



(12) 发明专利申请

(10) 申请公布号 CN 117083071 A

(43) 申请公布日 2023. 11. 17

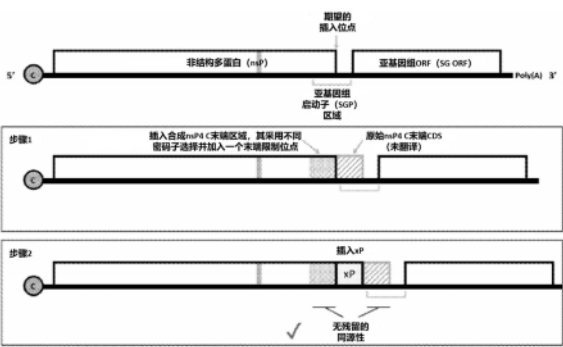
(21) 申请号 202280022669.3
(22) 申请日 2022.03.18
(30) 优先权数据
63/163,475 2021.03.19 US
(85) PCT国际申请进入国家阶段日
2023.09.19
(86) PCT国际申请的申请数据
PCT/US2022/020881 2022.03.18
(87) PCT国际申请的公布数据
W02022/198002 EN 2022.09.22
(71) 申请人 蒂巴生物技术有限公司
地址 美国马萨诸塞州
(72) 发明人 J·S·查哈尔 J·S·马克帕特兰

(74) 专利代理机构 北京润平知识产权代理有限公司 11283
专利代理师 刘依云
(51) Int.Cl.
A61K 35/76 (2015.01)

权利要求书4页 说明书49页
序列表147页 附图30页

(54) 发明名称
人工甲病毒衍生的RNA复制子表达系统

(57) 摘要
描述了合成的甲病毒衍生复制子表达系统，其包含编码至少一种经过修饰的非结构蛋白的核酸序列和编码至少一种异源蛋白的合成核酸序列。提供了在细胞中产生至少一种异源蛋白的方法，或通过施用和/或表达合成的甲病毒衍生复制子表达系统在受试者中诱导免疫应答的方法。



1. 一种合成的甲病毒衍生复制子核酸分子,其包含:(i) 第一核酸,其编码甲病毒非结构蛋白nsP1、nsP2、nsP3和nsP4,并且包含在如SEQ ID NO:17所示的甲病毒基因组序列中从nt 503至nt 658、nt 658至nt 1620、nt 1620至nt 2560、nt 2560至nt 3954、nt 3954至nt 4120、nt 6381至nt 7083和nt 6966至nt 7526的区域内的任意位置处引入的至少一个沉默突变;以及(ii) 第二核酸,其包含经过修饰的亚基因组开放阅读框(ORF)。

2. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,经过修饰的亚基因组ORF包含编码第一异源蛋白的序列。

3. 根据权利要求2所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码经过修改的nsP4的多核苷酸,所述经过修改的nsP4包含第二异源蛋白。

4. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 503至nt 658的区域内的任意位置处引入的至少一个沉默突变。

5. 根据权利要求4所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:21所示的参考序列具有至少90%同一性的序列的多核苷酸。

6. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 658至nt 1620的区域内的任意位置处引入的至少一个沉默突变。

7. 根据权利要求6所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:25所示的参考序列具有至少90%同一性的序列的多核苷酸。

8. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 1620至nt 2560的区域内的任意位置处引入的至少一个沉默突变。

9. 根据权利要求8所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:24所示的参考序列具有至少90%同一性的序列的多核苷酸。

10. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 3694至nt 3954的区域内的任意位置处引入的至少一个沉默突变。

11. 根据权利要求10所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:22所示的参考序列具有至少90%同一性的序列的多核苷酸。

12. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 2560至nt 3954的区域内的任意位置处引入的至少一个沉默突变。

13. 根据权利要求12所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:26所示的参考序列具有至少90%同一性的序列的多核苷酸。

14. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 3954至nt 4120的区域内的任意位置处引入的至少一个沉默突变。

15. 根据权利要求14所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP2和nsP3的接合物并且含有与如SEQ ID NO:23所示的参考序列具有至少90%同一性的序列的多核苷酸。

16. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在

甲病毒基因组中从nt 6381至nt 7083的区域内的任意位置处引入的至少一个沉默突变。

17. 根据权利要求16所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP4并且含有与如SEQ ID NO:20所示的参考序列具有至少90%同一性的序列的多核苷酸。

18. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 658至nt 1620和nt 2560至nt 3954的区域内的任意位置处引入的至少一个沉默突变。

19. 根据权利要求18所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:25所示的参考序列具有至少90%同一性的序列的多核苷酸,以及编码nsP2并且含有与如SEQ ID NO:26所示的参考序列具有至少90%同一性的序列的多核苷酸。

20. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 3694至nt 3954和nt 6381至nt 7083的区域内的任意位置处引入的至少一个沉默突变。

21. 根据权利要求20所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP2并且含有与如SEQ ID NO:22所示的参考序列具有至少90%同一性的序列的多核苷酸,以及编码nsP4并且含有与如SEQ ID NO:20所示的参考序列具有至少90%同一性的序列的多核苷酸。

22. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:21所示的参考序列具有至少90%同一性的序列的多核苷酸;编码nsP1并且含有与如SEQ ID NO:25所示的参考序列具有至少90%同一性的序列的多核苷酸,编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:24所示的参考序列具有至少90%同一性的序列的多核苷酸;编码nsP2并且含有与如SEQ ID NO:26所示的参考序列具有至少90%同一性的序列的多核苷酸;编码nsP2和nsP3的接合物并且含有与如SEQ ID NO:23所示的参考序列具有至少90%同一性的序列的多核苷酸;以及编码nsP4并且含有与如SEQ ID NO:20所示的参考序列具有至少90%同一性的序列的多核苷酸。

23. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码所述非结构蛋白并且含有与如SEQ ID NO:30所示的参考序列具有至少90%同一性的多核苷酸序列的多核苷酸。

24. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,第一核酸分子包含在甲病毒基因组中从nt 6966到nt 7523的区域内的任意位置处引入的至少一个沉默突变。

25. 根据权利要求24所述的合成的甲病毒衍生复制子核酸分子,第一核酸分子包含与如SEQ ID NO:18所示的参考序列具有至少90%同一性的多核苷酸序列。

26. 根据权利要求1所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含引入编码nsP4蛋白的多核苷酸的至少一个沉默突变。

27. 根据权利要求26所述的合成的甲病毒衍生复制子核酸分子,第一核酸分子包含与如SEQ ID NO:19所示的参考序列具有至少90%同一性的多核苷酸序列。

28. 根据权利要求3所述的合成的甲病毒衍生复制子核酸分子,其中,第一异源蛋白或第二异源蛋白包含细胞代谢、免疫或信号传导功能的病毒衍生的或哺乳动物细胞衍生的调

节剂。

29. 根据权利要求3所述的合成的甲病毒衍生复制子核酸分子, 其中, 第一异源蛋白或第二异源蛋白影响细胞抗病毒先天免疫应答并增强甲病毒复制子RNA分子的复制。

30. 根据权利要求3所述的合成的甲病毒衍生复制子核酸分子, 其中, 第一异源蛋白或第二异源蛋白包含自病毒病原体分离或衍生的抗原蛋白。

31. 根据权利要求30所述的合成的甲病毒衍生复制子核酸分子, 其中, 病毒病原体选自: 小核糖核酸病毒、腺病毒、鼻病毒、轮状病毒、西尼罗病毒、寨卡病毒、疱疹病毒或冠状病毒 (CoV) 组成的组中。

32. 根据权利要求31所述的合成的甲病毒衍生复制子核酸分子, 其中, 病毒病原体为小核糖核酸病毒。

33. 根据权利要求3所述的合成的甲病毒衍生复制子核酸分子, 其中, 第二异源蛋白包含3C蛋白酶蛋白, 并且经过修改的nsP4包含与如SEQ ID NO:40所示的参考序列具有至少90%同一性的氨基酸序列。

34. 根据权利要求33所述的合成的甲病毒衍生复制子核酸分子, 其中, 经过修改的nsP4由与如SEQ ID NO:32所示的参考序列具有至少90%同一性的多核苷酸序列编码。

35. 根据权利要求3所述的合成的甲病毒衍生复制子核酸分子, 其中, 第二核酸包含编码01 ManisaP1蛋白的多核苷酸。

36. 根据权利要求35所述的合成的甲病毒衍生复制子核酸分子, 其中, 所述多核苷酸包含与如SEQ nt NO:33所示的参考序列具有至少90%同一性的序列。

37. 根据权利要求3所述的合成的甲病毒衍生复制子核酸分子, 其中, 第一异源蛋白包含01 ManisaP1蛋白并且第二异源蛋白包含3C蛋白酶。

38. 根据权利要求3所述的合成的甲病毒衍生复制子核酸分子, 其中, 第二核酸包含编码抗原蛋白的多核苷酸。

39. 根据权利要求3所述的合成的甲病毒衍生复制子核酸分子, 其中, 第二异源蛋白包含STING蛋白, 所述经过修改的nsP4包含与如SEQ ID NO:42所示的参考序列具有至少90%同一性的氨基酸序列。

40. 根据权利要求3所述的合成的甲病毒衍生复制子核酸分子, 其中, 第一异源蛋白包含抗原蛋白并且第二异源蛋白包含STING蛋白。

41. 一种甲病毒衍生RNA复制子表达系统, 其包含权利要求1-40中任一项所述的合成的甲病毒衍生复制子核酸分子和宿主细胞。

42. 一种药学上可接受的组合物, 其包含权利要求1-40中任一项所述的甲病毒衍生复制子和药学上可接受的载体或赋形剂。

43. 一种在细胞中产生至少一种异源蛋白的方法, 其包括在宿主细胞中表达权利要求1-40中任一项所述的合成的甲病毒衍生复制子核酸分子。

44. 一种预防、抑制或治疗受试者的疾病或病症的症状的方法, 其包括: 提供合成的甲病毒衍生复制子核酸分子, 其包含:

(i) 第一核酸, 其编码甲病毒非结构蛋白nsP1、nsP2、nsP3和nsP4, 并且包含在如SEQ ID NO:17所示的甲病毒基因组序列中从nt 503至nt 658、nt 658至nt 1620、nt 1620至nt 2560、nt 2560至nt 3954、nt 3954至nt 4120、nt 6381至nt 7083和nt 6966至nt 7526的区

域内的任意位置处引入的至少一个沉默突变;以及

(ii) 第二核酸,其包含经过修饰的亚基因组开放阅读框(ORF);以及
将治疗有效量的合成的甲病毒衍生复制子核酸分子施用给受试者。

45. 根据权利要求44所述的方法,其中,第一核酸包含与选自由SEQ ID NOS:20-30组成的组中的参考序列具有至少90%同一性的至少一个序列。

46. 根据权利要求44所述的方法,其中,经过修饰的亚基因组ORF包含编码第一异源蛋白的序列。

47. 根据权利要求44所述的方法,其中,第一核酸包含编码经过修改的nsP4的多核苷酸,所述经过修改的nsP4包含第二异源蛋白。

48. 根据权利要求47所述的方法,其中,第一异源蛋白或第二异源蛋白包含细胞代谢、免疫或信号传导功能的病毒衍生的或哺乳动物细胞衍生的调节剂。

49. 根据权利要求47所述的方法,其中,第一异源蛋白或第二异源蛋白影响细胞抗病毒先天免疫应答并增强甲病毒复制子RNA分子的复制。

50. 根据权利要求47所述的方法,其中,第一异源蛋白或第二异源蛋白包含自病毒病原体分离或衍生的抗原蛋白。

51. 根据权利要求50所述的方法,其中,病毒病原体选自由:小核糖核酸病毒、腺病毒、鼻病毒、轮状病毒、西尼罗病毒、寨卡病毒、疱疹病毒或冠状病毒(CoV)组成的组中。

52. 根据权利要求51所述的方法,其中,病毒病原体为小核糖核酸病毒。

53. 根据权利要求47所述的方法,其中,第一异源蛋白包含O1Manisa P1蛋白并且第二异源蛋白包含3C蛋白酶。

54. 根据权利要求47所述的方法,其中,第一异源蛋白或第二异源蛋白包含STING蛋白。

55. 根据权利要求47所述的方法,其中,第二异源蛋白包含STING蛋白并且第一异源蛋白包含抗原蛋白。

56. 根据权利要求44所述的方法,其中,施用步骤使得受试者的疾病得到预防、治疗、严重性减轻或病程发展减缓。

57. 根据权利要求44所述的方法,其中,疾病为癌症。

58. 根据权利要求57所述的方法,其中,癌症为黑素瘤或上皮性肿瘤。

59. 根据权利要求58所述的方法,其中,该方法还包括向受试者共同施用一种或多种另外的抗癌疗法。

60. 根据权利要求59所述的方法,其中,另外的抗癌疗法选自由外科手术、化学疗法、放射疗法、温热疗法、免疫疗法、激素疗法、激光疗法、抗血管生成疗法及其任何组合组成的组中。

61. 根据权利要求44所述的方法,其中,疾病为FMD。

62. 根据权利要求44所述的方法,其中,受试者为哺乳动物。

63. 根据权利要求62所述的方法,其中,哺乳动物为偶蹄动物或人。

64. 根据权利要求63所述的方法,其中,偶蹄动物选自:牛、水牛、绵羊、山羊、猪、羚羊、鹿、野牛、刺猬、大象、美洲驼和羊驼组成的组中。

人工甲病毒衍生的RNA复制子表达系统

[0001] 相关申请的交叉引用

[0002] 本申请要求2021年3月19日提交的美国临时申请No.63/163,475的权益,其通过引用并入本文,如同已完整阐述。

[0003] 与本申请一起电子提交的名为“SEQUENCE LISTING”的序列表创建于2022年3月18日,大小为413,759字节,其通过引用并入本文,如同已完整阐述。

发明领域

[0004] 本发明涉及源自甲病毒基因组的自我复制RNA分子,以及通过突变和以其他方式编辑初级核苷酸序列来提高其安全性和基因表达能力的方法。本发明还涉及产生和使用此类工程化RNA分子来治疗和预防人类或动物疾病的方法。

背景技术

[0005] 自我复制RNA,或“复制子”制子”,已被用作体外和体内转基因表达的载体。甲病毒属的病毒被用作复制子RNA设计的模板,因为它们的基因组布局在不同物种之间是一致的,并且适合分子生物学操作。源自许多不同种类的甲病毒(例如委内瑞拉马脑炎病毒、辛德比斯病毒、塞姆利基森林病毒)的基于甲病毒的复制子已被用作体外和体内转基因表达的载体。这些载体依赖于其中编码的非结构病毒蛋白的活性来介导在结构基因开放阅读框(ORF)位置编码的外源蛋白的表达,所述结构基因开放阅读框(ORF)是从复制生命周期后期合成的亚基因组(SG)mRNA翻译而来的。

[0006] 复制子是先天免疫反应的强效诱导剂,使其可用于疫苗和其他免疫调节用途。甲病毒复制子RNA通常基于常见病毒种类,例如委内瑞拉马脑炎病毒(VEEV),该病毒在世界大部分地区自然存在并流行。当复制子载体的序列与野生型基因组相同时,与野生型病毒重组的风险更大。RNA病毒的一级序列对基因组复制、翻译和包装成病毒体所需的二级结构和非蛋白质编码功能的保护很敏感。

[0007] 对复制子RNA的免疫反应很难调整,因为最初的干扰素反应是在病毒生命周期的早期阶段启动的,此时只有非结构多蛋白(nsP)以有效的细胞内浓度表达。目前,将额外的调控基因产物纳入复制子中的修饰必须通过在SG ORF中插入来进行,这导致其仅在病毒复制周期的后期阶段表达,还会破坏该ORF中编码的任何其他转基因序列。如果转基因可以从nsP ORF中顺式表达,或在nsP ORF之外顺式表达,确保它们在复制生命周期的早期表达,那么RNA复制子的作用就会大得多,因为此时关键的生物过程,例如模式识别受体(PRR)的触发或蛋白质的翻译后加工,都可以通过传递其中编码的外源转基因来有意调节。此外,如果nsP区域的编码序列与天然存在的病毒基因组的序列同一性较低,那么RNA复制子在现场使用中将更安全。

发明内容

[0008] 在一个方面,本发明涉及合成的甲病毒衍生的自我复制或复制子核酸(RNA)分子。

合成的甲病毒衍生复制子核酸分子(synthetic alphavirus-derived replicon nucleic acid molecule)包含第一核酸,其编码甲病毒非结构蛋白nsP1、nsP2、nsP3和nsP4,并且包含在如SEQ ID NO:17所示的甲病毒基因组序列中从nt 503至nt 658、nt 658至nt 1620、nt 1620至nt 2560、nt 2560至nt 3954、nt 3954至nt 4120、nt 6381至nt 7083和nt 6966至nt 7526的区域内的任意位点引入的至少一个沉默突变。合成的甲病毒衍生复制子核酸分子还包含第二核酸,其包含亚基因组启动子(SGP)和经过修饰的亚基因组开放阅读框(SG ORF)。

[0009] 在一个方面,本发明涉及甲病毒衍生RNA复制子表达系统,其包含任何一种合成的甲病毒衍生复制子核酸分子。在分子中,非结构蛋白4(nsP4)的C末端通过不可裂解、可白裂解或可蛋白水解裂解的肽连接序列与异源蛋白融合。合成的甲病毒衍生复制子核酸分子还编码nsP4融合蛋白下游(即3')的功能性SGP序列,以驱动经过修饰的SG ORF的表达。在一个方面,nsP4编码序列中存在至少一个沉默突变,使其中包含的SGP同源序列丧失功能并且与nsP4融合蛋白ORF下游编码的功能性SGP不同。

[0010] 在一个方面,本发明涉及甲病毒衍生RNA复制子表达系统。该甲病毒衍生RNA复制子表达系统包含本发明所述的任何一种合成的甲病毒衍生复制子核酸分子,其被封装或配制用于以病毒、病毒样颗粒、脂质体或脂质、聚合物或树枝状聚合物纳米颗粒的形式递送至生物体细胞中。

[0011] 在一个方面,本发明涉及疫苗。该疫苗包含本发明所述的任意一种合成的甲病毒衍生复制子核酸分子。

[0012] 在一个方面,本发明涉及治疗药物(drug)或药品(medicine)。该治疗药物或药品包含本发明所述的任意一种合成的甲病毒衍生复制子核酸分子。

[0013] 在一个方面,本发明涉及在细胞中产生至少一种异源蛋白的方法。该方法包括在细胞中表达本发明所述的任意一种合成的甲病毒衍生复制子核酸分子。

[0014] 在一个方面,本发明涉及预防、抑制或治疗受试者的疾病(disease)或病症(condition)的症状的方法,其包括向受试者施用治疗有效量的本发明所述的任意一种合成的甲病毒衍生复制子核酸分子。

附图说明

[0015] 当结合附图阅读时,将更好地理解以下详细描述的本发明的优选实施方式。为了说明本发明,在附图中示出了特定的实施方式。然而,应当理解的是,本发明不限于附图中所示的精确布置和手段。在附图中:

[0016] 图1是说明野生型甲病毒基因组非结构多蛋白(nsP1-nsP2-nsP3-nsP4)的基因表达和翻译后加工步骤的示意图。

[0017] 图2是说明甲病毒基因组非结构多蛋白的基因表达步骤的示意图,该多蛋白通过与nsP4(nsP1-nsP2-nsP3-nsP4-xP)框内的附属物人工修饰以表达另外的外源蛋白("xP")。

[0018] 图3是说明克隆可修饰的nsP4开放阅读框(ORF)同时保持完整的经过延伸的亚基因组启动子(SGP、复制域(duplicated region),以条纹框表示)的步骤的示意图。

[0019] 图4是说明克隆可修饰的nsP4开放阅读框(ORF)同时保持完整的经过延伸的亚基因组启动子(SGP)的最佳步骤的示意图。在此图中,灰色框表示合成的nsP4 C末端区域和xP区域,条纹框表示原始nsP4 C末端CDS(非翻译)。

[0020] 图5是编码构建体8 (C08) 的复制子序列模板的DNA质粒的示意图。质粒从T7启动子转录产生C08复制子RNA。C08的5'UTR、nsP1-4 ORF和SGP如SEQ ID NO:27所示。

[0021] 图6是描绘了用于选择甲病毒基因组区域的示例性过程示意图,所述区域可通过应用沉默突变进行修改(alter),从而降低与野生型病毒同源性,同时不降低复制子的复制功能。在此图中,圈出的是可以在不扰乱二级结构氨基酸编码序列的情况下修改的核酸碱基。修改此类核酸碱基以产生沉默突变被称为密码子适应过程。

[0022] 图7是编码构建体2 (C02) 的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸503-658位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C02复制子RNA。C02的5'UTR、nsP1-4 ORF和SGP如SEQ ID NO:21所示。

[0023] 图8是编码构建体6 (C06) 的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸658-1620位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C06复制子RNA。C06的5'UTR、nsP1-4 ORF和SGP如SEQ ID NO:25所示。

[0024] 图9是编码构建体5 (C05) 的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸1620-2560位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C05复制子RNA。C05的5'UTR、nsP1-4 ORF和SGP如SEQ ID NO:24所示。

[0025] 图10是编码构建体3 (C03) 的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸3694-3954位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C03复制子RNA。C03的5'UTR、nsP1-4 ORF和SGP如SEQ ID NO:22所示。

[0026] 图11是编码构建体7 (C07) 的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸2560-3954位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C07复制子RNA。C07的5'UTR、nsP1-4 ORF和SGP如SEQ ID NO:26所示。

[0027] 图12是编码构建体4 (C04) 的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸3954-4120位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C04复制子RNA。C04的5'UTR、nsP1-4 ORF和SGP如SEQ ID NO:23所示。

[0028] 图13是编码构建体1 (C01) 的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸6381-7083位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C01复制子RNA。C01的5'UTR、nsP1-4 ORF和SGP如SEQ ID NO:20所示。

[0029] 图14是编码构建体9 (C09) 的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸658-1620位和2560-3954位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C09复制子RNA。C09的5'UTR、nsP1-4 ORF和SGP如SEQ ID NO:28所示。

[0030] 图15是编码构建体10 (C10) 的复制子序列模板的DNA质粒的示意图。在此图中,DNA

质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸3694-3954位和6381-7083位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C10复制子RNA。C10的5' UTR、nsP1-4 ORF和SGP如SEQ ID NO 29所示。

[0031] 图16是编码构建体11(C11)的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5所示类似,不同之处仅在于修改了基因组核苷酸503-658位、658-1620位、1620-2560位、2560-3954位、3954-4120位和6381-7083位(灰色框)的密码子以降低与野生型病毒的同源性。质粒从T7启动子转录产生C11复制子RNA。C11的5' UTR、nsP1-4 ORF和SGP如SEQ ID NO:30所示。

[0032] 图17A-17D示出了SEAP表达测试的结果,以证明存在于密码子适应复制子RNA中的编码SEAP的SG ORF的功能,并且通过比色试验定量SEAP浓度。

[0033] 图17A是示出了SEAP在C01、C02、C03、C04、C05、C06、C07和C08复制子RNA上的表达测试结果与野生型(WT,类似编码SEAP的非密码子适应的复制子RNA)对比的柱状图,该测试是通过对转染所示时间的BHK细胞的条件上清液取样来进行的,转染时间为4小时(左侧柱)、16小时(中间柱)和24小时(右侧柱)。

[0034] 图17B和17C是示出了SEAP表达测试结果的柱状图,该测试通过对用指定复制子RNA(图17B-C09,图17C-C10)转染16小时的BHK细胞的条件上清液取样并通过比色法在650nm波长处测量吸光度来定量SEAP表达。据观察,密码子适应构建体的表达比野生型(WT)对照更强。

[0035] 图17D是示出了构建体C03、C05和C06与对照野生型RNA复制子(VEEVrep-SEAP;WT)和未处理(Untransf.)小鼠相比的通过发光试验(RLU)定量的SEAP表达结果的柱状散点图。在治疗中,每组四只小鼠注射1 μ g所示RNA,并于第二天收集血清。通过发光试验以相对发光单位(RLU)测量血清SEAP浓度。柱状图:带有误差条的组的平均RLU=S.D。

[0036] 图18是VEEV基因组的比例图,其中通过本发明所述的RNA折叠分析和实验确定的非结构多蛋白区域被标记为灰色框,其富集于易于沉默突变的非结构区域。

[0037] 图19是示出图4所示的为实验而修改的nsP4复制子序列的一个示例的示意图,其中附加到nsP4蛋白的外源蛋白(xP)是绿色荧光蛋白(GFP;灰色框)。在此图中,SG ORF中的结构多蛋白被报告基因SEAP取代。此复制子RNA被称为VEEVrep-nsP4[GFP]-SEAP,RNA序列如SEQ ID NO:31所示。

[0038] 图20A-20C示出了体外细胞中由工程化的经nsP4修饰的VEEV复制子介导的基因表达。BHK是用在非结构蛋白ORF中编码GFP和在SG ORF的SEAP编码SEAP的实验性经nsP4修饰的复制子(VEEVrep-nsP4[GFP]-SEAP)转染的细胞。

[0039] 图20A示出了荧光显微镜检查的结果,证明当用nsP4(GFP)修饰的VEEV复制子RNA(VEEVrep-nsP4[GFP]-SEAP)进行转染时存在多个GFP阳性细胞,并且在未转染的对照细胞中不存在GFP表达。

[0040] 图20B是用GFP特异性单克隆抗体转染的BHK细胞裂解物的免疫印迹照片,其中GAPDFI检测物用作上样对照,其确认了VEEVrep-nsP4[GFP]-SEAP的两个独立分离的克隆对GFP的表达。

[0041] 图20C是示出通过对转染的BHK细胞的条件上清液取样并通过比色试验定量SEAP表达来测试编码SEAP的SG ORF的功能的柱状图。

[0042] 图21A-21B示出了由工程化的经nsP4修饰的VEEV复制子介导的基因表达。此复制子RNA结合了如本发明的实施方式中所描述的密码子适应和nsP4 C末端修饰。

[0043] 图21A是显示荧光显微镜检查结果的照片,示出了与明场(BF)显微镜(左图)相比,当用密码子适应的、经nsP4(GFP)-修饰的VEEV-SEAP复制子RNA(C13)进行转染时存在多个GFP阳性细胞(上图,右),且在未转染的对照细胞(下图,右)中不存在GFP表达。

[0044] 图21B是示出了发光试验(RLU)结果的图表,该发光试验与从未注射的对照小鼠收集的血清进行比较,对从注射了野生型VEEV-SEAP复制子RNA的小鼠(WT)或注射了在非结构化多蛋白基因的非结构区中携带优化密码子的经nsP4(GFP)-修饰的VEEV复制子RNA的小鼠(经修饰的)收集的血清中的SEAP浓度进行定量。血清于注射后第1天、第3天和第5天收集。

[0045] 图22A-22B示出了经修饰以编码小核糖核酸病毒P1结构多蛋白(其在SG ORF中编码)所需的3Cpro蛋白酶的甲病毒复制子RNA的基因表达。

[0046] 图22A是示出了对甲病毒复制子的nsP4编码序列进行修饰以编码小核糖核酸病毒P1结构多蛋白(01 Manisa P1多蛋白)所需的3Cpro蛋白酶的构建体的示意图,称为VEEVrep-nsP4[3C]-P1。此RNA的5' UTR、nsP1-4、3Cpro和SGP区域的序列如SEQ ID NO:32所示。在SG ORF中编码的P1多蛋白的氨基酸序列如SEQ ID NO:40所示。

[0047] 图22B是对用以下RNA转染的BHK细胞裂解物进行的免疫印迹的照片:泳道1,VEEVrep-nsP4[3C]-P1;泳道2,复制子RNA与泳道1相似,但携带插入突变,引入提前终止密码子并切除3Cpro多肽的表达,作为阴性对照(VEEVrep-nsP4[3Cmut]-P1);泳道3,VEEVrep-P1(类似于VEEVrep-nsP4[3C]-P1,但缺少nsP4融合的3Cpro);泳道4,将VEEVrep-P1与1/40质量的编码3Cpro的mRNA组合转染,作为阳性对照。参考该图,只有在VEEVrep-nsP4[3C]-P1的介导下,才能将~100kDa的P1多蛋白成功蛋白水解加工成免疫原性~35kDa的VP0衣壳片段。可以看到代表完整VP4+VP2+VP3的不完全裂解中间产物,其表观分子量为~55kDa。使用VP2特异性抗体来探测印迹。M,分子量参照标记;C,未经处理的阴性对照细胞。

[0048] 图23A-23D示出了经修饰以编码人STING蛋白的甲病毒复制子RNA的基因表达数据,该甲病毒复制子RNA被突变为对IFN刺激活性具有组成型活性,并在SG ORF中编码报告基因SEAP。

[0049] 图23A是构建体VEEVrep-nsP4[STING]-SEAP的示意图,该构建体是包含编码人STING蛋白的编码序列修饰的复制子RNA,其被突变为对IFN刺激活性具有组成型活性,并且在SG ORF中编码报告基因SEAP。此RNA的5' UTR、nsP1-4、STING和SGP区域的序列如SEQ ID NO:34所示。

[0050] 图23B是在用以下RNA转染的HEK-Lucia™空白细胞上进行的免疫印迹的照片:泳道1,VEEVrep-SEAP;泳道2,VEEVrep-nsP4[STING]-SEAP;泳道3,无转染对照。

[0051] 图23C是示出与对照(未转染细胞)相比,在VEEVrep-SEAP和VEEVrep-nsP4[STING]-SEAP复制子RNA中的SG ORF中编码的SEAP报告基因的表达结果(通过比色试验中650nm处的吸光度测量)的柱状图。

[0052] 图23D是示出与对照(未转细胞)相比由VEEVrep-SEAP SEAP和VEEVrep-nsP4[STING]-SEAP复制子RNA诱导的IFN信号传导活性的替代测量结果的柱状图。

[0053] 图24A-24D示出了在B16黑色素瘤中来自经修饰以编码人STING蛋白的甲病毒复制子RNA的基因表达结果,该甲病毒复制子RNA被突变为对IFN刺激活性具有组成型活性,并且

在SG ORF中编码报告基因SEAP。

[0054] 图24A是示出与未转染的对照细胞相比,用1 μ g VEEVrep-nsP4[GFP]-SEAP复制子RNA转染后约1天,在12孔培养皿培养的B16黑色素瘤细胞中的GFP表达结果的照片。

[0055] 图24B是示出与未转染的对照细胞相比,在用1 μ g指定复制子RNA:VEEVrep-nsP4[STING]-SEAP、VEEVrep-nsP4[GFP]-SEAP和VEEVrep-SEAP转染后,在12孔培养皿中培养的B16黑色素瘤细胞中SEAP表达结果的柱状图。

[0056] 图24C是示出与未转染的对照细胞相比,在用1 μ g指定复制子RNA:VEEVrep-nsP4[STING]-SEAP和VEEVrep-SEAP转染~1天后培养的B16黑色素瘤细胞(通过ELISA夹心法测定培养上清)的IFN- β 分泌测量结果的柱状图。

[0057] 图24D是与健康未转染的对照细胞(最左图)相比,用2 μ g VEEVrep-nsP4[GFP]-SEAP或VEEVrep-nsP4[STING]-SEAP转染1天后B16细胞的光学显微镜检查结果的照片。

[0058] 图25A-25D示出了在TC-1癌细胞中来自经修饰以编码人STING蛋白的甲病毒复制子RNA的基因表达,该甲病毒复制子RNA被突变为对IFN刺激活性具有组成型活性,并在SG ORF中编码报告基因SEAP。

[0059] 图25A是示出与未转染的对照细胞相比,用2 μ g指定复制子RNA:VEEVrep-nsP4[GFP]-SEAP、VEEVrep-nsP4[STING]-SEAP转染2天后,在12孔培养皿中培养的TC-1癌细胞中SEAP表达结果的柱状图。在本实验中,通过培养基的比色试验对SEAP表达进行定量,显示在SG ORF中编码的SEAP基因的成功表达。

[0060] 图25B是示出与未转染的对照细胞相比,用2 μ g指定复制子RNA:VEEVrep-nsP4[GFP]-SEAP或VEEVrep-nsP4[STING]-SEAP转染1天后,培养的TC-1癌细胞(通过ELISA夹心法测定培养上清)的IFN- β 分泌测量结果的柱状图。

[0061] 图25C是示出用1 μ g VEEVrep-nsP4[GFP]-SEAP转染后约3天,转染后的TC-1癌细胞团聚物中GFP表达结果的照片。参考此图,与未处理的对照细胞中没有GFP表达(下图,右)相比,通过在构建体nsP4 C末端编码荧光蛋白,在整个多细胞结构中观察到大量GFP表达。

[0062] 图25D是示出与未处理的对照细胞相比,用4 μ g指定复制子RNA VEEVrep-nsP4[STING]-SEAP转染1天后,TC-1癌细胞的光学显微镜检查结果的照片。参考此图,观察到未处理的细胞(对照)已增殖并填满培养皿;转染VEEVrep-nsP4[STING]-SEAP后观察到生长抑制和细胞死亡。

[0063] 图26A-26G示出了经修饰与野生型同源性降低的VEEV复制子RNA以及来自这些构建体的表达数据。

[0064] 图26A是编码构建体12(VEEVrepHK-SEAP)的复制子序列模板的DNA质粒的示意图。参考此图,DNA质粒与图5所示的类似,不同之处仅在于修改了基因组核苷酸705-3406位(灰色框)的密码子以降低与野生型病毒的同源性,而不考虑可能存在的任何二级结构元件。

[0065] 图26B是编码构建体13(VEEVrepHK(Δ U)-SEAP)的复制子序列模板的DNA质粒的示意图。参考此图,DNA质粒与图58中所示的类似,不同之处仅在于修改了基因组核苷酸658-3359位(灰色框)的密码子以降低与野生型病毒的同源性,而不考虑可能存在的任何二级结构元件,并且优先选择规避尿嘧啶碱基的密码子。

[0066] 图26C是编码构建体14(VEEVrepES-SEAP)的复制子序列模板的DNA质粒的示意图。参考此图,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸4120-6965位

(灰色框)的密码子以降低与野生型病毒的同源性,而不考虑可能存在的任何二级结构元件。

[0067] 图26D是编码构建体15 (VEEVrepES (Δ U) -SEAP) 的复制子序列模板的DNA质粒的示意图。参考此图,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸4120-6965位(灰色框)的密码子以降低与野生型病毒的同源性,而不考虑可能存在的任何二级结构元件,并且优先选择规避尿嘧啶核苷碱基的密码子。

[0068] 图26E是示出与未转染的对照细胞相比,在用1 μ g VEEVrep SEAP、VEEVrepHK SEAP或VEEVrepHK (Δ U) -SEAP RNA转染16小时后,在12孔培养皿中培养的BHK细胞中SEAP表达结果的柱状图。在本实验中,通过对培养基的比色试验来定量SEAP的表达,其显示在SG ORF中编码的SEAP基因的成功表达。

