的嵌合构建体。

[0076] 本发明以拉沙病毒株Josiah的G蛋白为例。该蛋白质的该序列可以例如作为UniProtKB P08669数据库条目或作为NCBI NP_694870.1数据库条目访问。

[0077] 在替代实施方案中,设想的沙粒病毒是其中G蛋白的蛋白质序列与拉沙病毒株Josiah的G蛋白(如上述引用的数据库条目中所公开的)具有至少70、至少80、至少90、至少95或至少99%同一性的序列同一性的病毒。

[0078] 本发明的构建体允许编码的插入物正确呈现到ER腔并进行蛋白水解加工。如拉沙G蛋白所示,插入片段编码的蛋白质缺乏N末端信号序列和GP2跨膜结构域。为了保留所需的拓扑结构,例如WNV的两个跨膜结构域与糖蛋白序列c末端融合。基于该原理,任何免疫原性蛋白质都可以通过本发明的载体呈递,该蛋白质缺乏N末端膜靶向结构域,并且在C末端含有靶向膜的序列,然后是胞质的序列,以允许与NS1蛋白之前的跨膜的膜连接。

[0079] 现在进一步描述本发明的实施方案,其中使用黄病毒作为主链并且使用拉沙病毒的G蛋白作为插入物。

[0080] 不同沙粒病毒的G蛋白之间的高度序列同一性对本领域技术人员在相关序列中鉴定与拉沙病毒G蛋白中存在的序列元件相对应的序列元件来说没有问题。

[0081] 黄病毒具有长度约11,000个核苷酸的正单链RNA基因组。基因组包含5'非翻译区(UTR)、长开放阅读框(ORF)和3'UTR。ORF编码三种结构蛋白(衣壳[C]、前体膜[prM]和包膜[E])和七种非结构蛋白(NS1、NS2A、NS2B、NS3、NS4A、NS4B和NS5)。结构蛋白与基因组RNA一起形成病毒颗粒。非结构蛋白参与病毒多蛋白加工、复制、病毒粒子组装和逃避宿主免疫反应。C蛋白C末端的信号肽(C信号肽;也称为C锚定结构域)通过协调信号肽序列的N末端(通过胞质中的病毒NS2B/NS3蛋白酶)和C末端(通过内质网[ER]腔中的宿主信号酶)的连续切割来调节黄病毒包装。

[0082] 正义单链基因组被翻译成单个多蛋白,被病毒和宿主蛋白共翻译和翻译后切割成三个结构蛋白[衣壳(C)、前膜(prM)、包膜(E)]和七个非结构蛋白(NS1、NS2A、NS2B、NS3、NS4A、NS4B、NS5)。结构蛋白负责形成病毒粒子的(球形)结构,启动病毒粒子粘附、内化和病毒RNA释放到细胞中,从而启动病毒生命周期。另一方面,非结构蛋白负责病毒复制、调节和逃避受感染细胞中的免疫反应,以及将病毒传播给蚊子。结构蛋白和非结构蛋白之间的分子内和分子间相互作用在病毒感染和发病机制中起关键作用。

[0083] E蛋白在其C末端包含两个跨膜序列,表示为TM1和TM2。

[0084] NS1通过对应于E的最后24个氨基酸的信号序列转移到ER腔中,并通过ER驻留宿主信号肽酶的切割在其氨基末端从E释放(Nowak等人(1989) *Virology* 169,365-376)。NS1在其C末端包含8-9个氨基酸的信号序列,其中包含蛋白酶的识别位点(Muller&Young (2013) *Antiviral Res.* 98,192-208)

[0085] 本发明的构建体是嵌合病毒,其中拉沙G蛋白插入在E蛋白和NS1蛋白之间的边界。然而,在G蛋白插入物的N末端和C末端提供了额外的序列元件。

[0086] 本发明涉及包含活的、感染性的、减毒的黄病毒的序列的多核苷酸,其中编码沙粒病毒G蛋白的至少一部分的核苷酸序列插入所述黄病毒的E基因和NS1基因之间的基因间区域,使得表达嵌合病毒,其特征在于所述黄病毒E蛋白的C末端和所述黄病毒NS1蛋白的N末端的编码序列按以下顺序包含:

[0087] -允许通过信号肽酶从E蛋白中蛋白水解加工G蛋白的序列元件。

[0088] -缺乏其信号肽的G蛋白和GP2跨膜蛋白,和

[0089] -黄病毒E蛋白的两个TM结构域

[0090] 为了允许在其氨基末端从黄病毒E蛋白中蛋白水解加工沙粒病毒G蛋白并允许在其C末端从黄病毒NS1蛋白中蛋白水解加工沙粒病毒G蛋白,提供了作为信号肽酶底物的序列元件。这些可以在长度和序列上有所不同,并且可以短至一个氨基酸,如上文引用的Jang等人所示。关于信号传导蛋白酶的合适识别位点的讨论见Nielsen等人(1997) *Protein Eng.* 10,1-6。

[0091] 通常,在G蛋白的C末端,将使用NS1蛋白的N末端处的信号肽(或允许蛋白水解加工的片段)。

[0092] 通常,在G蛋白的N末端,引入黄病毒主链的NS1蛋白的相同的信号肽(或片段)。

[0093] 本发明同样涉及包含活的、感染性的、减毒的黄病毒序列的多核苷酸。在此,将编码至少一部分沙粒病毒G蛋白的核苷酸序列插入所述黄病毒的E基因和NS1基因之间的基因间区域。提供了额外的序列,使得当嵌合病毒被表达时,从E蛋白的C末端到NS1蛋白的信号肽的N末端的编码序列按以下顺序包含:

[0094] 黄病毒NS1基因的另一信号肽(或其可切割片段),E蛋白的C末端以及NS1蛋白的N末端。

[0095] 缺乏功能性信号肽和GP2结构域的跨膜序列的沙粒病毒G蛋白。该G蛋白位于NS1信号肽的C末端。G蛋白的C末端是黄病毒TM1和TM2跨膜结构域的序列。这些TM序列的C末端跟随NS1蛋白,包括其天然信号肽序列。

[0096] 因此,G蛋白和TM结构域在N末端和C末端由NS1序列侧接。在实施例公开的实施方案中,两个NS1的蛋白质和DNA序列是相同的。

[0097] 在典型的实施方案中,两个NS1信号序列都具有序列DQGCAINFG[SEQ ID NO:10]。

[0098] 由于该重复序列的存在,本发明的构建体不显示重组。可以引入序列修饰或使用来自不同黄病毒的NS1序列来避免相同序列的存在,只要编码的肽仍然是处理这些NS1 N末端信号序列的蛋白酶的靶标。

[0099] 在典型的实施方案中,如实施例中所公开的,G蛋白是拉沙病毒的,优选地是拉沙病毒的Josiah株。

[0100] 为了促进在哺乳动物宿主中产生病毒,G蛋白的核苷酸序列进行了密码子优化。

[0101] 提出可以在G蛋白和C末端尾部中引入微小的序列修饰而不会丧失这些序列元件的功能。例如,其中疏水侧链保留在跨膜结构域中的氨基酸取代,或具有足够长度以允许跨膜结构域在胞质结构域的N末端和C末端正确定位的胞质结构域的截短形式。

[0102] 已经发现G蛋白的功能性信号肽的存在导致负选择压力,由此包含其信号肽的G蛋白的一部分被缺失或突变。因此,本发明的构建体通常通过部分或完全去除该序列或通过引入使信号蛋白无功能的突变而包含有缺陷的G蛋白信号。

[0103] 位于G蛋白的C末端和NS1的N末端的TM结构域通常属于黄病毒,通常来自E蛋白,更典型的是E蛋白的TM结构域。在优选的实施方案中,E蛋白的这些TM结构域来自与形成主链的病毒不同的黄病毒。本发明的实施例描述了西尼罗河病毒E蛋白的TM1和TM2结构域。这些结构域具有序列GGMSWITQGLLGALLLWGMINARD[SEQ ID NO:14]和RSIAMTFLAVGGVLLFLSVNVHA[SEQ ID NO:15]。

[0104] 在下面的实施例部分和示意图中,所有序列元件形成一个连续序列,没有任何中间序列元件。认为在这些序列元件之间,只要蛋白质在ER腔或胞质溶胶中的定位不受干扰并且蛋白水解加工得以维持,就可以存在额外的氨基酸。

[0105] 上述核苷酸序列可以是病毒本身的核苷酸序列,也可以指载体中的序列。用于克隆黄病毒和嵌合形式的合适载体是细菌人工染色体等技术,如下文更详细描述。

[0106] 本发明的方法和化合物具有药用价值,由此病毒或编码该病毒的载体可用于针对含有克隆在黄病毒中的G蛋白的沙粒病毒进行疫苗接种。此外,来自黄病毒的蛋白质同样提供保护,使得本发明的化合物可用于使用单一病毒或DNA疫苗针对黄病毒和沙粒病毒进行疫苗接种。

[0107] 细菌人工染色体的使用,尤其是本发明人在W02014174078中公开的诱导型BACS的使用,特别适用于RNA病毒的cDNA(例如本发明的嵌合构建体)的高产量、高质量扩增。

[0108] 本公开内容BAC中描述的BAC包括:

[0109] -用于将所述BAC扩增至每个细菌细胞超过10个拷贝的诱导型细菌ori序列,和

[0110] -病毒表达盒,其包含RNA病毒基因组的cDNA并包含用于在哺乳动物细胞中转录所述病毒cDNA和用于将转录的RNA加工成感染性RNA病毒的顺式调节元件。

[0111] 如本发明的情况,RNA病毒基因组是RNA病毒基因组和沙粒病毒G蛋白的嵌合病毒cDNA构建体。

[0112] 在这些BACS中,病毒表达盒包含正链RNA病毒基因组的cDNA,通常是

[0113] -RNA聚合酶驱动的启动子,位于所述cDNA的5'末端之前,用于启动所述cDNA的转录,和

[0114] -在所述cDNA的3'末端之后用于RNA自切割的元件,用于在设定位置切割所述病毒

cDNA的RNA转录物。

[0115] BAC还可以包含酵母自主复制序列,用于穿梭到酵母中并将所述细菌人工染色体维持在酵母中。酵母ori序列的一个示例是2 μ 质粒起点或ARS1(自主复制序列1)或其功能同源衍生物。

[0116] 本发明第一方面的RNA聚合酶驱动的启动子可以是RNA聚合酶II启动子,例如巨细胞病毒立即早期(CMV-IE)启动子,或猿猴病毒40启动子或其功能同源衍生物。

[0117] RNA聚合酶驱动的启动子同样可以是RNA聚合酶I或III启动子。

[0118] BAC还可以包含用于RNA自切割的元件,例如丁型肝炎病毒基因组核酶的cDNA或功能同源的RNA元件。

[0119] 将DNA配制成疫苗制剂是本领域已知的并且在例如“DNA Vaccines”*Methods in Molecular Medicine* Vol 127, (2006) Springer Saltzman, Shen和Brandsma (Eds.) Humana Press, Totoma, N.J.的第6至10章,以及在*Alternative vaccine delivery methods*, P 1200-1231, of *Vaccines* (第6版) (2013) (Plotkin等人, Eds.)的第61章中有详细描述。关于适用于制备DNA疫苗的可接受的载体、稀释剂、赋形剂和佐剂的详细信息也可以在WO2005042014中找到,如下所示。

[0120] “可接受的载体、稀释剂或赋形剂”是指用于人类和/或兽医学,特别是免疫疗法的可接受的附加物质。

[0121] 例如,可接受的载体、稀释剂或赋形剂可以是可安全用于全身或局部施用的固体或液体填充剂、稀释剂或包封物质。取决于特定的施用途径,可以使用本领域众所周知的各种载体。这些载体可以选自包括以下的组:糖、淀粉、纤维素及其衍生物、麦芽、明胶、滑石、硫酸钙和碳酸盐、植物油、合成油、多元醇、海藻酸、磷酸盐缓冲溶液、乳化剂、等渗盐水和盐例如无机酸盐,包括盐酸盐、溴化物和硫酸盐,有机酸例如乙酸盐、丙酸盐和丙二酸盐,以及无热原水。

[0122] 描述药学上可接受的载体、稀释剂和赋形剂的有用参考文献是Remington's *Pharmaceutical Sciences* (Mack Publishing Co. N.J. USA, 2 (091), 其通过引用并入本文。

[0123] 可以采用任何安全的施用途径为患者提供DNA疫苗。例如,可以使用口服、直肠、肠胃外、舌下、口腔、静脉内、关节内、肌肉内、真皮内、皮下、吸入、眼内、腹膜内、脑室内、透皮等。肌肉内和皮下注射可以是合适的,例如,用于施用免疫治疗组合物、蛋白质疫苗和核酸疫苗。还预期微粒轰击或电穿孔对于核酸疫苗的递送可能特别有用。

[0124] 剂型包括片剂、分散剂、悬浮液、注射剂、溶液、糖浆剂、锭剂、胶囊、栓剂、气雾剂、透皮贴剂等。这些剂型还可以包括专门为此目的设计的注射或植入控制释放装置或经过修饰以这种方式另外起作用的其他形式的植入物。治疗剂的控制释放可以通过例如用疏水聚合物包被来实现,所述疏水聚合物包括丙烯酸树脂、蜡、高级脂肪醇、聚乳酸和聚乙醇酸以及某些纤维素衍生物例如羟丙基甲基纤维素。此外,可以通过使用其他聚合物基质、脂质体和/或微球来实现控制释放。

[0125] 适用于口服或肠胃外施用的DNA疫苗可以呈离散单元,例如胶囊、小袋或片剂,每一个都含有预定量的质粒DNA,呈粉末或颗粒,或呈溶液或在水性液体中的悬浮液,非水性液体、水包油乳液或油包水液体乳液。此类组合物可通过任何药学方法制备,但所有方法都包括将如上所述的一种或多种试剂与构成一种或多种必需成分的载体结合的步骤。通

常,通过将DNA质粒与液体载体或细碎的固体载体或两者均匀且紧密地混合,然后,如果需要,将产品成型为所需的外观来制备组合物。

[0126] 上述组合物可以以与剂型相容的方式施用,并且以有效的量施用。施用给患者的剂量应足以在适当的时间段内对患者产生有益的反应。待施用的一种或多种试剂的量可取决于待治疗的受试者,包括其年龄、性别、体重和一般健康状况,这些因素将取决于从业者的判断。

[0127] 此外,DNA疫苗可以通过细菌转导递送,如使用用所述DNA质粒转化的沙门氏菌减毒活株,例如Darji等人(2000)FEMS Immunol Med Microbiol 27,341-349和Cicin-Sain等人(2003)J Virol 77,8249-8255,给出作为参考。

[0128] 通常,DNA疫苗用于人类的预防性或治疗性免疫接种,但对于某些病毒,也可用于脊椎动物(通常是哺乳动物、鸟类和鱼类),包括家养动物,例如家畜和伴侣动物。设想对作为病毒活库(人畜共患病)的动物进行疫苗接种,例如猴子、狗、小鼠、大鼠、鸟类和蝙蝠。

[0129] 在某些实施方案中,疫苗可包括佐剂,即增强疫苗组合物的免疫原性和/或功效的一种或多种物质。然而,生命疫苗最终可能受到可刺激与病毒复制无关的先天免疫反应的佐剂的损害。合适的佐剂的非限制性实例包括角鲨烷和角鲨烯(或其他动物来源的油);嵌段共聚物;去污剂,例如Tween-80;Quilla,矿物油,例如Drakeol或Marcol,植物油,例如花生油;棒状杆菌来源的佐剂,例如短小棒状杆菌(*Corynebacterium parvum*);丙酸杆菌来源的佐剂,例如痤疮丙酸杆菌(*Propionibacterium acne*);牛分枝杆菌(*Bacille Calmette*和*Guerin*或BCG);白细胞介素,例如白细胞介素2和白细胞介素12;单核因子,例如白细胞介素1;肿瘤坏死因子;干扰素,例如 γ 干扰素;组合,例如皂甙-氢氧化铝或Quil-A氢氧化铝;脂质体;ISCOMt)和ISCOMATRIX(B)佐剂;分枝杆菌细胞壁提取物;合成糖肽,例如胞壁酰二肽或其他衍生物;阿夫立定(Avridine);脂质A衍生物;硫酸葡聚糖;DEAE-葡聚糖或磷酸铝;羧基聚亚甲基,例如Carbopol'EMA;丙烯酸共聚物乳液,例如Neocryl A640;牛痘或动物痘病毒蛋白;亚病毒颗粒佐剂,例如霍乱毒素,或其混合物。

实施例

[0130] 实施例1YFV17D/拉沙构建体

[0131] 将来自Josiah株的拉沙糖蛋白前体(LASV-GPC)插入YF-E/NS1之间,以生成两个构建体,如下所示(图1):

[0132] 1) PLLAV-YFV17D-LASV-GPC:N末端信号肽序列(SSP)和GP2跨膜结构域(TM)缺失的拉沙糖蛋白。LASV糖蛋白切割位点发生突变(R246A)以保持前体GPC(GP1和GP2连接)。引入这些点突变R207C和G360C(共价结合GP1和GP2)和E329P(在Hastie等人(2017)Science 356,923-928中描述)以改善稳定性。这种带有突变的拉沙-GPC与WNV的跨膜结构域(TM1和TM2)融合,以保持复制YFV17D所需的多蛋白拓扑结构并允许LASV-GPC的正确表达。此外,在LASV-GPC序列之前引入了编码YF-NS1前9个氨基酸的序列,以便正确处理抗原。

[0133] 2) PLLAV-YFV17D-LASV-GPCs:与上述构建体类似的构建体,但在此构建体中,恢复切割位点(R246A突变恢复为R246R)。其余突变相似,引入突变R207C和G360C(共价结合GP1和GP2)和E329P(在Hastie等人(2017)Science 356,923-928中描述)以改善稳定性。

[0134] 实施例2构建体#1PLLAV-YFV17D-LASV-GPC

[0135] 将PLLAV-YFV17D-LASV-GPC转染到BHK21J细胞中,观察到典型的CPE,并且与YFV17D的噬斑表型相比,从中收获的病毒上清液形成明显更小的噬斑(图2A)。因此,产生的转基因病毒(YFV17D-LASV-GPC)进一步减毒,与YFV17D相比,病毒产量至少低10倍。

[0136] PLLAV-YFV17D-LASV-GPC的稳定性是通过进行RT-PCR来检测在YFV17D-LASV-GPC连续传代期间收获的病毒样品中的转基因插入物来测定的(图2B)。RT-PCR产物的测序表明,在BHK21J细胞中至少在第5代之前可以检测到没有突变的LASV-GPC插入物。

[0137] 实施例3PLLAV-YFV17D-LASV-GPC在AG129小鼠中的免疫原性

[0138] 在AG129小鼠体内评估了PLLAV-YFV17D-LASV-GPC和衍生的减毒活病毒(LAV)的免疫原性。动物(n=9/组)接种了25 μ g的PLLAV-YFV17D-LASV-GPC或375PFU的YFV17D-LASV-GPC(图3)。YFV和LASV特异性抗体反应通过间接免疫荧光测定法(IIFA)进行量化,细胞介导的免疫反应通过ELISPOT进行量化(图4)。

[0139] 每天监测接种疫苗的小鼠的发病率/死亡率,并在基线和以两周间隔采集血液进行血清学分析。该疫苗是安全的,因为在任何接种疫苗的小鼠中均未观察到不良反应。一些动物(9只小鼠中的4只)在第一次接种PLLAV或LAV YFV17D-LASV-GPC两周后使用与第一次疫苗接种相同的剂量和途径加强免疫(图3)。

[0140] YFV17D-LASV-GPC(PLLAV或LAV)的免疫原性分析显示,在疫苗接种后14天,分别接种PLLAV或LAV的3只和1只小鼠中出现了针对LASV的特异性抗体。值得注意的是,对于LASV,目前认为CD8⁺T细胞反应是负责提供针对LASV感染的保护的主要决定因素。因此,在疫苗接种后4个月对两个组的T细胞反应进行了分析。该分析表明,在接种YFV17D-LASV-GPC(LAV)的所有小鼠中以及接种PLLAV版本后的9只小鼠中的7只小鼠出现了针对LASV和YFV的T细胞反应(图4)。因此,这些T细胞反应可以被认为赋予免疫和保护免受LASV感染。

[0141] 实施例4构建体#2PLLAV-YFV17D-LASV-GPCcs(切割位点)

[0142] 产生了类似于上述构建体的第二个构建体。在该PLLAV-YFV17D-LASV-GPCcs中恢复了GP1和GP2之间的自然切割位点。将该构建体转染到BHK21J细胞中,观察到典型的CPE,并且与YFV17D的噬斑表型相比,从中收获的病毒上清液形成明显更小的噬斑(图5A)。因此,与之前的YFV17D/LASV构建体类似,得到的转基因病毒(YFV17D-LASV-GPC)进一步减毒,与YFV17D相比,病毒产量至少低10倍。

[0143] 可以确认LASV-GPC与YFV多蛋白共表达(图5B),表明LASV-GPC的正确表达和折叠。

[0144] 为了评估该构建体的免疫原性,AG129小鼠皮下(s.c.)接种YFV17D-LASV-GPCcs(LAV)并在接种后28天测定T细胞反应(图6)。分析表明,针对LASV和YFV都有强烈的特异性T细胞反应。这些结果表明,该疫苗可以作为针对两种病毒LASV和YFV的二价疫苗。该疫苗是安全的,因为在任何接种疫苗的小鼠中均未观察到不良反应。

[0145] 本申请中描述的序列

[0146] 构建体#1:PLLAV-YFV17D-LASV-GPC

[0147] (信号肽缺失、跨膜结构域gp2缺失,切割位点突变(r256a)和突变r207c,e329p和g360c)

[0148] -末端YF-E(氨基酸1-40)

[0149] -SNS1(氨基酸41-49)的前27个核苷酸(9个氨基酸)加粗下划线

[0150] -lasv-gp1结构域[无信号肽以及突变tgt(r207c)](氨基酸50-250)