[0069] 图26F是示出与未转染的对照相比,在用1 μ g VEEVrep-SEAP、VEEVrepES-SEAP或VEEVrepES (Δ U) -SEAP RNA转染16小时后,在12孔培养皿中培养的BHK细胞中SEAP表达结果的柱状图。

[0070] 图26G是示出与对照(来自未处理的小鼠的血清)相比,注射2 μ g LNP配制的VEEVrep-SEAP、VEEVrepHK-SEAP或VEEVrepHK (Δ U) -SEAP后(通过化学发光试验)测量的血清中SEAP体内表达结果的柱状图。

[0071] 图27A-27D示出了正常HFF细胞中的VEEV复制子RNA表达结果。

[0072] 图27A是示出与未转染的对照细胞相比,在用2或4 μ g指定复制子RNA (VEEVrep-nsP4[STING]-SEAP、VEEVrep-SEAP) 转染后1天,在12孔培养皿中培养的正常HFF细胞中SEAP表达结果的柱状图。在此实验中,通过对培养基的比色试验来定量SEAP的表达,其显示在SG ORF中编码的SEAP基因的成功表达。

[0073] 图27B是用4 μ g VEEVrep-nsP4[GFP]-SEAP转染后约4天,胰蛋白酶解离的正常人HFF细胞中GFP表达结果的照片,其显示来自nsP4 C末端编码的GFP蛋白在这些健康的非癌细胞中没有有效的基因表达。

[0074] 图27C是示出来自经转染的HFF的IFN- β 活性的测量结果的柱状图。在转染后1天收集来自12孔培养皿中的HFF的条件培养基,与未转染的HFF细胞对照,以及无培养基对照进行比较,转染指定RNA (VEEVrep-nsP4[STING]-SEAP、VEEVrep-nsP4[GFP]-SEAP、VEEVrep-SEAP) 的条件培养基应用于HEK-Lucia空白细胞以测量其对培养基中I型IFN的应答。在本实验中,通过化学发光试验在第二天检测上清液中的IFN应答报告基因Lucia。

[0075] 图27D是与未转染的对照相比,用4 μ g指定复制子RNA (VEEVrep-nsP4[GFP]-SEAP和VEEVrep-nsP4[STING]-SEAP) 转染后4天HFF细胞的光学显微镜检查结果的照片。

具体实施方式

[0076] 除非另有明确说明,否则权利要求书和说明书的相应部分中使用的词语“一种(a)”和“一个(one)”被定义为包括一个或多个所引用的项目。此术语包括上述具体提及的词语、其派生词以及类似含义的词语。短语“至少一个”后跟两个或更多个项目的列表,例如“A、B或C”,是指A、B或C中的任何单独一个及其任何组合。

[0077] 甲病毒

[0078] 如本发明所用,“甲病毒”是指披膜病毒科(Togaviridae)家族的包膜单链正义RNA

病毒。甲病毒属包含大约30个成员,其可以感染人类和其他动物。甲病毒颗粒的直径通常为70nm,往往呈球形或略带多形性,并具有40nm等距核衣壳(isometric nucleocapsid)。

[0079] 甲病毒基因组是大约12000nt长的ssRNA分子,其编码两个开放阅读框(ORF),即非结构性(ns)和结构性(本发明也称为亚基因组或SG,ORF)。ns ORF从基因组5'端附近开始,占基因组总长度的前三分之二。它编码非结构蛋白(nsP),即nsP1、nsP2、nsP3和nsP4,这些蛋白作为单一多蛋白前体产生,通过蛋白水解加工裂解成成熟蛋白。nsP通过细胞质中典型的帽结构依赖性翻译(cap-dependent translation)来表达。感染后,甲病毒基因组RNA被翻译成nsP多蛋白,该多蛋白被裂解为四种蛋白:nsP1、nsP2、nsP3和nsP4(图1),它们共同作用以催化基因组复制和第二个甲病毒ORF的转录/翻译。nsP1的大小约为60kDa且可能具有甲基转移酶活性,并参与病毒加帽反应(viral capping reaction)。nsP2的大小约为90kDa且可能具有解旋酶和蛋白酶活性,而nsP3约为60kDa并包含三个结构域:宏结构域、中央(或甲病毒属独具的)结构域和高变结构域(HVD)。nsP4大小约为70kDa并包含核心RNA依赖性RNA聚合酶(RdRp)催化结构域。

[0080] 结构性的SG ORF大约占基因组长度的三分之一,本质上,此ORF编码病毒体结构蛋白:核心核衣壳蛋白C,以及作为异二聚体结合的包膜蛋白P62和E1。病毒包膜锚着的表面糖蛋白负责受体识别并通过膜融合进入靶细胞。

[0081] 直到nsP产物介导基因组的RNA依赖性RNA转录以产生互补的负义链(-RNA)模板后,才会发生此SG ORF的翻译。nsP产物从-RNA模板的3'端开始完成全长转录以重建原始基因组序列,从而完成基因组复制的循环。基因组复制的开始代表了从病毒复制生命周期的“早期”阶段到“晚期”阶段的转变。由于基因组复制是通过构成了病原体相关分子模式(PAMP)的dsRNA中间体的产生而发生的,因此细胞内模式识别受体(PRR)会触发强烈的先天免疫反应。从-RNA模板,nsP产物还介导来自内部RNA依赖性RNA聚合酶启动子(“亚基因组启动子”或SGP)的转录,该启动子映射到(当从原始有义链RNA基因组的角度读取时)紧邻第二个ORF上游的遗传区域以产生仅编码第二个ORF的较小的或“亚基因组的”(SG)mRNA。SG mRNA的转录比全长基因组的转录更有利,其导致感染细胞中SG mRNA拷贝数高,并因此其中编码的结构基因的产量也在高水平。

[0082] 如本发明所用,RNA复制子是指从甲病毒属(披膜病毒科家族)基因组而来的序列衍生的核酸分子,其在引入合适的细胞内环境时能够自我复制。由于其自我复制,RNA复制子可以大幅放大所编码蛋白质的产生,从而导致所需蛋白质在体内持续翻译。术语“复制子RNA”和“RNA复制子”在本文中可互换使用以指代这些RNA分子。

[0083] 在一种实施方式中,RNA复制子可以是源自甲病毒属物种的RNA复制子。甲病毒属物种可包括但不限于东部马脑炎病毒(Eastern equine encephalitis virus,EEEV)、委内瑞拉马脑炎病毒(Venezuelan equine encephalitis virus,VEEV)、大沼泽地病毒(Everglades virus,EVEV)、穆坎博热病毒(Mucambo virus,MUCV)、塞姆利基森林病毒(Semliki forest virus,SFV)、皮苏纳病毒(Pixuna virus,PIXV)、米德尔堡病毒(Middleburg virus,MIDV)、基孔肯雅病毒(Chikungunya virus,CHIKV)、奥尼永-尼永病毒(O'Nyong-Nyong virus,ONNV)、罗斯河病毒(Ross River virus,RRV)、巴尔马森林病毒(Barmah Forest virus,BF)、盖塔病毒(Getah virus,GET)、鹭山病毒(Sagiyama virus,SAGV)、贝巴鲁病毒(Bebaru virus,BEBV)、马亚罗病毒(Mayaro virus,MAYV)、乌纳病毒

(Una virus,UNAV)、辛德比斯病毒(Sindbis virus,SINV)、奥拉病毒(Aura virus,AURAV)、沃塔罗阿病毒(Whataroa virus,WHAV)、巴班基病毒(Babanki virus,BABV)、孜拉加奇病毒(Kyzylagach virus,KYZV)、西部马脑炎病毒(Western equine encephalitis virus,WEEV)、高地J病毒(Highland J virus,HJV)、摩根堡病毒(Fort Morgan virus,FMV)、恩杜茂病毒(Ndumu,NDUV)和虫溪病毒(Buggy Creek virus)。

[0084] 在一种实施方式中, RNA复制子可以是自通过自然选择产生的RNA病毒基因组衍生的复制子。

[0085] 在一种实施方式中, 本发明公开的RNA复制子可以是工程化的、合成的或重组的RNA复制子。RNA复制子可以在体外合成或修饰。修饰可以包括但不限于以下程序: 进行化学或酶促技术处理, 例如通过使用化学核酸合成、使用用于复制、聚合、核酸外切消化、核酸内切消化、连接、逆转录、转录、碱基修饰(例如甲基化)或核酸分子的重组的酶。RNA复制子可以通过使用合理的设计方法来工程化, 该方法选择性地保持关键序列元件、附加功能并减少与危险的、流行的病毒的同源性。RNA复制子可以与在自然界中不结合的核苷酸序列结合。可以使用分子克隆技术来操作RNA复制子, 使得其相对于天然存在的核苷酸序列具有一种或多种序列改变或重排。

[0086] 基于甲病毒的RNA复制子可用作体外和体内基因表达的载体。甲病毒的结构基因在可以用所需的重组蛋白替换的单个亚基因组(SG)开放阅读框(ORF)中编码。非结构(nsP)编码序列位于单独的前一个ORF中, 其携带必要的生化功能, 可介导细胞内复制子RNA的复制以及SG ORF的翻译。当nsP编码序列与流行的甲病毒序列几乎相同时, 会产生在现场与病原病毒基因组重组的可能性, 从而存在安全风险。为了防止与病原病毒序列重组, 也可以修饰nsP序列。nsP的修饰具有挑战性, 因为修改一级氨基酸序列可能会影响生化功能, 而且修改一级核苷酸序列可能会破坏复制子翻译和转录所需的关键二级结构。

[0087] 合成的甲病毒衍生复制子核酸分子

[0088] 在一种实施方式中, 提供了合成的甲病毒衍生复制子核酸分子。合成的甲病毒衍生复制子核酸分子可包含第一核酸, 其编码甲病毒非结构蛋白nsP1、nsP2、nsP3和nsP4。第一核酸可包含从5'非翻译区(UTR)的起点到亚基因组启动子(SGP)的末端范围内的序列。第一核酸可包含在如SEQ ID NO:17所示的甲病毒基因组序列中从nt 503至nt 658、nt 658至nt 1620、nt 1620至nt 2560、nt 2560至nt 3954、nt 3954至nt 4120、nt 6381至nt 7083和nt 6966至nt 7526的区域内的任意位置处引入的至少一个沉默突变。合成的甲病毒衍生复制子核酸分子还可以包含第二核酸序列, 其包含经过修饰的亚基因组(SG)开放阅读框(ORF)。经过修饰的SG ORF可以包含编码第一异源蛋白的序列。第一核酸序列可包含与nsP4 C末端融合编码的第二异源蛋白。nsP4编码序列可以含有在从nt 6966至nt 7526的区域内的任意位置处引入的至少一个沉默突变。

[0089] 如本发明所用, 术语“异源蛋白”是指任意在生物体中非天然存在的蛋白质。

[0090] 如本发明所用, 术语“沉默突变”是指不修改所编码蛋白质的氨基酸序列的核苷酸碱基序列的改变。例如, 沉默突变可以是将第一密码子改变为编码相同氨基酸的第二密码子的点突变。沉默突变可包括一个或多于一个核苷酸的改变。

[0091] 如本发明所用, 术语“密码子优化”是指用于通过适应宿主生物体的密码子偏好来改善基因表达并增加感兴趣基因的翻译效率的过程。

[0092] 如本发明所用,术语“密码子适应”是指用于通过沉默突变修改密码子的过程,其经过选择从而不会影响通过RNA折叠的生物信息学分析确定的二级结构基序(motif)。

[0093] 在一种实施方式中,第一核酸可包含在甲病毒基因组中从nt 503至nt 658的区域内的任意位置处引入的至少一个沉默突变。如本发明所用,给出的所有基因组位置均参考委内瑞拉马脑炎病毒的特立尼达驴病毒株,其完整基因组参考序列L01442见NCBI GenBank数据库(SEQ ID NO:17;Kinney,R.M.,Johnson,B.J.,Brown,V.L.and Trent,D.W.1986, Nucleofide sequence of the 26S mRNA of the virulent Trinidad donkey strain of Venezuelan equine encephalitis virus and deduced sequence of the encoded structural proteins,Virology152(2),400-413,其通过引用并入本发明,如同已完全阐述)。

[0094] 携带此沉默突变的第一核酸可以包含编码nsP1的多核苷酸,并且包含与如SEQ ID NO:21所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98、99%或100%的同一性的序列。

[0095] 在一种实施方式中,第一核酸可以包含在如SEQ ID NO:17所示的甲病毒基因组序列中从nt 658至nt 1620的区域内的任意位置处引入的至少一个沉默突变的序列。携带此沉默突变的第一核酸可以包含编码nsP1的多核苷酸,并且包含与如SEQ ID NO:25所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0096] 在一种实施方式中,第一核酸可以包含在甲病毒基因组序列(SEQ ID NO:17)中从nt 1620至nt 2560的区域内的任意位置处引入的至少一个沉默突变的序列。携带沉默突变的第一核酸可以包含编码nsP1和nsP2的接合物(junction),并且包含与如SEQ ID NO:24所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0097] 在一种实施方式中,第一核酸可以包含在甲病毒基因组序列(SEQ ID NO:17)中从nt 3694至nt 3954的区域内的任意位置处引入的至少一个沉默突变的序列。包含沉默突变的第一核酸可以包含编码nsP1和nsP2的接合物的多核苷酸,并且包含与如SEQ ID NO:22所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0098] 在一种实施方式中,第一核酸可以包含在甲病毒基因组序列(SEQ ID NO:17)中从nt2560到nt 3954的区域内的任意位置处引入的至少一个沉默突变的序列。包含沉默突变的核酸可包含编码nsP1和nsP2的接合物的多核苷酸,并且包含与如SEQ ID NO:26所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0099] 在一种实施方式中,第一核酸可以包含在甲病毒基因组序列(SEQ ID NO:17)中从nt 3954至nt 4120的区域内的任意位置引入的至少一个沉默突变的序列。包含沉默突变的第一核酸可包含编码nsP2和nsP3的接合物的多核苷酸,并且包含与如SEQ ID NO:23所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96、97%、98%、99%或100%的同一性的序列。

[0100] 在一种实施方式中,第一核酸可以包含在甲病毒基因组序列(SEQ ID NO:17)中从

nt 6381到nt 7083的区域内的任意位置处引入的至少一个沉默突变的序列。包含沉默突变的第一核酸可包含编码nsP4的多核苷酸,并且包含具有与如SEQ ID NO:20所示的参考序列至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0101] 在一种实施方式中,第一核酸序列可以包含在甲病毒基因组序列(SEQ ID NO:17)中从nt 6966到nt 7526的区域内的任意位置处引入的至少一个沉默突变的序列。包含沉默突变的第一核酸可包含编码nsP4的C末端区域的多核苷酸,并且包含与如SEQ ID NO:18或19所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0102] 在一种实施方式中,第一核酸可以包含在基因组的一个以上区域中具有沉默突变的序列。该核酸可包含在甲病毒基因组序列(SEQ ID NO:17)中从nt 658至nt 1620的区域内和nt 2560至nt 3954的区域内的任意位置处引入的一个以上沉默突变。包含沉默突变的第一核酸可包含编码nsP1的多核苷酸,并且包含与如SEQ ID NO:25所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列,以及,包含编码nsP2的多核苷酸,并且包含与如SEQ ID NO:26所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0103] 在一种实施方式中,第一核酸可以包含在甲病毒基因组序列(SEQ ID NO:17)中从nt 3694至nt 3954的区域内和nt 6381至nt 7083的区域内的任意位置处引入的一个以上沉默突变的序列。包含这些沉默突变的第一核酸可以包含编码nsP2的多核苷酸,并且包含与如SEQ ID NO:22所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列,以及,包含编码nsP4的多核苷酸,并且包含与如SEQ ID NO:20所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0104] 在一种实施方式中,第一核酸可以包含在甲病毒基因组序列(SEQ ID NO:17)中,在从nt 503到nt 658的区域内、在从nt 658到nt 1620的区域内、在从nt 1620到nt 2560的区域内、在从nt 2560至nt 3954的区域内、在从nt 3954至nt 4120的区域内以及在从nt 6381至nt 7083的区域内的任意位置处引入的一个以上沉默突变的序列。包含这些沉默突变的第一核酸可包含编码nsP1的多核苷酸,并且包含具有与如SEQ ID NO:21所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列;编码nsP1的多核苷酸,并且包含与如SEQ ID NO:25所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列;编码nsP1和nsP2的接合物的多核苷酸,并且包含与如SEQ ID NO:24所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列;编码nsP2的多核苷酸,并且包含与如SEQ ID NO:26所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列;编码nsP2和nsP3的接合物的多核苷酸,并且包含与如SEQ ID NO:23所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的

同一性的序列；以及，编码nsP4的多核苷酸，并且包含与如SEQ ID NO:20所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。包括这些沉默突变的第一核酸可以包含编码非结构蛋白的多核苷酸，其包含与如SEQ ID NO:30所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0105] 确定两个氨基酸序列或两个核酸序列的百分比同一性可以包括调整 and 比较两个序列中相应位置处的氨基酸残基或核苷酸。如果两个序列中的所有位置都被相同的氨基酸残基或核苷酸占据，则称该序列100%同一。百分比同一性通过Smith Waterman算法测量 (Smith TF, Waterman MS1981 "Identification of Common Molecular Subsequences," J Mol Biol 147:195-197, 其通过引用并入本文，如同已完全阐述)。

[0106] 在一种实施方式中，第一核酸可以包含表1中列出的甲病毒之一的nsP区域的序列，其中nsP1编码序列中的一个以上的沉默突变在第85位氨基酸之后的位置处产生，使得10-30%的核苷酸与野生型nsP1编码序列不同。

[0107] 表1甲病毒及其对应的可以如本发明所述通过沉默突变进行修饰以降低与其同源的野生型基因组的同源性的nsP区域。

[0108]

抗原 复合体	物种	亚型	参考序列 (NCBI GenBank)	nsP 蛋白参 考序列	nsP 氨基酸位置			
					nsP1	nsP2	nsP3	nsP4
巴 马 森 林 病 毒 复 合 体								
	巴 马 森 林 病 毒		NC_001786	P87515	1-533	534-13 31	1332-1 801	1802- 2411
东 部 马 脑 炎 病 毒 复 合 体								
	东 部 马 脑 炎 病 毒		NC_003899	NP_632021	1-532	533-13 26	1327-1 878	1880- 2493
米 德 尔 堡 病 毒 复 合 体								
	米 德 尔 堡 病 毒		NC_024887	YP_009058 892	1-537	538-13 36	1327-1 794	1795- 2411
恩 杜 茂 病 毒 复 合 体								
	恩 杜 茂 病 毒		NC_016959	YP_008888 544	1-538	539-13 32	1333-1 822	1823- 2433
塞 姆 利 基 森 林 病 毒 复 合 体								
	贝 巴 鲁 病 毒		NC_016962	YP_008901 140	1-535	536-13 34	1335-1 845	1846- 2456
	基 孔 肯 雅 病 毒		NC_004162	Q8JUX6	1-535	536-13 33	1334-1 863	1864- 2474
	盖 塔 病 毒		NC_006558	YP_164438	1-534	535-13 32	1333-1 856	1857- 2467
	马 亚 罗 病 毒		NC_003417	NP_579968	1-538	539-13 34	1335-1 819	1820- 2437

[0109]

		亚型：乌纳病毒	NC_043403	YP_009665988	1-560	561-1359	1360-1916	N/A
	奥尼永-尼永病毒		NC_001512	NP_041254	1-535	536-1333	1334-1903	1904-2514
		伊博-奥拉病毒 (Igbo-Oravirus)	MF409176	O90370	1-535	536-1333	1334-1902	1903-2513
	罗斯河病毒		NC_001544	NP_062879	1-533	534-1331	1332-1869	1870-2480
		亚型：鹭山病毒	MW410934	Q9JGL0	1-534	535-1332	1333-1856	1857-2467
委内瑞拉马脑脊髓炎复合体 (Venezuelan equine encephalitis complex)								
	卡巴索病毒 (Cabassou virus)		NC_038670	YP_009507794	1-535	536-1329	1330-1844	1845-2451
	大沼泽地病毒		NC_038671	YP_009507796	1-535	536-1329	1330-1871	1872-2478
	佩德拉斯湖病毒 (Mosso-das Pedras virus)		NC_038857	YP_009508088	1-535	536-1329	1330-1899	1900-2505
	穆坎博热病毒		NC_038672	YP_009507798	1-535	536-1329	1330-1848	1849-2455
	皮苏纳病毒		NC_038673	YP_009507800	1-535	536-1329	1330-1870	1871-2477
	里约黑病毒 (Rio Negro virus)		NC_038674	YP_009507802	1-535	536-1329	1330-1874	1875-2480
	委内瑞拉马脑炎病毒		NC_001449	NP_040822	1-535	536-1329	1330-1886	1887-2492
西部马脑炎 (WEE) 复合体								
	奥拉病毒		NC_003900	NP_632023	1-539	540-1345	1346-1889	1890-2499
	辛德比斯病毒		NC_001547	P03317	1-540	541-1347	1348-1903	1904-2513
	奥科尔伯病毒 (Ockelbo virus)		M69205	AAA96972	1-540	541-1347	1348-1905	1906-2515
	沃塔罗阿病毒		NC_016961	YP_008888546	1-540	541-1347	1348-1861	1862-2471
WEE 复合体中的重组物种								
	摩根堡病毒		NC_013528	YP_003324587	1-533	534-1327	1328-1849	1850-2457
	高地 J 病毒		NC_012561	YP_002802299	1-533	534-1327	1328-1842	1843-2450
	西部马脑炎病毒		NC_003908	NP_640330	1-533	534-1327	1328-1859	1860-2467
未分类的	伊拉特病毒 (Eilat virus)		NC_018615	YP_008901141	1-543	544-1352	1353-1998	1999-2415
	姆维尼-伦加甲病毒		LC361437	BBC45634	1-543	545-1352	1354-1975	1976-2438

[0110]

(Mwini-lunga alpha-virus)							
昏睡病病毒 (Sleeping disease virus)		NC_003433	NP_598184	1-563	564-14 20	1421-1 984	1985- 2593
鲑鱼胰腺病 病毒 (Salmon pan-creatic disease virus)		NC_003930	NP_647496	1-562	563-14 21	1422-1 992	1993- 2601
南部象海豹 病毒 (Southern elephant seal virus)		NC_016960	YP_008888 545	1-548	549-13 43	1344-1 832	1833- 2442
托纳特病毒 (Tonate virus)		NC_038675	YP_009507 804	1-535	536-13 29	1330-1 894	1895- 2501
卡因瓜病毒 (Caaingua virus)		MK353339	QBM15857	1-535	536-13 61	1362-2 028	2029- 2636
江豚甲病毒 (Harbor porpoise alpha-virus)		N/A	QJE50387	1-538	539-13 60	1361-1 879	1880- 2489

[0111] 在一种实施方式中,第一核酸可以包含表1中列出的甲病毒之一的nsP区域的序列,其中在nsP2编码序列中产生一个以上沉默突变,使得10-30%的核苷酸不同于野生型nsP2编码序列。

[0112] 在一种实施方式中,第一核酸可以包含表1中列出的甲病毒之一的nsP区域的序列,其中在nsP3编码序列中产生一个以上沉默突变,使得10-30%的核苷酸不同于野生型nsP3编码序列。

[0113] 在一种实施方式中,第一核酸可以包含表1中列出的甲病毒之一的nsP区域的序列,其中在该nsP的最后30个氨基酸之前的位置产生nsP4编码序列中的一个以上沉默突变,使得10-30%的核苷酸与野生型nsP4编码序列不同。

[0114] 在一种实施方式中,第一核酸可以包含表1中列出的甲病毒之一的nsP区域的序列,其中在该nsP的最后30个氨基酸内的位置处产生nsP4编码序列中的一个以上沉默突变,使得10-30%的核苷酸与野生型nsP4编码序列不同并降低与内源SGP的同源性。

[0115] 在一种实施方式中,基于表1中的物种之一的甲病毒属衍生复制子核酸分子可包含第二核酸,其包含编码第一异源蛋白的多核苷酸。第一异源蛋白可以是替代由甲病毒亚基因组(SG) ORF编码的甲病毒结构蛋白的蛋白质。

[0116] 在一种实施方式中,甲病毒衍生复制子核酸分子可包含编码第二异源蛋白的多核苷酸。第二异源蛋白可以是插入经修改的nsP1、nsP2、nsP3或nsP4蛋白中的蛋白质。

[0117] 在一种实施方式中,经修改的nsP蛋白可以是经修改的nsP4蛋白。经修改的nsP4可以包含与其C末端区域融合的第二异源蛋白。编码异源C末端区域的多核苷酸序列可以在其3'末端包含编码甲病毒亚基因组启动子的序列。该序列可以插入nsP4编码序列区域的下游(即3')并且可以翻译成蛋白质。或者,此插入序列可以不被翻译成蛋白质。插入此区域的非翻译序列可以是短发夹RNA(shRNA)。

[0118] 在一种实施方式中,经修改的nsP4可以包含GFP蛋白作为第二异源蛋白。经修改的nsP4可包含与如SEQ ID NO:12或41所示的参考序列具有至少70%、72%、75%、80%、85%、

90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的氨基酸序列。经修改的nsP4可以由SEQ ID NO:31所示的多核苷酸序列编码。经修改的nsP4可包含具有与如SEQ ID NO:15所示的参考序列至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的氨基酸序列。

[0119] (在SG ORF中的)第一异源蛋白或(在nsP ORF中的)第二异源蛋白可以是病毒病原体分离或衍生的任意抗原蛋白。病毒病原体可选自腺病毒科(Adenoviridae)、乳多空病毒科(Papovaviridae)、细小病毒科(Parvoviridae)、疱疹病毒科(Herpesviridae)、痘病毒科(Poxviridae)、指环病毒科(Anelloviridae)、多形性包膜病毒科(Pleolipoviridae)、呼肠病毒科(Reoviridae)、小核糖核酸病毒科(Picomaviridae)、杯状病毒科(Caliciviridae)、披膜病毒科(Togaviridae)、沙粒病毒科(Arenaviridae)、黄病毒科(Flaviviridae)、正粘病毒科(Orthomyxoviridae)、副粘液病毒科(Paramyxoviridae)、布尼亚病毒科(Bunyaviridae)、弹状病毒科(Rhabdoviridae)、丝状病毒科(Filoviridae)、冠状病毒科(Coronaviridae)、星状病毒科(Astroviridae)、玻那病毒科(Bomaviridae)、动脉炎病毒科(Arteriviridae)、肝炎病毒科(Hepeviridae)、逆转录病毒科(Retroviridae)、花椰菜病毒科(Caulimoviridae)或肝脱氧核糖核酸病毒科(Hepadnaviridae)的家族成员。例如,病毒病原体可以为腺病毒、鼻病毒、轮状病毒、西尼罗病毒、寨卡病毒、疱疹病毒、丝状病毒或冠状病毒(CoV)。冠状病毒可以为MERS CoV、SARS-CoV和SARS-CoV-2中的一种。

[0120] 如本发明所用,术语“抗原蛋白”是指包含一种或多种触发免疫应答的抗原决定簇的蛋白质。免疫反应可能涉及抗体的产生,或特定免疫活性细胞的激活,或两者兼而有之。抗原蛋白可以是病原体或癌细胞的结构成分。抗原蛋白可以在宿主中合成、重组产生,或者可以衍生自生物样品,包括但不限于组织样品、细胞或生物液体。

[0121] 抗原蛋白可以是但不限于寄生虫抗原蛋白、细菌抗原蛋白、肿瘤抗原蛋白、环境抗原蛋白、治疗性抗原蛋白或过敏原。抗原蛋白可以是包含一种或多种本发明所述的病毒病原体的抗原决定簇的蛋白质。

[0122] 在一种实施方式中,第一异源蛋白或第二异源蛋白可以是酶。该酶可以源自哺乳动物并可调节细胞代谢或细胞信号传导。该酶可以是细胞或病毒蛋白酶。病毒蛋白酶可以是小核糖核酸病毒3C蛋白酶(3Cpro)。如本发明所用,术语“3C^{pro}”、“3C蛋白酶”、“3Cpro”或“3CP”是指在小核糖核酸病毒物种中发现的半胱氨酸蛋白酶,其活性位点内通常含有保守的半胱氨酸-组氨酸-天冬氨酸/谷氨酸催化三联体或含有半胱氨酸-组氨酸二联体。催化三联体或二联体通常形成电荷中继网络,极化并激活攻击底物的亲核体(通常为半胱氨酸),形成共价中间体,然后水解以再生游离酶。3C^{pro}可在多聚蛋白的结构区和非结构区中进行成熟裂解且对谷氨酰胺-甘氨酸/丝氨酸/丙氨酸/缬氨酸/组氨酸/精氨酸和谷氨酸-丝氨酸/甘氨酸/精氨酸/甲硫氨酸中具有显著的底物偏好。3Cpro蛋白酶可源自口蹄疫病毒(FMDV)。

[0123] FMDV 3C蛋白酶是分子量约23.1-kDa并且含有213个氨基酸的半胱氨酸蛋白酶。FMDV 3C蛋白酶活性位点的半胱氨酸-组氨酸-天冬氨酸催化三联体由残基H46、D84和C163形成。

[0124] 在一种实施方式中,经修改的nsP4可以包含第二异源蛋白,其为3C蛋白酶。经修改的nsP4可包含与如SEQ ID NO:40所示的参考序列具有至少70%、72%、75%、80%、85%、

90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的氨基酸序列。编码包含3C蛋白酶的改变的nsP4多核苷酸序列可包含具有与SEQ ID NO:32中所示参考序列至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性。

[0125] 在一种实施方式中,合成的甲病毒衍生复制子核酸分子可以进一步包含第一异源蛋白,其为FMDV P1前体多肽或其他感兴趣的多肽。FMDV P1可衍生自FMDV血清变型01 Manisa,并且在本发明中称为01 Manisa P1多蛋白。01 Manisa P1多蛋白可由包含与如SEQ ID NO:33所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的多核苷酸序列编码。

[0126] 在一种实施方式中,第一异源蛋白可以是FMDV P1前体多肽(01 Manisa P1多蛋白),并且第二异源蛋白可以是3C蛋白酶。

[0127] 第一或第二异源蛋白可以是病毒衍生的或哺乳动物细胞衍生的细胞代谢、免疫或信号传导功能的调节剂。这种异源蛋白的表达可能会影响细胞抗病毒先天免疫反应,从而增强甲病毒复制子RNA分子的复制。

[0128] 在一种实施方式中,第二异源蛋白可以是模式识别受体(PRR)蛋白。PRR蛋白可以是STING(干扰素基因刺激因子)蛋白。STING蛋白可以是野生型STING蛋白。包含STING蛋白的经修改的nsP4可包含与如SEQ ID NO:42所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的氨基酸序列。包含STING蛋白的经修改的nsP4蛋白可以由与如SEQ ID NO:34所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%同一性的多核苷酸序列编码。第一异源蛋白可包含抗原蛋白并且第二异源蛋白可包含STING蛋白。

[0129] 第一异源蛋白或第二异源蛋白可以是适合保护受试者免受疾病侵害的任何免疫原性多肽。所述疾病可以是癌症或由微生物、细菌、原生动物、寄生虫或病毒病原体引起的疾病。

[0130] 第一异源蛋白可以是与第二异源蛋白相同的蛋白质。

[0131] 在一种实施方式中,合成的甲病毒衍生复制子核酸分子可以包含翻译中断序列,例如2A序列。翻译中断序列允许蛋白质翻译的中断,并在不一定被蛋白酶识别的位点(例如经修改的nsP4蛋白中的位点)进行有效切割。2A序列可以在编码第二异源蛋白质的序列之前插入到编码经修改的nsP4的核酸序列中。2A序列的插入可以导致通过2A造成的翻译的中断以及不具有N端Met残基的多肽的产生。2A序列可以是源自但不限于明脉扁刺蛾病毒(Thosea asigna virus)(SEQ ID NO:43)、猪疱疹病毒-1(porcine teschovirus-1)(SEQ ID NO:44)、口蹄疫病毒(SEQ ID NO:45)、A型马鼻炎病毒(equine rhinitis A virus)(SEQ ID NO:46)或其他小核糖核酸病毒科或Permutotetraviridae科病毒的2A的氨基酸序列。

[0132] 在一种实施方式中,合成的甲病毒衍生复制子核酸分子可包含指导编码第一异源蛋白和/或第二异源蛋白的序列表达的亚基因组(SGP)启动子。亚基因组启动子可以是源自甲病毒属物种(参见表1)的26S亚基因组启动子。编码异源蛋白的序列也可以处于内部核糖体进入位点(IRES)的控制之下。

[0133] 本发明所述的合成的甲病毒衍生复制子核酸分子可包含一种或多种经修饰的核

苷酸。经修饰的核苷酸可以是5-甲基尿苷、5-甲基胞苷、假尿苷或N6-甲基腺苷。

[0134] 合成的甲病毒衍生复制子核酸分子可以具有3' poly-A尾巴。

[0135] 在一种实施方式中,可以对合成的甲病毒衍生复制子核酸分子的RNA序列进行密码子优化以提高翻译效率。RNA分子可以通过本领域已知的任何方法进行修饰以增强稳定性和/或翻译(效率)。RNA分子可以通过添加例如至少30个腺苷残基的polyA尾巴来修饰。RNA分子可以通过用经修饰的核糖核苷酸在5末端进行加帽修饰(例如7-甲基鸟苷帽),这种修饰可以在RNA合成期间掺入,也可由在RNA转录后通过酶工程的方式加入。

[0136] 在一种实施方式中,提供了包含编码合成的甲病毒衍生复制子核酸分子的DNA序列的核酸。该核酸可以是例如DNA质粒或线性化DNA质粒的片段。该核酸还可以包含可操作地连接至DNA序列的5'末端的启动子,例如T7启动子。根据本发明公开的内容,核酸可用于使用本领域已知的方法生产本申请的RNA复制子。例如,合成的甲病毒衍生复制子核酸分子可通过体内或体外核酸的转录获得。

[0137] 在一种实施方式中,DNA序列可以是DNA质粒。DNA质粒可包含编码nsP1并包含与如SEQ ID NO:2所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。

[0138] DNA质粒可包含编码nsP1并且包含与如SEQ ID NO:6所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。

[0139] DNA质粒可包含编码nsP1和nsP2接合物并且包含与如SEQ ID NO:5所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。

[0140] DNA质粒可包含编码nsP1和nsP2接合物并且包含与如SEQ ID NO:3所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。

[0141] DNA质粒可包含编码nsP1和nsP2接合物并且包含与如SEQ ID NO:7所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。

[0142] DNA质粒可包含编码nsP2和nsP3接合物并且包含与如SEQ ID NO:4所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。

[0143] DNA质粒可包含编码nsP4并包含与如SEQ ID NO:1所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。

[0144] DNA质粒可包含编码nsP1,并且包含与如SEQ ID NO:6所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸,以及编码nsP2,并且包含与如SEQ ID NO:7所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。

[0145] DNA质粒可包含编码nsP2,并且包含与如SEQ ID NO:3所示的参考序列具有至少

70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸,以及编码nsP4,并且包含与如SEQ ID NO:1所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。