- [0151] -切割位点突变的gcaagattgcta (r256a) [SEQ ID NO:16] (氨基酸247-250)
- [0152] -lasv-p2[无tm以及突变cca (e329p) ,tgt (g360c)] (氨基酸251-418)
- [0153] -WNV TM1 (氨基酸418-442) 下划线
- [0154] -WNV TM2 (氨基酸443-465) 下划线
- [0155] -开始yf-ns1 (氨基酸466-527)
- [0156] SEQ ID NO:1DNA
- [0157] SEQ ID NO:2蛋白质

AAGGTCATCATGGGGCGGTACTTATATGGGTTGGCATCAACACAAGAAACATGACAATG
K V I M G A V L I W V G I N T R N M T M 20
TCCATGAGCATGATCTTGGTAGGAGTGATCATGATGTTTTTGTCTCTAGGAGTTGGcGCc
S M S M I L V G V I M M F L S L G V G A 40
40
GACCAGGGCTGGCGGATAAATTTCCGGTaccagtcctttataaaaggggtttatgatgcttcag
D Q G C A I N F G T S L Y K G V Y E L Q 60
41 49 51
actctggaaactaaacatgggagacactcaatgatgaccatgacctctctctctgcacaaagaac
T L E L N M E T L N M T M P L S C T K N 80
aacagtcacattatataaattgggtgggcaatgagacaggactagaactgaccttgaccaac
N S H H Y I M V G N E T G L E L T L T N 100
acgagcatttataatcacamaattttgcaatctgtctgatgcccacaaaaagaacctctat
T S I I N H K F C N L S D A H K K N L Y 120
gaccacgctcttatgagcatabaacttcaacttccacttctccatcccccaacttcaatcag
D H A L M S I I S T F H L S I P N F N Q 140
tatgagcmaatgagctgctgatttttaattgggggaaagattagtggtgcagtacaacctgagt
Y E A M S C D F N G G K I S V Q Y N L S 160
cacagcttatgtggggatgacagcccaaccattgtggtactgttgcacaaatgggtgtttacag
H S Y A G D A A N H C G T V A N G V L Q 180
acttttatgaggatggcttgggtgggagctacattgctctctgactcagggcIgtggcaac
T F M R M A W G G S Y I A L D S G C G N 200
R207C
tgggactgttattatgactagttatcaatctctgataatccccaaaatacaacctggggaagat
W D C I M T S Y Q Y L I I Q N T T W E D 220
cactgccaattctcgagaccatctcccatcggttatctctgggctcctctccacaaaggact
H C Q F S R P S P I G Y L G L L S Q R T 240
agagatatttatattagtgAGATTGCTAGGCACATTCAATGGACACTGTTCAGATTCT
R D I Y I S A R L L G T F T W T L S D S 260
R256A
247 250
GAAGGTAARGACACACCAGGGGATATTGTCTGACCAGGTGGATGCTAATTGAGGGCTGAA
E G K D T P G G Y C L T R W M L I E A E 280
CTAAAATGCTTCGGGAACACAGCTGTGGCABAAATGTAATGAGAAGCATGATGAGGAATTT
L K C F G N T A V A K C N E K H D E E F 300
TGTGACATGCTGAGGGCTGTTTGACTTCAACAAACAAGCCATTCAAAGGTTGAAAGCTccA
C D M L R L F D F N K Q A I Q R L K A P 320
E329P
GCACAAATGAGCATTCAAGTTGATCAACAAAGCAGTAAATGCTTTGATAAATGACCAACTT
A Q M S I Q L I N K A V N A L I N D Q L 340
ATAATGAAGAACCATCTACGGGACATCATGtGtATTCCATACTGTAATTACAGCAAGTAT
I M K N H L R D I M C I P Y C N Y S K Y 360
G360C
TGGTACTTCAACCACAACTACTGGGAGAACATCACTGCCCAAATGTTGGCTGTATCA
W Y L N H T T T G R T S L P K C W L V S 380
AATGGTTCACTTGAACGAGACCCACTTTTCTGATGATATTGAACAACAAGCTGACAAT
N G S Y L N E T H F S D D I E Q Q A D N 400
ATGATCACTGAGATGTTACAGAAGGAGTATATGGAGAGGCAGGGGAAGACACCAGGGGG
M I T E M L Q K E Y M E R Q G K T P G G 420
418 419
ATGCTCTGGATCACACAGGGACTTCTGGGAGCTCTTCTGTTGTTGGATGGGAATCAATGCC
M S W I T Q G L L G A L L L W M G I N A 440
CGTGACAGGTCAAITGCTAIGACGIIITCTTGGGTTGGAGGAGITTTIGCTCTTCTTTCG
R D R S I A M T F L A V G G V L L F L S 460
GTCAACGTCATGCTGATCAAGGATGGCCATCAACTTTGGCAAGAGAGAGCTCAAGTGC
V N V H A D Q G C A I N F G K R E L K C 480
465 466 474 475
GGAGATGGTATCTTCAATATTAGAGACTCTGATGACTGGCTGAACAAGTACTCATACTAT
G D G I F I F R D S D D W L N K Y S Y Y 500
CCAGAAGATCCTGTGAAGCTTGCATCAATAGTGAAGGCTCTTTTGAAGAAGGGGAAGTGT
P E D P V K L A S I V K A S F E E G K C 520
GGCCTAAATTCAGTTGACTCC
G L N S V D S 527
527

[0158]

- [0159] 构建体#2: PLLAV-YFV17D-LASV-GP CCS
- [0160] 信号肽缺失, 跨膜结构域gp2缺失, 切割位点恢复 (R246R) 以及突变R207C, E329P和 G360C)
- [0161] -末端YFE(氨基酸1-40)
- [0162] -前27个核苷酸ns1(9个氨基酸)(氨基酸41-49)

- [0163] -lasv-gp1[无信号肽以及突变tgt (r207c)] (氨基酸50至250)
- [0164] -切割位点恢复agaagattgcta (r256r) [SEQ ID NO:17]
- [0165] -lasv-gp2[无tm以及突变cca (e329p),tgt (g360c)] (氨基酸251至418)
- [0166] -WNV-TM1 (氨基酸419-442)
- [0167] -WNV-TM2 (氨基酸462-465)
- [0168] -开始yf-ns1 (氨基酸466-527)
- [0169] SEQ ID NO:3
- [0170] SEQ ID NO:4

AAGGTCATCATGGGGCGGTACTTATATGGGTTGGCATCAACACAGA AACATGACAATG
K V I M G A V L I W V G I N T R N M T M 20

TCCATGAGCATGATCTTGGTAGGAGTGCATGATGTTTTTGTCTCTAGGAGTTGGcGCc
S M S M I L V G V I M M F L S L G V G A 40

40

GACCAGGGCTGCGGATAAATTCGGTaccagtcctttataaaaggggtttatgagcttcag
D Q G C A I N F G T S L Y K G V Y E L Q 60

41 49 50

actctgggaactaaacatgggagacactcaatatgaccatgacctctctctctgacacaaagAAC
T L E L N M E T L N M T M P L S C T K N 80

aacagtcacatcattatataatgggtgggcaatgagacaggactagaactgacctgaccAAC
N S H H Y I M V G N E T G L E L T L T N 100

acgagcattatataacacaaaatgggcaatctgctctgagccccacaaaaagAACctctat
T S I I N H K F C N L S D A H K K N L Y 120

gaccacgctctttatgagcataaactctcaactttccactgtccatccccaaacttcaatcag
D H A L M S I I S T F H L S I P N F N Q 140

tatgaggcaatgagctgagatttttaaggggaaagattagtgctgagcagcacaacctgagt
Y E A M S C D F N G G K I S V Q Y N L S 160

cacagctatgctggggatgacagccaaccatgtggtactgttgcaaatgggtggtttacag
H S Y A G D A A N H C G T V A N G V L Q 180

acttttatgaggatggcttgggggtgggagctacattgctcttgactcaggcIgtggcAAC
T F M R M A W G G S Y I A L D S G C G N 200

tgggactgtattatgactagttatcaatatctgataatccaaaaatacaacctgggaaagat
W D C I M T S Y Q Y L I I Q N T T W E D 220

cactgccaattctcgagaccatctcccatcgggttatctcgggctcctctccacaaaggact
H C Q F S R P S P I G Y L G L L S Q R T 240

agagatatttatattagtagAGATTGCTAGGCACATTCACATGGACACTGTCAGATTCT
R D I Y I S R R L L L G T F T W T L S D S 260

[0171] R256
247 250

GAAGGTAAGACACACCAGGGGATATTGTCTGACCAGGTGGATGCTAATTGAGGCTGAA
E G K D T P G G Y C L T R W M L I E A E 280

CTAAAATGCTTCGGGAACACAGCTGTGGCAAAATGTAATGAGAAGCATGATGAGGAATT
L K C F G N T A V A K C N E K H D E E F 300

TGTGACATGCTGAGGCTGTTTGACTTCAACAACAAGCCATTCAAAGGTTGAAAGCTccA
C D M L R L F D F N K Q A I Q R L K A p 320

GCACAAATGAGCATTCAAGTTCATCAACAAGCAGTAAATGCTTTGATAAATGACCACTT
A Q M S I Q L I N K A V N A L I N D Q L 340

ATAATGAAGAACCATCTACGGGACATCATGgGtATTCACATCTGTAATTACAGCAAGTAT
I M K N H L R D I M c I P Y C N Y S K Y 360

TGGTACCTCAACCAACAACACTACTGGGAGAACATCACTGCCCAAAATGTTGGCTTGTATCA
W Y L N H T T T G R T S L P K C W L V S 380

AATGGTTTCACTATGAACGAGACCCACTTTTCTGATGATATTGAACAACAAGCTGACAAT
N G S Y L N E T H F S D D I E Q Q A D N 400

ATGATCACTGAGATGTTACAGAAGGAGTATATGGAGAGGCAGGGGAAGACACCAGGAGGG
M I T E M L Q K E Y M E R Q G K T P G G 420

ATGCTCTGGATCACACAGGGACTTCTGGGAGCTCTTCTGTTTGGATGGGAATCAATGCC
M S W I T Q G L L G A L L L W M G I N A 440

CGTGACAGGTCAATTGCTATGACGTTTTCTGCGGTTGGAGGAGTTTTTGTCTCTTCTTTCG
R D R S I A M T F L A V G G V L L F L S 460

GTCACGTCATGCTGATCAAGGATGCGCCATCAACTTGGCAAGAGAGCTCAAGTGC
V N V H A D Q G C A I N F G K R M L K C 480

GGAGATGGTATCTTCATATTTAGAGACTCTGATGACTGGCTGAACAAGTACTACTACTAT
G D G I F I F R D S D D W L N K Y S Y Y 500

CCAGAAGATCCTGTGAAGCTTGCATCAATAGTGAAAGCCTCTTTTGAAGAAGGGAAGTGT
P E D P V K L A S I V K A S F E E G K C 520

GGCCTAAATTCAGTTGACTCC
G L N S V D S 527

[0172] 缺失的信号肽 (SSP) 的核苷酸序列和氨基酸序列

[0173] SEQ ID NO:5

[0174] SEQ ID NO:6

[0175] atgggacaaatagtgacattcttccaggaagtgcctcatgtaatagaagaggtgatgaac
 M G Q I V T F F Q E V P H V I E E V M N 20
 attgttctcattgcactgtctgtactagcagtggtgaaaggtctgtacaattttgcaacg
 I V L I A L S V L A V L K G L Y N F A T 40
 tgtggccttgttggtttggtcactttcctcctgttgtgtggtaggtcttgcaca
 C G L V G L V T F L L L C G R S C T 58

[0176] 缺失的LASV-GP2跨膜结构域和胞质尾的核苷酸和氨基酸序列:

[0177] SEQ ID NO:7

[0178] SEQ ID NO:8

[0179] TTGGGTCTAGTTGACCTCTTTGTGTTCCAGCACAAAGTTTCTATCTTATTAGCATCTTCTCTT
 L G L V D L F V F S T S F Y L I S I F L 20
 CACCTAGTCAAAATACCAACTCATAGGCATATTGTAGGCAAGTCGTGTCCCAACCTCAC
 H L V K I P T H R H I V G K S C P K P H 40
 AGATTGAATCATATGGGCATTGTTCTCTGTGGACTCTACAAACAGCCTGGTGTGCCTGTG
 R L N H M G I C S C G L Y K Q P G V P V 60
 AAATGGAAGAGA
 K W K R 64

[0180] 拉沙Josiah株G蛋白序列SEQ ID NO:9

[0181] (氨基酸1-58:信号序列)

[0182] (氨基酸59-259:GP1结构域)

[0183] (氨基酸260-437:GP2结构域)

[0184] (氨基酸438-481:跨膜结构域和胞质尾)

[0185] MGQIVTFFQE VPHVIEVMN IVLIALS VLA VLKGLYNFAT CGLVGLVTFL 50
 LLCGRSCTTS LYKGVYELQT LELNMETLNM TMPLSCTRNN SHHYIMVGNE 100
 58
 TGLELTLTNT SIINHRCNL SDAHKNLYD HALMSIISTF HLSIPNFNQY 150
 EAMSCDFNGG KISVQYNLSH SYAGDAANHC GTVANGVLQT FMRMAWGGSY 200
 IALDSGRGNW DCIMTSYQYL IIQNTTWEDH CQFSRPSPIG YLGLLSQRTR 250
 DIYISRRLLG TPTWTLS DSE GKDTPGGYCL TRWMLIEAEL KCFGNTAVAK 300
 259
 CNEKHDEEFC DMLRLDFDNK QAIQRLKAEA QMSIQLINKA VNALINDQLI 350
 MKNHLRDMG IPYCNYSKYW YLNHTTTGRT SLPKRCWLVS N GSYLNETHFS 400
 DDIEQQADNM ITEMLQKEYM ERQKTPPLGL VDLFVVFSTSF YLISIFLHLV 450
 437
 KIPTHRHRIVG KSCP KPHRLN HMGICSCGLY KQPGVPVKWK R 481

[0186] NS1信号序列[SEQ ID NO:10]

[0187] DQGCAINFG

[0188] 连接YFV NS1-Lassa GP1结构域[SEQ ID NO:11]

[0189] AINFG TSLYK

[0190] 连接拉沙GP2结构域-WNV TM1结构域[SEQ ID NO:12]

[0191] QGKTP GGMSW

[0192] 连接WNV TM2结构域-YFV NS1[SEQ ID NO:13]

[0193] VNVHA DQGCA

[0194] WNV TM1序列[SEQ ID NO:14]

[0195] GGMSWITQGLL GALLLW MGINARD

[0196] WNV TM2序列[SEQ ID NO:15]

[0197] RSIAMTFLAVGGVLLF LSVNVHA

序列表

<110> 勒芬天主教大学

<120> 拉沙病毒疫苗

<130> ZL919098

<150> EP19197202.5

<151> 2019-09-13

<160> 17

<170> PatentIn version 3.5

<210> 1

<211> 1581

<212> DNA

<213> 人工序列

<220>

<223> YFV拉沙chimer

<220>

<221> CDS

<222> (1) .. (1581)

<400> 1

aag gtc atc atg ggg gcg gta ctt ata tgg gtt ggc atc aac aca aga 48

Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile Asn Thr Arg

1 5 10 15

aac atg aca atg tcc atg agc atg atc ttg gta gga gtg atc atg atg 96

Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val Ile Met Met

20 25 30

ttt ttg tct cta gga gtt ggc gcc gac cag ggc tgc gcg ata aat ttc 144

Phe Leu Ser Leu Gly Val Gly Ala Asp Gln Gly Cys Ala Ile Asn Phe

35 40 45

ggt acc agt ctt tat aaa ggg gtt tat gag ctt cag act ctg gaa cta 192

Gly Thr Ser Leu Tyr Lys Gly Val Tyr Glu Leu Gln Thr Leu Glu Leu

50 55 60

aac atg gag aca ctc aat atg acc atg cct ctc tcc tgc aca aag aac 240

Asn Met Glu Thr Leu Asn Met Thr Met Pro Leu Ser Cys Thr Lys Asn

65 70 75 80

aac agt cat cat tat ata atg gtg ggc aat gag aca gga cta gaa ctg 288

Asn Ser His His Tyr Ile Met Val Gly Asn Glu Thr Gly Leu Glu Leu

85 90 95

acc ttg acc aac acg agc att att aat cac aaa ttt tgc aat ctg tct 336

Thr Leu Thr Asn Thr Ser Ile Ile Asn His Lys Phe Cys Asn Leu Ser

gat gcc cac aaa aag aac ctc tat gac cac gct ctt atg agc ata atc	100	105	110	384
Asp Ala His Lys Lys Asn Leu Tyr Asp His Ala Leu Met Ser Ile Ile				
tca act ttc cac ttg tcc atc ccc aac ttc aat cag tat gag gca atg	115	120	125	432
Ser Thr Phe His Leu Ser Ile Pro Asn Phe Asn Gln Tyr Glu Ala Met				
agc tgc gat ttt aat ggg gga aag att agt gtg cag tac aac ctg agt	130	135	140	480
Ser Cys Asp Phe Asn Gly Gly Lys Ile Ser Val Gln Tyr Asn Leu Ser				
cac agc tat gct ggg gat gca gcc aac cat tgt ggt act gtt gca aat	145	150	155	528
His Ser Tyr Ala Gly Asp Ala Ala Asn His Cys Gly Thr Val Ala Asn				
ggt gtg tta cag act ttt atg agg atg gct tgg ggt ggg agc tac att	165	170	175	576
Gly Val Leu Gln Thr Phe Met Arg Met Ala Trp Gly Gly Ser Tyr Ile				
gct ctt gac tca ggc tgt ggc aac tgg gac tgt att atg act agt tat	180	185	190	624
Ala Leu Asp Ser Gly Cys Gly Asn Trp Asp Cys Ile Met Thr Ser Tyr				
caa tat ctg ata atc caa aat aca acc tgg gaa gat cac tgc caa ttc	195	200	205	672
Gln Tyr Leu Ile Ile Gln Asn Thr Thr Trp Glu Asp His Cys Gln Phe				
tcg aga cca tct ccc atc ggt tat ctc ggg ctc ctc tca caa agg act	210	215	220	720
Ser Arg Pro Ser Pro Ile Gly Tyr Leu Gly Leu Leu Ser Gln Arg Thr				
aga gat att tat att agt gca aga ttg cta ggc aca ttc aca tgg aca	225	230	235	768
Arg Asp Ile Tyr Ile Ser Ala Arg Leu Leu Gly Thr Phe Thr Trp Thr				
ctg tca gat tct gaa ggt aaa gac aca cca ggg gga tat tgt ctg acc	245	250	255	816
Leu Ser Asp Ser Glu Gly Lys Asp Thr Pro Gly Gly Tyr Cys Leu Thr				
agg tgg atg cta att gag gct gaa cta aaa tgc ttc ggg aac aca gct	260	265	270	864
Arg Trp Met Leu Ile Glu Ala Glu Leu Lys Cys Phe Gly Asn Thr Ala				
gtg gca aaa tgt aat gag aag cat gat gag gaa ttt tgt gac atg ctg	275	280	285	912
Val Ala Lys Cys Asn Glu Lys His Asp Glu Glu Phe Cys Asp Met Leu				
agg ctg ttt gac ttc aac aaa caa gcc att caa agg ttg aaa gct cca	290	295	300	960
Arg Leu Phe Asp Phe Asn Lys Gln Ala Ile Gln Arg Leu Lys Ala Pro				

305	310	315	320	
gca caa atg agc att cag ttg atc aac aaa gca gta aat gct ttg ata				1008
Ala Gln Met Ser Ile Gln Leu Ile Asn Lys Ala Val Asn Ala Leu Ile				
	325	330	335	
aat gac caa ctt ata atg aag aac cat cta cgg gac atc atg tgt att				1056
Asn Asp Gln Leu Ile Met Lys Asn His Leu Arg Asp Ile Met Cys Ile				
	340	345	350	
cca tac tgt aat tac agc aag tat tgg tac ctc aac cac aca act act				1104
Pro Tyr Cys Asn Tyr Ser Lys Tyr Trp Tyr Leu Asn His Thr Thr Thr				
	355	360	365	
ggg aga aca tca ctg ccc aaa tgt tgg ctt gta tca aat ggt tca tac				1152
Gly Arg Thr Ser Leu Pro Lys Cys Trp Leu Val Ser Asn Gly Ser Tyr				
	370	375	380	
ttg aac gag acc cac ttt tct gat gat att gaa caa caa gct gac aat				1200
Leu Asn Glu Thr His Phe Ser Asp Asp Ile Glu Gln Gln Ala Asp Asn				
385	390	395	400	
atg atc act gag atg tta cag aag gag tat atg gag agg cag ggg aag				1248
Met Ile Thr Glu Met Leu Gln Lys Glu Tyr Met Glu Arg Gln Gly Lys				
	405	410	415	
aca cca gga ggg atg tcc tgg atc aca cag gga ctt ctg gga gct ctt				1296
Thr Pro Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu Gly Ala Leu				
	420	425	430	
ctg ttg tgg atg gga atc aat gcc cgt gac agg tca att gct atg acg				1344
Leu Leu Trp Met Gly Ile Asn Ala Arg Asp Arg Ser Ile Ala Met Thr				
	435	440	445	
ttt ctt gcg gtt gga gga gtt ttg ctc ttc ctt tcg gtc aac gtc cat				1392
Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser Val Asn Val His				
	450	455	460	
gct gat caa gga tgc gcc atc aac ttt ggc aag aga gag ctc aag tgc				1440
Ala Asp Gln Gly Cys Ala Ile Asn Phe Gly Lys Arg Glu Leu Lys Cys				
465	470	475	480	
gga gat ggt atc ttc ata ttt aga gac tct gat gac tgg ctg aac aag				1488
Gly Asp Gly Ile Phe Ile Phe Arg Asp Ser Asp Asp Trp Leu Asn Lys				
	485	490	495	
tac tca tac tat cca gaa gat cct gtg aag ctt gca tca ata gtg aaa				1536
Tyr Ser Tyr Tyr Pro Glu Asp Pro Val Lys Leu Ala Ser Ile Val Lys				
	500	505	510	
gcc tct ttt gaa gaa ggg aag tgt ggc cta aat tca gtt gac tcc				1581
Ala Ser Phe Glu Glu Gly Lys Cys Gly Leu Asn Ser Val Asp Ser				