[0146] DNA质粒可包含编码nsP1,并且包含与如SEQ ID NO:2所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸;编码nsP1,并且包含与如SEQ ID NO:6所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸,编码nsP1和nsP2的接合物,并且包含与如SEQ ID NO:5所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸,编码nsP2,并且包含与如SEQ ID NO:7所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸;编码nsP2和nsP3接合物,并且包含与如SEQ ID NO:4所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸;以及编码nsP4,并且包含与如SEQ ID NO:1所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列的多核苷酸。包括这些沉默突变的第一核酸可以包含编码非结构蛋白的多核苷酸,其包含与如SEQ ID NO:11所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的序列。

[0147] 在一种实施方式中,包含编码甲病毒衍生复制子核酸分子的DNA序列的核酸可以含有包含编码第一异源蛋白的多核苷酸的第二核酸。第一异源蛋白可以是替代由甲病毒SG ORF编码的甲病毒结构蛋白的蛋白质。

[0148] 在一种实施方式中,核酸可以包含编码第二异源蛋白的多核苷酸。第二异源蛋白可以是插入到经修改的nsP1、nsP2、nsP3或nsP4蛋白中的蛋白质。

[0149] 在一种实施方式中,经修改的nsP蛋白可以是经修改的nsP4蛋白。经修改的nsP4可以包含第二异源蛋白。编码经修改的nsP4的多核苷酸可包含编码第一C末端区域的多核苷酸序列和编码第二C末端区域的多核苷酸序列。编码第二C末端区域的多核苷酸序列可包含编码甲病毒亚基因组启动子的序列。编码亚基因组启动子的序列可以与如SEQ ID NO:14所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性。

[0150] 在一种实施方式中,经修改的nsP4可以包含GFP蛋白作为第二异源蛋白。经修改的nsP4可包含与如SEQ ID NO:12所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的氨基酸序列。经修改的nsP4可以由如SEQ ID NO:13所示的多核苷酸序列编码。经修改的nsP4可包含与如SEQ ID NO:15所示的参考序列具有至少70%、72%、75%、80%、85%、90%、91%、92%、93%、94%、95%、96%、97%、98%、99%或100%的同一性的氨基酸序列。经修改的nsP4可以由如SEQ ID NO:16所示的多核苷酸序列编码。

[0151] 甲病毒衍生RNA复制子表达系统

[0152] 在一种实施方式中,提供了甲病毒衍生的RNA复制子表达系统。甲病毒衍生RNA复制子表达系统可包含本发明所述的任意一种合成的甲病毒衍生复制子核酸分子以及一个或多个宿主细胞(a host cell or cells)。

[0153] 宿主细胞可以是但不限于仓鼠卵巢(CHO)细胞、肿瘤细胞系、BHK细胞、人细胞系,例如HEK293细胞、PER.C6细胞。宿主细胞可以为酵母、真菌、昆虫细胞、动物或植物细胞。可以使用来自人类或任意哺乳动物物种的未转化的原代正常细胞,例如成纤维细胞、干细胞或者淋巴或髓系细胞。

[0154] 在一种实施方式中,在宿主细胞中第一异源蛋白和/或第二异源蛋白的生产包括合成的甲病毒衍生的复制子核酸分子的引入,所述复制子核酸分子包含编码将在宿主细胞中表达的第一异源蛋白质和/或第二异源蛋白质的核酸序列,在适合核酸分子表达的条件下培养宿主细胞并允许第一异源蛋白和/或第二异源蛋白在宿主细胞中表达。

[0155] 为了在宿主细胞中表达,包含编码第一异源蛋白和/或第二异源蛋白的核酸序列的合成的甲病毒衍生复制子核酸分子可以包含在表达盒中,所述表达盒包含核酸序列的表达所需的或促进表达的调控序列。调节序列可包括启动子、增强子、启动子和/或多聚腺苷酸化信号分子等。本领域已知的各种启动子可用于在宿主细胞中表达本发明公开的核酸。启动子可以是组成型或诱导型启动子。启动子可以源自原核或真核生物,或者可以人工设计。宿主细胞可以在可商购的合适用于培养表达第一异源蛋白和/或第二异源蛋白的细胞的培养基中培养。

[0156] 一种实施方式提供了含有合成的甲病毒衍生复制子核酸分子的宿主细胞,所述复制子核酸分子包含编码3C蛋白酶的核酸序列。宿主细胞还可以包含编码FMDV P1前体多肽或其他目的多肽的一种或多种多核苷酸构建体。

[0157] 生产至少一种异源蛋白的方法

[0158] 在一种实施方式中,提供了在细胞中生产至少一种异源蛋白的方法。该方法可以包括在细胞中表达本发明所述的任意一种合成的甲病毒衍生复制子核酸分子。

[0159] 本发明描述了如何成功编辑甲病毒复制子的nsP基因区域,以降低与野生型病毒的同源性,并将新的生物学功能纳入此类人工复制子的自我扩增的早期阶段,该早期阶段在亚基因组编码的转基因翻译之前。这可以通过识别nsP区域中可以在不破坏保守的结构元件的情况下安全编辑的关键核苷酸片段,并且将额外的序列附加到nsP的天然末端来实现。本发明展示了如何可以在不损害复制子的自扩增功能的情况下进行此类修饰。因此,本发明提供了一种产生更安全的复制子的方法,所述复制子配备有额外的功能以修饰宿主细胞的细胞和免疫学参数。

[0160] 在一种实施方式中,该方法可以包括使用合适的培养基培养宿主细胞,并在宿主细胞中表达3C蛋白酶。该方法还可以包括将FMDV P1前体多肽(或其他目标多肽)表达和/或加工成FMDV病毒蛋白。该方法可包括培养宿主细胞并回收病毒蛋白VP0、VP1、VP2、VP3或VP4,或3C蛋白酶的其他切割产物。

[0161] 在一种实施方式中,该方法可以包括培养宿主细胞并表达含有合成的甲病毒衍生复制子核酸分子,该核酸分子包含编码作为第一异源蛋白的O1 Manisa P1多蛋白和作为第二异源蛋白的3C蛋白酶的核酸序列。

[0162] 在一种实施方式中,该方法可以包括培养含有翻译中断序列(例如2A序列)的合成

的甲病毒衍生复制子核酸分子。翻译中断序列允许蛋白质翻译的中断,并在不一定被3C蛋白酶识别的位点(例如经修改的nsP4蛋白质中的位点)进行有效切割。2A序列可以在编码第二异源蛋白的序列之前插入到编码经修改的nsP4的核酸序列中。2A序列的插入可以导致由2A造成的翻译的中断以及不具有N末端Met残基的多肽的产生。

[0163] 组合物

[0164] 在一种实施方式中,本发明描述的合成的甲病毒衍生复制子核酸分子或异源蛋白可以在组合物中施用。

[0165] 在一种实施方式中,组合物可以用于治疗自身免疫性疾病或过敏性疾病。该治疗可包括任意一种包含如本发明所述的异源蛋白的甲病毒衍生的复制子核酸分子,其表达抑制或以其他方式减少针对所需靶向抗原的适应性免疫应答的因子。

[0166] 在一种实施方式中,组合物可以包含用于表达靶向特异性抗原的合成的甲病毒衍生复制子核酸分子。靶向特异性抗原可以是肿瘤抗原。癌症特异性抗原可以衍生自以下任一种:自靶细胞分离的总mRNA、一种或多种特异性靶向mRNA分子、靶细胞的蛋白质裂解物、来自靶细胞的特异性蛋白,或合成的靶向特异性肽或蛋白质以及编码靶向特异性抗原或其衍生肽的合成mRNA或DNA。非限制性实例为以下癌症或肿瘤特异性抗原:WT1、MUC1、LMP2、HPV E6 E7、EGFRvIII、HER-2/neu、独特型(Idiotype)、MAGE A3、p53(非突变体)、NY-ESO-1、PSMA、GD2、CEA、MelanA/MART 1、Ras突变体、gp100、p53突变体、蛋白酶3(PR1)、Bcr-abl、酪氨酸酶、生存素、PSA、hTERT、肉瘤易位断点、EphA2、PAP、ML-IAP、AFP、EpCAM、ERG (TMPRSS2 ETS融合基因)、NA17、PAX3、ALK、雄激素受体、细胞周期蛋白B1、聚唾液酸、MYCN、RhoC、TRP-2、GD3、岩藻糖基GM1、间皮素、PSCA、MAGEA1、sLe(a)、CYP1B1、PLAC1、GM3、BORIS、Tn、GloboH、ETV6-AML、NY-BR-1、RGS5、SART3、STn、碳酸酐酶IX、PAX5、OY-TES1、精子蛋白17、LCK、HMWMAA、AKAP-4、SSX2、XAGE 1、B7H3、豆荚蛋白(Legumain)、Tie 2、Page4、VEGFR2、MAD-CT-1、FAP、PDGFR--、MAD-CT-2、Fos相关抗原1。

[0167] 在一种实施方式中,组合物可以用于治疗癌症。该治疗可包括任意一种包含本发明所述的异源蛋白质的甲病毒衍生复制子核酸分子,其表达抑制或以其他方式减少人或动物体内肿瘤细胞生长或增殖的因子。

[0168] 在一种实施方式中,组合物可包含3C蛋白酶。此类组合物可用于诱导或检测针对3C蛋白酶的免疫应答,例如直接针对3C蛋白酶的体液或细胞免疫应答。该组合物还可以包含前体多肽,例如FMDV P1前体,其含有被经修饰的FMDV 3C蛋白酶识别的位点。该组合物可以是可用于体外加工前体多肽的形式。

[0169] 在一种实施方式中,组合物可以是疫苗。疫苗可包含任意一种包含本发明所述的抗原蛋白的甲病毒衍生复制子核酸分子。术语“疫苗”是指含有活性成分的试剂或组合物,能够有效诱导受试者针对某种病原体或疾病产生一定程度的免疫力,从而至少降低与病原体或疾病感染相关的症状的严重程度、持续时间、其他表现的降低或消除症状。

[0170] 在一种实施方式中,本发明任一实施方式中公开的合成的甲病毒衍生复制子核酸分子可以封装在载体系统中用于体内施用。载体可以是但不限于阴离子脂质体、阳离子脂质体或树枝状聚合物(dendrimer)。术语“阴离子脂质体”是指包括含有阴离子基团的脂质的脂质体。阴离子脂质体可以由阴离子磷脂形成。术语“阳离子脂质体”是指全部或部分由带正电荷的脂质,或更具体地,由包含阳离子基团和亲脂部分的脂质制成的脂质体。用于阳

离子脂质体的阳离子脂质的带正电的分子提供了有利的结构特征。例如,阳离子脂质的亲脂性部分是疏水性的,因此可以引导其自身远离脂质体的水性内部并与其他非极性和疏水性物质缔合,或者相反,阳离子部分可以与极性分子和物质缔合,从而在阳离子脂质体的水性内部与之络合。带正电荷的脂质体可以与带负电荷的核酸分子相互作用以形成稳定的复合物。

[0171] 脂质体可以包括两性离子脂质。如本发明所用,术语“两性离子”是指同时含有正电荷和负电荷但具有净中性电荷的分子。

[0172] 脂质体可以由单一脂质或脂质混合物形成。脂质的亲水部分可以被聚乙二醇化,即通过聚乙二醇的共价连接进行修饰,以增加稳定性并防止脂质体的非特异性吸附 (Heyes et al. (2005) J Controlled Release 107:276-87,其通过引用并入本文,如同已完全阐述)。

[0173] 术语“树枝状聚合物”是指具有球体形状的高度分支化的大分子。树枝状聚合物分子的表面可以通过多种方式进行修饰,并且所得构建体的许多性质可以由其表面决定。树枝状聚合物可以被修饰以具有正表面电荷,即成为阳离子树枝状聚合物。阳离子树枝状聚合物可以与核酸形成暂时缔合。到达其目的地后,树枝状聚合物-核酸复合物可通过胞吞作用被摄入细胞中。

[0174] 单个树枝状聚合物-核酸复合物(本文也称为经修饰的树枝状聚合物纳米粒子(MDN Ps))的示例性尺寸在最长尺寸上可以在30nm至1,000nm的范围内。MDNP可具有30nm至450nm(含端值)、50nm至300nm(含端值)、或60nm至250nm(含端值)的平均尺寸。MDNPs可以是烷基改性的树枝状聚合物纳米颗粒。纳米颗粒的尺寸可能受到取代核心树枝状聚合物的烷基链的长度的影响。制备和配制改性树枝状聚合物纳米颗粒的方法可见W02021 207020(公开日为2021年10月14日);US20210330600(公开日为2021年10月28日);以及US20210338789(公开日为2021年11月4日)中的描述;所有这些内容均通过引用并入本文,如同已完整阐述。

[0175] 在一种实施方式中,组合物还可包含一种或多种佐剂。术语“佐剂”是指引起免疫系统刺激的一种或多种物质。佐剂可用于增强对用于使免疫受试者抵抗疾病的异源蛋白的免疫应答。一种或多种佐剂可以包括铝,例如氢氧化铝和/或磷酸铝;油乳液组合物或水包油组合物,例如角鲨烯-水乳液,例如MF59;皂苷制剂,例如QS21和免疫刺激复合物(ISCMS);细菌或微生物衍生物,例如单磷酸脂质A(MPL)、3-O-脱酰化MPL(3dMPL)、含有寡核苷酸的CpG基序、ADP-核糖基化细菌毒素或其突变体,例如大肠杆菌热不稳定肠毒素LT,霍乱毒素CT、核酸和其他TLR激动剂,例如聚(I:C);等等;真核蛋白,例如抗体或其片段,以及受体的配体,其在与受体细胞相互作用时刺激免疫反应。

[0176] 在一种实施方式中,本发明描述的组合物可以在没有佐剂的情况下使用。

[0177] 在一种实施方式中,组合物可以在药学上可接受的载体或赋形剂中施用。本发明所用的术语“药学上可接受的载体”是指药学上可接受的材料、组合物或媒介物,例如液体或固体填充剂、稀释剂、赋形剂、制造助剂(例如润滑剂、滑石粉、硬脂酸镁、硬脂酸钙或硬脂酸锌,或硬脂酸)或溶剂封装材料,其参与将主体化合物从一个器官或身体的一部分携带或运输至另一器官或身体的一部分。每种载体必须是“可接受的”,即与制剂中的其他成分相容并且不会对患者造成伤害。可用作药学上可接受的载体的材料的一些实例包括:(1)糖

类,如乳糖、葡萄糖和蔗糖;(2)淀粉类,如玉米淀粉和马铃薯淀粉;(3)纤维素及其衍生物,如羧甲基纤维素钠、甲基纤维素、乙基纤维素、微晶纤维素和醋酸纤维素;(4)黄芪胶粉;(5)麦芽;(6)明胶;(7)润滑剂,如硬脂酸镁、十二烷基硫酸钠和滑石粉;(8)赋形剂,例如可可脂和栓剂蜡;(9)油类,如花生油、棉籽油、红花油、芝麻油、橄榄油、玉米油和大豆油;(10)二醇类,例如丙二醇;(11)多元醇,例如甘油、山梨醇、甘露醇和聚乙二醇(PEG);(12)酯类,例如油酸乙酯和月桂酸乙酯;(13)琼脂;(14)缓冲剂,例如氢氧化镁和氢氧化铝;(15)海藻酸;(16)无热原水(pyrogen-free water);(17)等渗盐水;(18)林格氏溶液;(19)乙醇;(20)pH缓冲溶液;(21)聚酯、聚碳酸酯和/或聚酐;(22)增容剂,例如多肽和氨基酸(23)血清成分,例如血清白蛋白、HDL和LDL;(22)C2-C12醇类,例如乙醇;以及(23)用于药物制剂的其他无毒相容物质。润湿剂、着色剂、脱模剂、包衣剂、甜味剂、矫味剂、芳香剂、防腐剂和抗氧化剂也可以存在于制剂中。诸如“赋形剂”、“载体”、“药学上可接受的载体”或类似的术语在本发明中可互换使用。药学上可接受的载体和赋形剂是本领域已知的(Remington's Pharmaceutical Sciences,18th edition,A.R.Gennaro,Ed.,Mack Publishing Company,1990;Pharmaceutical Formulation Development of Peptides and Proteins,S.Frokjaer and L.Hovgaard,Eds.,Taylor&Francis,2000;Handbook of Pharmaceutical Excipients,3rd edition,A.Kibbe,Ed.,Pharmaceutical Press,2000)。本发明所述的合成的甲病毒衍生复制子核酸分子可以配制为无菌和/或冻干溶液并施用。无菌溶液可以通过无菌过滤或任何其他已知的方法来制备。然后将溶液冻干或填充到药物剂量容器中。溶液的pH通常可以在pH 3.0至9.5的范围内,例如pH 5.0至7.5。在一种实施方式中,可以添加稳定剂,例如白蛋白或另外的糖(例如蔗糖)。在一种实施方式中,可以添加去垢剂(detergent)。在一种实施方式中,包含本发明所述的任意复制子RNA的核酸混合物可以配制为可注射制剂。

[0178] 一种实施方式提供了通过本发明公开的方法产生的复制子RNA制剂在制备能够在有需要的患者中引发针对癌症的免疫应答的疫苗中的用途。

[0179] 在一种实施方式中,提供了在受试者中诱导免疫应答的方法。该方法可以包括向受试者施用治疗有效量的疫苗,该疫苗包含本发明所述的任意一种合成的甲病毒衍生复制子核酸分子。

[0180] 在一种实施方式中,提供了用于预防和/或治疗受试者的疾病或病症的方法。该方法可以包括利用本发明描述的任意一种组合物和疫苗。该方法可以包括向有需要的受试者施用治疗有效量的包含本发明如上所述的合成的甲病毒衍生复制子核酸分子的组合物或疫苗。

[0181] 本发明描述的组合物或疫苗的施用可以使用已知的施用途径进行。组合物或疫苗可以通过使用肠胃外给药途径来施用,例如皮内、肌内、皮下、经皮或粘膜给药途径,例如鼻内、口服等。在一种实施方式中,通过肌内注射施用组合物。如本发明所用,术语“施用”是指通过使所述组合物至少部分定位在期望部位处从而产生期望效果的方法或途径将组合物置入受试者体内。本发明所述的化合物或组合物可以通过本领域已知的任何适当途径施用,包括但不限于口服或肠胃外途径,包括静脉内、肌内、皮下、经皮、气道(气雾剂)、肺、鼻、直肠或局部(包括口腔和舌下)给药。

[0182] 示例性的施用模式包括但不限于注射、输注、滴注、吸入或吞食。“注射”包括但不

限于静脉注射、肌肉注射、动脉注射、鞘内注射、心室内注射、囊内注射、眼眶内注射、心内注射、皮内注射、腹腔注射、经气管注射、皮下注射、表皮下注射、关节内注射、囊下注射、蛛网膜下注射、椎管内注射、脑脊液内注射和皮内注射 (intrastemal injection) 以及输液。在一种实施方式中, 组合物可以通过静脉内输注或注射来施用。

[0183] 本发明描述的疫苗可以通过任何已知途径施用, 以诱导针对疫苗中的抗原的免疫应答。

[0184] 在一种实施方式中, 施用或递送的方法可以不限于上述方法, 并且可以使用用于细胞内递送的任何手段。

[0185] 治疗有效量是指有效预防、改善和/或治疗由病原体感染引起的疾病或病症的蛋白质或核酸分子的量。预防包括抑制或减少病原体的传播或者抑制或减少与病原体感染相关的一种或多种症状的发作、发展或进展。如本发明所用, 改善是指可见或可感知的疾病症状、病毒血症或由病原体诱导的感染的任何其他可测量的表现的减少。

[0186] 就治疗癌症而言, “治疗有效量” 是有效防止癌症进一步发展或转化生长、甚至实现癌症或实体瘤消退的量。

[0187] 治疗有效量的确定通常完全在本领域技术人员的能力范围之内。一般来说, 治疗有效量可以根据受试者的病史、年龄、状况、性别以及受试者的严重性和医疗状况的类型, 以及其他药剂的施用减轻了待治疗的疾病或病症的情况而变化。

[0188] 毒性和治疗功效可以通过细胞培养物或实验动物中的标准制药程序来确定, 例如, 用于确定LD50 (半数致死剂量) 和ED50 (半数治疗有效剂量)。毒性作用与治疗作用之间的剂量比是治疗指数, 可以用LD50/ED50的比值表示。优选表现出大治疗指数的组合物。如本发明所用, 术语ED表示有效剂量并且与动物模型结合使用。术语EC表示有效浓度并且与体外模型结合使用。

[0189] 从细胞培养试验和动物研究中获得的数据可用于制定用于人类的一系列剂量。此类化合物的剂量优选位于包括具有很小或无毒性的ED50的循环浓度范围内。剂量可以在此范围内变化, 这取决于所采用的剂型和所采用的施用途径。

[0190] 治疗有效剂量最初可以通过细胞培养试验进行估计。可以在动物模型中配制剂量, 以达到包括在细胞培养中确定的IC50 (即实现症状半最大抑制的治疗药物浓度) 的循环血浆浓度范围。血浆中的水平可以通过例如高效液相色谱法来测量。任何特定剂量的效果可以通过合适的生物试验来监测。

[0191] 剂量可由医生确定并根据需要进行调整以适应所观察到的治疗效果。一般而言, 可以施用组合物使得活性剂以1 μ g/kg至150mg/kg、1 μ g/kg至100mg/kg、1 μ g/kg至50mg/kg、1 μ g/kg至20mg/kg、1 μ g/kg至10mg/kg、1 μ g/kg至1mg/kg、100 μ g/kg至100mg/kg、100 μ g/kg至50mg/kg、100 μ g/kg至20mg/kg、100 μ g/kg至10mg/kg、100 μ g/kg至1mg/kg、1mg/kg至100mg/kg、1mg/kg至50mg/kg、1mg/kg至20mg/kg、1mg/kg至10mg/kg、10mg/kg至100mg/kg、10mg/kg至50mg/kg或10mg/kg至20mg/kg的计量给药。应理解的是, 本发明给出的范围包括所有中间范围, 例如, 范围1mg/kg至10mg/kg包括1mg/kg至2mg/kg、1mg/kg至3mg/kg、1mg/kg至4mg/kg、1mg/kg至5mg/kg、1mg/kg至6mg/kg、1mg/kg至7mg/kg、1mg/kg至8mg/kg、1mg/kg至9mg/kg、2mg/kg至10mg/kg、3mg/kg至10mg/kg、4mg/kg至10mg/kg、5mg/kg至10mg/kg、6mg/kg至10mg/kg、7mg/kg至10mg/kg、8mg/kg至10mg/kg、9mg/kg至10mg/kg等。还应当理解, 上述给出

的中间范围也在本发明的范围内,例如,在1mg/kg至10mg/kg的范围内,剂量范围例如2mg/kg至8mg/kg、3mg/kg至7mg/kg、4mg/kg至6mg/kg等。

[0192] 在一种实施方式中,可以以剂量施用该组合物,以使该活性剂的体内浓度在给药15分钟、30分钟、1小时、1.5小时、2小时、2.5小时、3小时、4小时、5小时、6小时、7小时、8小时、9小时、10小时、11小时、12小时或更多的给药时间后,小于500nM、小于400nM、小于300nM、小于250nM、小于200nM、小于150nM、小于100nM、小于50nM、小于25nM、小于20nM、小于10nM、小于5nM、小于1nM、小于0.5nM、小于0.1nM、小于0.05、小于0.01nM、小于0.005nM、小于0.001nM。

[0193] 关于治疗的持续时间和频率,通常由熟练的临床医生监测受试者以确定治疗何时提供治疗益处,并确定是否增加或减少剂量、增加或减少给药频率、停止治疗、恢复治疗或对治疗方案进行其他修改。给药方式可以从每周一次到每天一次,具体取决于许多临床因素,例如受试者对多肽的敏感性。可以每天或每第三、第四、第五或第六天施用所需剂量。所需剂量可以一次施用或分成亚剂量,例如2-4个亚剂量,并在一段时间内施用,例如在一天中以适当的间隔或其他适当的方案施用。此类亚剂量可以作为单位剂型施用。在一种实施方式中,施用可以是长期的,例如一次或在数周或数月内每天施用一剂量更多剂量。给药方案的实例可包括在1周、2周、3周、4周、1个月、2个月、3个月、4个月、5个月或6个月或更长时间的时间内每天施用、每天两次、每天三次或每天四次或更多次施用。

[0194] 在一种实施方式中,提供了制备针对疾病或病症的疫苗或治疗组合物的方法。该方法可包括提供本发明所述的疫苗或治疗组合物并将其配制为药学上可接受的组合物。疫苗组合物可包含可引起针对疾病的免疫应答的,有效量的异源免疫原性蛋白和/或编码该蛋白的核酸分子。

[0195] 在一种实施方式中,疾病可以由病毒病原体引起。病原体可以是病毒病原体。病毒病原体可以是腺病毒、鼻病毒、轮状病毒、西尼罗病毒、寨卡病毒、疱疹病毒或冠状病毒(CoV)。冠状病毒可以为MERS CoV、SARS-CoV和SARS-CoV-2病毒中的一种。

[0196] 在一种实施方式中,疾病可以是口蹄疫(FMD)。FMD是家养和野生偶蹄动物(包括家养和野生牛和猪)的急性且高度传染性的病毒性疾病。示例性的易感动物包括牛、水牛、绵羊、山羊、猪、羚羊、鹿和野牛、刺猬和大象;美洲驼和羊驼。在实验室实验中,小鼠、大鼠和鸡已通过人工手段成功感染,但据信它们在自然条件下不会感染该疾病。人类很少被感染。FMD的症状包括持续大约两到六天的高烧,然后口腔内和脚上出现水泡,这些水泡可能会破裂并导致跛行。口蹄疫对畜牧业具有严重影响,因为它具有高度传染性,受感染的动物可以通过气溶胶、通过接触受污染的农业设备、车辆、织物或饲料以及家养和野生捕食者传播。特别是由口蹄疫病毒感染引起的牛和猪的症状。口蹄疫爆发造成的经济损失是所有牲畜疾病中最高的。病原体口蹄疫病毒(FMDV)是小核糖核酸病毒科的家族成员,因此,病毒病原体可以是小核糖核酸病毒。如本发明所用,术语“小核糖核酸病毒”是指一种非包膜病毒,代表一个具有二十面体衣壳的小细胞质正链RNA(7.0-8.5kb)的病毒家族。它是一个大型病毒家族,包括多个属,含有重要的人类和动物病原体,例如肠道病毒(人类脊髓灰质炎病毒)、肝病毒(甲型肝炎病毒)和以口蹄疫为原型成员的口疮病毒(aphthoviruses)。FMDV存在七种不同的血清型(欧亚血清型A、O、C以及亚洲和南非领土[SAT]血清型SAT1、SAT2和SAT3)和反映显着遗传变异性的多个亚型。该科内的属包括口疮病毒属(Aphthovims)、水病毒属

(Aquamavirus)、禽肝病毒属(Aquamavirus)、心脏病毒属(Cardiovirus)、科萨维病毒属(Cosavirus)、二分病毒属(Dicipivirus)、肠道病毒属(Enterovirus)、埃博病毒属(Erbovirus)、肝病毒属(Hepatovirus)、柯布病毒属(Kobuvirus)、巨型病毒属(Megrivirus)、帕雷奇病毒属(Parechovirus)、鱼病毒属(Piscevirus)、唾液病毒属(Salivirus)、沙佩洛病毒属(Sapelovirus)、塞内卡病毒属(Senecavirus)、特沙病毒属(Teschovirus)和特雷莫病毒属(Tremovirus)。小核糖核酸病毒的病毒基因组通常含有一个开放阅读框,其编码包含结构蛋白区域P1和非结构蛋白区域P2和P3的单一多蛋白。成熟和功能性蛋白质从多蛋白中的释放主要由包括3C蛋白酶在内的病毒蛋白酶介导。

[0197] 提供疫苗组合物和免疫方案

[0198] 本发明描述的“疫苗”可以作为药物组合物提供。组合物可以包含药学上可接受的稀释剂、载体或赋形剂。该组合物可以包含进一步的活性成分。可以以单剂量给药或在初免-加强组合给药(prime-boost setting)中提供施用。初免-加强组合给药可以涉及“初免”和“加强”免疫方案,其中初免疫苗诱导的免疫反应可以通过加强疫苗加强。例如,在用编码第一异源蛋白的任意一种多核苷酸引发(至少一次)后,可以施用包含有效量的第二异源抗原蛋白或其亚基的加强疫苗以加强初免的免疫应答。

[0199] 本发明描述的组合物可以施用于受试者,例如人类受试者。用于单次施用的组合物中的异源免疫原性蛋白的总剂量可以是例如约0.01 μ g至约10mg,例如1 μ g-1mg,例如10 μ g-100 μ g。确定推荐剂量可以通过实验进行,并且对于本领域技术人员来说是常规的。

[0200] 术语“受试者”和“个体”在本文中可互换使用,并且是指人类或动物。通常,动物是脊椎动物,例如灵长类动物、啮齿类动物、家养或猎用动物。灵长类动物包括黑猩猩、食蟹猴(cynomologous monkeys)、蜘蛛猴和猕猴,例如恒河猴。啮齿动物包括小鼠、大鼠、土拨鼠、雪貂、兔子和仓鼠。家养和猎用动物包括牛、马、猪、鹿、野牛、水牛、猫科动物,例如家猫,犬科动物,例如狗、狐狸、狼、禽类,例如鸡、鹌鹑、鸵鸟和鱼,例如鳟鱼、鲑鱼和鲑鱼。患者或受试者包括上述的任何子集,例如上述所有,但不包括一个或多个群体或物种,例如人类、灵长类动物或啮齿类动物。在一种实施方式中,受试者可以是哺乳动物,例如灵长类动物,例如人类。术语“患者”和“受试者”在本发明中可互换使用。术语“患者”和“受试者”在本文中可互换使用。优选地,受试者是哺乳动物。哺乳动物可以是人、非人灵长类动物、小鼠、大鼠、狗、猫、马或牛,但不限于这些实例。

[0201] 下面的清单包括本发明的特定实施方式。但该清单不是限制性的并且不排除替代实施方式或本文另外描述的实施方式。以下实施方式清单中描述的百分比同一性是指所列举的序列沿着参考序列的全长的同一性。

[0202] 治疗癌症的方法

[0203] 一种实施方式提供了治疗癌症的方法。该方法可以包括向有需要的受试者施用治疗有效量的包含本发明描述的任何一种免疫原性蛋白的组合物。免疫原性蛋白可包含刺激先天免疫应答的模式识别受体或效应蛋白。免疫原性蛋白可以是细胞外信号分子,例如细胞因子、膜表面受体、膜通道、或整联蛋白或影响免疫细胞识别的其他糖蛋白。免疫原性蛋白质可包含STING蛋白。干扰素基因刺激剂(STING)在感染和癌症期间的先天免疫中发挥着核心作用。STING被2',3'-环-GMP-AMP(cGAMP)内源激活,cGAMP是一种由cGAMP合酶(cGAS)响应胞质DNA作为危险信号而合成的环状二核苷酸。STING的激活介导多方面的I型干扰素

(IFN- γ) 反应,其促进树突状细胞的成熟和迁移,并启动细胞毒性T淋巴细胞和自然杀伤细胞进行自发免疫反应。

[0204] 在一种实施方式中,用于治疗癌症的方法可以包括向有需要的受试者施用治疗有效量的本发明所述的任意一种免疫原性蛋白。

[0205] 在一种实施方式中,该方法还可以包括分析肿瘤生长的抑制。分析步骤可以包括观察受试者中肿瘤生长的超过约60%、70%、80%或约90%的抑制。

[0206] 多种已知的控释或缓释剂型、制剂和装置可适合与本发明公开的免疫原性组合物一起使用。实例包括但不限于美国专利Nos.3,845,770;3,916,899;3,536,809;3,598,123;4,008,719;5674,533;5,059,595;5,591,767;5,120,548;5,073,543;5,639,476;5,354,556;5,733,566和6,365,185 B1中描述的那些,所有这些都通过引用并入本文,就如已完全阐述。这些剂型可用于通过使用例如羟丙基甲基纤维素、其他聚合物基质、凝胶、渗透膜、渗透系统(例如**OROS**[®] (Alza Corporation, Mountain View, Calif. USA)),或其组合,以提供一种或多种活性成分的缓慢或受控释放,

[0207] 如本发明所用,术语“癌症”是指可能干扰身体器官和系统的正常功能的细胞不受控制的生长。癌症可以是原发性癌症,或转移性癌症,或两者兼有。从原来的位置迁移并播散到重要器官的癌症最终会因受影响器官的功能恶化而导致受试者死亡。转移是指与原发肿瘤位置不同的癌细胞或癌细胞群,是由于癌细胞从原发肿瘤扩散到身体其他部位而产生的结果。在诊断出原发性肿瘤块时,可以监测受试者是否存在转移灶,例如扩散过程中的癌细胞。

[0208] 如本发明所用,术语“癌症”还包括但不限于实体瘤和血源性肿瘤。术语癌症是指皮肤、组织、器官、骨、软骨、血液和血管的疾病。术语“癌症”包括恶性上皮肿瘤(来自肺、胃、十二指肠、结肠、直肠、乳腺、子宫、前列腺、膀胱等)。

[0209] 术语“癌症”进一步涵盖原发性癌症和转移性癌症。可以用本发明的方法治疗的癌症的实例包括但不限于实体瘤;黑色素瘤、脑癌,包括但不限于胶质瘤、胶质母细胞瘤、多形性胶质母细胞瘤 (GBM)、少突胶质细胞瘤、原始神经外胚层肿瘤、低、中和高级别星形细胞瘤、室管膜瘤(例如粘液乳头状室管膜瘤、乳头状室管膜瘤、室管膜下瘤、间变性室管膜瘤)、少突胶质细胞瘤、髓母细胞瘤、脑膜瘤、垂体腺瘤、神经母细胞瘤和颅咽管瘤;乳腺癌,包括但不限于导管原位癌、浸润(invasive)(或浸润性(infiltrating))导管癌、浸润(或浸润性)小叶癌、腺样囊性(或腺囊性)癌、低度腺鳞癌、髓样癌、粘液性(或胶质)癌、乳头状癌、小管癌、炎性乳腺癌、乳头佩吉特病、叶状瘤、三阴性乳腺癌、转移性乳腺癌;恶性上皮癌(carcinoma),包括在膀胱、乳腺、结肠、肾脏、肺、卵巢、胰腺、胃、子宫颈、甲状腺的恶性上皮癌,以及皮肤,包括鳞状细胞癌;其他肿瘤包括精原细胞瘤、四子细胞癌(tetratocarcinoma);中枢和周围神经系统肿瘤;以及其他肿瘤,包括但不限于异皮瘤、色素瘤、角化棘皮瘤、甲状腺滤泡癌和畸胎瘤。