	515	520	525
<210>	2		
<211>	527		
<212>	PRT		
<213>	人工序列		
<220>			
<223>	合成构建体		
<400>	2		
Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile Asn Thr Arg			
1	5	10	15
Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val Ile Met Met			
	20	25	30
Phe Leu Ser Leu Gly Val Gly Ala Asp Gln Gly Cys Ala Ile Asn Phe			
	35	40	45
Gly Thr Ser Leu Tyr Lys Gly Val Tyr Glu Leu Gln Thr Leu Glu Leu			
	50	55	60
Asn Met Glu Thr Leu Asn Met Thr Met Pro Leu Ser Cys Thr Lys Asn			
65	70	75	80
Asn Ser His His Tyr Ile Met Val Gly Asn Glu Thr Gly Leu Glu Leu			
	85	90	95
Thr Leu Thr Asn Thr Ser Ile Ile Asn His Lys Phe Cys Asn Leu Ser			
	100	105	110
Asp Ala His Lys Lys Asn Leu Tyr Asp His Ala Leu Met Ser Ile Ile			
	115	120	125
Ser Thr Phe His Leu Ser Ile Pro Asn Phe Asn Gln Tyr Glu Ala Met			
	130	135	140
Ser Cys Asp Phe Asn Gly Gly Lys Ile Ser Val Gln Tyr Asn Leu Ser			
145	150	155	160
His Ser Tyr Ala Gly Asp Ala Ala Asn His Cys Gly Thr Val Ala Asn			
	165	170	175
Gly Val Leu Gln Thr Phe Met Arg Met Ala Trp Gly Gly Ser Tyr Ile			
	180	185	190
Ala Leu Asp Ser Gly Cys Gly Asn Trp Asp Cys Ile Met Thr Ser Tyr			
	195	200	205
Gln Tyr Leu Ile Ile Gln Asn Thr Thr Trp Glu Asp His Cys Gln Phe			
	210	215	220
Ser Arg Pro Ser Pro Ile Gly Tyr Leu Gly Leu Leu Ser Gln Arg Thr			
225	230	235	240
Arg Asp Ile Tyr Ile Ser Ala Arg Leu Leu Gly Thr Phe Thr Trp Thr			

	245		250		255
Leu Ser Asp Ser	Glu Gly Lys Asp Thr	Pro Gly Gly Tyr Cys Leu Thr			
	260		265		270
Arg Trp Met Leu	Ile Glu Ala Glu Leu Lys Cys Phe Gly Asn Thr Ala				
	275		280		285
Val Ala Lys Cys	Asn Glu Lys His Asp Glu Glu Phe Cys Asp Met Leu				
	290		295		300
Arg Leu Phe Asp	Phe Asn Lys Gln Ala Ile Gln Arg Leu Lys Ala Pro				
305		310		315	320
Ala Gln Met Ser	Ile Gln Leu Ile Asn Lys Ala Val Asn Ala Leu Ile				
	325		330		335
Asn Asp Gln Leu	Ile Met Lys Asn His Leu Arg Asp Ile Met Cys Ile				
	340		345		350
Pro Tyr Cys Asn	Tyr Ser Lys Tyr Trp Tyr Leu Asn His Thr Thr Thr				
	355		360		365
Gly Arg Thr Ser	Leu Pro Lys Cys Trp Leu Val Ser Asn Gly Ser Tyr				
	370		375		380
Leu Asn Glu Thr	His Phe Ser Asp Asp Ile Glu Gln Gln Ala Asp Asn				
385		390		395	400
Met Ile Thr Glu	Met Leu Gln Lys Glu Tyr Met Glu Arg Gln Gly Lys				
	405		410		415
Thr Pro Gly Gly	Met Ser Trp Ile Thr Gln Gly Leu Leu Gly Ala Leu				
	420		425		430
Leu Leu Trp Met	Gly Ile Asn Ala Arg Asp Arg Ser Ile Ala Met Thr				
	435		440		445
Phe Leu Ala Val	Gly Gly Val Leu Leu Phe Leu Ser Val Asn Val His				
	450		455		460
Ala Asp Gln Gly	Cys Ala Ile Asn Phe Gly Lys Arg Glu Leu Lys Cys				
465		470		475	480
Gly Asp Gly Ile	Phe Ile Phe Arg Asp Ser Asp Asp Trp Leu Asn Lys				
	485		490		495
Tyr Ser Tyr Tyr	Pro Glu Asp Pro Val Lys Leu Ala Ser Ile Val Lys				
	500		505		510
Ala Ser Phe Glu	Glu Gly Lys Cys Gly Leu Asn Ser Val Asp Ser				
	515		520		525

<210> 3

<211> 1581

<212> DNA

<213> 人工序列

<220>
 <223> YFV拉沙Chimer
 <220>
 <221> CDS
 <222> (1) .. (1581)
 <400> 3
 aag gtc atc atg ggg gcg gta ctt ata tgg gtt ggc atc aac aca aga 48
 Lys Val Ile Met Gly Ala Val Leu Ile Trp Val Gly Ile Asn Thr Arg
 1 5 10 15
 aac atg aca atg tcc atg agc atg atc ttg gta gga gtg atc atg atg 96
 Asn Met Thr Met Ser Met Ser Met Ile Leu Val Gly Val Ile Met Met
 20 25 30
 ttt ttg tct cta gga gtt ggc gcc gac cag ggc tgc gcg ata aat ttc 144
 Phe Leu Ser Leu Gly Val Gly Ala Asp Gln Gly Cys Ala Ile Asn Phe
 35 40 45
 ggt acc agt ctt tat aaa ggg gtt tat gag ctt cag act ctg gaa cta 192
 Gly Thr Ser Leu Tyr Lys Gly Val Tyr Glu Leu Gln Thr Leu Glu Leu
 50 55 60
 aac atg gag aca ctc aat atg acc atg cct ctc tcc tgc aca aag aac 240
 Asn Met Glu Thr Leu Asn Met Thr Met Pro Leu Ser Cys Thr Lys Asn
 65 70 75 80
 aac agt cat cat tat ata atg gtg ggc aat gag aca gga cta gaa ctg 288
 Asn Ser His His Tyr Ile Met Val Gly Asn Glu Thr Gly Leu Glu Leu
 85 90 95
 acc ttg acc aac acg agc att att aat cac aaa ttt tgc aat ctg tct 336
 Thr Leu Thr Asn Thr Ser Ile Ile Asn His Lys Phe Cys Asn Leu Ser
 100 105 110
 gat gcc cac aaa aag aac ctc tat gac cac gct ctt atg agc ata atc 384
 Asp Ala His Lys Lys Asn Leu Tyr Asp His Ala Leu Met Ser Ile Ile
 115 120 125
 tca act ttc cac ttg tcc atc ccc aac ttc aat cag tat gag gca atg 432
 Ser Thr Phe His Leu Ser Ile Pro Asn Phe Asn Gln Tyr Glu Ala Met
 130 135 140
 agc tgc gat ttt aat ggg gga aag att agt gtg cag tac aac ctg agt 480
 Ser Cys Asp Phe Asn Gly Gly Lys Ile Ser Val Gln Tyr Asn Leu Ser
 145 150 155 160
 cac agc tat gct ggg gat gca gcc aac cat tgt ggt act gtt gca aat 528
 His Ser Tyr Ala Gly Asp Ala Ala Asn His Cys Gly Thr Val Ala Asn
 165 170 175

ggt gtg tta cag act ttt atg agg atg gct tgg ggt ggg agc tac att	576
Gly Val Leu Gln Thr Phe Met Arg Met Ala Trp Gly Gly Ser Tyr Ile	
180 185 190	
gct ctt gac tca ggc tgt ggc aac tgg gac tgt att atg act agt tat	624
Ala Leu Asp Ser Gly Cys Gly Asn Trp Asp Cys Ile Met Thr Ser Tyr	
195 200 205	
caa tat ctg ata atc caa aat aca acc tgg gaa gat cac tgc caa ttc	672
Gln Tyr Leu Ile Ile Gln Asn Thr Thr Trp Glu Asp His Cys Gln Phe	
210 215 220	
tcg aga cca tct ccc atc ggt tat ctc ggg ctc ctc tca caa agg act	720
Ser Arg Pro Ser Pro Ile Gly Tyr Leu Gly Leu Leu Ser Gln Arg Thr	
225 230 235 240	
aga gat att tat att agt aga aga ttg cta ggc aca ttc aca tgg aca	768
Arg Asp Ile Tyr Ile Ser Arg Arg Leu Leu Gly Thr Phe Thr Trp Thr	
245 250 255	
ctg tca gat tct gaa ggt aaa gac aca cca ggg gga tat tgt ctg acc	816
Leu Ser Asp Ser Glu Gly Lys Asp Thr Pro Gly Gly Tyr Cys Leu Thr	
260 265 270	
agg tgg atg cta att gag gct gaa cta aaa tgc ttc ggg aac aca gct	864
Arg Trp Met Leu Ile Glu Ala Glu Leu Lys Cys Phe Gly Asn Thr Ala	
275 280 285	
gtg gca aaa tgt aat gag aag cat gat gag gaa ttt tgt gac atg ctg	912
Val Ala Lys Cys Asn Glu Lys His Asp Glu Glu Phe Cys Asp Met Leu	
290 295 300	
agg ctg ttt gac ttc aac aaa caa gcc att caa agg ttg aaa gct cca	960
Arg Leu Phe Asp Phe Asn Lys Gln Ala Ile Gln Arg Leu Lys Ala Pro	
305 310 315 320	
gca caa atg agc att cag ttg atc aac aaa gca gta aat gct ttg ata	1008
Ala Gln Met Ser Ile Gln Leu Ile Asn Lys Ala Val Asn Ala Leu Ile	
325 330 335	
aat gac caa ctt ata atg aag aac cat cta cgg gac atc atg tgt att	1056
Asn Asp Gln Leu Ile Met Lys Asn His Leu Arg Asp Ile Met Cys Ile	
340 345 350	
cca tac tgt aat tac agc aag tat tgg tac ctc aac cac aca act act	1104
Pro Tyr Cys Asn Tyr Ser Lys Tyr Trp Tyr Leu Asn His Thr Thr Thr	
355 360 365	
ggg aga aca tca ctg ccc aaa tgt tgg ctt gta tca aat ggt tca tac	1152
Gly Arg Thr Ser Leu Pro Lys Cys Trp Leu Val Ser Asn Gly Ser Tyr	
370 375 380	

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

Pro Tyr Cys Asn Tyr Ser Lys Tyr Trp Tyr Leu Asn His Thr Thr Thr
 355 360 365
 Gly Arg Thr Ser Leu Pro Lys Cys Trp Leu Val Ser Asn Gly Ser Tyr
 370 375 380
 Leu Asn Glu Thr His Phe Ser Asp Asp Ile Glu Gln Gln Ala Asp Asn
 385 390 395 400
 Met Ile Thr Glu Met Leu Gln Lys Glu Tyr Met Glu Arg Gln Gly Lys
 405 410 415
 Thr Pro Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu Gly Ala Leu
 420 425 430
 Leu Leu Trp Met Gly Ile Asn Ala Arg Asp Arg Ser Ile Ala Met Thr
 435 440 445
 Phe Leu Ala Val Gly Gly Val Leu Leu Phe Leu Ser Val Asn Val His
 450 455 460
 Ala Asp Gln Gly Cys Ala Ile Asn Phe Gly Lys Arg Glu Leu Lys Cys
 465 470 475 480
 Gly Asp Gly Ile Phe Ile Phe Arg Asp Ser Asp Asp Trp Leu Asn Lys
 485 490 495
 Tyr Ser Tyr Tyr Pro Glu Asp Pro Val Lys Leu Ala Ser Ile Val Lys
 500 505 510
 Ala Ser Phe Glu Glu Gly Lys Cys Gly Leu Asn Ser Val Asp Ser
 515 520 525

<210> 5

<211> 174

<212> DNA

<213> 拉沙病毒

<220>

<221> CDS

<222> (1) .. (174)

<400> 5

atg gga caa ata gtg aca ttc ttc cag gaa gtg cct cat gta ata gaa 48
 Met Gly Gln Ile Val Thr Phe Phe Gln Glu Val Pro His Val Ile Glu
 1 5 10 15
 gag gtg atg aac att gtt ctc att gca ctg tct gta cta gca gtg ctg 96
 Glu Val Met Asn Ile Val Leu Ile Ala Leu Ser Val Leu Ala Val Leu
 20 25 30
 aaa ggt ctg tac aat ttt gca acg tgt ggc ctt gtt ggt ttg gtc act 144
 Lys Gly Leu Tyr Asn Phe Ala Thr Cys Gly Leu Val Gly Leu Val Thr
 35 40 45

ttc ctc ctg ttg tgt ggt agg tct tgc aca	174
Phe Leu Leu Leu Cys Gly Arg Ser Cys Thr	
50 55	
<210> 6	
<211> 58	
<212> PRT	
<213> 拉沙病毒	
<400> 6	
Met Gly Gln Ile Val Thr Phe Phe Gln Glu Val Pro His Val Ile Glu	
1 5 10 15	
Glu Val Met Asn Ile Val Leu Ile Ala Leu Ser Val Leu Ala Val Leu	
20 25 30	
Lys Gly Leu Tyr Asn Phe Ala Thr Cys Gly Leu Val Gly Leu Val Thr	
35 40 45	
Phe Leu Leu Leu Cys Gly Arg Ser Cys Thr	
50 55	
<210> 7	
<211> 192	
<212> DNA	
<213> 拉沙病毒	
<220>	
<221> CDS	
<222> (1) .. (192)	
<400> 7	
ttg ggt cta gtt gac ctc ttt gtg ttc agc aca agt ttc tat ctt att	48
Leu Gly Leu Val Asp Leu Phe Val Phe Ser Thr Ser Phe Tyr Leu Ile	
1 5 10 15	
agc atc ttc ctt cac cta gtc aaa ata cca act cat agg cat att gta	96
Ser Ile Phe Leu His Leu Val Lys Ile Pro Thr His Arg His Ile Val	
20 25 30	
ggc aag tcg tgt ccc aaa cct cac aga ttg aat cat atg ggc att tgt	144
Gly Lys Ser Cys Pro Lys Pro His Arg Leu Asn His Met Gly Ile Cys	
35 40 45	
tcc tgt gga ctc tac aaa cag cct ggt gtg cct gtg aaa tgg aag aga	192
Ser Cys Gly Leu Tyr Lys Gln Pro Gly Val Pro Val Lys Trp Lys Arg	
50 55 60	
<210> 8	
<211> 64	
<212> PRT	

<213> 拉沙病毒

<400> 8

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

<210> 9

<211> 491

<212> PRT

<213> 拉沙病毒

<400> 9

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

Arg Met Ala Trp Gly Gly Ser Tyr Ile Ala Leu Asp Ser Gly Arg Gly
 195 200 205
 Asn Trp Asp Cys Ile Met Thr Ser Tyr Gln Tyr Leu Ile Ile Gln Asn
 210 215 220
 Thr Thr Trp Glu Asp His Cys Gln Phe Ser Arg Pro Ser Pro Ile Gly
 225 230 235 240
 Tyr Leu Gly Leu Leu Ser Gln Arg Thr Arg Asp Ile Tyr Ile Ser Arg
 245 250 255
 Arg Leu Leu Gly Thr Phe Thr Trp Thr Leu Ser Asp Ser Glu Gly Lys
 260 265 270
 Asp Thr Pro Gly Gly Tyr Cys Leu Thr Arg Trp Met Leu Ile Glu Ala
 275 280 285
 Glu Leu Lys Cys Phe Gly Asn Thr Ala Val Ala Lys Cys Asn Glu Lys
 290 295 300
 His Asp Glu Glu Phe Cys Asp Met Leu Arg Leu Phe Asp Phe Asn Lys
 305 310 315 320
 Gln Ala Ile Gln Arg Leu Lys Ala Glu Ala Gln Met Ser Ile Gln Leu
 325 330 335
 Ile Asn Lys Ala Val Asn Ala Leu Ile Asn Asp Gln Leu Ile Met Lys
 340 345 350
 Asn His Leu Arg Asp Ile Met Gly Ile Pro Tyr Cys Asn Tyr Ser Lys
 355 360 365
 Tyr Trp Tyr Leu Asn His Thr Thr Thr Gly Arg Thr Ser Leu Pro Lys
 370 375 380
 Cys Trp Leu Val Ser Asn Gly Ser Tyr Leu Asn Glu Thr His Phe Ser
 385 390 395 400
 Asp Asp Ile Glu Gln Gln Ala Asp Asn Met Ile Thr Glu Met Leu Gln
 405 410 415
 Lys Glu Tyr Met Glu Arg Gln Gly Lys Thr Pro Leu Gly Leu Val Asp
 420 425 430
 Leu Phe Val Phe Ser Thr Ser Phe Tyr Leu Ile Ser Ile Phe Leu His
 435 440 445
 Leu Val Lys Ile Pro Thr His Arg His Ile Val Gly Lys Ser Cys Pro
 450 455 460
 Lys Pro His Arg Leu Asn His Met Gly Ile Cys Ser Cys Gly Leu Tyr
 465 470 475 480
 Lys Gln Pro Gly Val Pro Val Lys Trp Lys Arg
 485 490

<210> 10

<211> 9
 <212> PRT
 <213> 黄热病病毒
 <400> 10
 Asp Gln Gly Cys Ala Ile Asn Phe Gly
 1 5
 <210> 11
 <211> 10
 <212> PRT
 <213> 人工序列
 <220>
 <223> 连接YFV NS1- 拉沙GP1结构域
 <400> 11
 Ala Ile Asn Phe Gly Thr Ser Leu Tyr Lys
 1 5 10
 <210> 12
 <211> 10
 <212> PRT
 <213> 人工序列
 <220>
 <223> 连接拉沙GP2结构域- WNV TM1结构域
 <400> 12
 Gln Gly Lys Thr Pro Gly Gly Met Ser Trp
 1 5 10
 <210> 13
 <211> 10
 <212> PRT
 <213> 人工序列
 <220>
 <223> 连接WNV TM2结构域- YFV NS1
 <400> 13
 Val Asn Val His Ala Asp Gln Gly Cys Ala
 1 5 10
 <210> 14
 <211> 24
 <212> PRT
 <213> 西尼罗河病毒
 <400> 14
 Gly Gly Met Ser Trp Ile Thr Gln Gly Leu Leu Gly Ala Leu Leu Leu

1	5	10	15	
Trp Met Gly Ile Asn Ala Arg Asp				
	20			
<210>	15			
<211>	23			
<212>	PRT			
<213>	西尼罗河病毒			
<400>	15			
Arg Ser Ile Ala Met Thr Phe Leu Ala Val Gly Gly Val Leu Leu Phe				
1	5	10	15	
Leu Ser Val Asn Val His Ala				
	20			
<210>	16			
<211>	12			
<212>	DNA			
<213>	人工序列			
<220>				
<223>	突变的切割位点			
<400>	16			
gcaagattgc ta				12
<210>	17			
<211>	12			
<212>	DNA			
<213>	人工序列			
<220>				
<223>	恢复的切割位点			
<400>	17			
agaagattgc ta				12

1) PLLAV-YFV17D-LASV-GPC (突变的切割位点)

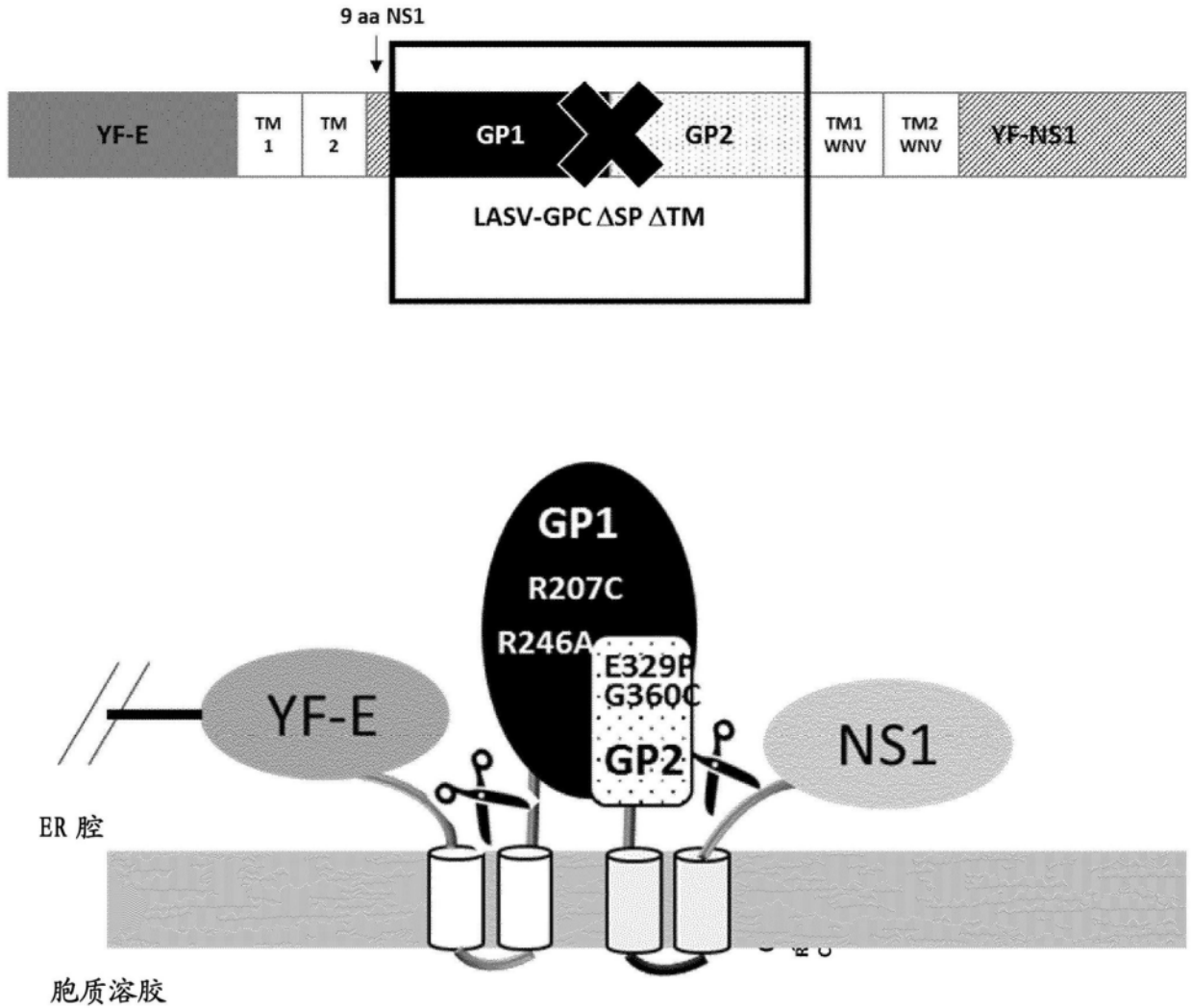


图1

2) PLLAV-YFV17D-LASV-GPCcs (切割位点)