[0210] 本发明公开的方法可用于治疗先前已接受过癌症治疗的患者以及先前未接受过癌症治疗的患者。事实上,本发明描述的方法和组合物可用于一线和二线癌症治疗。

[0211] 如本发明所用,术语“癌前病症”具有其普通含义,即无转移的不受调节的生长,并且包括各种形式的增生和良性肥大。因此,“癌前病症”是一种疾病、综合征或如果不治疗,可能导致癌症的发现物。这是一种与癌症风险显着增加相关的全身状态。恶性肿瘤前病变

是一种形态发生改变的组织,其中比表面正常的组织更容易发生癌症。恶性肿瘤前病变的实例包括但不限于口腔白斑、光化性角化病(日光性角化病)、巴雷特食管、萎缩性胃炎、前列腺良性增生、结肠或直肠癌前息肉、胃上皮发育不良、腺瘤性发育不良、遗传性非息肉病结肠癌综合征(HNPCC)、巴雷特食管、膀胱发育不良、宫颈癌前病变和宫颈发育不良。

[0212] 在一种实施方式中,癌症可以选自由以下组成的组中:乳腺癌;卵巢癌;脑癌;胃肠道癌;前列腺癌;恶性上皮癌、肺恶性上皮癌、肝细胞恶性上皮癌、睾丸癌;宫颈癌;子宫内膜癌;膀胱癌;头颈癌;肺癌;胃食管癌和妇科癌症。

[0213] 在一种实施方式中,癌症可以是乳腺癌,包括但不限于导管原位癌、浸润(或浸润性)导管癌、浸润(或浸润性)小叶癌、腺样囊性(或腺囊性)癌、低度腺鳞癌、髓样癌、粘液性(或胶质)癌、乳头状癌、小管癌、炎性乳腺癌、乳头佩吉特病、叶状瘤、三阴性乳腺癌、转移性乳腺癌。

[0214] 在一种实施方式中,本发明描述的方法可以涉及治疗患有或诊断为患有癌症的受试者。患有癌症的受试者可以由医生使用当前诊断癌症的方法来鉴定。表征这些病症并有助于诊断的癌症症状和/或并发症是本领域众所周知的,并且可以是但不限于肿瘤生长、含有癌细胞的器官或组织的功能受损等。可能有助于诊断的测试,例如癌症包括但不限于组织活检和组织学检查。癌症家族史或接触癌症危险因素(例如烟草产品、辐射等)也可能有助于确定受试者是否可能患有癌症或做出癌症诊断。

[0215] 下面的清单包括本发明的特定实施方式。但该清单不是限制性的并且不排除替代实施方式或本文另外描述的实施方式。以下实施方式清单中描述的百分比同一性是指所列举的序列沿着参考序列的全长的同一性。

[0216] 实施方式

[0217] 1、一种合成的甲病毒衍生复制子核酸分子,其包含:(i)第一核酸,其编码甲病毒非结构蛋白nsP1、nsP2、nsP3和nsP4,并且包含在如SEQ ID NO:17所示的甲病毒基因组序列中从nt 503至nt 658、nt 658至nt 1620、nt 1620至nt 2560、nt 2560至nt 3954、nt 3954至nt 4120、nt 6381至nt 7083和nt 6966至nt 7526的区域内的任意位置处引入的至少一个沉默突变;以及(ii)第二核酸,其包含经过修饰的亚基因组开放阅读框(ORF)。

[0218] 2、实施方式1所述的合成的甲病毒衍生复制子核酸分子,其中,经过修饰的亚基因组ORF包含编码第一异源蛋白的序列。

[0219] 3、实施方式1和2中任意一个或两个所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 503至nt 658的区域内的任意位置处引入的至少一个沉默突变。

[0220] 4、实施方式1-3中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:21所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0221] 5、实施方式1-4中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含在甲病毒基因组中从nt 658至nt 1620的区域内的任意位置处引入的至少一个沉默突变。

[0222] 6、实施方式1-5中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:25所示的参考序列具有至少90%同一

性的序列的多核苷酸。

[0223] 7、实施方式1-6中任意一个或多个所述的合成的甲病毒衍生的复制子核酸分子，其中，第一核酸包含在甲病毒基因组中从nt 1620至nt 2560的区域内的任意位置处引入的至少一个沉默突变。

[0224] 8、实施方式1-7中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:24所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0225] 9、实施方式1-8中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含在甲病毒基因组中从nt 3694至nt 3954的区域内的任意位置处引入的至少一个沉默突变。

[0226] 10、实施方式1-9中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:22所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0227] 11、实施方式1-10中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含在甲病毒基因组中从nt 2560至nt 3954的区域内的任意位置处引入的至少一个沉默突变。

[0228] 12、实施方式1-11中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:26所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0229] 13、实施方式1-12中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含在甲病毒基因组中从nt 3954至nt 4120的区域内的任意位置处引入的至少一个沉默突变。

[0230] 14、实施方式1-13中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含编码nsP2和nsP3的接合物并且含有与如SEQ ID NO:23所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0231] 15、实施方式1-15中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含在甲病毒基因组中从nt 6381至nt 7083的区域内的任意位置处引入的至少一个沉默突变。

[0232] 16、实施方式1-15中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含编码nsP4并且含有与如SEQ ID NO:20所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0233] 17、实施方式1-16中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含在甲病毒基因组中从nt 658至nt 1620和nt 2560至nt 3954的区域内的任意位点处引入的至少一个沉默突变。

[0234] 18、实施方式1-17中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，其中，第一核酸包含编码nsP1并且含有与如SEQ ID NO:25所示的参考序列具有至少90%同一性的序列的多核苷酸，以及编码nsP2并且含有与如SEQ ID NO:26所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0235] 19、实施方式1-18中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子，

其中,第一核酸包含在甲病毒基因组中从nt 3694至nt 3954和nt 6381至nt 7083的区域内的任意位置处引入的至少一个沉默突变。

[0236] 20、实施方式1-19中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP2并且含有与如SEQ ID NO:22所示的参考序列具有至少90%同一性的序列的多核苷酸,以及编码nsP4并且含有与如SEQ ID NO:20所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0237] 21、实施方式1-20中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:21所示的参考序列具有至少90%同一性的序列的多核苷酸;编码nsP1并且含有与如SEQ ID NO:25所示的参考序列具有至少90%同一性的序列的多核苷酸,编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:24所示的参考序列具有至少90%同一性的序列的多核苷酸;编码nsP2并且含有与如SEQ ID NO:26所示的参考序列具有至少90%同一性的序列的多核苷酸;编码nsP2和nsP3的接合物并且含有与如SEQ ID NO:23所示的参考序列具有至少90%同一性的序列的多核苷酸;以及编码nsP4并且含有与如SEQ ID NO:20所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0238] 22、实施方式1-21中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码所述非结构蛋白并且含有与如SEQ ID NO:30所示的参考序列具有至少90%同一性的多核苷酸序列的多核苷酸。

[0239] 23、实施方式1-22中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,第一核酸分子包含在甲病毒基因组的从nt 6966到nt 7523的区域内的任何位置处引入的至少一个沉默突变。

[0240] 24实施方式1-23中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,第一核酸分子包含与如SEQ ID NO:18所示的参考序列具有至少90%同一性的多核苷酸序列。

[0241] 25、实施方式1-24中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含引入编码nsP4蛋白的多核苷酸的至少一个沉默突变。

[0242] 26、实施方式1-25中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,第一核酸分子包含与如SEQ ID NO:19所示参考序列具有至少90%同一性的多核苷酸序列。

[0243] 27、实施方式1-26中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一核酸包含编码经过修改的nsP4的多核苷酸,所述经过修改的nsP4包含第二异源蛋白。

[0244] 28、实施方式1-27中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,编码经过修改的nsP4的多核苷酸包含编码第一C末端区域的多核苷酸序列和编码第二C末端区域的多核苷酸序列。

[0245] 29、实施方式1-28中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第二C末端区域可翻译成蛋白质。

[0246] 30、实施方式1-28中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第二C末端区域不可翻译成蛋白质。

[0247] 31、实施方式1-28中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,编码第一C末端区域的多核苷酸序列不同于编码第二C末端区域的多核苷酸序列。

[0248] 32、实施方式1-31中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,

其中,第二异源蛋白包含GFP蛋白,并且经过修改的nsP4包含与如SEQ ID NO:41所示的参考序列具有至少90%同一性的氨基酸序列。

[0249] 33、实施方式1-32中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,经过修改的nsP4由包含与如SEQ ID NO:31所示的参考序列具有至少90%同一性的核苷酸序列编码。

[0250] 34、实施方式1-33中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一异源蛋白或第二异源蛋白包含由病毒衍生的或由哺乳动物细胞衍生的细胞代谢、免疫或信号传导功能的调节剂。

[0251] 35、实施方式1-34中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一异源蛋白或第二异源蛋白影响细胞抗病毒先天免疫应答并增强甲病毒复制子RNA分子的复制。

[0252] 36、实施方式1-35中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一异源蛋白或第二异源蛋白包含自病毒病原体分离或衍生的抗原蛋白。

[0253] 37、实施方式1-36中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,病毒病原体选自:小核糖核酸病毒、腺病毒、鼻病毒、轮状病毒、西尼罗病毒、寨卡病毒、疱疹病毒或冠状病毒(CoV)组成的组中。

[0254] 38、实施方式1-37中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,病毒病原体为小核糖核酸病毒。

[0255] 39、实施例1-35中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第二异源蛋白包含3C蛋白酶蛋白,并且经过修改的nsP4包含与如SEQ ID NO:40所示的参考序列具有至少90%同一性的氨基酸序列。

[0256] 40、实施方式1-35和39中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,经过修改的nsP4由与如SEQ ID NO:32所示的参考序列具有至少90%同一性的多核苷酸序列编码。

[0257] 41、实施方式1-35和39-40中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一异源蛋白包含O1 Manisa P1蛋白。

[0258] 42、实施方式1-35和39-41中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一异源蛋白包含O1 Manisa P1蛋白并且第二异源蛋白包含3C蛋白酶。

[0259] 43、实施方式1-35中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第二异源蛋白包含STING蛋白,并且所述经过修改的nsP4包含与如SEQ ID NO:42所示的参考序列具有至少90%同一性的氨基酸序列。

[0260] 44、实施方式1-35和44中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子,其中,第一异源蛋白包含抗原蛋白并且第二异源蛋白包含STING蛋白。

[0261] 45、一种核酸,其包含编码合成的甲病毒衍生复制子核酸分子的DNA序列,其包含:(i) 第一核酸,其编码甲病毒非结构蛋白nsP1、nsP2、nsP3和nsP4,并且包含在如SEQ ID NO:17所示的甲病毒基因组序列中从nt 503至nt 658、nt 658至nt 1620、nt 1620至nt 2560、nt 2560至nt 3954、nt 3954至nt 4120、nt 6381至nt 7083和nt 6966至nt 7526的区域内的任意位置处引入的至少一个沉默突变;以及(ii) 第二核酸,其包含经过修饰的亚基因组开放阅读框(ORF)。

[0262] 46、实施方式45所述的核酸,其中,经过修饰的亚基因组ORF包含编码第一异源蛋白的序列。

[0263] 47、实施方式45和46中任意一个或多个所述的核酸,其中,第一核酸包含在甲病毒基因组中从nt 503至nt 658的区域内的任意位置处引入的至少一个沉默突变。

[0264] 48、实施方式45-47中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:2所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0265] 49、实施方式45-48中任意一个或多个所述的核酸,其中,第一核酸包含在甲病毒基因组中从nt 658至nt 1620的区域内的任意位置处引入的至少一个沉默突变。

[0266] 50、实施方式45-49中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:6所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0267] 51、实施方式45-50中任意一个或多个所述的核酸,其中,第一核酸包含在甲病毒基因组中从nt 1620至nt 2560的区域内的任意位置处引入的至少一个沉默突变。

[0268] 52、实施方式45-51中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:5所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0269] 53、实施方式45-52中任意一个或多个所述的核酸,其中,第一核酸包含在甲病毒基因组中从nt 3694至nt 3954的区域内的任意位置处引入的至少一个沉默突变。

[0270] 54、实施方式45-53中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:3所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0271] 55、实施方式45-54中任意一个或多个所述的核酸,其中,第一核酸包含在甲病毒基因组中从nt 2560至nt 3954的区域内的任意位置处引入的至少一个沉默突变。

[0272] 56、实施方式45-55中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:7所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0273] 57、实施方式45-57中任意一个或多个所述的核酸,其中,第一核酸包含在甲病毒基因组中从nt 3954至nt 4120的区域内的任意位置处引入的至少一个沉默突变。

[0274] 58、实施方式45-57中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP2和nsP3的接合物并且含有与如SEQ ID NO:4所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0275] 59、实施方式45-58中任意一个或多个所述的核酸,其中,第一核酸包含在甲病毒基因组中从nt 6381至nt 7083的区域内的任意位置处引入的至少一个沉默突变。

[0276] 60、实施方式45-59中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP4并且含有与如SEQ ID NO:1所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0277] 61、实施方式45-60中任意一个或多个所述的核酸,其中,第一核酸包含在甲病毒基因组中从nt 658至nt 1620和nt 2560至nt 3954的区域内的任意位点处引入的至少一个沉默突变。

[0278] 62、实施方式45-61中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:6所示的参考序列具有至少90%同一性的序列的多核苷酸,以及

编码nsP2并且含有与如SEQ ID NO:7所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0279] 63、实施方式45-63中任意一个或多个所述的核酸,其中,第一核酸包含在甲病毒基因组中从nt 3694至nt 3954和nt 6381至nt 7083的区域内的任意位置处引入的至少一个沉默突变。

[0280] 64、实施方式45-63中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP2并且含有与如SEQ ID NO:3所示的参考序列具有至少90%同一性的序列的多核苷酸,以及编码nsP4并且含有与如SEQ ID NO:1所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0281] 65、实施方式45-64中任意一个或多个所述的核酸,其中,第一核酸包含编码nsP1并且含有与如SEQ ID NO:2所示的参考序列具有至少90%同一性的序列的多核苷酸;编码nsP1并且含有与如SEQ ID NO:6所示的参考序列具有至少90%同一性的序列的多核苷酸,编码nsP1和nsP2的接合物并且含有与如SEQ ID NO:5所示的参考序列具有至少90%同一性的序列的多核苷酸;编码nsP2并且含有与如SEQ ID NO:7所示的参考序列具有至少90%同一性的序列的多核苷酸;编码nsP2和nsP3的接合物并且含有与如SEQ ID NO:4所示的参考序列具有至少90%同一性的序列的多核苷酸;以及编码nsP4并且含有与如SEQ ID NO:1所示的参考序列具有至少90%同一性的序列的多核苷酸。

[0282] 66、实施方式45-65中任意一个或多个所述的核酸,其中,第一核酸包含编码非结构蛋白并且含有与如SEQ ID NO:11所示的参考序列具有至少90%同一性的多核苷酸序列的多核苷酸。

[0283] 67、实施方式45-66中任意一个或多个所述的核酸,其中,第一核酸包含编码经过修改的nsP4的多核苷酸,所述经过修改的nsP4包含第二异源蛋白。

[0284] 68、实施方式45-67中任意一个或多个所述的核酸,其中,编码经修改的nsP4的多核苷酸包含编码第一C末端区域的多核苷酸序列和编码第二C末端区域的多核苷酸序列。

[0285] 69、实施方式45-68中任意一个或多个所述的核酸,其中,第二C末端区域包含亚基因组启动子。

[0286] 70、实施方式45-69中任意一个或多个所述的核酸,其中,亚基因组启动子由与如SEQ ID NO:14所示的参考序列具有至少90%同一性的序列编码。

[0287] 71、实施方式45-70中任意一个或多个所述的核酸,其中,第二C末端区域可翻译成蛋白质。

[0288] 72、实施方式45-70中任意一个或多个所述的核酸,其中,第二C末端区域不可翻译成蛋白质。

[0289] 73、实施方式45-70中任意一个或多个所述的核酸,其中,编码第一C末端区域的多核苷酸序列不同于编码第二C末端区域的多核苷酸序列。

[0290] 74、实施方式45-73中任意一个或多个所述的核酸,其中,第二异源蛋白包含GFP蛋白,并且经过修改的nsP4包含与如SEQ ID NO:12所示的参考序列具有至少90%同一性的氨基酸序列。

[0291] 75、实施方式45-74中任意一个或多个所述的核酸,其中,经过修改的nsP4由如SEQ ID NO:13所示的多核苷酸序列编码。

[0292] 76、实施方式45-75中任意一个或多个所述的核酸,其中,第二异源蛋白包含GFP蛋白,并且经过修改的nsP4包含与如SEQ ID NO:15所示的参考序列具有至少90%同一性的氨基酸序列。

[0293] 77、实施方式45-76中任意一个或多个所述的核酸,其中,经过修改的nsP4由如SEQ ID NO:16所示的多核苷酸序列编码。

[0294] 78、实施方式45-73中任意一个或多个所述的核酸,其中,第一异源蛋白或第二异源蛋白包含由病毒衍生的或由哺乳动物细胞衍生的细胞代谢、免疫或信号传导功能的调节剂。

[0295] 79、实施方式45X73中任意一个或多个所述的核酸,其中,第一异源蛋白或第二异源蛋白影响细胞抗病毒先天免疫应答并增强甲病毒复制子RNA分子的复制。

[0296] 80、实施方式45-73中任意一个或多个所述的核酸,其中,第一异源蛋白或第二异源蛋白包含自病毒病原体分离或衍生的抗原蛋白。

[0297] 81、实施方式45-73和80中任意一个或多个所述的核酸,其中,病毒病原体选自由:小核糖核酸病毒、腺病毒、鼻病毒、轮状病毒、西尼罗病毒、寨卡病毒、疱疹病毒或冠状病毒(CoV)组成的组中。

[0298] 82、实施方式45-73和80-81中任意一个或多个所述的核酸,其中,病毒病原体为小核糖核酸病毒。

[0299] 83、实施方式45-73和80-82中任意一个或多个所述的核酸,其中,第一异源蛋白包含O1 ManisaP1蛋白,和/或第二异源蛋白包含3C蛋白酶。

[0300] 84、实施方式45-73中任意一个或多个所述的核酸,其中,第一异源蛋白和/或第二异源蛋白包含STING蛋白并且第一异源蛋白质包含抗原蛋白。

[0301] 85、一种甲病毒衍生RNA复制子表达系统,其包含实施方式1-44中任意一个或多个所述的合成的甲病毒衍生复制子核酸分子和宿主细胞。

[0302] 86、一种药学上可接受的组合物,其包含实施方式1-44中任意一个或多个所述的甲病毒衍生复制子和药学上可接受的载体或赋形剂。

[0303] 87、一种在细胞中产生至少一种异源蛋白的方法,包括在宿主细胞中表达实施方式1-44中任意一个或多个所述的合成甲病毒衍生的复制子核酸分子。

[0304] 88、一种预防、抑制或治疗受试者的疾病或病症的症状的方法,包括:提供合成的甲病毒属衍生的复制子核酸分子,其包含:

[0305] (i) 第一核酸,其编码甲病毒非结构蛋白nsP1、nsP2、nsP3和nsP4,并且包含在如SEQ ID NO:17所示的甲病毒基因组序列中从nt 503至nt 658、nt 658至nt 1620、nt 1620至nt 2560、nt 2560至nt 3954、nt 3954至nt 4120、nt 6381至nt 7083和nt 6966至nt 7526的区域内的任意位置处引入的至少一个沉默突变;以及

[0306] (ii) 第二核酸,其包含经过修饰的亚基因组开放阅读框(ORF);以及

[0307] 施用治疗有效量的合成的甲病毒衍生复制子核酸分子予以受试者。

[0308] 89、实施方式88所述的方法,其中,经过修饰的亚基因组ORF包含编码第一异源蛋白的序列。

[0309] 90、实施方式88-89中任意一个或两个所述的方法,其中,第一核酸包含与选自由SEQ ID NOS:20-30组成的组中的参考序列具有至少90%同一性的至少一个序列。

[0310] 91、实施方式88-90中任意一个或多个所述的方法,其中,第一核酸包含编码经过修改的nsP4的多核苷酸,所述经过修改的nsP4包含第二异源蛋白。

[0311] 92、实施方式88-91中任意一个或多个所述的方法,其中,编码经修改的nsP4的多核苷酸包含编码第一C末端区域的多核苷酸序列和编码第二C末端区域的多核苷酸序列。

[0312] 93、实施方式88-92中任意一个或多个所述的方法,其中,第二C末端区域可翻译成蛋白质。

[0313] 94、实施方式88-92中任意一个或多个所述的方法,其中,第二C末端区域不可翻译成蛋白质。

[0314] 95、实施方式88-92中任意一个或多个所述的方法,其中,编码第一C末端区域的多核苷酸序列不同于编码第二C末端区域的多核苷酸序列。

[0315] 96、实施方式88-95中任意一个或多个所述的方法,其中,第一异源蛋白或第二异源蛋白包含由病毒衍生的或由哺乳动物细胞衍生的细胞代谢、免疫或信号传导功能的调节剂。

[0316] 97、实施方式88-95中任意一个或多个所述的方法,其中,第一异源蛋白或第二异源蛋白影响细胞抗病毒先天免疫应答并增强甲病毒复制子RNA分子的复制。

[0317] 98、实施方式88-95中任意一个或多个所述的方法,其中,第一异源蛋白或第二异源蛋白包含自病毒病原体分离或衍生的抗原蛋白。

[0318] 99、实施方式88-95和98中任意一个或多个所述的方法,其中,病毒病原体选自由:小核糖核酸病毒、腺病毒、鼻病毒、轮状病毒、西尼罗病毒、寨卡病毒、疱疹病毒或冠状病毒(CoV)组成的组中。

[0319] 100、实施方式88-95和98-99中任意一个或多个所述的方法,其中,病毒病原体为小核糖核酸病毒。

[0320] 101、实施方式88-95和98-100中任意一个或多个所述的方法,其中,第一异源蛋白包含O1 ManisaP1蛋白并且第二异源蛋白包含3C蛋白酶。

[0321] 102、实施方式88-95中任意一个或多个所述的方法,其中,第一异源蛋白或第二异源蛋白包含STING蛋白。

[0322] 103、实施方式88-95和102中任意一个或多个所述的方法,其中,第二异源蛋白包含STING蛋白并且第一异源蛋白包含抗原蛋白。

[0323] 104、实施方式88-104中任意一个或多个所述的方法,其中,施用步骤使得受试者的疾病得到预防、治疗、严重性减轻或病程发展减缓。

[0324] 105、实施方式88-95和103-104中任意一个或多个所述的方法,其中,疾病为癌症。

[0325] 106、实施方式88-95和103-105中任意一个或多个所述的方法,其中,癌症为黑色素瘤或上皮性肿瘤。

[0326] 107、实施方式88-95和103-106中任意一个或多个所述的方法,其中,该方法进一步包括向受试者共同施用一种或多种另外的抗癌疗法。

[0327] 108、实施方式88-95和103-107中任意一个或多个所述的方法,其中,另外的抗癌疗法选自由外科手术、化学疗法、放射疗法、温热疗法、免疫疗法、激素疗法、激光疗法、抗血管生成疗法及其任意组合组成的组中。

[0328] 109、实施方式88-95和98-101中任意一个或多个所述的方法,其中,疾病为FMD。

[0329] 110、实施方式88-109中任意一个或多个所述的方法,其中,受试者为哺乳动物。

[0330] 111、实施方式88-110中任意一个或多个所述的方法,其中,哺乳动物为偶蹄动物或人。

[0331] 112、实施方式88-111中任意一个或多个所述的方法,其中,偶蹄动物选自:牛、水牛、绵羊、山羊、猪、羚羊、鹿、野牛、刺猬、大象、美洲驼和羊驼组成的组中。

[0332] 本发明的进一步的实施方式可以通过用来自本发明的任何一个或多个其他实施方式中的一种或多种要素补充实施方式和/或用来自一个或多个其他实施方式中的一种或多种要素替换来自一个实施方式的一种或多种要素来形成。

[0333] 实施例

[0334] 提供以下非限制性实施方式来说明特定的实施方式。贯穿全文的实施方式可以用来自下面的一个或多个实施例中的一个或多个细节来补充,和/或来自一个实施方式中的一个或多个元件可以用来自下面的一个或多个实施例中的一个或多个细节替代。

[0335] 实施例1 nsP4修饰复制子的设计

[0336] 图1说明了甲病毒基因组非结构多蛋白(nsP;nsP1-nsP2-nsP3-nsP4)的基因表达步骤。此图表明,非结构多蛋白的表达是基因组合成和随后亚基因组ORF(SG ORF)表达的先决条件,在天然病毒中SG ORF编码结构多蛋白。在图中,显示出nsP包括野生型VEEV中存在的早期终止密码子,该密码子导致nsP-1-2-3多蛋白的翻译终止;nsP4通过基因组5682位处的琥珀终止密码子(UGA)的低频翻译通读进行翻译,其结合了精氨酸,半胱氨酸或色氨酸。随后对多蛋白进行蛋白水解,产生分离的nsP1、nsP2、nsP3和nsP4蛋白。

[0337] 病毒感染的“早期阶段”是指在发生基因组复制之前,主要以nsP多蛋白的表达为特征的感染后时间。“晚期”是由基因组复制发生后的基因表达模式定义的,亚基因组多蛋白的表达是主要的病毒基因产物。

[0338] 本发明报道了成功编辑甲病毒复制子的nsP基因区域,以降低与野生型病毒的同源性,并将新的生物学功能纳入此类人工复制子的自我扩增(在亚基因组编码的转基因翻译之前)的早期阶段。这是通过识别nsP区域中的关键核苷酸片段来实现的,这些核苷酸片段可以安全地进行编辑而不破坏保守的结构元件,并将额外的序列附加到nsP的天然末端。本发明显示了如何在不损害复制子的自扩增功能的情况下进行此类修饰。因此,本发明提供了一种产生更安全的复制子的新方法,所述复制子配备有额外的功能以修饰宿主细胞的细胞和免疫学参数。

[0339] 甲病毒复制子目前主要用作疫苗载体,其中目标抗原在亚基因组ORF中编码,从而取代通常组成病毒颗粒的天然甲病毒结构蛋白。目前,如果需要以额外多肽(例如报告基因、免疫调节因子或其他宿主细胞功能调节因子)的形式提供辅助生物学功能,则辅助因子必须(i)通过编码序列的融合在与第一个转基因相同的亚基因组ORF中编码,(ii)在单独的RNA转录/翻译启动子(例如复制的亚基因组启动子(SGP)或内部核糖体进入位点(IRES)的控制下编码,或(iii)通过将额外的核酸或蛋白质种类与复制子RNA产物混合来从复制子中分离。方法(i)需要对功能性多肽进行工程设计,而对于某些抗原而言,在不改变其主要序列的情况下这并不总是可能的。方法(ii)除了所需的辅助ORF之外,还需要添加大量额外的核苷酸序列以产生SGP或IRES。方法(i)和使用额外的SGP作为(ii)中的方法仅允许辅助因子在病毒复制周期的后期与亚基因组ORF同时表达,从而使其无法在早期影响复制子功能。

方法(iii)需要生产完全额外的生物制品,并且必须通过某种方式解决与复制子RNA的共同递送。目前,使用的复制子nsP序列与流行的地方性甲病毒基本相同,在与另一个密切相关的甲病毒(例如流行的致病性野生型病毒菌株)基因组共同感染的情况下,会增加重组的风险。这种重组有可能产生有生命力的复制病毒,从而造成环境安全问题。与野生型天然病毒株同源性降低的复制子将具有安全性提高的优点。

[0340] 通过特异性编辑甲病毒基因组的天然编码序列,故意保留关键的一级和二级结构元件完整,解决了(a)与天然病毒的序列同源性和(b)包括辅助早期生物功能的方法的挑战。通过在编辑之前分析区域以检测预计携带特定功能的保守一级和二级序列元件(例如,如图6所示),可以修改nsP编码区的天然密码子以产生沉默突变,从而产生相同的蛋白质产物并且不损害复制子的功能。扩展该技术,可以编辑nsP多蛋白的C末端,以通过融合到nsP编码序列中允许附加新的转基因x,而不损害复制子的功能。对nsP ORF 3'区域的必要非编码元件进行复制,以重新创建SGP并确保SG ORF表达,并对上游副本进行沉默密码子修改,以创建新的nsP-x融合蛋白并消除重复位点之间的假性同源性,可以抑制标准细菌生产菌株中此类构建体的DNA模板的增殖。

[0341] 图2示出了甲病毒基因组非结构多蛋白的基因表达步骤,该多蛋白通过与nsP4(nsP1-nsP2-nsP3-nsP4-xP)框内的附属物人工修饰以表达另外的外源蛋白(“xP”)。该图显示,外源蛋白xP被引入编码nsP多蛋白的nsP4片段的基因,导致nsP4和xP蛋白通过接合点连接在一起进行翻译和蛋白水解。nsP4-xP接合点可能包含不可裂解的、可自我裂解的(例如,2A肽)或蛋白裂解的(例如,nsP2蛋白酶或弗林蛋白酶识别位点)肽接头序列。

[0342] 图3示出了可用于对甲病毒复制子RNA序列进行遗传修饰以产生如图2中所描述的构建体的步骤的示意图。如图3所示,所需的插入位置在亚基因组启动子(SGP)区域内,步骤1涉及SGP区域的复制。步骤2涉及在重复区域插入编码xP蛋白的序列。由于SGP的复制,该方法导致近端相同序列,从而使构建体在遗传上不稳定。这个问题的解决方式是在上游重复元件中实施沉默密码子修改,从而消除有问题的同源性。

[0343] 图4示出了克隆可修饰的nsP4开放阅读框(ORF),同时保持完整的延伸亚基因组启动子的最佳步骤。如图4所示,步骤1涉及使用不同的密码子优化选择,将合成的nsP4末端区域插入SGP区域(灰色框),并加入末端限制位点。步骤2涉及在合成的nsP4末端区域和原始nsP4 C末端编码序列(非翻译编码序列CDS,条纹框)之间插入编码xP蛋白的序列。该方法不会产生任何近端相同的序列延伸,并产生遗传稳定的构建体。这种方法消除了有问题的重复序列元素,从而DNA模板可以很容易地在常见菌株中繁殖和制造。对nsP4的这种附加转基因x的修饰可以与本发明所述的nsP区域中的其他沉默突变结合进行,以产生进一步改进的复制子RNA。

[0344] 实施例2基因组45-260位核苷酸内nsP1同源性降低的VEEV复制子RNA修饰

[0345] 为了证明甲病毒nsP对任意核苷酸序列修改的敏感性,在对应于基因组核苷酸45-260位的nsP1蛋白的起始处替换了一系列沉默突变,如图5所示。图5是编码构建体8(C08)的复制子序列模板的DNA质粒的示意图。

[0346] 在此图中,复制子序列源自甲病毒中的委内瑞拉马脑炎病毒(VEEV)。该DNA编码氨苄青霉素抗性基因(AmpR),以促进在细菌细胞培养物中的特异性生产,编码复制起点(ori)以允许在细菌细胞中繁殖,工程化VEEV复制子的DNA模板(包含病毒的基因组5'UTR、nsPs、

亚基因组启动子和5'UTR、亚基因组ORF中的目的基因、基因组3'UTR和poly(A)尾巴)可通过从T7启动子开始的流式转录将其转录为RNA。修改基因组核苷酸45-260位(灰色框)的密码子以减少与野生型病毒的同源性。这个45-260nt编码区包含细胞内RNA复制所需的甲病毒保守序列元件(CSE)。已知复制子功能所需的CSE已通过该构建体中的沉默突变(密码子适应)在二级结构水平上被破坏,因此它作为阴性对照。总的来说,沉默突变使nsP1 CDS与野生型的序列同一性降低了19.04%。质粒从T7启动子转录产生C08复制子RNA。C08的5'UTR、nsP1-40RF和SGP如SEQ ID NO:27所示。

[0347] C08区包括公认的甲病毒保守序列元件(CSE),这是一个含有51nt的核苷酸跨度,折叠成有序的茎环。该二级结构基序是基因组复制所必需的,从而从亚基因组ORF转基因表达。BHK细胞的体外转染证实,该RNA(C08)无法介导分泌胚胎碱性磷酸酶的亚基因组编码的报告基因的表达(如图17A中所示)。

[0348] 实施例3基因组503-658位核苷酸内nsP1同源性降低的VEEV复制子RNA修饰

[0349] 与实施例1中描述的构建体C08相反,在CSE下游的区域中进行不同的nsP1 ORF修饰,其中如图6所示,基于热力学参数对局部RNA折叠进行生物信息学检查用于选择区域非结构化核苷酸,因此一系列沉默突变在整个区域被取代以产生命名为C02的构建体,如图7所示。

[0350] 图6是描绘用于选择甲病毒基因组区域的示意图,该区域可通过使用降低与野生型病毒同源性的沉默突变修改,而不降低复制子的复制功能。在此图中示出了野生型VEEV基因组(NCBI GenBank数据库中的参考序列L01442)中横跨BsiWI至EcoNI限制片段的区域的RNA二级结构,该结构是通过计算77计时的最小杂交自由能预测的。图中框出的是涉及局部(即,在大约100nt窗口内)发夹结构的多核苷酸延伸,这些结构是完整且未发生突变的,以保存基因组的总体二级结构。图中圈出了可以修改以产生沉默突变同时保留局部序列的无序性质的核酸位置的例子。

[0351] 图7是编码构建体2(C02)的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸503-658(灰色框)的密码子以降低与野生型病毒的同源性。总的来说,沉默突变使nsP1 CDS与野生型的序列同一性降低了14.09%。质粒从T7启动子转录产生C02复制子RNA。C02的5'UTR、nsP1-40RF和SGP如SEQ ID NO:21所示。

[0352] 据观察,BHK细胞的体外转染证实了C02构建体的RNA能够介导亚基因组ORF编码的报告基因(分泌型胚胎碱性磷酸酶)的表达(图17A)。

[0353] 实施例4基因组658-1620位核苷酸内nsP1同源性降低的VEEV复制子RNA修饰

[0354] 如实施例2中描述的构建体C02一样,基于热力学参数对局部RNA折叠进行生物信息学检查用于选择非结构化核苷酸区域,并且相应地在该区域中替换一系列沉默突变,以创建命名为C06的构建体。图8是编码构建体6(C06)的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸658-1620(灰色框)中的密码子以降低与野生型病毒的同源性。总的来说,沉默突变使nsP1 CDS与野生型的序列同一性降低了8.96%。质粒从T7启动子转录产生C06复制子RNA。C06的5'UTR、nsP1-40RF和SGP如SEQ ID NO:25所示。

[0355] 据观察,BHK细胞的体外转染证实了该RNA(构建体C06)能够介导亚基因组ORF编码

的报告基因(分泌型胚胎碱性磷酸酶)的表达(图17A)。注射该构建体的小鼠表现出与注射野生型VEEV-SEAP的小鼠同等或更高的平均SEAP血清表达(图17D)。

[0356] 实施例5基因组1620-2560位核苷酸内nsP1-nsP2接合物区域中同源性降低的VEEV复制子RNA修饰

[0357] 如实施例2中描述的构建体C02一样,基于热力学参数对局部RNA折叠进行生物信息学检查用于选择nsP1和nsP2接合处的非结构化核苷酸区域,并且相应地在该区域中替换一系列沉默突变,以创建命名为C05的构建体。图9是编码构建体5(C05)的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸1620-2560(灰色框)中的密码子以降低与野生型病毒的同源性。总的来说,沉默突变使nsP1 CDS与野生型的序列同一性降低了15.94%,与nsP2CDS的序列同一性降低了0.8%。