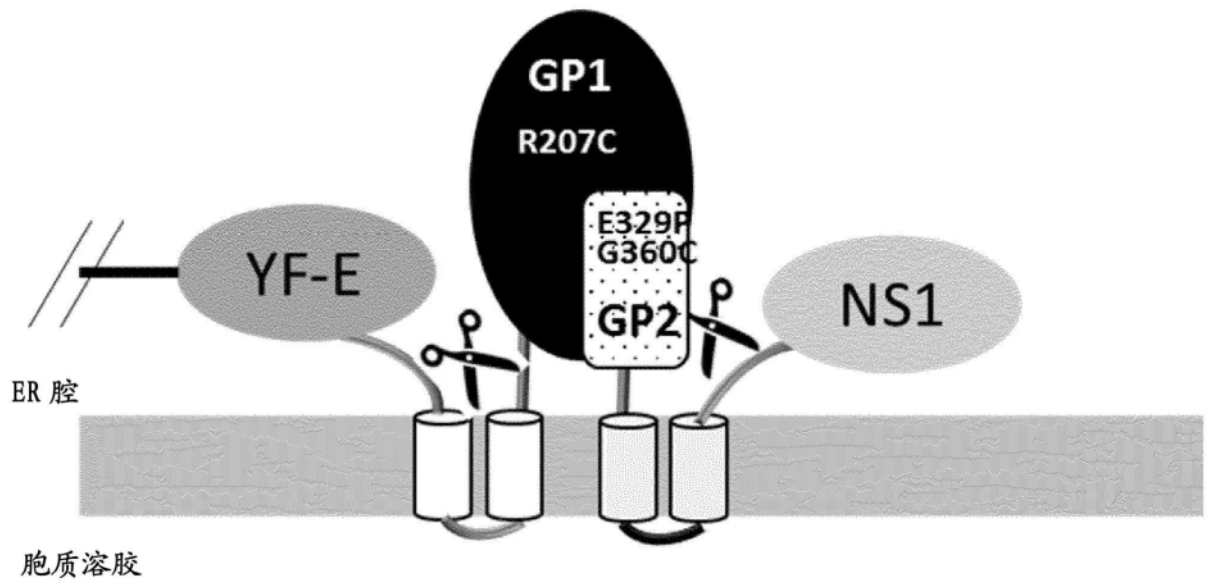
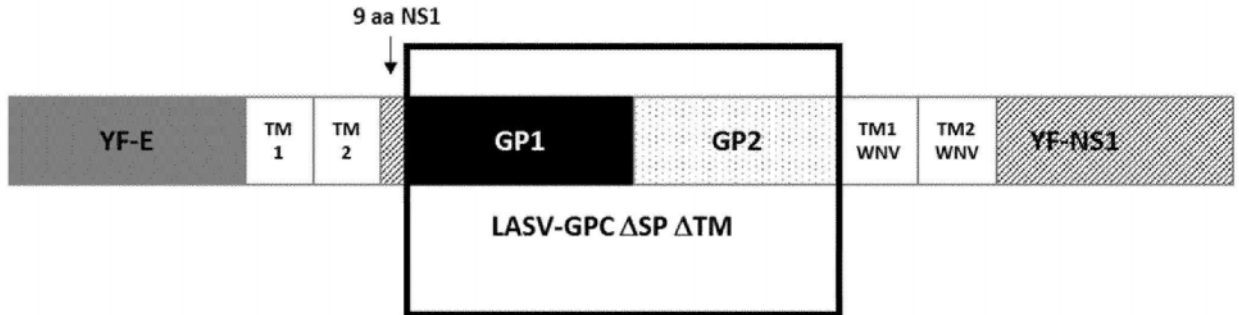
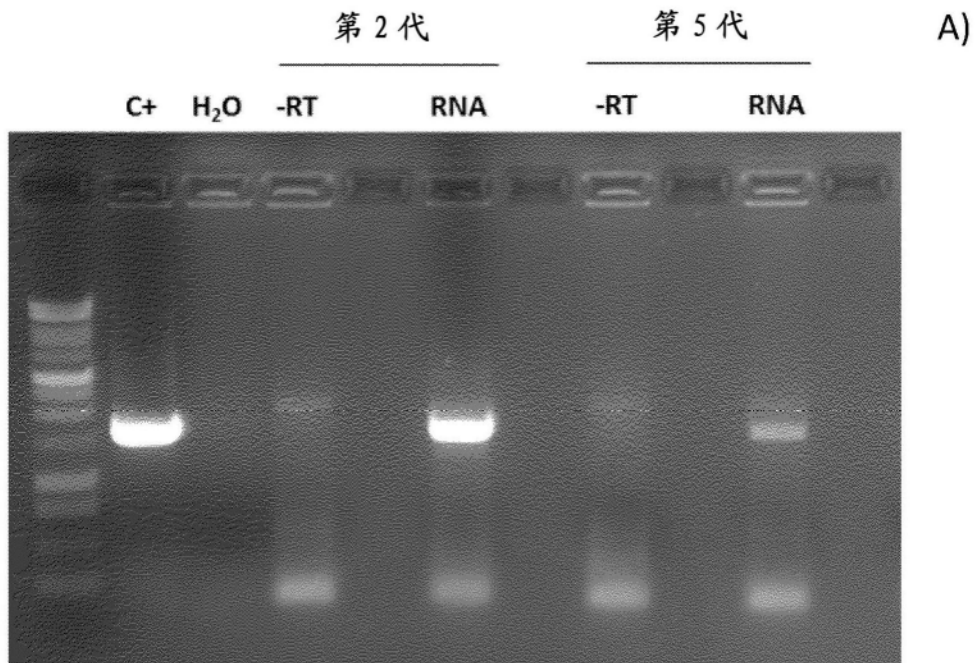


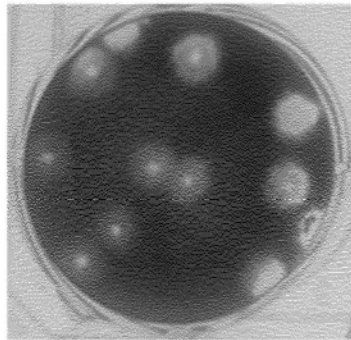
图1 (续)

BHK-21J



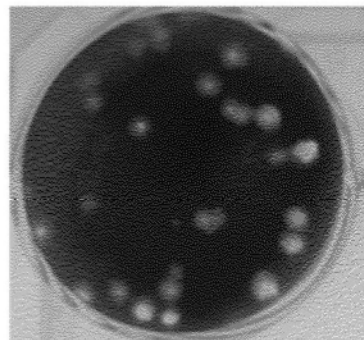
B)

YF17D



8×10^4 pfu/ml

YF17D-LASV-GPC



1.5×10^3 pfu/ml

TCID50/ml: $1,46 \times 10^6$

$7,90 \times 10^4$

图2

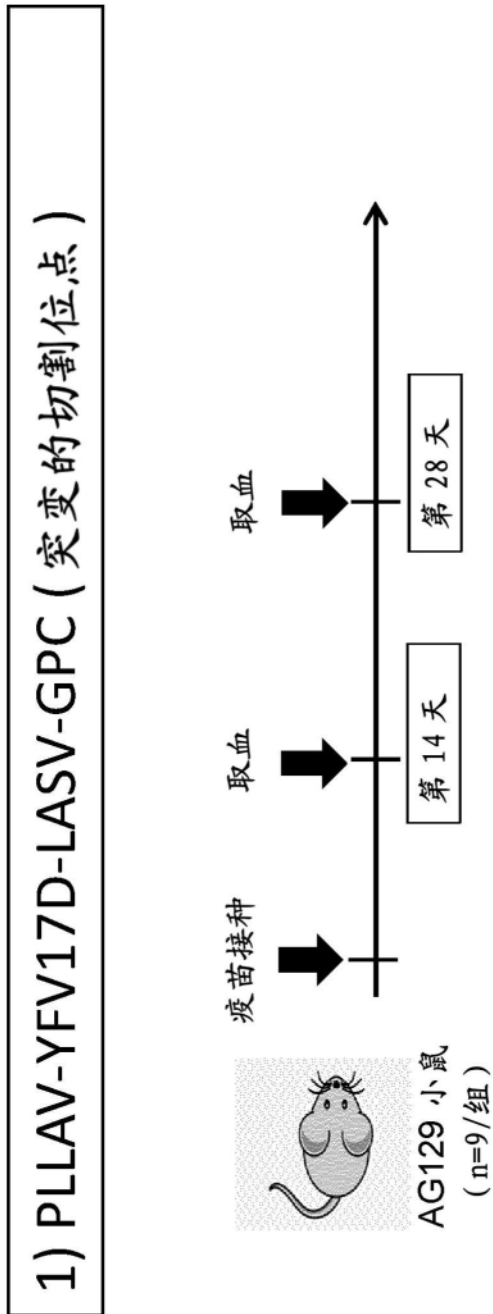


图3

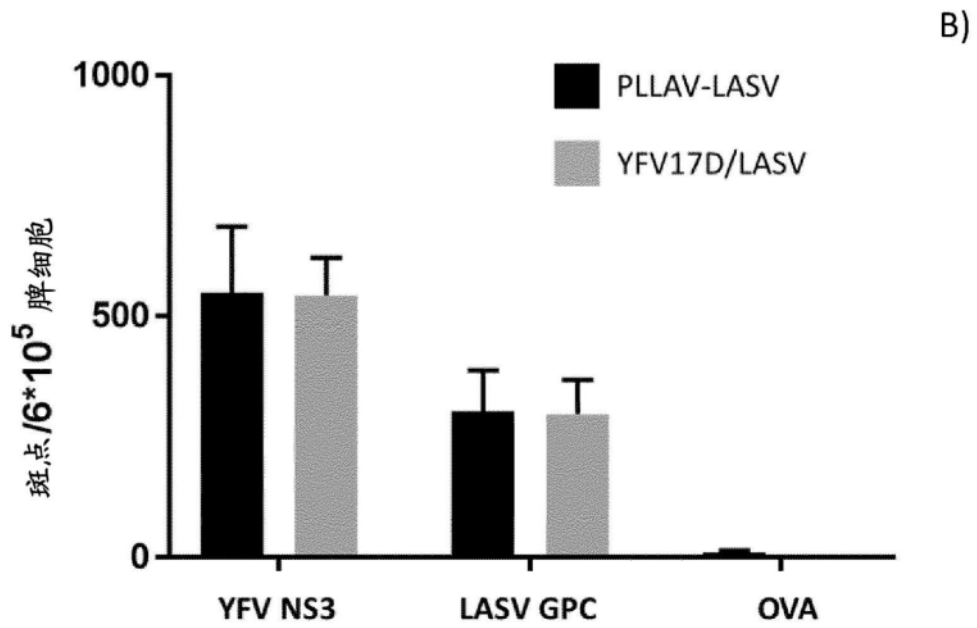
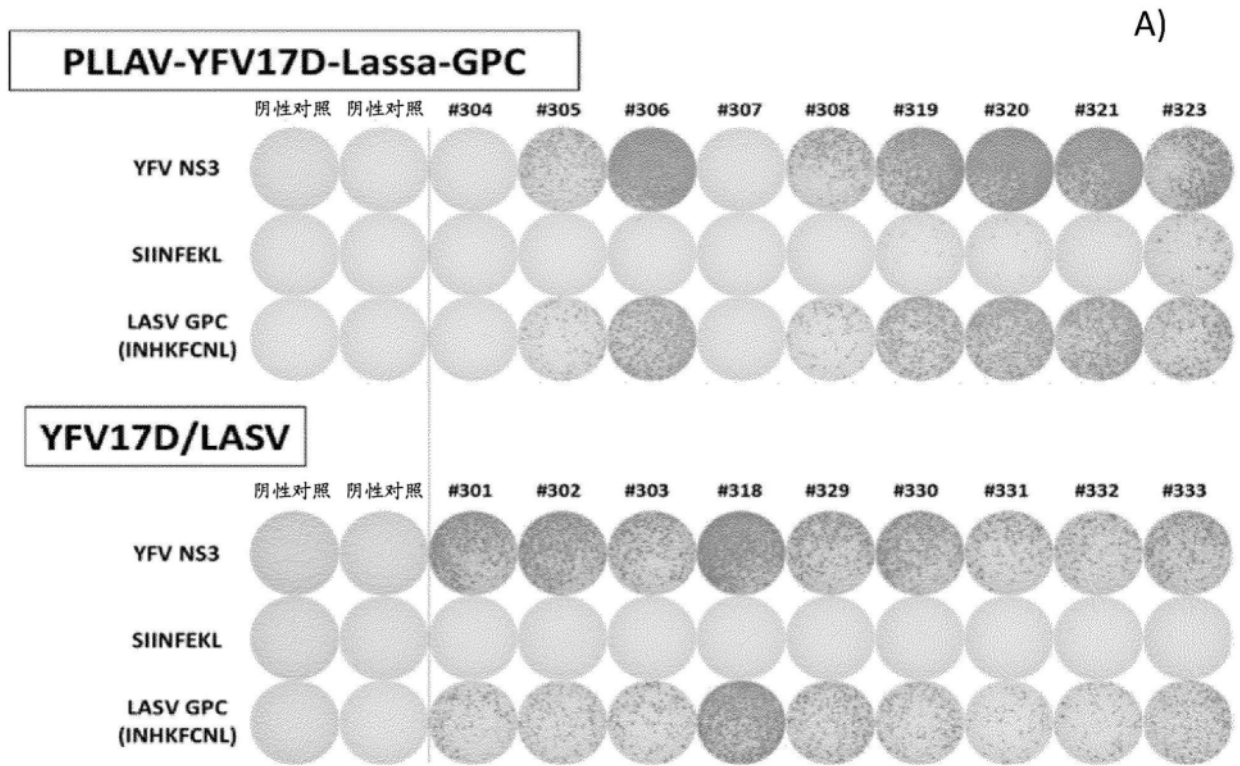


图4

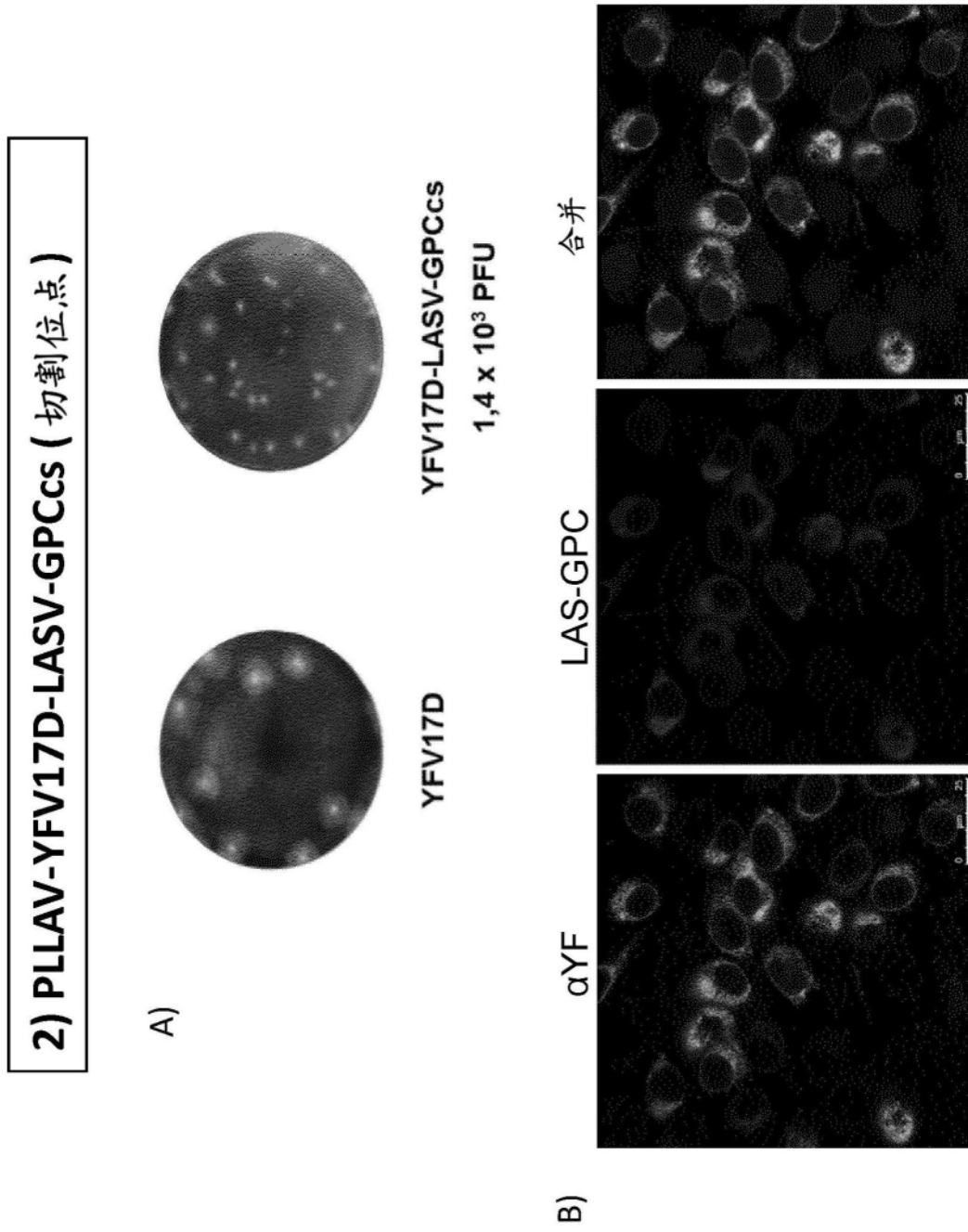


图5

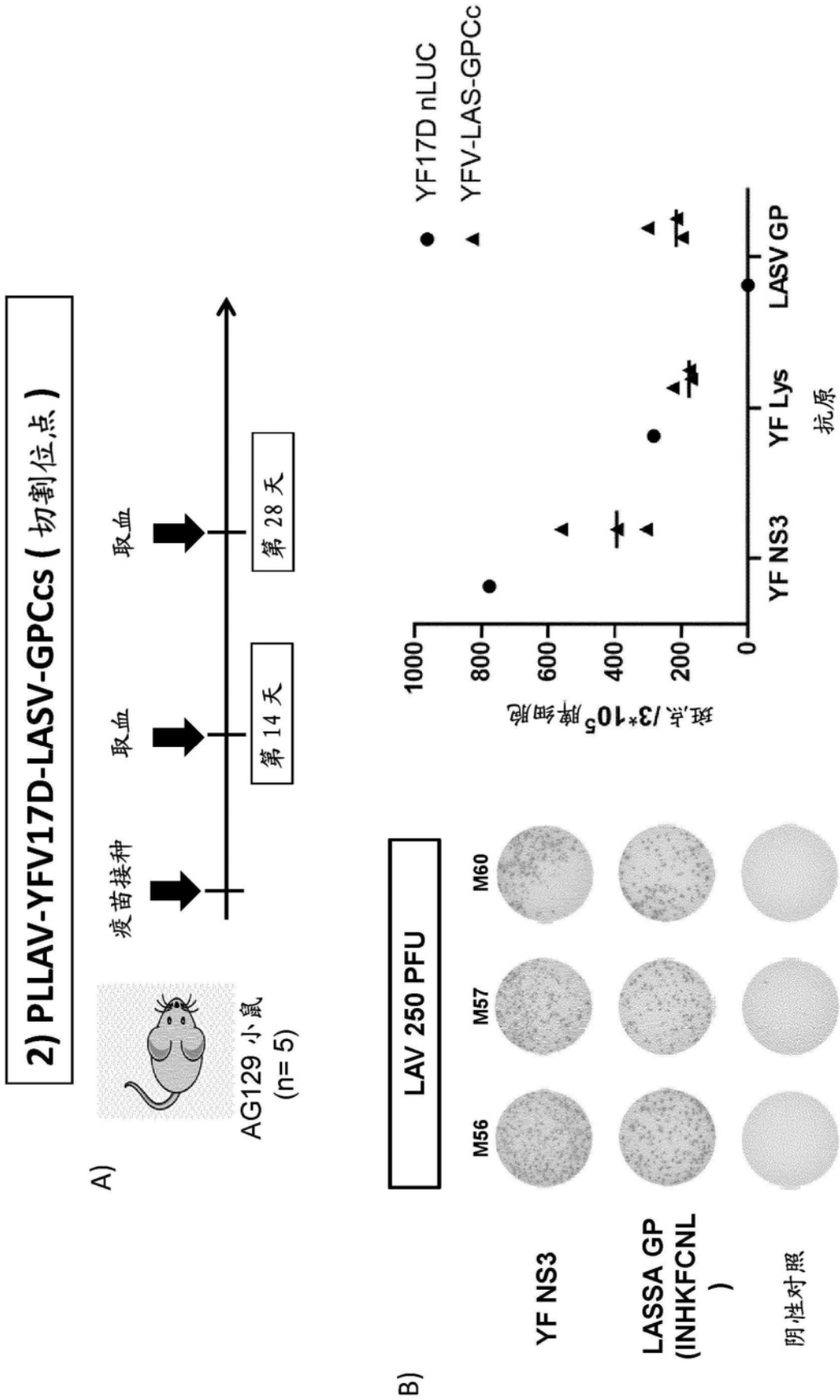


图6