[0358] 质粒从T7启动子转录产生C05复制子RNA。C05的5' UTR、nsP1-4 ORF和SGP如SEQ ID NO:24所示。

[0359] BHK细胞的体外转染证实了该RNA(构建体C05)能够介导亚基因组ORF编码的报告基因(分泌型胚胎碱性磷酸酶)的表达(图17A)。注射该构建体的小鼠表现出与注射野生型VEEV-SEAP的小鼠同等或更高的平均SEAP血清表达(图17D)。

[0360] 实施例6基因组3694-3954位核苷酸内nsP2同源性降低的VEEV复制子RNA修饰
[211]如实施例2中描述的构建体C02一样,基于热力学参数对局部RNA折叠进行生物信息学检查用于选择非结构化核苷酸区域,并且相应地在该区域中替换一系列沉默突变,以创建命名为C03的构建体。图10是编码构建体3(C03)的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸3694-3954(灰色框)中的密码子以降低与野生型病毒的同源性。总的来说,沉默突变使nsP2 CDS与野生型的序列同一性降低了3.19%。

[0361] 质粒从T7启动子转录产生C03复制子RNA。C03的5' UTR、nsP1-4 ORF和SGP如SEQ ID NO:22所示。

[0362] BHK细胞的体外转染证实了RNA能够介导亚基因组ORF编码的报告基因(分泌型胚胎碱性磷酸酶)的表达(图17A)。据观察,注射构建体C03的小鼠表现出与注射野生型VEEV-SEAP的小鼠同等或更高的平均SEAP血清表达(图17D)。

[0363] 实施例7基因组2560-3954位核苷酸内nsP2同源性降低的VEEV复制/子RNA修饰

[0364] 鉴于实施例5中观察到的成功,同源性降低的区域在5'方向进一步延伸。再次基于基因组核苷酸2560-3954位的热力学参数对局部RNA折叠进行生物信息学检查,并且相应地用沉默突变取代非结构化核苷酸区域,以创建命名为C07的构建体。图11是编码构建体7(C07)的复制子序列模板的DNA质粒的示意图。在该图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸2560-3954(灰色框)中的密码子以降低与野生型病毒的同源性。总的来说,沉默突变使nsP2 CDS与野生型的序列同一性降低了15.97%。质粒从T7启动子转录产生C07复制子RNA。C07的5' UTR、nsP1-4 ORF和SGP如SEQ ID NO:26所示。

[0365] BHK细胞的体外转染证实了该RNA(构建体C07)能够介导亚基因组ORF编码的报告基因(分泌型胚胎碱性磷酸酶)的表达(图17A)。

[0366] 实施例8基因组3954-4120位核苷酸内nsP1-nsP2接合物区域中同源性降低的VEEV

复制子RNA修饰

[0367] 如实施例2中描述的构建体C02一样,基于热力学参数对局部RNA折叠进行生物信息学检查用于选择nsP1和nsP2接合处的非结构化核苷酸区域,并且相应地在该区域中替换一系列沉默突变,以创建命名为C04的构建体。图12是编码构建体4(C04)的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸3954-4120(灰色框)中的密码子以降低与野生型病毒的同源性。总的来说,沉默突变使nsP2 CDS与野生型的序列同一性降低了4.61%、与nsP3 CDS的序列同一性降低了0.12%。质粒从T7启动子转录产生C04复制子RNA。C04的5' UTR、nsP1-4 ORF和SGP如SEQ ID NO:23所示。

[0368] BHK细胞的体外转染证实了该RNA能够介导亚基因组ORF编码的报告基因(分泌型胚胎碱性磷酸酶)的表达(图17A)。

[0369] 实施例9基因组6381-7083位核苷酸内nsP4同源性降低的VEEV复制子RNA修饰

[0370] 如实施例2中描述的构建体C02一样,基于热力学参数对局部RNA折叠进行生物信息学检查用于选择非结构化核苷酸区域,并且相应地在该区域中替换一系列沉默突变,以创建命名为C01的构建体。图13是编码构建体1(C01)的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸6381-7083(灰色框)中的密码子以降低与野生型病毒的同源性。总的来说,沉默突变使nsP4 CDS与野生型的序列同一性降低了4.69%。质粒从T7启动子转录产生C01复制子RNA。C01的5' UTR、nsP1-4ORF和SGP如SEQ ID NO:20所示。

[0371] BHK细胞的体外转染证实了该RNA(C01)能够介导亚基因组ORF编码的报告基因(分泌型胚胎碱性磷酸酶)的表达(图17A)。

[0372] 实施例10由C06的nsP1同源性降低和C07的nsP2同源性降低修饰的单个VEEV复制子RNA

[0373] 鉴于在实施例4和7中观察到的成功,将同源性降低的区域组合在单个克隆中,以创建命名为C09的构建体。图14是编码构建体9(C09)的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸658-1620和2560-3954(灰色框)中的密码子以降低与野生型病毒的同源性。总的来说,沉默突变使nsP1 CDS与野生型的序列同一性降低了8.96%,nsP2 CDS的序列同一性降低了15.97%。质粒从T7启动子转录产生C09复制子RNA。C09的5' UTR、nsP1-4ORF和SGP如SEQ ID NO:28所示。

[0374] 图17B的柱状图示出了SEAP表达测试的结果,该测试通过对用指定复制子RNA转染16小时的BHK细胞的条件下清液取样并通过比色法在650nm波长处吸光度来定量SEAP表达。

[0375] 如图17B所示,BHK细胞的体外转染证实该RNA(C09)能够介导亚基因组ORF编码的报告基因(分泌型胚胎碱性磷酸酶)的表达。

[0376] 实施例11由C03的nsP2同源性降低和C01的nsP4同源性降低修饰的单个VEEV复制子RNA

[0377] 鉴于在实施例6和9中观察到的成功,将同源性降低的区域组合在单个克隆中,以创建命名为C10的构建体。图15是编码构建体10(C10)的复制子序列模板的DNA质粒的示意图。在此图中,DNA质粒与图5中所示的类似,不同之处仅在于修改了基因组核苷酸3694-3954位和6381-7083位(灰色框)中的密码子以降低与野生型病毒的同源性。总的来说,沉默

突变使nsP2 CDS与野生型的序列同一性降低了3.19%，与nsP4 CDS的序列同一性降低了4.69%。质粒从T7启动子转录产生C10复制子RNA。C10的5' UTR、nsP1-4 ORF和SGP如SEQ ID NO 29所示。

[0378] 图17C是示出了SEAP表达测试的结果的柱状图，该测试通过对用指定复制子RNA转染16小时的BHK细胞的条件下清液取样并通过比色法在650nm波长处吸光度来定量SEAP表达。

[0379] 如图17C所示，BHK细胞的体外转染证实该RNA (C09) 能够介导亚基因组ORF编码的报告基因 (分泌型胚胎碱性磷酸酶) 的表达。

[0380] 实施例12在所有nsP的区域中用同源性降低修饰单个VEEV复制子RNA

[0381] 考虑到先前实施例中描述的方法的普遍成功，先前被分析且同源性降低以创建构建体C01、C02、C04、C05、C06和C07的区域被重新集中分析，并在确定的非结构化核苷酸延伸中替换沉默突变，以创建命名为C11的构建体。图16是编码构建体11 (C11) 的复制子序列模板的DNA质粒的示意图。在此图中，DNA质粒与图5中所示的类似，不同之处仅在于修改了基因组核苷酸503-658、658-1620、1620-2560、2560-3954、3954-4120和6381-7083 (灰色框) 中的密码子以降低与野生型病毒的同源性。在此实施例中，与野生型病毒相比，所使用的沉默突变导致nsP CDS的同源性总体降低，降低程度相当于所有先前突变区域的总和。质粒从T7启动子转录产生C11复制子RNA。C11的5' UTR、nsP1-4 ORF和SGP如SEQ ID NO:30所示。

[0382] 图17A-17D示出了SEAP表达测试的结果，以证明密码子适应复制子RNA中存在的编码SEAP的亚基因组ORF的功能。图17A是示出了SEAP表达测试的结果的柱状图，该测试是通过与对照、野生型 (WT) 和未转染 (无转染) 细胞相比，对转染了4小时 (左测柱)、16小时 (中间柱) 和24小时 (右侧柱) 的BHK细胞的条件下清液取样进行检测的。图17B和17C是显示SEAP表达测试的结果的柱状图，其通过对用指定复制子RNA (图17B-C09，以及图17C-C10) 转染16小时的BHK细胞的条件下清液进行取样，并通过在比色试验中测量650nm处的吸光度来量化SEAP表达。正如该数据所证明的，生物信息学分析确定的区域中nsP密码子的修饰不会对亚基因组ORF中编码的转基因的体外表达产生负面影响。图17D示出了小鼠中SEAP表达的结果。在这些实验中，每组四只小鼠通过肌肉注射1μg指定的密码子适应复制子进行给药。收集血清并通过发光试验定量SEAP的相对表达。图17D显示了每只动物的平均RLU、误差条、每组的标准差。两只未注射的小鼠 (标记为“误差条、每组的标.”) 用作阴性对照以指示背景信号。如该数据所证明的，通过生物信息学分析鉴定的区域中nsP密码子的修饰不会对体内亚基因组ORF中编码的转基因的表达产生负面影响。

[0383] 总的来说，通过该分析揭示了nsP的广泛区域包含非结构化核苷酸片段，这些片段易于发生不影响基因表达的沉默突变。图18是VEEV基因组的比例图，其中通过RNA折叠分析和本发明所述的实验确定的非结构多蛋白的区域富集于易于沉默突变的非结构区域，被标记为灰色框。本发明的实施方式公开了跨越这些区域的沉默突变，使得通过基本局部比对搜索 (BLAST) 得到的序列与天然存在的VEEV型基因组nsP编码区域相差高达10%的非同一性。

[0384] 实施例13复制具有nsP4 C末端编码区的VEEV复制子RNA以引入C末端融合位点并恢复亚基因组启动子 (SGP)

[0385] 通过常规克隆直接插入任意附加蛋白 (“过常规”) 以融合nsP4和xP ORF会破坏组成

SGP的核苷酸序列元件。由于介导亚基因组RNA的转录,从而介导亚基因组ORF的后期翻译需要SGP,因此这种将xP融合到nsP4 C末端的方法将破坏复制子介导亚基因组基因座转基因表达的能力。图3示出了通过复制SGP区域(包括nsP4 C末端的编码区域)来规避该问题的过程。通过掺入产生常规克隆的限制性位点的核苷酸,然后,上游重复区域可以用作xP框内插入的靶标。图3示出了可遵循的简单步骤的实例,以克隆可修饰的nsP4开放阅读框(ORF),同时保持完整的延伸亚基因组启动子(SGP,条纹框)。然而,这种简单化的方法留下了两个具有相同序列的区域,当在制造所需的细菌菌株中繁殖该DNA模板序列时,该序列会导致遗传不稳定。

[0386] 当该构建体在实践中被克隆为DNA模板时,代表编码nsP4 C末端的密码子的相同序列使得标准细菌菌株(DH5-a)无法保持重复序列完整。从用图3中概述的方法构建的DNA模板中分离的所有克隆都携带消除nsP4 C-末端重复的缺失。

[0387] 实施例14具有nsP4 C末端编码区重复和同源性降低以引入C末端融合位点,从而恢复亚基因组启动子(SGP),并消除了重复序列元件的VEEV复制子RNA

[0388] 图4示出了在保持完整的延伸亚基因组启动子(SGP,条纹框)的同时克隆可修饰的nsP4开放阅读框(ORF)所遵循的最佳步骤的实例。这种方法消除了有问题的重复序列元素,DNA模板可以很容易地在常见菌株中繁殖和制造。图4示出了将xP融合到nsP4C末端,同时保持完整的SGP区域的最佳过程。上游重复序列的密码子首先被修改,以消除与下游重复序列的同源性,下游重复序列保持野生型序列以确保亚基因组ORF附近的完整SGP。位于任意xP的3'的下游SGP的序列如SEQ ID NO:14所示。然后通过常规分子克隆将上游同源性降低的重复区域(以SEQ ID NO:18所示的序列为例)融合在nsP4末端框内插入的任意xP。最终nsP4编码RNA序列的实例如SEQ ID NO:19所示。图19是说明为了实验而修饰的nsP4复制子序列的示意图,其中附加到nsP4蛋白的外源蛋白(xP)是绿色荧光蛋白(GFP)。在此图中,SG ORF中的结构多蛋白被报告基因SEAP取代。在此图中,nsP包括编码GFP的基因,其前面是作为xP的自裂解2A肽,亚基因组ORF包括分泌型胚胎碱性磷酸酶(SEAP)报告基因。由这些修改产生的nsP1-nsP2-nsP3-nsP4-2A融合多蛋白氨基酸序列可以进一步与目标xP(例如GFP)融合,如SEQ ID NO:12所示。编码该工程化的nP-2A ORF的示例性DNA序列如SEQ ID NO:13所示。nsP1-nsP2-nsP3-nsP4-2A-GFP氨基酸序列如SEQ ID NO:15所示,编码该工程化的nsP-2A-GFP ORF的示例性DNA序列如SEQ ID NO:16所示。经过总的nsP4-2A-GFP修饰的复制子RNA构建体被命名为C16。该复制子RNA被称为VEEVrep-nsP4[GFP]-SEAP,其5'UTR、nsP1-4-GFP融合ORF和SGP的相应RNA序列如SEQ ID NO:31所示。

[0389] 图20A-20C示出了体外细胞中由工程化的nsP4修饰的VEEV复制子介导的基因表达。BHK细胞用实验性nsP4修饰的复制子转染,该复制子(VEEVrep-nsP4[GFP]-SEAP)编码非结构蛋白ORF中的GFP和亚基因组ORF中的SEAP。图20A示出了荧光显微镜检查的结果,证明当用nsP4(GFP)修饰的VEEV-SEAP复制子RNA进行转染时存在多个GFP阳性细胞(右上),并且在未转染的对照细胞中不存在GFP表达(右下)。图20B是用GFP特异性单克隆抗体转染的BHK细胞裂解物的免疫印迹(顶部条带),以GAPDH检测用作负载对照(下方条带),其显示与未转染对照(NoTx)相比,通过两个独立分离的nsP4(GFP)修饰的VEEV复制子RNA克隆(克隆1和克隆2)确认了GFP表达。图20C是显示编码SEAP的亚基因组ORF的功能的柱状图,该实验通过对转染的BHK细胞的条件上清液取样并通过比色试验定量SEAP表达。参考此图,与未转染对照

(其中仅观察到试验背景吸光度)相比,克隆1和克隆2都显示出高SEAP表达。如图20A-20C所示,该RNA在转染至BHK细胞中时同时驱动GFP表达(图20A和29B)和SEAP表达(图20C)。这证实了nsP4介导的基因组复制转录和SGP介导的翻译都得到了维持,同时,从独立于亚基因组ORF的早期翻译位点成功产生了额外的辅助功能蛋白(GFP)。

[0390] 实施例15同时包含nsP4C末端xP融合和同源性降低的nsP的单个VEEV复制子RNA

[0391] 鉴于实施例9和12中的成功,生成了组合了C09的同源性降低的区域和C16的nsP4-2A-GFP修饰的VEEV复制子,以创建命名为C13的新构建体。图21A-21B示出了由该工程化的nsP4修饰的VEEV复制子介导的基因表达。该复制子RNA结合了如本发明中所述的密码子适应和nsP4 C末端修饰。

[0392] 图21A是示出了荧光显微镜检查结果的图片,显示了当用密码子适应的nsP4(GFP)修饰的VEEV-SEAP复制子RNA(C13)进行转染时的多个GFP阳性细胞(上图,右),以及在未转染的对照细胞中不存在GFP表达(下图,右)。

[0393] 图21B是显示发光试验(RLU)结果的图表,该发光试验与从未注射的对照小鼠收集的血清相比,对从注射了野生型VEEVrep-SEAP复制子RNA(WT)或注射了在非结构多蛋白基因的非结构化区域中携带优化的密码子的经nsP4(GFP)修饰的VEEV(经修饰的)复制子RNA(类似于VEEVrep-nsP4[GFP]-SEAP)的小鼠收集的血清中的SEAP浓度进行定量。于注射后第1天、第3天和第5天收集血清。在这些实验中,小鼠(每组三只)通过肌肉注射标准VEEV-SEAP复制子RNA或等剂量在非结构多蛋白基因的非结构区域中携带优化的密码子的经nsP4(GFP)修饰的VEEV复制子RNA。在指定时间收集血清并通过发光试验定量SEAP的浓度。示出了每组小鼠的平均RLU、误差条、标准差。两只未注射的小鼠用作阴性对照以指示背景信号。如该数据所证明的,通过此处描述的方法对nsP4基因的修饰不会对体内亚基因组ORF中编码的转基因的表达产生负面影响。

[0394] 如图21A所示,观察到该RNA能够在体外转染的细胞中表达GFP。如图21B所示,还观察到在小鼠中注射C13导致与亲代VEEV-SEAP复制子RNA相等的SEAP血清浓度,表明与野生型病毒的同源性降低,以及通过与nsP4融合而结合的额外转基因有效载荷并不会妨碍体内基因表达功效。

[0395] 实施例16编码与nsP4 C末端融合的功能性酶,并在SG ORF中分离转基因的复制子RNA

[0396] 按照实施例14中描述的设计,可以将酶安装在nsP4 C末端(即,xP可以是功能性酶)。如上文概述和SEQ ID NOS:18和SEQ ID NO:14所示,SGP可以通过复制来维持,将野生型SGP序列置于xP融合体的下游(在本实施例中为酶的下游),并修改nsP4 C末端编码区的上游重复序列,以保持正确的氨基酸序列,同时降低与野生型SGP的同源性。这确保了构建体的遗传稳定性并避免了同源驱动的重组,否则会导致xP的删除(如实施例13中所例示的,如果不进行密码子改变,就会发生这种情况)。

[0397] 图22A-22B示出了经修饰以编码小核糖核酸病毒P1结构多蛋白(其在SG ORF中编码)的蛋白水解加工所需的3Cpro蛋白酶的甲病毒复制子RNA的基因表达。

[0398] 图22A示出了对甲病毒复制子的nsP4编码序列进行修饰以编码小核糖核酸病毒P1结构多蛋白(01 Manisa P1多蛋白)所需的3Cpro蛋白酶的构建体的示意图,称为VEEVrep-nsP4[3C]-P1。该RNA的5' UTR、nsP1-4、3Cpro和SGP区域的序列如SEQ ID NO:32所示。编码SG

ORF中的P1结构多蛋白的RNA序列如SEQ ID NO:33所示。nsPORF中编码的nsP4-2A-3Cpro融合多蛋白的氨基酸序列如SEQ ID NO:40所示。

[0399] 该图说明了示例性VEEV复制子RNA的设计,其中使用本发明描述的方法通过2A自裂解肽将C末端融合至病毒蛋白酶。本实施例中的病毒蛋白酶是小核糖核酸病毒(口蹄疫病毒(FMDV)) A12株的3Cpro。3Cpro酶是插入复制子NPS4修饰区域的良好候选酶的一个例子。3Cpro是一种高度加工酶,当外源递送至细胞时,与其底物相比,仅需要相对较低的分子拷贝数。3Cpro活性是小核糖核酸病毒衣壳多蛋白(称为P1多蛋白)的翻译后加工所必需的。除了在低表达水平下发挥全部功能外,由于蛋白酶固有的毒性,3Cpro的低拷贝数也是理想的选择。当3Cpro在体外高水平表达时,细胞代谢停滞且活力下降。人们已尝试创建FMDV疫苗,其中编码P1多蛋白的复制子RNA和编码3Cpro的mRNA共同配制,目的是将两种分子递送至同一细胞并允许P1的加工(反式互补)。在这种反式互补疫苗设计中,P1-2A与3Cpro的比例需要为~30-40:1,以尽量减少3Cpro的毒性。虽然存在允许将多个RNA分子共封装和递送到同一细胞的递送平台,但由于这些比率要求,不可能100%的纳米颗粒含有至少一个编码3Cpro的mRNA拷贝。因此,反式互补方法最大限度地降低了疫苗产品的潜在功效,因为单个纳米颗粒如果不具有至少一个拷贝的两种RNA分子,将无法有效生成经过加工的衣壳,从而无法促进体液免疫反应。

[0400] 鉴于3Cpro酶具有高度细胞毒性,并且必须维持在低稳态浓度,以避免通过其蛋白水解活性最终破坏细胞过程,因此它在nsP4 C末端区域编码,作为FMDV疫苗构建体的一部分来介导表达。P1多蛋白包含VP0、VP2和VP3亚基,这些亚基必须被3Cpro切割成单独的多肽链,以便组装成衣壳结构。为了制造功能性小核糖核酸病毒疫苗,这些衣壳结构必须组装起来才能被免疫系统识别,从而驱动中和抗体的产生。P1多蛋白的VP0片段还包含VP4和VP2蛋白,它们通过3Cpro独立过程进行裂解,并且不是衣壳组装所必需的。为了制造有用的FMDV疫苗,FMDV 01 Manisa株的P1多蛋白被编码在VEEV的复制子的SG ORF中。如图22A中所示,称为VEEVrep-nsP4[3C]-P1的复制子通过在nsP4 C末端进行编码而提供顺式3Cpro活性。另一个复制子VEEVrep-nsP4[3C^{mut}]-P1包含一个会中止3Cpro序列的翻译的插入突变,以用作平行对照。作为额外的对照,制备了具有野生型nsP4和SGP区域的复制子,其仅编码SG ORF中的P1多蛋白(VEEVrep-P1)。这些RNA通过体外转录和转录后加帽合成。用每孔1μg的每种P1编码复制子转染12孔培养皿中的BHK细胞,并于约1天后制备细胞裂解物用于通过免疫印迹进行分析。

[0401] 图22B是对用以下RNA转染的BHK细胞裂解物进行的免疫印迹的照片:泳道1,VEEVrep-nsP4[3C]-P1;泳道2,复制子RNA与泳道1相似,但携带插入突变,引入提前终止密码子并切除3Cpro多肽的表达,作为阴性对照(VEEVrep-nsP4[3C^{mut}]-P1);泳道3,VEEVrep-P1(类似于VEEVrep-nsP4[3C]-P1,但缺少nsP4融合的3Cpro);泳道4,将VEEVrep-P与1/40质量的编码3Cpro的mRNA组合转染,作为阳性对照。参考该图,只有在VEEVrep-nsP4[3C]-P1(泳道1)的介导下,才能将~100kDa的P1多蛋白成功蛋白水解加工成~35kDa的免疫原性VP0衣壳片段。可以看到代表完整VP4+VP2+VP3的不完全裂解中间产物,其表观分子量为~55kDa。使用VP2特异性抗体来探测印迹。M,分子量参照标记;C,未经处理的阴性对照细胞。该图显示了用FMDV VP2特异性单克隆抗体探测后样品的免疫印迹结果。缺乏3Cpro活性的VEEVrep-nsP4[3C^{mut}]-P1和VEEVrep-P1,仅产生全长完整P1多肽,视为~100kDa带。

VEEVrep-nsP4[3C]-P1成功产生了所需的完全加工的VP0片段,表明衣壳加工按预期进行。值得注意的是,处理效果优于阳性对照(泳道4),其中VEEVrep-P1与常规mRNA分子中反式编码的单独3Cpro以40:1的质量比(已知当作为疫苗施用时在猪中产生保护性抗体滴度的比例)组合转染。由于缺乏加工中间片段VP4+2+3,因此卓越的加工效率是显而易见的。这证明了使用本发明中解释的方法在nsP4 C末端区域编码功能性蛋白(在本特定实施例中为酶)以引发所需的生物化学结果的实用性。

[0402] 实施例17编码与nsP4 C末端融合的免疫调节因子,并在SG ORF中分离转基因的复制子RNA

[0403] 为了产生期望的结果,增加或减少细胞中复制子RNA活性引发的免疫反应是有用的。例如,可以通过在癌细胞中诱导强烈的先天免疫反应来治疗癌症。

[0404] 图23A-23D示出了来自经修饰以编码人STING蛋白的甲病毒复制子RNA的基因表达数据,所述甲病毒复制子RNA被突变为对于IFN刺激活性具有组成型活性,并且在SG ORF中编码报道基因SEAP。

[0405] 图23A是构建体VEEVrep-nsP4[STING]-SEAP的示意图,其为包括甲病毒复制子的nsP4编码序列的修饰,以编码人STING蛋白的复制子RNA,其被突变为对于IFN刺激活性具有组成型活性,并且报道基因SEAP在SG ORF中编码。该RNA的5'UTR、nsP1-4、STING和SGP区域的序列如SEQ ID NO:34所示。产生的经翻译的nsP4-2A-STING融合多肽的氨基酸序列如SEQ ID NO:42所示。该图示出了复制子RNA构建体,其中通过与nsP4 C末端融合来编码有效的免疫调节蛋白STING(突变为组成型活性)。在癌症治疗的情况下,SG ORF中可以编码肿瘤相关抗原或其他免疫调节因子,以驱动针对所需癌症靶点的特异性细胞免疫。为了说明STING修饰的nsP4的功能,在此实施例中,SG ORF编码报告基因SEAP,其可以轻松地从肿瘤细胞培养物中进行定量。为了证明该设计的STING和SEAP编码复制子(称为VEEVrep-nsP4[STING]-SEAP)在人类细胞中可以发挥作用,使用商业转染试剂(TransIT-mRNA,Mirus)与1 μ g VEEVrep-nsP4[STING]-SEAP或作为对照的非nsP4修饰的VEEVrep-SEAP转染HEK-LuciaTM空白细胞。约1天后,对上清液培养基进行取样并裂解细胞以通过免疫印迹分析STING含量。如图。图23B是在用以下RNA转染的HEK-LuciaTM空白细胞上进行的免疫印迹的照片:泳道1,VEEVrep-SEAP;泳道2,VEEVrep-nsP4[STING]-SEAP;泳道3,无转染对照。在此图中,常见的背景条带用星号标记(*,可能是内源野生型STING),并且使用人STING特异性抗体来探测印迹。据观察,只有VEEVrep-nsP4[STING]-SEAP表现出一条额外的条带,代表nsP4 C末端区域编码的组成型活性STING蛋白。

[0406] 图23C是示出与对照(未转染细胞)相比,在VEEVrep-SEAP和VEEVrep-nsP4[STING]-SEAP复制子RNA中的SG ORF中编码的SEAP报道基因的表达结果(通过比色试验中650nm处的吸光度测量)的柱状图。据观察,尽管用免疫调节多肽对nsP4 C末端进行了新的修饰,但VEEVrep-nsP4[STING]-SEAP亚基因组启动子活性得以保留,并且转基因表达完整。参见该图,在比色试验中对培养物上清液中的SEAP进行定量,该基因被证明由两种构建体表达,证明VEEVrep-nsP4[STING]-SEAP构建体成功地进行了复制和SGP驱动的转基因表达。SEAP表达较低可能是由于STING蛋白触发的强烈先天免疫反应诱导HEK细胞活力降低。I型IFN信号传导活性可以通过HEK-LuciaTM空白细胞中分泌的荧光素酶的上清液浓度来测量,在这些细胞中,荧光素酶是在包含IFN刺激基因(ISG)54启动子的IFN诱导型启动子的控制

下进行编码的,IFN刺激基因 (ISG) 54启动子通过多聚体干扰素刺激反应元件 (ISRE) 增强。如图23D示出与对照 (未转染细胞) 相比,由VEEVrep-SEAP和VEEVrep-nsP4[STING]-SEAP复制子RNA诱导的IFN信号传导活性的替代测量结果的柱状图。HEK-Lucia™空白细胞编码受IFN诱导型启动子控制的分泌型荧光素酶报告基因,IFN诱导型启动子包含由多聚体IFN刺激反应元件 (ISRE) 增强的IFN刺激基因 (ISG) 54启动子。据观察,VEEVrep-nsP4[STING]-SEAP RNA触发有效的IFN信号传导,而未修饰的VEEVrep-SEAP RNA则不会。

[0407] 参考图23D,在分泌型荧光素酶的发光试验中,仅VEEVrep-nsP4[STING]-SEAP显示出I型IFN信号激活,证明活性STING蛋白是有功能的并且对转染的人细胞赋予期望的效果。STING是一种免疫调节蛋白,可通过IFN信号级联具有免疫刺激和免疫抑制两种特性,具体取决于施用的组织和细胞类型。因此,这种嵌入STING或其他免疫调节蛋白的方法是一种引发或减弱适应性免疫反应的方法,其使得工程复制子不仅可用于治疗传染病和癌症,还可通过在需要时触发免疫耐受来治疗自身免疫性疾病和过敏。

[0408] 实施例18黑色素瘤的治疗

[0409] 使用本发明描述的方法工程化的表达来自nsP4编码区的辅助因子的复制子可用于治疗例如黑色素瘤等癌症。图24A-24D示出了在B16黑色素瘤中来自甲病毒复制子RNA的基因表达和生物学效应,所述甲病毒复制基RNA经修饰以在nsP ORF中编码GFP或人STING蛋白 (突变为对IFN刺激活性具有组成型活性),并在SG ORF中编码报告基因SEAP。

[0410] 图24A是示出与未转染的对照细胞相比,用1μg VEEVrep-nsP4[GFP]-SEAP转染后约1天,在12孔培养皿中培养的B16黑色素瘤细胞中的GFP表达结果的照片。由本发明所述构建体的nsP4 C末端编码荧光蛋白,观察到大量的GFP表达。图24A描绘了B16黑色素瘤细胞中,由RNA VEEVrep-nsP4[GFP]-SEAP介导的GFP成功表达,其证明了nsP4-融合蛋白在肿瘤细胞系中具有有效表达的能力。

[0411] 图24B是示出与未转染的对照细胞相比,在用1μg指定复制子RNA:VEEVrep-nsP4[STING]-SEAP、VEEVrep-nsP4[GFP]-SEAP和VEEVrep-SEAP转染后,在12孔培养皿中培养的B16黑色素瘤细胞中SEAP表达结果的柱状图。在实验中,SEAP表达定量在转染后1天,通过对培养基的比色试验进行,其显示SG ORF中编码的SEAP基因成功表达。参考图24B,无论nsP4融合蛋白的性质如何,SG ORF (SEAP) (代表抗原或其他癌症治疗的基因序列的转基因位置) 中编码的报告基因序列都能在相同细胞中有效表达。

[0412] 图24C是示出与未转染的对照细胞相比,在用1μg指定复制子RNA:VEEVrep-nsP4[STING]-SEAP和VEEVrep-SEAP转染~1天后培养的B16黑色素瘤细胞 (通过ELISA夹心法测定培养上清) 的IFN-β分泌测量结果的柱状图。参考该图,观察到VEEVrep-nsP4[STING]-SEAP RNA成功诱导这些癌细胞中的IFN-β分泌,其水平比VEEVrep-SEAP高约73%。这表明与nsP4融合表达的STING蛋白成功诱导了癌细胞的先天免疫反应,这与临床环境中肿瘤的成功清除相关。

[0413] 图24D是与健康未转染的对照细胞 (对照;最左图) 相比,用2μg VEEVrep-nsP4[GFP]-SEAP或VEEVrep-nsP4[STING]-SEAP转染1天后B16细胞的光学显微镜检查的照片。参考该图,未转染的对照细胞已增殖并填满培养皿;转染VEEVrep-nsP4[GFP]-SEAP后观察到中度生长失败和细胞病理变化;转染VEEVrep-nsP4[STING]-SEAP后观察到最大的生长抑制和细胞死亡。

[0414] 参考图24D,当转染1天后通过明场显微镜检查B16细胞时,未处理的细胞(对照)已如预期那样增殖并充满培养皿;相反,在转染VEEVrep-nsP4[GFP]-SEAP后观察到中度生长失败和细胞病理变化,并且在转染VEEVrep-nsP4[STING]-SEAP后观察到更大程度的生长抑制和细胞死亡。总的来说,数据表明nsP4修饰的复制子可以在黑色素瘤细胞中复制,驱动来自nsP4 C末端编码ORF和SG ORF的基因表达,当免疫调节蛋白与nsP4融合时激活更高的先天免疫反应,并且抑制黑色素瘤细胞的生长和存活。在本实施例中描述的所有B16黑色素瘤细胞实验中,细胞在12孔培养皿中,在DMEM+10%灭活胎牛血清(IFS)中生长。为了分析基因表达(GFP、SEAP)或免疫激活(IFN- β),用1 μ g指定RNA转染细胞,并在1天后进行分析。为了通过显微镜测量细胞增殖和活力,用2 μ gRNA转染细胞,并在1天后捕获图像。

[0415] 实施例19HPV诱发的上皮肿瘤的治疗

[0416] 为了证明本发明中公开的工程化复制子的普遍适用性,在另一种肿瘤细胞类型中证明了nsP4修饰的复制子RNA的功能:TC-1上皮肿瘤细胞,该细胞表现出由HPV E6和E7癌基因转化驱动的恶性生长。

[0417] 图25A-25D示出了在TC-1癌细胞中来自经修饰以编码人STING蛋白的甲病毒复制子RNA的基因表达,该甲病毒复制子RNA被突变为对IFN刺激活性具有组成型活性,并在SG ORF中编码报告基因SEAP。

[0418] 图25A示出与未转染的对照细胞相比,用2 μ g指定复制子RNA:VEEVrep-nsP4[GFP]-SEAP、VEEVrep-nsP4[STING]-SEAP转染2天后,在12孔培养皿中培养的TC-1癌细胞中SEAP表达结果的柱状图。在本实验中,通过培养基的比色试验对SEAP表达进行定量,显示在SG ORF中编码的SEAP基因的成功表达。参考图25A,如在B16黑素瘤(实施例18)中所见,SEAP的表达证实了SG ORF对于VEEVrep-nsP4[STING]-SEAP和VEEVrep-nsP4[GFP]-SEAP RNA在这种不相关癌症类型中的活性。

[0419] 图25B是示出与未转染的对照细胞相比,用2 μ g指定复制子RNA:VEEVrep-nsP4[GFP]-SEAP或VEEVrep-nsP4[STING]-SEAP转染1天后,培养的TC-1癌细胞(通过ELISA夹心法测定培养上清)的IFN- β 分泌测量结果的柱状图。参考图25B,nsP4编码的组成型活性STING融合蛋白比不含STING的复制子诱导更多的IFN- β 分泌。

[0420] 图25C是示出用1 μ g VEEVrep-nsP4[GFP]-SEAP转染后约3天,转染后的TC-1癌细胞团聚物中GFP表达结果的照片。参考此图,与未处理的对照细胞中没有GFP表达(下图,右)相比,通过在构建体nsP4 C末端编码荧光蛋白,在整个多细胞结构中观察到大量GFP表达(上图,右)。参考图25C,在从TC-1培养物分离的细胞团块中,证实VEEVrep-nsP4[GFP]-SEAP的GFP表达在整个结构中广泛存在。

[0421] 图25D是示出与未处理的对照细胞相比,用4 μ g指定复制子RNA:VEEVrep-nsP4[STING]-SEAP转染1天后,TC-1癌细胞的光学显微镜检查结果的照片。参考此图,观察到未处理的细胞(对照)已增殖并填满培养皿;转染VEEVrep-nsP4[STING]-SEAP后观察到生长抑制和细胞死亡。

[0422] 如图25D所示,正如对B16黑色素瘤细胞所观察到的一样,用VEEVrep-nsP4[STING]-SEAP治疗严重抑制了增殖和活力。

[0423] 这些数据证实经nsP4修饰的复制子在不同类型的癌症治疗中具有广泛的适用性。在本实施例中描述的所有TC-1癌细胞实验中,细胞在12孔培养皿中,在DMEM+10% IFS中生

长。为了通过显微镜分析GFP表达,用1 μ g RNA转染细胞,用刮刀刮擦细胞以释放团聚物,并在3天后成像。为了通过IFN- γ 过为了通过镜分测量免疫激活,用2 μ g指定的RNA转染细胞,并在1天后分析培养基。为了通过显微镜测量细胞增殖和活力,用4 μ g RNA转染细胞,并在1天后捕获图像。用2 μ g指定的RNA转染2天后,通过对培养基进行比色试验来测量SEAP表达结果。

[0424] 实施例20在未经生物信息学检查的情况下,经同源性降低的修饰的VEEV复制子RNA

[0425] 生成了另外四个同源性降低的复制子RNA。26A-26G示出了经修饰与野生型同源性降低的VEEV复制子RNA以及来自这些构建体的表达数据。VEEVrepHK-SEAP是通过在酸658和3359位核苷之间的nsP1和nsP2接合处使用一系列沉默突变而产生的(图26A);核苷酸位置参考SEQ ID NO:17所示的VEEV基因组序列。图26A是编码构建体12(C12;VEEVrepHK-SEAP)的复制子序列模板的DNA质粒的示意图。参考此图,DNA质粒与图5所示的类似,不同之处仅在于修改了基因组核苷酸658-3359位(灰色框)的密码子以降低与野生型病毒的同源性,而不考虑可能存在的任何二级结构元件。C12质粒转录产生的RNA的5' UTR、nsP1-4编码区和SGP如SEQ ID NO:35所示。

[0426] VEEVrepHK(Δ U)-SEAP在相同区域内进行类似突变,但通过仅选择尽可能省略尿嘧啶核苷碱基的密码子(图26B)。图26B是构建体13(C13;VEEVrepHK(Δ U)-SEAP)的示意图。参考该图,DNA质粒与图5所示的类似,不同之处仅在于修改了基因组核苷酸658-3359位(灰色框)的密码子以降低与野生型病毒的同源性,而不考虑任何二级结构元件的可能存在,并且优先选择避开尿嘧啶碱基的密码子。C13质粒转录产生的RNA的5' UTR、nsP1-4编码区和SGP如SEQ ID NO:36所示。

[0427] VEEVrepES-SEAP是通过在核苷酸4120和6965位之间的nsP3和nsP4接合处应用一系列沉默突变而生成的。图26C是构建体14(C14;VEEVrepES-SEAP)的示意图。参考该图,DNA质粒与图5所示的类似,不同之处仅在于修改了基因组核苷酸4120-6965位(灰色框)中的密码子以降低与野生型病毒的同源性,而不考虑任何二级结构元件的可能存在。C14质粒转录产生的RNA的5' UTR、nsP1-4编码区和SGP如SEQ ID NO:37所示。

[0428] VEEVrepES(Δ U)-SEAP在相同区域进行类似突变,但是通过仅选择尽可能省略尿嘧啶核苷碱基的密码子(图26D)。图26D是构建体15(C15;VEEVrepES(Δ U)-SEAP)的示意图。参考此图,DNA质粒与图5所示的相似,不同之处仅在于修改了基因组核苷酸4120-6965位(灰色框)的密码子以降低与野生型病毒的同源性,不考虑任何二级结构元件的可能存在,并且优先选择避开尿嘧啶碱基的密码子。C15质粒转录产生的RNA的5' UTR、nsP1-4编码区和SGP如SEQ ID NO:38所示。

[0429] 这些构建体在没有考虑序列中存在的二级结构元件的情况下生成,以评估减少与野生型病毒同源性的典型非生物信息学方法的性能。

[0430] 图26E示出与未转染的对照细胞相比,在用1 μ g VEEVrep-SEAP、VEEVrepHK-SEAP或VEEVrepHK(p相比,在用典型非生物信转染16小时后,在12孔培养皿中培养的BHK细胞中SEAP表达结果的柱状图。在本实验中,通过对培养基的比色试验来定量SEAP的表达,其显示在SG ORF中编码的SEAP基因的成功表达。

[0431] 图26F示出与未转染的对照相比,在用1 μ g VEEVrep-SEAP、VEEVrepES-SEAP或

VEEVrepES (SEAPAP过对培养基的转染16小时后,在12孔培养皿中培养的BHK细胞中SEAP表达结果的柱状图。在此实验中,通过对培养基的比色试验定量SEAP表达,其显示VEEVrepES-SEAP或VEEVrepES (SEAP,P过对构建体的SG ORF中编码的SEAP基因没有表达。

[0432] BHK细胞的体外转染证实VEEVrepHK-SEAP和VEEVrepHK (SEAP,P过对能够介导亚基因组ORF编码的报道基因(分泌型胚胎碱性磷酸酶)的表达(图26E),但VEEVrepES-SEAP和VEEVrepES (SEAP,P过对不能介导(图26F)。这表明如果不检查一级RNA序列来识别二级结构元件,就无法通过对任意编码片段的沉默突变来可靠地获得功能性复制子。VEEVrepFIK-SEAP和VEEVrepHK (-SEAP获得功似乎代表了一个偶然的发现,表明原始方法可以自发地产生工作复制子,然而进一步的表征表明,这两种构建体都失去了在体内的SEAP表达活性,该表达活性如通过在小鼠(n=5)肌内注射2 μ g LNP配制的RNA 1天后,通过化学发光试验测定血清SEAP浓度来测量。图26G是示出与对照(来自未处理的小鼠的血清)相比,注射2 μ g LNP配制的VEEVrep-SEAP、VEEVrepHK-SEAP或VEEVrepHK (SEAP作复制子后(通过化学发光试验)测量的血清中SEAP体内表达结果的柱状图。在此实验中,小鼠(每组5只)通过肌内(i.m.)注射施用指定的RNA,并在第二天采集血清样本,并通过化学发光试验对SEAP进行定量。绘制每组小鼠的平均RLU;误差线,标准差。来自两只未经治疗的小鼠(n=2)的血清用作阴性对照组。

[0433] 虽然体内性能得到证实,但表达低于标准水平,并且凸显了本发明内容中描述的生物信息学方法的优越性,与野生型复制子相比,其在体内产生了等同或更好的基因表达(参见实施例15)。

[0434] 实施例21工程复制子是安全的,不会伤害健康的人体细胞

[0435] 为了证明本发明中公开的工程化复制子的安全性,确认了nsP4修饰的复制子RNA的功能在非癌性人类细胞中较低。人包皮成纤维细胞(HFF)是一种通常用于研究的正常细胞类型,代表健康的非转化细胞生物学特性,同时还是多种核酸和病毒物种的有效转染和感染宿主。图27A-27D示出了正常HFF细胞中的VEEV复制子RNA表达结果和影响。图27A示出了与未转染的对照细胞相比,在用2 μ g或4 μ g指定复制子RNA (VEEVrep-nsP4[STING]-SEAP,或VEEVrep-SEAP)转染后1天,在12孔培养皿中培养的正常HFF细胞中SEAP表达结果的柱状图。在此实验中,通过对培养基的比色试验来定量SEAP的表达。

[0436] 参考该图,用2 μ g或4 μ g在SG ORF中编码SEAP报道基因的非工程复制子(VEEVrep-SEAP)转染12孔培养皿中的HFF,导致转染后1天在通过对培养基进行的比色试验测得的SEAP表达水平低;相反,含有nsP4编码的组成型活性STING融合蛋白的复制子(VEEVrep-nsP4[STING]-SEAP)不会诱导SEAP分泌,表明工程复制子在健康细胞中的复制和转基因表达受到抑制。

[0437] 图27B是用47B表明工程复制子在健康细胞中的复制和转基因表达受转染后约4天,胰蛋白酶解离的正常人HFF细胞中GFP表达结果的照片,其显示来自nsP4 C末端编码的GFP蛋白在这些健康的非癌细胞中没有有效的基因表达。参考该图,与上面测试的癌细胞类型(参见实施例18和19)不同,在用VEEVrep-nsP4[GFP]-SEAP转染的HFF细胞中没有观察到GFP表达。

[0438] 图27C是示出来自经转染的HFF的IFN- γ 活性的测量结果的柱状图。在以1胞对、2胞对和4胞对剂量转染指定RNA (VEEVrep-nsP4[STING]-SEAP、VEEVrep-nsP4[GFP]-SEAP、

VEEVrep-SEAP) 1天后收集来自12孔培养皿中的HFF的条件培养基。从未转染的细胞收集培养基作为阴性对照(对照HFF)。收集的培养基用于HEK-Lucia空白细胞培养,以测量其对于培养基中的I型IFN的相应。未处理的HEK-Lucia空白细胞作为额外的阴性对照(无培养基对照)包含在内。在本实验中,通过化学发光试验在第二天检测HEK-Lucia空白细胞上清液中的IFN应答报告基因Lucia。如本发明中所述的工程化的复制子VEEVrep-nsP4[GFP]-SEAP和VEEVrep-nsP4[STING]-SEAP在健康HFF型细胞中不刺激I型IFN先天免疫应答。参见图27C,基于用条件HFF培养基处理IFN- β 敏感性HEK-Lucia空白细胞的结果,工程化复制子VEEVrep-nsP4[STING]-SEAP和VEEVrep-nsP4[GFP]-SEAP也不会这些正常细胞中驱动I型IFN信号传导的急剧增加。

[0439] 图27D是与未转染的对照相比,用4 μ g指定复制子RNA(VEEVrep-nsP4[GFP]-SEAP和VEEVrep-nsP4[STING]-SEAP)转染后4天HFF细胞的光学显微镜检查结果的图片。经观察,未处理的细胞(对照)已增殖填满培养皿并显示出健康的纤维形态。RNA转染的细胞表现出相同的增殖和健康的形态。如图27D所示,没有观察到对细胞生长、形态或活力的有害影响。为了通过显微镜测量细胞增殖并评估健康状况,用2 μ g或4 μ g RNA转染细胞,并在3天后捕获图像。即使在用4 μ g的高剂量RNA转染时,所有条件下的细胞都增殖并保持完整,具有预期的纤维形态(图27D),这表明这些RNA在健康细胞中的施用是安全的,并且表现出针对非癌旁观细胞的非靶向免疫原性可忽略不计。这与它们对癌细胞的作用形成对比,其中nsP4融合蛋白和SG ORF均强烈表达并且能够阻碍癌细胞生长并触发先天免疫应答(参见实施例18和19)。

[0440] 参考文献

[0441] US20140079734A1——对甲病毒复制子进行修饰,其为了影响复制效果而将基因组5'末端的一个组分复制到亚基因组区域。

[0442] W02014170493A2——对复制子的不同的修饰方法,其提及为了调节SGP而在nsP4的3'区域的突变。

[0443] J Virol.2006Oct;80(20):9962-9969——描述了nsP4的功能和蛋白的C末端标签,表明通过修饰这些端点,结构可能会得到保留。

SEQUENCE LISTING

- <110> 蒂巴生物技术有限公司
 <120> 人工甲病毒衍生的RNA复制子表达系统
 <130> TIB-PT012W0
 <150> 63/163,475
 <151> 2021-03-19
 <160> 46
 <170> PatentIn version 3.5
 <210> 1
 <211> 7482
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, C01 nsP coding sequence DNA
 <400> 1

```

atggagaaag ttcacgttga catcgaggaa gacagcccat tcctcagagc tttgcagcgg      60
agcttcccgc agtttgaggt agaagccaag caggtcactg ataatgacca tgctaatgcc      120
agagcgtttt cgcactctggc ttcaaaactg atcgaaacgg aggtggaccc atccgacacg      180
atccttgaca ttggaagtgc gcccggccgc agaatgtatt ctaagcacia gtatcattgt      240
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg      300
aagaaaaact gtaaggaaat aactgataag gaattggaca agaaaatgaa ggagctcgcc      360
gccgtcatga gcgaccctga cctggaaact gagactatgt gcctccacga cgacgagtcg      420
tgtcgctacg aagggcaagt cgctgtttac caggatgtat acgcggttga cggaccgaca      480
agtctctatc accaagccaa taaggaggtt agagtcgcct actggatagg ctttgacacc      540
acccctttta tgtttaagaa cttggctgga gcatatccat catactctac caactgggcc      600
gacgaaaccg tgtaacggc tcgtaacata ggcctatgca gctctgacgt tatggagcgg      660
tcacgtagag ggatgtccat tcttagaaag aagtatttga aaccatccaa caatgttcta      720
ttctctgttg gctcgaccat ctaccacgag aagagggact tactgaggag ctggcacctg      780
ccgtctgtat ttcacttacg tggcaagcaa aattacacat gtcggtgtga gactatagtt      840
agttgcgacg ggtacgtcgt taaaagaata gctatcagtc caggcctgta tgggaagcct      900
tcaggctatg ctgctacgat gcaccgcgag ggattcttgt gctgcaaagt gacagacaca      960
ttgaacgggg agagggtctc ttttcccggtg tgcacgtatg tgccagctac attgtgtgac     1020
caaatgactg gcatactggc aacagatgtc agtgcgagc acgcgcaaaa actgctgggtt     1080
gggctcaacc agcgtatagt cgtcaacggc cgcaccaga gaaacaccaa taccatgaaa     1140
aattaccttt tgcccgtagt ggcccaggca tttgctaggt gggcaaagga atataaggaa     1200
gatcaagaag atgaaaggcc actaggacta cgagatagac agttagtcat ggggtgttgt     1260
tgggctttta gaaggcacia gataacatct atttataagc gcccgatac ccaaaccatc     1320
atcaaagtga acagcgattt ccactcattc gtgctgccc gtagtaggcag taacacattg     1380

```

gagatcgggc	tgagaacaag	aatcaggaaa	atgttagagg	agcacaagga	gccgtcacct	1440
ctcattaccg	ccgaggacgt	acaagaagct	aagtgcgcag	ccgatgaggc	taaggaggtg	1500
cgtgaagccg	aggagttgcg	cgcagctcta	ccacctttgg	cagctgatgt	tgaggagccc	1560
actctggaag	ccgatgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagt	gaaaaattat	cttgcattcca	ccctctcgct	1740
gaacaagtca	tagtgataac	acactctggc	cgaaaagggc	gttatgccgt	ggaaccatac	1800
catggtaaag	tagtggtgcc	agagggacat	gcaatacccg	tccaggactt	tcaagctctg	1860
agtgaagtg	ccaccattgt	gtacaacgaa	cgtgagttcg	taaacaggta	cctgcaccat	1920
attgccacac	atggaggagc	gctgaacact	gatgaagaat	attacaaaac	tgtcaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtcaactg	ggctagggct	cacaggcgag	ctggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccag	taccaaccat	aggggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatatgtga	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgacaca	agtcttccac	2520
aaaagcatct	ctcgccgttg	cactaaatct	gtgacttcgg	tcgtctcaac	cttgttttac	2580
gacaaaaaaaa	tgagaacgac	gaatccgaaa	gagactaaga	ttgtgattga	cactaccggc	2640
agtaccaaac	ctaagcagga	cgatctcatt	ctcacttggt	tcagagggtg	ggtgaagcag	2700
ttgcaaatag	attacaaagg	caacgaaata	atgacggcag	ctgcctctca	agggtgacc	2760
cgtaaagggtg	tgtatgccgt	tcggtacaag	gtgaatgaaa	atcctctgta	cgcaccacc	2820
tcagaacatg	tgaacgtcct	actgaccgcg	acggaggacc	gcacgtgtg	gaaaacacta	2880
gccggcgacc	catggataaa	aacactgact	gccaaagtacc	ctgggaattt	cactgccacg	2940
atagaggagt	ggcaagcaga	gcacgatgcc	atcatgaggc	acatcttgga	gagaccggac	3000
cctaccgacg	tcttccagaa	taaggcaaac	gtgtgttggg	ccaaggcttt	agtgccggtg	3060
ctgaagaccg	ctggcataga	catgaccact	gaacaatgga	acactgtgga	ttattttgaa	3120
acggacaaaag	ctcactcagc	agagatagta	ttgaaccaac	tatgcgtgag	gttcttttga	3180
ctcgatctgg	actccggtct	atcttctgca	cccactgttc	cgttatccat	taggaataat	3240
cactgggata	actccccgtc	gcctaacatg	tacgggtgta	ataaagaagt	ggtccgtcag	3300
ctctctcgca	ggtaccacac	actgcctcgg	gcagttgcc	ctggaagagt	ctatgacatg	3360
aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	ctttagtcct	ccaccataat	gaacaccac	agagtgactt	ttcttcattc	3480
gtcagcaaat	tgaaggcgag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatggttg	actggttgtc	agaccggcct	gaggctacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gaccccatat	3660
aaataccatc	actatcagca	gtgtgaagac	catgccatta	agcttagcat	gttgaccaag	3720

aaagcttgtc	tgcatctgaa	tcccggcgga	acctgtgtca	gcataggtta	tggttacgct	3780
gacagggcca	gcgaaagcat	cattggtgct	atagcgcggc	agttcaagtt	ttcccgggta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	tggttacgat	3900
cgcaaggccc	gtacgcacaa	tccttacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
ggggtgtgcg	gagcgctgta	taagaaattc	ccggaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaagggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgacag	aacctgatgc	agagctggtg	agggtgcatc	cgaagagttc	tttggctgga	4560
aggaagggct	acagcacaag	cgatggcaaa	actttctcat	atttggaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgcccttgctt	gtgcatccat	4800
gccatgactc	cagaaaagagt	acagcgccta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgctcatcct	ttccattgcc	gaagtataga	atcactggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatte	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgaccac	5160
caggtgctgc	aagtcgaggc	agacattcac	gggccgcctt	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgtccgc	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagttttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgctac	gcactcctag	caggtcggtc	tcgagaacca	gcctggtctc	caaccgccca	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgtacc	gaacctgca	tcctgttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaaagagaa	ctttccgact	6060

```

gtggcttctt actgtattat tccagagtac gatgcctatt tggacatggt tgacggagct 6120
tcatgctgct tagacactgc cagtttttgc cctgcaaagc tgcgagctt tccaaagaaa 6180
cactcctatt tggaaacccac aatacgatcg gcagtgcctt cagcgatcca gaacacgctc 6240
cagaacgtcc tggcagctgc cacaaaaaga aattgcaatg tcacgcaaat gagagaattg 6300
cccgtattgg attcggcggc ctttaatgtg gaatgcttca agaagtacgc ctgcaacaac 6360
gagtactggg agacattcaa agagaacccc atccggctga ccgaggaaaa cgtggtcaac 6420
tacatcacca agctgaaggg ccccaaagcc gccgctctgt ttgccaagac acacaacctg 6480
aacatgctgc aggacatccc catggacaga ttcgtgatgg acctgaagcg ggacgtgaaa 6540
gtgacccttg gcaccaagca caccgaggaa cggcctaagg tgcaagtgat ccaggccgct 6600
gatcctctgg ccacagccta tctgtgtggc atccacagag aactcgtgcg gagactgaat 6660
gccgtgctgc ttccgaacat tcataactg tttgatattg cggtgaaga ctttgacgct 6720
attatagccg agcattcca gcctggggat tgtgttctgg aaactgacat cgcgtcgttt 6780
gataaaaagtg aggacgacgc catggtcttg accgcgtga tgattctgga agatctcgga 6840
gtggacgccg agctgctgac actgattgaa gccgcctttg gcgagatcag cagcatccat 6900
ctgcctacca agaccaagtt caagttcggc gccatgatga aatctggaat gttcctcaca 6960
ctgtttgtga acacagtcac taacattgta atcgcaagca gagtgttgag agaacggcta 7020
acaggcagcc cttgtgcagc attcattgga gatgacaata tcgtgaaagg agtcaaatac 7080
gacaaattaa tggcagacag gtgcgccacc tggttgaata tggaaagtaa gattatagat 7140
gctgtggtgg gcgagaaaagc gccttatttc tgtggagggt ttattttgtg tgactccgtg 7200
accggcacag cgtgccgtgt ggcagacccc ctaaaaaggc tgtttaagct tggcaaacct 7260
ctggcagcag acgatgaaca tgatgatgac aggagaaggg cattgcatga agagtcaaca 7320
cgctggaacc gagtgggtat tctttcagag ctgtgcaagg cagtagaatc aaggtatgaa 7380
accgtaggaa cttccatcat agttatggcc atgactactc tagctagcag tgttaaatca 7440
ttcagctacc tgagaggggc ccctataact ctctacggct aa 7482

```

<210> 2

<211> 7482

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C02 nsP coding sequence DNA

<400> 2

```

atggagaaaag ttcacgttga catcgaggaa gacagcccat tcctcagagc tttgcagcgg 60
agcttcccg c agtttgaggt agaagccaag caggctactg ataatgacca tgctaattgcc 120
agagcgtttt cgcactctggc ttcaaaactg atcgaaacgg aggtggacct atccgacacg 180
atccttgaca ttggaagtgc gcccggccgc agaattgtatt ctaagcacia gtatcattgt 240
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg 300
aagaaaaact gtaaggaaat aactgataag gaattggaca agaaaatgaa ggagctcgcc 360
gccgtcatga gcgacctga cctggaaact gagactatgt gcctccacga cgacgagtcg 420
tgtcgctacg aagggaaggt cgctgtttac caggatgtat atgccgtgga tggccctaca 480

```

agcctgtacc	accaggccaa	caagggcgtc	agagtggcct	actggatcgg	cttcgacacc	540
acacctttca	tgttcaagaa	cctggctggc	gcttacccca	gctacagcac	aaactgggcc	600
gacgaaaccg	tgtaacggc	tcgtaacata	ggcctatgca	gctctgacgt	tatggagcgg	660
tcacgtagag	ggatgtccat	tcttagaaaag	aagtatttga	aaccatccaa	caatgttcta	720
ttctctgttg	gctcgaccat	ctaccacgag	aagagggact	tactgaggag	ctggcacctg	780
ccgtctgtat	ttcacttacg	tggcaagcaa	aattacacat	gtcgggtgtga	gactatagtt	840
agttgcgacg	ggtagctcgt	taaaagaata	gctatcagtc	caggcctgta	tgggaagcct	900
tcaggctatg	ctgtacgat	gcaccgag	ggattcttgt	gctgcaaagt	gacagacaca	960
ttgaacgggg	agagggctctc	ttttcccg	tgcacgtatg	tgccagctac	attgtgtgac	1020
caaatgactg	gcatactggc	aacagatgtc	agtgcggacg	acgcgcaaaa	actgctggtt	1080
gggctcaacc	agcgtatagt	cgtcaacgg	cgcaccaga	gaaacaccaa	taccatgaaa	1140
aattaccttt	tgcccgtagt	ggcccaggca	tttgctaggt	gggcaaagga	atataaggaa	1200
gatcaagaag	atgaaaggcc	actaggacta	cgagatagac	agttagtcat	ggggtgttgt	1260
tgggctttta	gaaggcacia	gataacatct	atttataagc	gcccggatac	ccaaccatc	1320
atcaaagtga	acagcgattt	ccactcattc	gtgctgcccc	ggataggcag	taacacattg	1380
gagatcgggc	tgagaacaag	aatcaggaaa	atgttagagg	agcacaagga	gccgtcacct	1440
ctcattaccg	ccgaggacgt	acaagaagct	aagtgcgcag	ccgatgaggc	taaggaggtg	1500
cgtgaagccg	aggagttg	cgcagctcta	ccacctttgg	cagctgatgt	tgaggagccc	1560
actctggaag	ccgatgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagt	gaaaaattat	cttgcatcca	ccctctcgt	1740
gaacaagtca	tagtgataac	acactctggc	cgaaaagggc	gttatgccgt	ggaaccatac	1800
catggtaaag	tagtggtgcc	agagggacat	gcaatacccg	tccaggactt	tcaagctctg	1860
agtgaaagtg	ccaccattgt	gtacaacgaa	cgtgagttcg	taaacaggta	cctgcacat	1920
attgccacac	atggaggagc	gctgaacact	gatgaagaat	attacaaaac	tgtcaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtcaactg	ggctagggct	cacaggcgag	ctgggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccag	taccaaccat	agggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatatattga	cgaagctttt	gcttgatcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgacaca	agtcttccac	2520
aaaagcatct	ctcgccgttg	cactaaatct	gtgacttcgg	tcgtctcaac	cttgttttac	2580
gacaaaaaaa	tgagaacgac	gaatccgaaa	gagactaaga	ttgtgattga	cactaccggc	2640
agtaccaaac	ctaagcagga	cgatctcatt	ctcacttggt	tcagagggtg	ggtgaagcag	2700
ttgcaaatag	attacaaaagg	caacgaaata	atgacggcag	ctgcctctca	agggtgacc	2760
cgtaaagggtg	tgtatgccgt	tcggtacaag	gtgaatgaaa	atcctctgta	cgcaccacc	2820

tcagaacatg	tgaacgtcct	actgacccgc	acggaggacc	gcatcgtgtg	gaaaacacta	2880
gccggcgacc	catggataaa	aacactgact	gccaaagtacc	ctgggaattt	cactgccacg	2940
atagaggagt	ggcaagcaga	gcatgatgcc	atcatgaggc	acatcttgga	gagaccggac	3000
cctaccgacg	tcttccagaa	taaggcaaac	gtgtgttggg	ccaaggcttt	agtgccggtg	3060
ctgaagaccg	ctggcataga	catgaccact	gaacaatgga	acactgtgga	ttattttgaa	3120
acggacaaaag	ctcactcagc	agagatagta	ttgaaccaac	tatgcgtgag	gttctttgga	3180
ctcgatctgg	actccggtct	atcttctgca	cccactgttc	cgttatccat	taggaataat	3240
cactgggata	actccccgtc	gcctaacatg	tacgggctga	ataaagaagt	ggtccgtcag	3300
ctctctcgca	ggtacccaca	actgcctcgg	gcagttgcc	ctggaagagt	ctatgacatg	3360
aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	ctttagtcct	ccaccataat	gaacacccac	agagtgactt	ttcttcattc	3480
gtcagcaaat	tgaagggcag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatggttg	actggttgtc	agaccggcct	gaggtacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gaccccatat	3660
aaataccatc	actatcagca	gtgtgaagac	catgccatta	agcttagcat	gttgaccaag	3720
aaagcttgtc	tgcattctgaa	tcccggcgga	acctgtgtca	gcataggtta	tggttacgct	3780
gacagggcca	gcgaaagcat	cattggtgct	atagcgcggc	agttcaagtt	ttcccgggta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	tgggtacgat	3900
cgcaaggccc	gtacgcacaa	tccttacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
gggggtgtgcg	gagcgtgta	taagaaattc	ccggaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgcacg	aacctgatgc	agagctggtg	agggtgcac	cgaagagttc	tttggctgga	4560
aggaagggct	acagcacaa	cgatggcaaa	actttctcat	atttggaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgccttgctt	gtgcatccat	4800
gccatgactc	cagaaaagagt	acagcgcta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgtcatcct	ttccattgcc	gaagtataga	atcactggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatc	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgacccac	5160

caggtgctgc	aagtcgaggc	agacattcac	gggccgcct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgcctccg	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagttttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgctcac	gcactcctag	caggtcggtc	tcgagaacca	gcctgggtctc	caaccgcgcca	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgcctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgtacc	gaaccctgca	tcctgttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatggt	tgacggagct	6120
tcatgtctgt	tagacactgc	cagtttttgc	cctgcaaagc	tgcgcagctt	tccaaagaaa	6180
cactcctatt	tggaaaccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240
cagaacgtcc	tggcagctgc	cacaaaaaga	aattgcaatg	tcacgcaa	gagagaattg	6300
cccgtattgg	attcggcggc	ctttaatgtg	gaatgcttca	agaaatatgc	gtgtaataat	6360
gaatatggg	aaacgtttaa	agaaaacccc	atcaggctta	ctgaagaaaa	cgtggttaaat	6420
tacattacca	aattaaaagg	acaaaaagct	gctgctcttt	ttgcgaagac	acataatttg	6480
aatatgttgc	aggacatacc	aatggacagg	tttgtaatgg	acttaaagag	agacgtgaaa	6540
gtgactccag	gaacaaaaca	tactgaagaa	cggcccaagg	tacaggtgat	ccaggctgcc	6600
gatccgctag	caacagcgta	tctgtgcgga	atccaccgag	agctggttag	gagattaaat	6660
gcggctcctgc	ttccgaacat	tcatacactg	tttgatatgt	cggctgaaga	ctttgacgct	6720
attatagccg	agcacttcca	gcctggggat	tgtgttctgg	aaactgacat	cgcgtcgttt	6780
gataaaaagt	aggacgacgc	catggctctg	accgcgttaa	tgattctgga	agacttaggt	6840
gtggacgcag	agctgttgac	gctgattgag	gcggctttcg	gcgaaatttc	atcaatacat	6900
ttgcccacta	aaactaaatt	taaattcgga	gccatgatga	aatctggaat	gttcctcaca	6960
ctgtttgtga	acacagtcac	taacattgta	atcgcaagca	gagtgttgag	agaacggcta	7020
accggatcac	catgtgcagc	attcattgga	gatgacaata	tcgtgaaagg	agtcaaatcg	7080
gacaaattaa	tggcagacag	gtgcgccacc	tggttgaata	tggaaagtcaa	gattatagat	7140
gctgtggtgg	gcgagaaaagc	gccttatttc	tgtggagggt	ttattttgtg	tgactccgtg	7200
accggcacag	cgtgccgtgt	ggcagacccc	ctaaaaaggc	tgtttaagct	tggcaaacct	7260
ctggcagcag	acgatgaaca	tgatgatgac	aggagaaggg	cattgcatga	agagtcaaca	7320
cgttggaacc	gagtgggtat	tctttcagag	ctgtgcaagg	cagtagaatc	aaggtatgaa	7380
accgtaggaa	cttccatcat	agttatggcc	atgactactc	tagctagcag	tgttaaatca	7440
ttcagctacc	tgagaggggc	ccctataact	ctctacggct	aa		7482

<210>	3	
<211>	7482	
<212>	DNA	
<213>	Artificial Sequence	
<220>		
<223>	Synthetic construct, C03 nsP coding sequence DNA	
<400>	3	
atggagaaaag	ttcacgttga	catcgaggaa gacagcccat tcctcagagc tttgcagcgg 60
agcttccccgc	agtttgaggt	agaagccaag caggtcactg ataatgacca tgctaatgcc 120
agagcgtttt	cgcacatcggc	ttcaaaaactg atcgaaacgg aggtggaccc atccgacacg 180
atccttgaca	ttggaagtgc	gccccgccgc agaatgtatt ctaagcacia gtatcattgt 240
atctgtccga	tgagatgtgc	ggaagatccg gacagattgt ataagtatgc aactaagctg 300
aagaaaaact	gtaaggaaat	aactgataag gaattggaca agaaaatgaa ggagctcgcc 360
gccgtcatga	gcgaccctga	cctggaaaact gagactatgt gcctccacga cgacgagtcg 420
tgctgctacg	aagggaaggt	cgctgtttac caggatgtat acgcggttga cggaccgaca 480
agtctctatc	accaagccaa	taaggaggtt agagtcgcct actggatagg ctttgacacc 540
acccctttta	tgtttaagaa	cttggtctgga gcatatccat catactctac caactgggcc 600
gacgaaaccg	tgtaaacggc	tcgtaacata ggcctatgca gctctgacgt tatggagcgg 660
tcacgtagag	ggatgtccat	tcttagaaaag aagtatttga aaccatccaa caatgttcta 720
ttctctgttg	gctcgaccat	ctaccacgag aagagggact tactgaggag ctggcacctg 780
ccgtctgtat	ttcacttacg	tggcaagcaa aattacacat gtcggtgtga gactatagtt 840
agttgcgacg	ggtacgtcgt	taaaagaata gctatcagtc caggcctgta tgggaagcct 900
tcaggctatg	ctgctacgat	gcaccgcgag ggattcttgt gctgcaaagt gacagacaca 960
ttgaacgggg	agagggtctc	ttttcccggtg tgcacgtatg tgccagctac attgtgtgac 1020
caaatgactg	gcatactggc	aacagatgtc agtgcgagac acgcgcaaaa actgctggtt 1080
gggctcaacc	agcgtatagt	cgtcaacggc cgcaccaga gaaacaccaa taccatgaaa 1140
aattaccttt	tgcccgtagt	ggcccaggca tttgctaggt gggcaaagga atataaggaa 1200
gatcaagaag	atgaaaggcc	actaggacta cgagatagac agttagtcac ggggtgttgt 1260
tgggctttta	gaaggcacia	gataacatct atttataagc gcccggatac ccaaaccatc 1320
atcaaagtga	acagcgattt	ccactcattc gtgctgcccc ggataggcag taacacattg 1380
gagatcgggc	tgagaacaag	aatcaggaaa atgttagagg agcacaagga gccgtcacct 1440
ctcattaccg	ccgaggacgt	acaagaagct aagtgcgcag ccgatgaggc taaggaggtg 1500
cgtgaagccg	aggagttgcg	cgcagctcta ccacctttgg cagctgatgt tgaggagccc 1560
actctggaag	ccgatgtcga	cttgatgtta caagaggctg gggccggctc agtggagaca 1620
cctcgtggct	tgataaaggt	taccagctac gctggcgagg acaagatcgg ctcttacgct 1680
gtgctttctc	cgcaggctgt	actcaagagt gaaaaattat cttgcatcca ccctctcgt 1740
gaacaagtca	tagtgataac	acactctggc cgaaaagggc gttatgccgt ggaaccatac 1800
catggtaaag	tagtgggtcc	agaggacat gcaatacccg tccaggactt tcaagctctg 1860
agtgaaagtg	ccaccattgt	gtacaacgaa cgtgagttcg taaacaggta cctgcaccat 1920

attgccacac	atggaggagc	gctgaacact	gatgaagaat	attacaaaac	tgtcaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtcaactg	ggctagggct	cacaggcgag	ctgggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccaag	taccaaccat	aggggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtgggtgagcg	ccaagaaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatatattga	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgacaca	agtcttccac	2520
aaaagcatct	ctcgccgttg	cactaaatct	gtgacttcgg	tcgtctcaac	cttgttttac	2580
gacaaaaaaaa	tgagaacgac	gaatccgaaa	gagactaaga	ttgtgattga	cactaccggc	2640
agtaccaaac	ctaagcagga	cgatctcatt	ctcacttggt	tcagagggtg	ggtgaagcag	2700
ttgcaaatag	attacaaagg	caacgaaata	atgacggcag	ctgcctctca	agggtgacc	2760
cgtaaagggtg	tgtatgccgt	tcggtacaag	gtgaatgaaa	atcctctgta	cgcaccacc	2820
tcagaacatg	tgaacgtcct	actgaccgcg	acggaggacc	gcatcgtgtg	gaaaacacta	2880
gccggcgacc	catggataaa	aacactgact	gccaaagtacc	ctgggaattt	cactgccacg	2940
atagaggagt	ggcaagcaga	gcatgatgcc	atcatgaggc	acatcttgga	gagaccggac	3000
cctaccgacg	tcttccagaa	taaggcaaac	gtgtgttggg	ccaaggcttt	agtgccggtg	3060
ctgaagaccg	ctggcataga	catgaccact	gaacaatgga	acactgtgga	ttattttgaa	3120
acggacaaaag	ctcactcagc	agagatagta	ttgaaccaac	tatgcgtgag	gttctttgga	3180
ctcgatctgg	actccggtct	attttctgca	cccactgttc	cgttatccat	taggaataat	3240
cactgggata	actccccgtc	gcctaacatg	tacgggctga	ataaagaagt	ggtccgtcag	3300
ctctctcgca	ggtaccacac	actgcctcgg	gcagttgcc	ctggaagagt	ctatgacatg	3360
aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	cttttagtcct	ccaccataat	gaacaccac	agagtgactt	ttcttcattc	3480
gtcagcaaat	tgaagggcag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatggttg	actggttgtc	agaccggcct	gaggctacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gacccctac	3660
aagtaccacc	actaccagca	gtgcgaggac	cacgccatca	agctgagcat	gctgaccaag	3720
aaggcctgcc	tgcactctgaa	tcccggcgga	acctgtgtca	gcataggtta	tggttacgct	3780
gacagggcca	gcgaaaagcat	cattggtgct	atagcgcggc	agttcaagtt	ttcccgggta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	cggctacgac	3900
agaaaaggccc	gtacgcacaa	tcettacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
gggggtgtgcg	gagcgctgta	taagaaattc	ccgaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260

atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgcag	aacctgatgc	agagctgggtg	agggtgcatc	cgaagagttc	tttggctgga	4560
aggaagggct	acagcacaag	cgatggcaaa	actttctcat	atttggaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgccttgctt	gtgcatccat	4800
gccatgactc	cagaaaagag	acagcgcta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgtcatcct	ttccattgcc	gaagtataga	atcactgggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatc	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgaccac	5160
caggtgctgc	aagtcgaggc	agacattcac	gggccgcct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgtccgc	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagttttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgtcac	gcactcctag	caggtcggtc	tcgagaacca	gcctggtctc	caaccgccca	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtgggtgttg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgccctga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgcctacc	gaaccctgca	tctgtttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatgg	tgacggagct	6120
tcatgctgct	tagacactgc	cagtttttgc	cctgcaaagc	tgcgcagctt	tccaaagaaa	6180
cactcctatt	tggaaaccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240
cagaacgtcc	tggcagctgc	cacaaaaaga	aattgcaatg	tcacgcaaat	gagagaattg	6300
cccgtattgg	attcggcggc	ctttaatgtg	gaatgcttca	agaaatatgc	gtgtaataat	6360
gaatatggg	aaacgtttaa	agaaaacccc	atcaggctta	ctgaagaaaa	cgtggttaaat	6420
tacattacca	aattaaaagg	acaaaaagct	gctgctcttt	ttgcgaagac	acataatttg	6480
aatatgttgc	aggacatacc	aatggacagg	tttgtaatgg	acttaaagag	agacgtgaaa	6540
gtgactccag	gaacaaaaca	tactgaagaa	cggcccaagg	tacaggtgat	ccaggctgcc	6600

```

gatccgctag caacagcgta tctgtgcgga atccaccgag agctgggttag gagattaaat 6660
gcggtcctgc ttccgaacat tcatacactg tttgatatgt cggctgaaga ctttgacgct 6720
attatagccg agcacttcca gcctggggat tgtgttcttg aaactgacat cgcgtcgttt 6780
gataaaagtg aggacgacgc catggctctg accgcgttaa tgattctgga agacttaggt 6840
gtggacgcag agctgttgac gctgattgag gcggctttcg gcgaaatttc atcaatacat 6900
ttgcccacta aaactaaatt taaattcgga gccatgatga aatctggaat gttcctcaca 6960
ctgtttgtga acacagtcac taacattgta atcgcaagca gagggttgag agaacggcta 7020
accggatcac catgtgcagc attcattgga gatgacaata tcgtgaaagg agtcaaatac 7080
gacaaattaa tggcagacag gtgcgccacc tgggtgaata tggaagtcaa gattatagat 7140
gctgtgggtg gcgagaaaag gccttatttc tgtggagggt ttattttgtg tgactccgtg 7200
accggcacag cgtgccgtgt ggcagacccc ctaaaaggc tgtttaagct tggcaaacct 7260
ctggcagcag acgatgaaca tgatgatgac aggagaaggc cattgcatga agagtcaaca 7320
cgctggaacc gagggttat tctttcagag ctgtgcaagg cagtagaatc aaggtatgaa 7380
accgtaggaa cttccatcat agttatggc atgactactc tagctagcag tgttaaatca 7440
ttcagctacc tgagaggggc ccctataact ctctacggct aa 7482

```

<210> 4

<211> 7482

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C04 nsP coding sequence DNA

<400> 4

```

atggagaaag ttcacgttga catcgaggaa gacagcccat tcctcagagc tttgcagcgg 60
agcttccccg agtttgaggt agaagccaag caggctactg ataatgacca tgctaattgc 120
agagcggttt cgcatctggc ttcaaaactg atcgaaacgg aggtggaccc atccgacagc 180
atccttgaca ttggaagtgc gcccggccgc agaattgtatt ctaagcacia gtatcattgt 240
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg 300
aagaaaaact gtaaggaaat aactgataag gaattggaca agaaaatgaa ggagctcgcc 360
gccgtcatga gcgaccctga cctggaaact gagactatgt gcctccacga cgacgagtcg 420
tgtcgctacg aagggaagt cgctgtttac caggatgtat acgcggttga cggaccgaca 480
agtctctatc accaagccaa taaggaggtt agagtcgcct actggatagg ctttgacacc 540
acccctttta tgtttaagaa cttggctgga gcataatcat catactctac caactgggcc 600
gacgaaaccg tgtaaacggc tcgtaacata ggcctatgca gctctgacgt tatggagcgg 660
tcacgtagag ggatgtccat tcttagaaag aagtatttga aaccatccaa caatgttcta 720
ttctctgttg gctcgacat ctaccacgag aagagggact tactgaggag ctggcacctg 780
ccgtctgtat ttacttacg tggcaagcaa aattacacat gtcggtgtga gactatagtt 840
agttgcgacg ggtacgtcgt taaaagaata gctatcagtc caggcctgta tgggaagcct 900
tcaggctatg ctgctacgat gcaccgcgag ggattcttgt gctgcaaagt gacagacaca 960
ttgaacgggg agagggtctc ttttcccgtg tgcacgtatg tgccagctac attgtgtgac 1020

```

caa	atg	actg	gcata	ctggc	aac	agatg	tc	agt	gcgg	acg	cga	aaaa	act	gct	ggtt	1080		
ggg	ctca	aacc	agc	gtatag	t	cgt	caac	gg	cgc	acc	caga	gaa	acac	taa	tac	atgaaa	1140	
aatt	accttt	t	gccc	gtagt	g	ggc	ccagg	ca	ttt	gctag	gt	ggg	caa	agga	atata	aaggaa	1200	
gat	caaga	aag	atg	aaagg	cc	act	agg	acta	c	gag	atag	ac	agt	tag	tcat	ggggtgt	1260	
tggg	ctttta	ga	agg	cacaa	gata	acat	ct	att	tata	agc	gccc	gata	c	caa	accat	c	1320	
atc	aaagt	ga	acag	cgttt	cc	act	catt	c	gtg	ctg	ccca	gga	tag	ggcag	ta	acacatt	1380	
gag	atcg	gggc	tgaga	acaag	aat	cagg	aaa	atg	ttag	agg	ag	caca	agga	gcc	gtcac	ct	1440	
ctc	attacc	g	ccg	agg	acg	t	aca	aga	ag	ct	aagt	gcg	cag	cc	gatg	aggc	taaggagg	1500
cgt	gaag	ccg	agg	agtt	gcg	cgc	ag	ct	ct	t	gg	cag	ct	gat	gt	t	gaggag	1560
act	ctg	gaag	ccg	atg	tcga	ctt	gat	gtta	ca	ag	agg	ctg	ggg	ccg	gctc	agt	ggag	1620
cct	ctg	ggct	tgata	aaag	t	tacc	ag	ctac	g	ctg	ggc	gagg	aca	ag	atcg	g	ctcttac	1680
gtg	ctttctc	cgc	ag	gctgt	act	caag	ag	t	gaaaa	att	at	ctt	g	cat	cca	ccctctc	gct	1740
ga	acaag	tca	tag	tgata	ac	a	act	ctgg	c	gaaa	agg	gc	gtt	atg	ccg	gga	accata	1800
cat	ggt	aaag	tag	tgg	tgc	ag	agg	gac	at	g	caat	acc	g	tcc	agg	actt	tcaagct	1860
agt	gaa	ag	t	ccacc	attgt	gt	taca	acg	aa	cgt	gag	ttcg	t	aa	acag	gta	cctgc	1920
att	gcc	acac	atg	gagg	agc	g	ct	ga	ac	act	gat	ga	aga	at	att	acaaa	c	1980
agc	gag	cacg	acg	gcga	ata	cct	gt	acg	ac	atc	gac	agga	aac	agt	gcg	t	caagaa	2040
ctag	t	actg	ggct	agg	gct	cac	agg	c	gag	ctggt	ggat	c	ctc	c	ttcca	tga	attcg	2100
tac	gag	ag	t	tgaga	acacg	acc	ag	ccg	ct	c	ttac	caag	t	acca	acc	at	agggtgt	2160
ggc	gtg	ccag	gat	cagg	caa	gt	ctg	gcat	c	att	aaa	agc	cag	t	cac	caa	aaa	2220
gtg	gtg	agc	g	cca	agaa	ga	aa	act	gtg	ca	gaa	att	tata	g	ggg	acg	tcaa	2280
ggg	ctg	gacg	t	caat	gccag	a	act	gtg	gac	t	cag	t	gct	t	tga	atg	gatg	2340
gtag	ag	accc	tgt	atatt	ga	cga	ag	cttt	g	ctt	gt	cat	g	cag	gt	act	ct	2400
atag	cc	atta	taag	ac	ctaa	aa	agg	cag	t	g	ct	gcg	ggg	at	cc	caa	aca	2460
ttta	ac	atga	tgt	gc	ctgaa	agt	gc	atttt	aa	cc	ac	gaga	ttt	g	ca	aca	agt	2520
aaa	ag	cat	ctc	gcc	gtt	c	act	aa	at	ct	gt	act	tcg	g	t	ct	caac	2580
gac	aaaa	aaaa	tgaga	acg	ac	ga	at	ccg	aaa	gag	act	aaga	ttg	t	gatt	ga	cact	2640
agt	acca	aaac	cta	ag	cagga	c	gat	ct	catt	ct	c	act	ttg	t	tcag	agg	gtg	2700
ttg	caaat	ag	att	acaaa	agg	ca	acg	aa	ata	atg	ac	ggc	ag	ctg	c	ct	ctca	2760
cgt	aaag	gtg	tgt	atg	ccg	t	cg	gt	taca	ag	gtg	aat	gaaa	at	c	ct	ctg	2820
tcaga	ac	atg	tga	acg	tc	ct	act	g	accc	gc	acg	gagg	ac	g	cat	cg	tgtg	2880
gcc	ggc	gac	cat	ggata	aaa	a	a	act	g	act	g	ccaa	gtac	ctg	gga	attt	cact	2940
atag	agg	ag	gg	caag	caga	g	cat	gat	gcc	at	cat	gagg	c	ac	at	ctt	gga	3000
cct	acc	gacg	t	ctt	ccagaa	ta	agg	caa	ac	gtg	tgt	ttgg	g	cca	agg	cttt	agt	3060
ctga	ag	accg	ctg	gcat	aga	cat	g	acc	act	ga	aca	atgga	a	ac	act	gtgga	t	3120
acg	gac	aaaag	ctc	act	cagc	ag	ag	atag	ta	ttg	a	acca	ac	tat	gcg	tgag	gtt	3180
ctc	gat	ctgg	act	ccg	gtct	at	ttt	ctg	ca	ccc	act	gttc	cgt	t	at	ccat	tag	3240
cact	ggg	gata	act	cccc	gtc	gc	cta	ac	atg	tac	ggg	ctga	ata	aa	aga	agt	ggt	3300
ctct	ctc	gca	ggt	accc	caca	act	gc	ctcg	g	gc	agt	tgcca	ctg	ga	ag	agt	ctat	3360

aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	ctttagtcct	ccaccataat	gaacacccac	agagtgactt	ttcttcatte	3480
gtcagcaaat	tgaagggcag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatggttg	actggttgtc	agaccggcct	gaggctacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gaccccatat	3660
aaataccatc	actatcagca	gtgtgaagac	catgccatta	agcttagcat	gttgaccaag	3720
aaagcttgtc	tgcatctgaa	tcccggcgga	acctgtgtca	gcataggtta	tggttacgct	3780
gacagggcca	gcgaaaagcat	cattgggtgct	atagcgcggc	agttcaagtt	ttcccgggta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	tggttacgat	3900
cgcaaggccc	gtacgcacaa	cccctacaag	ctgagcagca	ccctgaccaa	catctacacc	3960
ggcagcagac	tgcacgaagc	cggatgtgca	ccctcatatc	atgtgggtgcg	aggggatatt	4020
gccacagcca	cagaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
gggggtgtgcg	gagcgtgta	taagaaattc	ccggaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtacag	aacctgatgc	agagctggtg	agggtgcatc	cgaagagttc	tttggctgga	4560
aggaagggct	acagcacaa	cgatggcaaa	actttctcat	atttgaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgcccttgctt	gtgcatccat	4800
gccatgactc	cagaaagagt	acagcgccta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgctcatcct	ttccattgcc	gaagtataga	atcactgggtg	tgcaagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatte	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgacccac	5160
caggtgctgc	aagtcgaggc	agacattcac	gggccgccct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgtccgc	gcacaagaac	accgtcactt	gcacccagca	gggcctgctc	gagaaccagc	5460
ctagtttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgtcac	gcactcctag	caggtcggtc	tcgagaacca	gcctggtctc	caacccgcca	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700

aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgcctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatccccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgtctacc	gaaccctgca	tcctgttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatggt	tgacggagct	6120
tcatgtctgct	tagacactgc	cagtttttgc	cctgcaaagc	tgcgagctt	tccaaagaaa	6180
cactcctatt	tggaacccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240
cagaacgtcc	tggcagctgc	cacaaaaaga	aattgcaatg	tcacgcaa	gagagaattg	6300
cccgtattgg	attcggcggc	ctttaatgtg	gaatgcttca	agaaatatgc	gtgtaataat	6360
gaatattggg	aaacgtttaa	agaaaacccc	atcaggetta	ctgaagaaaa	cgtggtaaat	6420
tacattacca	aattaaaagg	accaaaagct	gctgctcttt	ttgcgaagac	acataatttg	6480
aatatgttgc	aggacatacc	aatggacagg	tttghtaatgg	acttaaagag	agacgtgaaa	6540
gtgactccag	gaacaaaaca	tactgaagaa	cggcccaagg	tacaggtgat	ccaggctgcc	6600
gatccgctag	caacagcgta	tctgtgcgga	atccaccgag	agctggttag	gagattaaat	6660
gcggctctgc	ttccgaacat	tcatacactg	tttgatatgt	cggctgaaga	ctttgacgct	6720
attatagccg	agcacttcca	gcctggggat	tgtgttcttg	aaactgacat	cgcgtcgttt	6780
gataaaagtg	aggacgacgc	catggctctg	accgcgttaa	tgattctgga	agacttaggt	6840
gtggacgcag	agctgttgac	gctgattgag	gcggctttcg	gcgaaatttc	atcaatacat	6900
ttgcccacta	aaactaaatt	taaattcgga	gccatgatga	aatctggaat	gttcctcaca	6960
ctgtttgtga	acacagtcac	taacattgta	atcgcaagca	gagtgttgag	agaacggcta	7020
accggatcac	catgtgcagc	attcattgga	gatgacaata	tcgtgaaagg	agtcaaatcg	7080
gacaaattaa	tggcagacag	gtgcgccacc	tggttgaata	tggaagtcaa	gattatagat	7140
gctgtggtgg	gcgagaaagc	gccttatttc	tgtggagggt	ttattttgtg	tgactccgtg	7200
accggcacag	cgtgccgtgt	ggcagacccc	ctaaaaaggc	tgtttaagct	tggcaaacct	7260
ctggcagcag	acgatgaaca	tgatgatgac	aggagaaggg	cattgcatga	agagtcaaca	7320
cgttggaacc	gagtgggtat	tctttcagag	ctgtgcaagg	cagtagaatc	aaggtatgaa	7380
accgtaggaa	cttccatcat	agttatggcc	atgactactc	tagctagcag	tgttaaatca	7440
ttcagctacc	tgagaggggc	ccctataact	ctctacggct	aa		7482
<210>	5					
<211>	7482					
<212>	DNA					
<213>	Artificial Sequence					
<220>						
<223>	Synthetic construct, C05 nsP coding sequence DNA					
<400>	5					
atggagaaag	ttcacgttga	catcgaggaa	gacagcccat	tcctcagagc	tttgcagcgg	60
agcttcccg	agtttgaggt	agaagccaag	caggtcactg	ataatgacca	tgctaattgcc	120

agagcgtttt	cgcatctggc	ttcaaaactg	atcgaaacgg	aggtggaccc	atccgacacg	180
atccttgaca	ttggaagtgc	gccccccgc	agaatgtatt	ctaagcacia	gtatcattgt	240
atctgtccga	tgagatgtgc	ggaagatccg	gacagattgt	ataagtatgc	aactaagctg	300
aagaaaaact	gtaaggaaat	aactgataag	gaattggaca	agaaaatgaa	ggagctcgcc	360
gccgtcatga	gcgaccctga	cctggaaact	gagactatgt	gcctccacga	cgacgagtcg	420
tgtcgctacg	aagggaagt	cgctgtttac	caggatgtat	acgcggttga	cggaccgaca	480
agtctctatc	accaagccaa	taaggaggtt	agagtcgcct	actggatagg	ctttgacacc	540
acccctttta	tgtttaagaa	cttggctgga	gcatatccat	catactctac	caactgggcc	600
gacgaaaccg	tgtaaacggc	tcgtaacata	ggcctatgca	gctctgacgt	tatggagcgg	660
tcacgtagag	ggatgtccat	tcttagaaag	aagtatttga	aaccatccaa	caatgttcta	720
ttctctgttg	gctcgaccat	ctaccacgag	aagagggact	tactgaggag	ctggcacctg	780
ccgtctgtat	ttcacttacg	tggcaagcaa	aattacacat	gtcgggtgtga	gactatagtt	840
agttgcgacg	ggtagtcgt	taaaagaata	gctatcagtc	caggcctgta	tgggaagcct	900
tcaggctatg	ctgctacgat	gcaccgcgag	ggattcttgt	gctgcaaagt	gacagacaca	960
ttgaacgggg	agagggtctc	ttttcccggt	tgcacgtatg	tgccagctac	attgtgtgac	1020
caaataactg	gcatactggc	aacagatgtc	agtgcggacg	acgcgcaaaa	actgctggtt	1080
gggctcaacc	agcgtatagt	cgtaacgggt	cgcaccacga	gaaacaccaa	taccatgaaa	1140
aattaccttt	tgcccgtagt	ggcccaggca	tttgctaggt	gggcaaagga	atataaggaa	1200
gatcaagaag	atgaaaggcc	actaggacta	cgagatagac	agttagtcat	ggggtgttgt	1260
tgggctttta	gaaggcacia	gataacatct	atttataagc	gcccggatac	ccaaaccatc	1320
atcaaagtga	acagcgattt	ccactcattc	gtgctgcccc	ggataggcag	taacacattg	1380
gagatcgggc	tgagaacaag	aatcaggaaa	atgttagagg	agcacaagga	gccgtcacct	1440
ctcattaccg	ccgaggacgt	acaagaagct	aagtgcgcag	ccgatgaggc	taaggaggtg	1500
cgtgaagccg	aggagttgcg	cgcagctcta	ccacctttgg	cagctgatgt	tgaggagccc	1560
actctggaag	ccgatgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagc	gagaagctga	gctgcattca	ccctctggcc	1740
gagcaagtga	tcgtgatcac	acacagcggc	cggaaaggga	gatatgccgt	ggaaccttat	1800
cacggcaagg	tgggtgtgcc	tgaggacac	gctattccag	tgcaggactt	tcaggccctg	1860
agcgagtctg	ccaccatcgt	gtacaacgag	cgcgagttcg	tgaacagata	cctgcaccac	1920
attgccacac	acggcggagc	cctgaacacc	gacgaagagt	actacaagac	cgtgaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtcaactg	ggctagggtc	cacaggcgag	ctggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccag	taccaaccat	aggggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggat	agcgtgctgc	tgaacggctg	caagcacccc	2340
gtagagaccc	tgtatatgtga	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcctaagca	gtgcggcttc	2460

ttcaacatga	tgtgcctgaa	ggtgcacttc	aaccacgaga	tctgcacca	ggtgttccac	2520
aaaagcatct	ctcgccgttg	cactaaatct	gtgacttcgg	tcgtctcaac	cttgttttac	2580
gacaaaaaaaa	tgagaacgac	gaatccgaaa	gagactaaga	ttgtgattga	cactaccggc	2640
agtaccaaac	ctaagcagga	cgatctcatt	ctcacttggt	tcagagggtg	ggtgaagcag	2700
ttgcaaatag	attacaaagg	caacgaaata	atgacggcag	ctgcctctca	agggtgacc	2760
cgtaaagggtg	tgtatgccgt	tcggtacaag	gtgaatgaaa	atcctctgta	cgcaccacc	2820
tcagaacatg	tgaacgtcct	actgacccgc	acggaggacc	gcatcgtgtg	gaaaacacta	2880
gccggcgacc	catggataaa	aacactgact	gccaaagtacc	ctgggaattt	cactgccacg	2940
atagaggagt	ggcaagcaga	gcatgatgcc	atcatgaggc	acatcttggg	gagaccggac	3000
cctaccgacg	tcttccagaa	taaggcaaac	gtgtgttggg	ccaaggcttt	agtgccggtg	3060
ctgaagaccg	ctggcataga	catgaccact	gaacaatgga	acactgtgga	ttattttgaa	3120
acggacaaaag	ctcactcagc	agagatagta	ttgaaccaac	tatgcgtgag	gttctttgga	3180
ctcgaatctg	actccgggtc	atcttctgca	cccactgttc	cgttatccat	taggaataat	3240
cactgggata	actccccgtc	gcctaacatg	tacgggctga	ataaagaagt	ggtccgtcag	3300
ctctctcgca	ggtacccaca	actgcctcgg	gcagttgcc	ctggaagagt	ctatgacatg	3360
aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	ctttagtcct	ccaccataat	gaacacccac	agagtgactt	ttcttcattc	3480
gtcagcaaat	tgaagggcag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatgggtg	actggttgtc	agaccggcct	gaggctacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gaccccatat	3660
aaataccatc	actatcagca	gtgtgaagac	catgccatta	agcttagcat	gttgaccaag	3720
aaagcttgtc	tgcatctgaa	tcccggcgga	acctgtgtca	gcataggtta	tggttacgct	3780
gacagggcca	gcgaaagcat	cattggtgct	atagcgcggc	agttcaagtt	ttcccggtta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	tgggtacgat	3900
cgcaaggccc	gtacgcacaa	tccttacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
ggggtgtgcg	gagcgtgtga	taagaaattc	ccgaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaagggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaagggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgacag	aacctgatgc	agagctggtg	agggtgcatc	cgaagagttc	tttggctgga	4560
aggaagggtg	acagcacaag	cgatggcaaa	actttctcat	atttggaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgccttgctt	gtgcatccat	4800

gccatgactc	cagaaagagt	acagcgccta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgtctatcct	ttccattgcc	gaagtataga	atcactgggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatctc	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgaccac	5160
cagggtgctgc	aagtcgaggc	agacattcac	gggccgccct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgtccgc	gcacaagaac	accgtcactt	gcaccagca	gggctgtctc	gagaaccagc	5460
ctagtttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgctac	gcactcctag	caggtcggtc	tcgagaacca	gcctgggtctc	caaccgcga	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtgggtgttg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgcctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgtctacc	gaaccctgca	tcctgttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggctgcagt	ggaagcctgt	aacgccatgt	tgaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatgg	tgacggagct	6120
tcatgtctgct	tagacactgc	cagtttttgc	cctgcaaagc	tgcgagctt	tccaaagaaa	6180
cactcctatt	tggaaccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240
cagaacgtcc	tggcagctgc	cacaaaaaga	aattgcaatg	tcacgcaaat	gagagaattg	6300
cccgtattgg	attcggcggc	ctttaatgtg	gaatgcttca	agaaatatgc	gtgtaataat	6360
gaatattggg	aaacgtttta	agaaaacccc	atcaggetta	ctgaagaaaa	cgtggttaaat	6420
tacattacca	aattaaaagg	accaaaagct	gctgctcttt	ttgcgaagac	acataatttg	6480
aatatgttgc	aggacatacc	aatggacagg	tttgtaatgg	acttaaagag	agacgtgaaa	6540
gtgactccag	gaacaaaaac	tactgaagaa	cggcccaagg	tacaggtgat	ccaggctgcc	6600
gatccgctag	caacagcgta	tctgtgcgga	atccaccgag	agctggttag	gagattaaat	6660
gcggtcctgc	ttccgaacat	tcatacactg	tttgatatgt	cggctgaaga	ctttgacgct	6720
attatagccg	agcatttcca	gcctggggat	tgtgttctgg	aaactgacat	cgcgtcgttt	6780
gataaaaagt	aggacgacgc	catggctctg	accgcgttaa	tgattctgga	agacttaggt	6840
gtggacgcag	agctgttgac	gctgattgag	gcggctttcg	gcgaaatttc	atcaatacat	6900
ttgcccacta	aaactaaatt	taaattcgga	gccatgatga	aatctggaat	gttcttcaca	6960
ctgtttgtga	acacagtcac	taacattgta	atcgcaagca	gagtgttgag	agaacggcta	7020
accggatcac	catgtgcagc	attcattgga	gatgacaata	tcgtgaaagg	agtcaaatcg	7080
gacaaaattaa	tggcagacag	gtgcgccacc	tggttgaata	tggaagtcaa	gattatagat	7140

gctgtggtgg gcgagaaagc gccttatttc tgtggagggt ttattttgtg tgactccgtg	7200
accggcacag cgtgccgtgt ggagacccc ctaaaaaggc tgtttaagct tggcaaacct	7260
ctggcagcag acgatgaaca tgatgatgac aggagaaggc cattgcatga agagtcaaca	7320
cgctggaacc gagtgggtat tctttcagag ctgtgcaagg cagtagaatc aaggtatgaa	7380
accgtaggaa cttccatcat agttatggcc atgactactc tagctagcag tgttaaataca	7440
ttcagctacc tgagaggggc ccctataact ctctacggct aa	7482

<210> 6

<211> 7482

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C06 nsP coding sequence DNA

<400> 6

atggagaaag ttcacgttga catcgaggaa gacagcccat tcctcagagc tttgcagcgg	60
agcttcccgc agtttgaggt agaagccaag caggtcactg ataatgacca tgctaatagcc	120
agagcgTTTT cgcattctggc ttcaaaactg atcgaaacgg aggtggaccc atccgacacg	180
atccttgaca ttggaagtgc gcccggccgc agaattgtatt ctaagcacia gtatcattgt	240
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg	300
aagaaaaact gtaaggaaat aactgataag gaattggaca agaaaatgaa ggagctcgcc	360
gccgtcatga gcgaccctga cctggaaact gagactatgt gcctccacga cgacgagtcg	420
tgctgctacg aagggaagt cgctgtttac caggatgtat acgcggttga cggaccgaca	480
agtctctatc accaagccaa taaggaggtt agagtcgcct actggatagg ctttgacacc	540
acccctttta tgtttaagaa cttggctgga gcatatccat catactctac caactgggcc	600
gacgaaaccg tgtaaacggc cagaaatata ggctgtgtga gcagcgacgt gatggaaaga	660
tccagacggg gcatgagcat cctgcggaag aagtacctga agcctagcaa caacgtgctg	720
ttcagcgtgg gcagcaccat ctaccacgag aagagggacc tgctgcggag ctggcatctg	780
ccttccgtgt ttcacctgag aggcaagcag aactacacct gtagatgca gacaatcgtg	840
tcctgcgacg gctacgtggt caagcggatc gccatttctc ctggcctgta cggcaagcct	900
tctggctatg ccgccaccat gcacagagaa ggctttctgt gttgcaaagt gaccgacaca	960
ctgaacggcg agcgggtgtc ctttctctgt tgtacctatg tgcccggcac actgtgcat	1020
cagatgacag gcattctggc caccgacgtg tcagccgacg atgcccagaa actgctcgtg	1080
ggcctgaacc agagaatcgt ggtcaacggc agaaccacg ggaacaccaa caccatgaag	1140
aactacctgc tgccctgtgt ggcccaggcc tttgccagat gggccaaaga gtacaaagag	1200
gatcaagagg acgagcggcc cctgggcctg agagatagac aactggtcat gggctgctgc	1260
tgggccttca gaaggcacia gatcaccagc atctacaagc ggcccagac acagaccatc	1320
atcaaagtga acagcgactt ccacagcttc gtgctgcctc ggatcggcag caacacactg	1380
gaaatcggcc tgcggacccg gatcagaaa atgctggaag aacacaaaga gccctctcca	1440
ctgatcaccg ccgaggatgt gcaagaggcc aaatgtgccg ccgacgaggc taaagaagtg	1500
cgcgaagccg aggaactgag agccgcactt cctcctctgg ccgccgatgt tgaagaaccc	1560

actctggaag	ccgacgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagt	gaaaaattat	cttgcatcca	ccctctcgct	1740
gaacaagtca	tagtgataac	acactctggc	cgaaaagggc	gttatgccgt	ggaaccatac	1800
catggtaaag	tagtgggtgcc	agagggacat	gcaatacccg	tccaggactt	tcaagctctg	1860
agtgaaaagt	ccaccattgt	gtacaacgaa	cgtgagttcg	taaacaggta	cctgcaccat	1920
attgccacac	atggaggagc	gctgaacact	gatgaagaat	attacaaaac	tgtcaagccc	1980
agcgagcacg	acggcgaaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtactctg	ggctagggct	cacaggcgag	ctgggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccaag	taccaacat	aggggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatatgtga	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgcacaca	agtcttccac	2520
aaaagcatct	ctcgccgttg	cactaaatct	gtgacttcgg	tcgtctcaac	cttgttttac	2580
gacaaaaaaaa	tgagaacgac	gaatccgaaa	gagactaaga	ttgtgattga	cactaccggc	2640
agtaccaaac	ctaagcagga	cgatctcatt	ctcacttggt	tcagagggtg	ggtgaagcag	2700
ttgcaaatag	attacaaagg	caacgaaata	atgacggcag	ctgcctctca	agggtgacc	2760
cgtaaagggtg	tgtatgccgt	tcggtacaag	gtgaatgaaa	atcctctgta	cgcacccacc	2820
tcagaacatg	tgaacgtcct	actgaccgcg	acggaggacc	gcatcgtgtg	gaaaacacta	2880
gccggcgacc	catggataaa	aacactgact	gccaaagtacc	ctgggaattt	cactgccacg	2940
atagaggagt	ggcaagcaga	gcatgatgcc	atcatgaggc	acatcttgga	gagaccggac	3000
cctaccgacg	tcttccagaa	taaggcaaac	gtgtgttggtg	ccaaggcttt	agtgccggtg	3060
ctgaagaccg	ctggcataga	catgaccact	gaacaatgga	acactgtgga	ttattttgaa	3120
acggacaaaag	ctcactcagc	agagatagta	ttgaaccaac	tatgcgtgag	gttctttgga	3180
ctcgatctgg	actccggtct	attttctgca	cccactgttc	cgttatccat	taggaataat	3240
cactgggata	actccccgtc	gcctaacatg	tacgggctga	ataaagaagt	ggtccgtcag	3300
ctctctcgca	ggtaaccaca	actgcctcgg	gcagttgcc	ctggaagagt	ctatgacatg	3360
aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	cttttagtcct	ccaccataat	gaacacccac	agagtgactt	ttcttcattc	3480
gtcagcaaat	tgaagggcag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatggttg	actggttgtc	agaccggcct	gaggtacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gacccatat	3660
aaataaccatc	actatcagca	gtgtgaagac	catgccatta	agcttagcat	gttgaccaag	3720
aaagcttgtc	tgcacttgaa	tcccggcgga	acctgtgtca	gcataggtta	tggttacgct	3780
gacagggcca	gcgaaaagcat	cattgggtgct	atagcgcggc	agttcaagtt	ttcccgggta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	tgggtacgat	3900

cgcaaggccc	gtacgcacaa	tccttacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
ggggtgtgcg	gagcgctgta	taagaaattc	ccggaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaagggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgacag	aacctgatgc	agagctggtg	agggtgcac	cgaagagttc	tttggctgga	4560
aggaagggt	acagcacaa	cgatggcaaa	actttctcat	atttggaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgccttgctt	gtgcatccat	4800
gccatgactc	cagaaagagt	acagcgcta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgtcatcct	ttccattgcc	gaagtataga	atcactggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatc	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgaccac	5160
caggtgctgc	aagtcgaggc	agacattcac	gggccgcct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaaacag	tattcaggaa	ccctccacat	5400
cccgtccgc	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagttttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgctcac	gcactcctag	caggtcggtc	tcgagaacca	gcctggtctc	caaccgcga	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cggtgtcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgccctga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgetacc	gaaccctgca	tcctgttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatggt	tgacggagct	6120
tcatgctgct	tagacactgc	cagtttttgc	cctgcaaagc	tgcgcagctt	tccaaagaaa	6180
cactcctatt	tggaaccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240

```

cagaacgtcc tggcagctgc cacaaaaaga aattgcaatg tcacgcaa at gagagaattg 6300
cccgtatttg attcggcggc ctttaatgtg gaatgcttca agaaatatgc gtgtaataat 6360
gaatattggg aaacgtttta agaaaacccc atcaggctta ctgaagaaaa cgtggtaaat 6420
tacattacca aattaaaagg accaaaagct gctgctcttt ttgcgaagac acataatttg 6480
aatatgttgc aggacatacc aatggacagg tttgtaatgg acttaaagag agacgtgaaa 6540
gtgactccag gaacaaaaca tactgaagaa cggcccaagg tacaggatgat ccaggctgcc 6600
gatccgctag caacagcgtg tctgtgcgga atccaccgag agctggtagg gagattaaat 6660
gcggtcctgc ttccgaacat tcatacactg tttgatattg cggctgaaga ctttgacgct 6720
attatagccg agcacttcca gcctggggat tgtgttctgg aaactgacat cgcgtcgttt 6780
gataaaaagt aggacgacgc catggctctg accgcgttaa tgattctgga agacttaggt 6840
gtggacgcag agctgttgac gctgattgag gcggctttcg gcgaaatttc atcaatacat 6900
ttgcccacta aaactaaatt taaattcgga gccatgatga aatctggaat gttcctcaca 6960
ctgtttgtga acacagtcac taacattgta atcgcaagca gagtgttgag agaacggcta 7020
accggatcac catgtgcagc attcattgga gatgacaata tcgtgaaagg agtcaaatcg 7080
gacaaattaa tggcagacag gtgcgccacc tggttgaata tggaagtcaa gattatagat 7140
gctgtgggtg gcgagaaagc gccttatttc tgtggagggt ttattttgtg tgactccgtg 7200
accggcacag cgtgccgtgt ggacagacccc ctaaaaaggc tgtttaagct tggcaaacct 7260
ctggcagcag acgatgaaca tgatgatgac aggagaaggg cattgcatga agagtcaaca 7320
cgctggaacc gagtgggtat tctttcagag ctgtgcaagg cagtagaatc aaggtatgaa 7380
accgtaggaa cttccatcat agttatggc atgactactc tagctagcag tgttaaatca 7440
ttcagctacc tgagaggggc ccctataact ctctacggct aa 7482

```

<210> 7

<211> 7482

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C07 nsP coding sequence DNA

<400> 7

```

atggagaaaag ttcacgttga catcgaggaa gacagcccat tcctcagagc tttgcagcgg 60
agcttcccg agtttgagggt agaagccaag caggctactg ataatgacca tgctaattgcc 120
agagcgtttt cgcacatcggc ttcaaaactg atcgaaacgg aggtggaccc atccgacacg 180
atccttgaca ttggaagtgc gcccgccgc agaatgtatt ctaagcaca gtatcattgt 240
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg 300
aagaaaaact gtaaggaaat aactgataag gaattggaca agaaaatgaa ggagctcgcc 360
gccgtcatga gcgaccctga cctggaaact gagactatgt gcctccacga cgacgagtcg 420
tgtcgctacg aagggcaagt cgctgtttac caggatgtat acgcggttga cggaccgaca 480
agtctctatc accaagccaa taaggaggt agagtcgcct actggatagg ctttgacacc 540
acccctttta tgtttaagaa cttggctgga gcatatccat catactctac caactgggcc 600
gacgaaaccg tgtaacggc tcgtaacata ggcctatgca gctctgacgt tatggagcgg 660

```

tcacgtagag	ggatgtccat	tcttagaaa	aagtatttga	aaccatccaa	caatgttcta	720
ttctctgttg	gctcgacat	ctaccacgag	aagagggact	tactgaggag	ctggcacctg	780
ccgtctgtat	ttcacttacg	tggcaagcaa	aattacacat	gtcgggtgtga	gactatagtt	840
agttgcgacg	ggtacgtcgt	taaaagaata	gctatcagtc	caggcctgta	tgggaagcct	900
tcaggctatg	ctgctacgat	gcaccgcgag	ggattcttgt	gctgcaaagt	gacagacaca	960
ttgaacgggg	agagggtctc	ttttcccgtg	tgcacgtatg	tgccagctac	attgtgtgac	1020
caaatgactg	gcatactggc	aacagatgtc	agtgcggacg	acgcgcaaaa	actgctggtt	1080
gggctcaacc	agcgtatagt	cgtcaacggg	cgcacccaga	gaaacaccaa	taccatgaaa	1140
aattaccttt	tgcccgtagt	ggcccaggca	tttgctaggt	gggcaaagga	atataaggaa	1200
gatcaagaag	atgaaaaggc	actaggacta	cgagatagac	agttagtcat	gggggtgttg	1260
tgggctttta	gaaggcacia	gataacatct	atttataagc	gcccggatac	ccaaaccatc	1320
atcaaagtga	acagcgattt	ccactcattc	gtgctgcccc	ggataggcag	taacacattg	1380
gagatcgggc	tgagaacaag	aatcaggaaa	atgttagagg	agcacaagga	gccgtcacct	1440
ctcattaccg	ccgaggacgt	acaagaagct	aagtgcgcag	ccgatgaggc	taaggaggtg	1500
cgtgaagccg	aggagttgcg	cgcagctcta	ccacctttgg	cagctgatgt	tgaggagccc	1560
actctggaag	ccgatgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagt	gaaaaattat	cttgcatcca	ccctctcgct	1740
gaacaagtca	tagtgataac	acactctggc	cgaaaagggc	gttatgccgt	ggaaccatac	1800
catggtaaag	tagtggtgcc	agaggacat	gcaatacccg	tccaggactt	tcaagctctg	1860
agtgaagtg	ccaccattgt	gtacaacgaa	cgtgagttcg	taaacaggta	cctgcaccat	1920
attgccacac	atggaggagc	gctgaacact	gatgaagaat	attacaaaac	tgtcaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtcactg	ggctagggct	cacaggcgag	ctggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccag	taccaaccat	agggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatattga	cgaagctttt	gcttgtcatg	caggctactt	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgacaca	agtcttccac	2520
aagagcatca	gcagacgggt	caccaagagc	gtgaccagcg	tgggtgtctac	cctgttctac	2580
gacaagaaga	tgcgagacg	aaaccccaaa	gagacaaaga	tcgtcatcga	caccaccggc	2640
agcaccaagc	ctaagcagga	cgatctgac	ctgacctgct	tcagaggctg	ggtcaagcag	2700
ctgcagatcg	actacaaggg	caacgagatc	atgaccgccg	ctgcctctca	gggcctgaca	2760
agaaaaggcg	tgtacgccgt	gcggtacaaa	gtgaacgaga	accctctgta	cgcccttacc	2820
agcgagcatg	tgaatgtgct	gctgaccggg	accgaggacc	ggatcgtttg	gaaaacactg	2880
gccggcgatc	cctggatcaa	gacctgaca	gccaaagtatc	ccggcaactt	caccgccacc	2940
atcgaggaat	ggcaggccga	gcacgatgcc	atcatgcggc	acatcctgga	aagaccgat	3000

cctaccgacg	tgttccagaa	caaggccaac	gtgtgctggg	ccaaagctct	ggtgcctgtg	3060
ctgaaaaccg	ccggcatcga	tatgaccacc	gagcagtgga	acaccgtgga	ctacttcgag	3120
acagacaagg	cccacagcgc	cgagatcgtg	ctgaatcagc	tgtgcgtgcg	gttcttcggc	3180
ctggatctgg	atagcggcct	gttctctgct	cctaccgtgc	ctctgagcat	ccggaacaac	3240
cactgggaca	acagcccctc	tcctaatatg	tacggcctga	acaaagaagt	cgtgcggcag	3300
ctgagcagaa	gatacccaca	gctgcctaga	gccgtggcca	caggcagagt	gtacgacatg	3360
aataccggca	cactgcggaa	ctacgacccc	agaatcaatc	tggtgcccgt	gaacagaagg	3420
ctgccccacg	ctctggttct	gcaccacaat	gagcaccctc	agagcgactt	cagcagcttc	3480
gtgtccaagc	tgaagggcag	aaccgtgctg	gttgtgggcg	agaagctgtc	tgtgcctggc	3540
aagatggtgg	actggctgag	cgatagacct	gaggccacct	ttagagccag	actggacctt	3600
ggaatccctg	gcgacgtgcc	caaatacgac	atcatcttcg	tgaacgtgcg	gacgccttac	3660
aagtaccacc	actaccagca	gtgcgaggac	cacgccatca	agctgagcat	gctgaccaag	3720
aaggcctgcc	tgcacctgaa	tcctggcggc	acctgtgtgt	ctatcggcta	cggctatgcc	3780
gacagagcca	gcgagtctat	catcggcgcc	attgccagac	agttcaagtt	cagcagagtg	3840
tgaagccca	agagcagcct	ggaagagaca	gaggtgctgt	tcgtgttcat	cggctatgac	3900
cggaaggccc	gtacgcacaa	tccttacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
gggggtgtgcg	gagcgtgta	taagaaattc	ccggaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgacag	aacctgatgc	agagctggtg	agggtgcatc	cgaagagttc	tttggctgga	4560
aggaagggct	acagcacaa	cgatggcaaa	actttctcat	atttggaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgcccttgctt	gtgcatccat	4800
gccatgactc	cagaaaaggt	acagcgcccta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgctcatcct	ttccattgcc	gaagtataga	atcactgggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatte	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgacccac	5160
caggtgctgc	aagtcgaggc	agacattcac	gggccgccct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340

atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgcctccgc	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagttttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgctcac	gcactcctag	caggtcggtc	tcgagaacca	gcctggtctc	caaccgcga	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgcctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgtctacc	gaaccttgca	tctgtttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatggt	tgacggagct	6120
tcatgctgct	tagacactgc	cagtttttgc	cctgcaaagc	tgcgcagctt	tccaaagaaa	6180
cactcctatt	tggaaaccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240
cagaacgtcc	tggcagctgc	cacaaaaaga	aattgcaatg	tcacgcaa	gagagaattg	6300
cccgatttgg	attcggcggc	ctttaatgtg	gaatgcttca	agaaatatgc	gtgtaataat	6360
gaatattggg	aaacgtttta	agaaaacccc	atcaggctta	ctgaagaaaa	cgtggtaaat	6420
tacattacca	aattaaaagg	acaaaaagct	gctgctcttt	ttgcgaagac	acataatttg	6480
aatatgttgc	aggacatacc	aatggacagg	tttgtaatgg	acttaaagag	agacgtgaaa	6540
gtgactccag	gaacaaaaca	tactgaagaa	cggcccaagg	tacaggtgat	ccaggctgcc	6600
gatccgctag	caacagcgta	tctgtgcgga	atccaccgag	agctggttag	gagattaaat	6660
gcggctcctgc	ttccgaacat	tcatacactg	tttgatatgt	cggctgaaga	ctttgacgct	6720
attatagccg	agcacttcca	gcctggggat	tgtgttctgg	aaactgacat	cgcgtcgttt	6780
gataaaagtg	aggacgacgc	catggctctg	accgcgttaa	tgattctgga	agacttaggt	6840
gtggacgcag	agctgttgac	gctgattgag	gcggctttcg	gcgaaatttc	atcaatacat	6900
ttgcccacta	aaactaaatt	taaattcgga	gccatgatga	aatctggaat	gttcctcaca	6960
ctgtttgtga	acacagtcac	taacattgta	atcgcaagca	gagtgttgag	agaacggcta	7020
accggatcac	catgtgcagc	attcattgga	gatgacaata	tcgtgaaagg	agtcaaatcg	7080
gacaaattaa	tggcagacag	gtgcgccacc	tggttgaata	tggaaagtaa	gattatagat	7140
gctgtggtgg	gcgagaaaagc	gccttatttc	tgtggagggt	ttattttgtg	tgactccgtg	7200
accggcacag	cgtgccgtgt	ggcagacccc	ctaaaaaggc	tgtttaagct	tggcaaacct	7260
ctggcagcag	acgatgaaca	tgatgatgac	aggagaaggg	cattgcatga	agagtcaaca	7320
cgttggaacc	gagtgggtat	tctttcagag	ctgtgcaagg	cagtagaatc	aaggtatgaa	7380
accgtaggaa	cttccatcat	agttatggcc	atgactactc	tagctagcag	tgttaaatca	7440
ttcagctacc	tgagaggggc	ccctataact	ctctacggct	aa		7482

<210> 8

<211> 7482

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C08 nsP coding sequence DNA

<400> 8

atggaaaagg tgcacgtgga catcgaagag gacagcccat tcctgagagc cctgcagaga	60
agcttccctc agttcgaggt ggaagccaaa caagtgaccg acaacgatca cgccaacgcg	120
agggccttct ctcacctcgc aagcaagctg atcgagacag aggtggaccc cagcgacacc	180
atcctggata ttgatctgc cctgccaga agaattgtatt ctaagcacia gtatcattgt	240
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg	300
aagaaaaact gtaaggaaat aactgataag gaattggaca agaaaatgaa ggagctcgcc	360
gccgtcatga gcgaccctga cctggaaact gagactatgt gcctccacga cgacgagtcg	420
tgctgctacg aagggaagt cgtgttttac caggatgtat acgcggttga cggaccgaca	480
agtctctatc accaagccaa taaggagtt agatgcgct actggatagg ctttgacacc	540
acccctttta tgtttaagaa cttggctgga gcatatccat catactctac caactgggcc	600
gacgaaaccg tgtaacggc tcgtaacata ggcctatgca gctctgacgt tatggagcgg	660
tcacgtagag ggatgtccat tcttagaaag aagtatttga aaccatccaa caatgttcta	720
ttctctgttg gctcgaccat ctaccacgag aagagggact tactgaggag ctggcacctg	780
ccgtctgtat ttcaattacg tggcaagcaa aattacacat gtcggtgtga gactatagtt	840
agttgcgacg ggtacgtcgt taaaagaata gctatcagtc caggcctgta tgggaagcct	900
tcaggctatg ctgctacgat gcaccgcgag ggattcttgt gctgcaaagt gacagacaca	960
ttgaacgggg agagggtctc tttcccgtg tgcacgtatg tgccagctac attgtgtgac	1020
caaattgactg gcatactggc aacagatgtc agtgcggacg acgcgcaaaa actgctggtt	1080
gggctcaacc agcgtatagt cgtcaacggc cgcaccacga gaaacaccaa taccatgaaa	1140
aattaccttt tgcccgtagt ggcccaggca tttgctaggt gggcaaagga atataaggaa	1200
gatcaagaag atgaaaggcc actaggacta cgagatagac agttagtcac ggggtgttgt	1260
tgggctttta gaaggcacia gataacatct atttataagc gcccgatac ccaaaccatc	1320
atcaaagtga acagcgattt ccaactcttc gtgctgcccc ggataggcag taacacattg	1380
gagatcgggc tgagaacaag aatcaggaaa atgttagagg agcacaagga gccgtcacct	1440
ctcattaccg ccgaggacgt acaagaagct aagtgcgcag ccgatgaggc taaggaggtg	1500
cgtgaagccg aggagttgcg cgcagctcta ccacctttgg cagctgatgt tgaggagccc	1560
actctggaag ccgatgtcga cttgatgtta caagaggctg gggccggctc agtggagaca	1620
cctctgggct tgataaaggt taccagctac gctggcgagg acaagatcgg ctcttacgct	1680
gtgctttctc cgcaggctgt actcaagagt gaaaaattat cttgcatcca ccctctcgct	1740
gaacaagtca tagtgataac aactctggc cgaaaagggc gttatgccgt ggaaccatac	1800
catggtaaag tagtggtgcc agaggacat gcaatacccg tccaggactt tcaagctctg	1860
agtgaagtg ccaccattgt gtacaacgaa cgtgagttcg taaacaggta cctgcacat	1920
attgccacac atggaggagc gctgaacact gatgaagaat attacaaaac tgtcaagccc	1980
agcgagcacg acggcgaata cctgtacgac atcgacagga aacagtgcgt caagaaagaa	2040
ctagtcactg ggctagggct cacaggcgag ctggtggatc ctcccttcca tgaattcgcc	2100

tacgagagtc	tgagaacacg	accagccgct	ccttaccaag	taccaacat	agggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatatgtg	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgacacac	agtcttccac	2520
aaaagcatct	ctcgccgttg	cactaaatct	gtgacttcgg	tcgtctcaac	cttgttttac	2580
gacaaaaaaaa	tgagaacgac	gaatccgaaa	gagactaaga	ttgtgattga	cactaccggc	2640
agtaccaaac	ctaagcagga	cgatctcatt	ctcacttggt	tcagagggtg	ggtgaagcag	2700
ttgcaaatag	attacaaaag	caacgaaata	atgacggcag	ctgcctctca	agggtgacc	2760
cgtaaagggtg	tgtatgccgt	tcggtacaag	gtgaatgaaa	atcctctgta	cgcaccacc	2820
tcagaacatg	tgaacgtcct	actgaccgcg	acggaggacc	gcacgtgtg	gaaaacacta	2880
gccggcgacc	catggataaa	aacactgact	gccaaagtacc	ctgggaattt	cactgccacg	2940
atagaggagt	ggcaagcaga	gcacgatgcc	atcatgaggc	acatcttgga	gagaccggac	3000
cctaccgacg	tcttccagaa	taaggcaaac	gtgtgttggg	ccaaggcttt	agtgccggtg	3060
ctgaagaccg	ctggcataga	catgaccact	gaacaatgga	acactgtgga	ttattttgaa	3120
acggacaaag	ctcactcagc	agagatagta	ttgaaccaac	tatgcgtgag	gttctttgga	3180
ctcgatctgg	actccggtct	attttctgca	cccactgttc	cgttatccat	taggaataat	3240
cactgggata	actccccgtc	gcctaacatg	tacgggctga	ataaagaagt	ggtccgtcag	3300
ctctctcgca	ggtaccacac	actgcctcgg	gcagttgcc	ctggaagagt	ctatgacatg	3360
aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	ctttagtcct	ccaccataat	gaacaccac	agagtgactt	ttcttcattc	3480
gtcagcaaat	tgaaggcgag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatgggtg	actgggtgtc	agaccggcct	gaggctacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gacccatat	3660
aaataccatc	actatcagca	gtgtgaagac	catgccatta	agcttagcat	gttgaccaag	3720
aaagcttggtc	tgcactgtga	tcccggcgga	acctgtgtca	gcataaggta	tggttacgct	3780
gacagggcca	gcgaaaagcat	cattgggtgct	atagcgcggc	agttcaagtt	ttcccgggta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	tgggtacgat	3900
cgcaaggccc	gtacgcacaa	tccttacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
ggggtgtgcg	gagcgtgtga	taagaaattc	ccgaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtec	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440

ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgcacg	aacctgatgc	agagctgggtg	agggtgcatc	cgaagagttc	tttggctgga	4560
aggaagggct	acagcacaag	cgatggcaaa	actttctcat	atttgaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgccttgctt	gtgcatccat	4800
gccatgactc	cagaaaagagt	acagcgccta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgctcatcct	ttccattgcc	gaagtataga	atcactgggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatct	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgaccac	5160
caggtgctgc	aagtcgaggc	agacattcac	gggccgcct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgtccgc	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagtttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgtcac	gcactcctag	caggtcggtc	tcgagaacca	gcctggtctc	caaccgccca	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgcctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgtacc	gaacctgca	tcctgttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatggt	tgacggagct	6120
tcatgtctgct	tagacactgc	cagtttttgc	cctgcaaagc	tgcgcagctt	tccaaagaaa	6180
cactcctatt	tggaaaccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240
cagaacgtcc	tggcagctgc	cacaaaaaga	aattgcaatg	tcacgcaaat	gagagaattg	6300
cccgtattgg	attcggcggc	ctttaatgtg	gaatgcttca	agaaatatgc	gtgtaataat	6360
gaatattggg	aaacgtttaa	agaaaacccc	atcaggctta	ctgaagaaaa	cgtggtaaat	6420
tacattacca	aattaaaagg	accaaaagct	gctgctcttt	ttgcgaagac	acataatttg	6480
aatatgttgc	aggacatacc	aatggacagg	tttgtaatgg	acttaaagag	agacgtgaaa	6540
gtgactccag	gaacaaaaca	tactgaagaa	cggcccaagg	tacaggtgat	ccaggctgcc	6600
gatccgctag	caacagcgta	tctgtgcgga	atccaccgag	agctggttag	gagattaaat	6660
gcggctctgc	ttccgaacat	tcatacactg	tttgatatgt	cggctgaaga	ctttgacgct	6720
attatagccg	agcacttcca	gcctggggat	tgtgttcttg	aaactgacat	cgcgtcgttt	6780

gataaaagtg aggacgacgc catggctctg accgcgttaa tgattctgga agacttaggt	6840
gtggacgcag agctgttgac gctgattgag gcggctttcg gcgaaatttc atcaatacat	6900
ttgcccacta aaactaaatt taaattcgga gccatgatga aatctggaat gttcctcaca	6960
ctgttttgta acacagtcac taacattgta atcgcaagca gagtggtgag agaacggcta	7020
accggatcac catgtgcagc attcattgga gatgacaata tcgtgaaagg agtcaaatac	7080
gacaaattaa tggcagacag gtgcgccacc tgggtgaata tggaagtcaa gattatagat	7140
gctgtggtgg gcgagaaaagc gccttatttc tgtggagggt ttattttgtg tgactccgtg	7200
accggcacag cgtgccgtgt ggcagacccc ctaaaaaggc tgtttaagct tggcaaacct	7260
ctggcagcag acgatgaaca tgatgatgac aggagaaggc cattgcatga agagtcaaca	7320
cgctggaacc gagtgggtat tctttcagag ctgtgcaagg cagtagaatc aaggtatgaa	7380
accgtaggaa cttccatcat agttatggc atgactactc tagctagcag tgttaaatca	7440
ttcagctacc tgagaggggc ccctataact ctctacggt aa	7482

<210> 9

<211> 7482

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C09 nsP coding sequence DNA

<400> 9

atggagaaag ttcacgttga catcgaggaa gacagcccat tcctcagagc tttgcagcgg	60
agcttcccg c agtttgaggt agaagccaag caggtcactg ataatgacca tgctaattgcc	120
agagcgtttt cgcatctggc ttcaaaactg atcgaaacgg aggtggacct atccgacacg	180
atccttgaca ttggaagtgc gcccgccgc agaatgtatt ctaagcaca gtatcattgt	240
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg	300
aagaaaaact gtaaggaaat aactgataag gaattggaca agaaaatgaa ggagctcgcc	360
gccgtcatga gcgaccctga cctggaaact gagactatgt gcctccacga cgacgagtcg	420
tgctgctacg aagggaagt cgctgtttac caggatgtat acgcggttga cggaccgaca	480
agtctctatc accaagccaa taaggaggt agagtcgcct actggatagg ctttgacacc	540
acccctttta tgtttaagaa cttggctgga gcatatccat catactctac caactgggcc	600
gacgaaaccg tgtaacggc cagaaatatc ggctgtgtga gcagcgacgt gatggaaaga	660
tccagacggg gcatgagcat cctgcggaag aagtacctga agcctagcaa caacgtgctg	720
ttcagcgtgg gcagcaccat ctaccacgag aagagggacc tgctgcggag ctggcatctg	780
ccttccgtgt ttcacctgag aggcaagcag aactacacct gtagatgcga gacaatcgtg	840
tcctgcgacg gctacgtggt caagcggatc gccatttctc ctggcctgta cggcaagcct	900
tctggctatg ccgccaccat gcacagagaa ggctttctgt gttgcaaagt gaccgacaca	960
ctgaacggcg agcgggtgtc ctttcctgtg tgtacctatg tgcccgcac actgtgcgat	1020
cagatgacag gcattctggc caccgacgtg tcagccgacg atgccagaa actgctcgtg	1080
ggcctgaacc agagaatcgt ggtcaacggc agaaccacg ggaacaccaa caccatgaag	1140
aactacctgc tgctgtggt ggcccaggcc tttgccagat gggccaaaga gtacaaagag	1200

gatcaagagg	acgagcggcc	cctgggcctg	agagatagac	aactggtcat	gggctgctgc	1260
tgggccttca	gaaggcacia	gatcaccagc	atctacaagc	ggcccagacac	acagaccatc	1320
atcaaagtga	acagcgactt	ccacagcttc	gtgctgcctc	ggatcggcag	caacacactg	1380
gaaatcggcc	tgcggacccg	gatcagaaag	atgctggaag	aacacaaaga	gccctctcca	1440
ctgatcaccg	ccgaggatgt	gcaagaggcc	aaatgtgccg	ccgacgaggc	taaagaagtg	1500
cgcgaagccg	aggaactgag	agccgcactt	cctcctctgg	ccgccgatgt	tgaagaaccc	1560
actctggaag	ccgacgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagt	gaaaaattat	cttgcattca	ccctctcgct	1740
gaacaagtca	tagtgataac	acactctggc	cgaaaagggc	gttatgccgt	ggaaccatac	1800
catggtaaaag	tagtggtgcc	agaggacat	gcaatacccg	tccaggactt	tcaagctctg	1860
agtgaaagtg	ccaccattgt	gtacaacgaa	cgtgagttcg	taaacaggta	cctgcaccat	1920
attgccacac	atggaggagc	gctgaacact	gatgaagaat	attacaaaac	tgtcaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtactcg	ggctagggct	cacaggcgag	ctggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccaag	taccaaccat	agggtgttat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatatatga	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgacaca	agtcttccac	2520
aagagcatca	gcagacggtg	caccaagagc	gtgaccagcg	tggtgtctac	cctgttctac	2580
gacaagaaga	tgcggacgac	aaaccccaaa	gagacaaaga	tcgtcatcga	caccaccggc	2640
agcaccaagc	ctaagcagga	cgatctgatc	ctgacctgct	tcagaggctg	ggtcaagcag	2700
ctgcagatcg	actacaaggg	caacgagatc	atgaccgccg	ctgcctctca	ggccctgaca	2760
agaaaaggcg	tgtacgccgt	gcggtacaaa	gtgaacgaga	accctctgta	cgcccctacc	2820
agcgagcatg	tgaatgtgct	gctgaccggg	accgaggacc	ggatcgtttg	gaaaacactg	2880
gccggcgatc	cctggatcaa	gacctgaca	gccaaagtac	ccggcaactt	caccgccacc	2940
atcgaggaaat	ggcaggccga	gcacgatgcc	atcatgcggc	acatcctgga	aagacccgat	3000
cctaccgacg	tgttccagaa	caaggccaac	gtgtgctggg	ccaaagctct	ggtgcctgtg	3060
ctgaaaaccg	ccggcatcga	tatgaccacc	gagcagtggg	acaccgtgga	ctacttcgag	3120
acagacaagg	cccacagcgc	cgagatcgtg	ctgaatcagc	tgtgcgtgcg	gttcttcggc	3180
ctggatctgg	atagcggcct	gttctctgct	cctaccgtgc	ctctgagcat	ccggaacaac	3240
cactgggaca	acagcccctc	tcctaatatg	tacggcctga	acaaagaagt	cgtgcggcag	3300
ctgagcagaa	gatacccaca	gctgcctaga	gccgtggcca	caggcagagt	gtacgacatg	3360
aataccggca	cactgcggaa	ctacgacccc	agaatcaatc	tggtgcccgt	gaacagaagg	3420
ctgccccacg	ctctggttct	gcaccacaat	gagcaccctc	agagcgactt	cagcagcttc	3480
gtgtccaagc	tgaagggcag	aaccgtgctg	gttgtgggcg	agaagctgtc	tgtgcctggc	3540

aagatggtgg	actggctgag	cgatagaccc	gaggccacct	ttagagccag	actggacctt	3600
ggaatccctg	gcgacgtgcc	caaatacgac	atcatcttcg	tgaacgtgcg	gacgccttac	3660
aagtaccacc	actaccagca	gtgcgaggac	cacgccatca	agctgagcat	gctgaccaag	3720
aaggcctgcc	tgcacctgaa	tcctggcggc	acctgtgtgt	ctatcggcta	cggctatgcc	3780
gacagagcca	gcgagtctat	catcggcgcc	attgccagac	agttcaagtt	cagcagagtg	3840
tgcaagccca	agagcagcct	ggaagagaca	gaggtgctgt	tcgtgttcat	cggctatgac	3900
cggaaggccc	gtacgcacaa	tccttacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
ggggtgtgcg	gagcgctgta	taagaaattc	ccggaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgcag	aacctgatgc	agagctggtg	agggtgcatc	cgaagagttc	tttggctgga	4560
aggaagggtc	acagcacaag	cgatggcaaa	actttctcat	atttgaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgccttgctt	gtgcatccat	4800
gccatgactc	cagaaagagt	acagcgccta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgctcatcct	ttccattgcc	gaagtataga	atcactggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatte	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgaccac	5160
cagggtgctgc	aagtcgaggc	agacattcac	gggccgccct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgtccgc	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagttttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgtcac	gacttcctag	caggtcggtc	tcgagaacca	gcctggtctc	caaccgccea	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgcctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880

```

ataacagcta gacgtattct gcaaggccta gggcattatt tgaaggcaga aggaaaagtg      5940
gagtgtctacc gaaccttgca tcctgttcct ttgtattcat ctagtgtgaa ccgtgccttt      6000
tcaagcccca aggtcgcagt ggaagcctgt aacgccatgt tgaaagagaa ctttccgact      6060
gtggctttctt actgtattat tccagagtac gatgcctatt tggacatggt tgacggagct      6120
tcatgtctgct tagacactgc cagtttttgc cctgcaaagc tgcgagcgtt tccaaagaaa      6180
cactcctatt tggaaccacac aatacgatcg gcagtgcctt cagcgatcca gaacacgctc      6240
cagaacgtcc tggcagctgc cacaaaaaga aattgcaatg tcacgcaaat gagagaattg      6300
cccgtattgg attcggcggc ctttaatgtg gaatgcttca agaaatatgc gtgtaataat      6360
gaatatattgg aaacgtttaa agaaaacccc atcaggetta ctgaagaaaa cgtggtaaat      6420
tacattacca aattaaaagg accaaaagct gctgctcttt ttgcgaagac acataatttg      6480
aatatgttgc aggacatacc aatggacagg tttgtaatgg acttaaagag agacgtgaaa      6540
gtgactccag gaacaaaaca tactgaagaa cggcccaagg tacaggatgat ccaggctgcc      6600
gatccgctag caacagcgta tctgtgcgga atccaccgag agctgggttag gagattaaat      6660
gcggtcctgc ttccgaacat tcatacactg tttgatatgt cggctgaaga ctttgacgct      6720
attatagccg agcacttcca gcctggggat tgtgttcttg aaactgacat cgcgtcgttt      6780
gataaaagtg aggacgacgc catggctctg accgcgttaa tgattctgga agacttaggt      6840
gtggacgcag agctgttgac gctgattgag gcggctttcg gcgaaatttc atcaatacat      6900
ttgcccacta aaactaaatt taaattcgga gccatgatga aatctggaat gttcctcaca      6960
ctgttttgta acacagtcac taacattgta atcgcaagca gagtgttgag agaacggcta      7020
accggatcac catgtgcagc attcattgga gatgacaata tcgtgaaagg agtcaaatec      7080
gacaaattaa tggcagacag gtgcgccacc tggttgaata tggaagtcaa gattatagat      7140
gctgtgggtgg gcgagaaagc gccttatttc tgtggagggt ttattttgtg tgactccgtg      7200
accggcacag cgtgccgtgt ggcagacccc ctaaaaaggc tgtttaagct tggcaaacct      7260
ctggcagcag acgatgaaca tgatgatgac aggagaaggg cattgcatga agagtcaaca      7320
cgctggaacc gagtgggtat tctttcagag ctgtgcaagg cagtagaatc aaggtatgaa      7380
accgtaggaa cttccatcat agttatggcc atgactactc tagctagcag tgttaaatca      7440
ttcagctacc tgagaggggc ccctataact ctctacggct aa                        7482

```

<210> 10

<211> 7482

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C10 nsP coding sequence DNA

<400> 10

```

atggagaaag ttcacgttga catcgaggaa gacagcccat tcctcagagc tttgcagcgg      60
agcttcccgc agtttgaggt agaagccaag caggtcactg ataatgacca tgctaatgcc      120
agagcgtttt cgcactctggc ttcaaaactg atcgaaacgg aggtggacct atccgacacg      180
atccttgaca ttggaagtgc gcccgccgcg agaatgtatt ctaagcacia gtatcattgt      240
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg      300

```

aagaaaaact	gtaaggaaat	aactgataag	gaattggaca	agaaaatgaa	ggagctcgcc	360
gccgtcatga	gcgaccctga	cctggaaact	gagactatgt	gcctccacga	cgacgagtcg	420
tgtcgctacg	aagggcaagt	cgctgtttac	caggatgtat	acgcggttga	cggaccgaca	480
agtctctatc	accaagccaa	taagggagtt	agagtcgcct	actggatagg	ctttgacacc	540
acccctttta	tgtttaagaa	cttggctgga	gcataatccat	catactctac	caactgggcc	600
gacgaaaccg	tgttaacggc	tcgtaacata	ggcctatgca	gctctgacgt	tatggagcgg	660
tcacgtagag	ggatgtccat	tcttagaaaag	aagtatttga	aaccatccaa	caatgttcta	720
ttctctgttg	gctcgaccat	ctaccacgag	aagagggact	tactgaggag	ctggcacctg	780
ccgtctgtat	ttcacttacg	tggcaagcaa	aattacacat	gtcgggtgtga	gactatagtt	840
agttgcgacg	ggtacgtcgt	taaaagaata	gctatcagtc	caggcctgta	tgggaagcct	900
tcaggctatg	ctgctacgat	gcaccgcgag	ggattcttgt	gctgcaaagt	gacagacaca	960
ttgaacgggg	agagggctct	ttttcccgctg	tgcacgtatg	tgccagctac	attgtgtgac	1020
caaatgactg	gcatactggc	aacagatgtc	agtgcggacg	acgcgcaaaa	actgctggtt	1080
gggctcaacc	agcgtatagt	cgtaacgggt	cgcaccacga	gaaacaccaa	taccatgaaa	1140
aattaccttt	tgcccgtagt	ggcccaggca	tttgctaggt	gggcaaagga	atataaggaa	1200
gatcaagaag	atgaaaaggcc	actaggacta	cgagatagac	agttagtcac	ggggtgttgt	1260
tgggctttta	gaaggcacia	gataacatct	atttataagc	gcccggatac	ccaaaccatc	1320
atcaaagtga	acagcgattt	ccactcattc	gtgctgcccc	ggataggcag	taacacattg	1380
gagatcgggc	tgagaacaag	aatcaggaaa	atgttagagg	agcacaagga	gccgtcacct	1440
ctcattaccg	ccgaggacgt	acaagaagct	aagtgcgcag	ccgatgaggc	taaggaggtg	1500
cgtgaagccg	aggagttgcg	cgcagctcta	ccacctttgg	cagctgatgt	tgaggagccc	1560
actctggaag	ccgatgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagt	gaaaaattat	cttgcatcca	ccctctcgct	1740
gaacaagtca	tagtgataac	acactctggc	cgaaaagggc	gttatgccgt	ggaaccatac	1800
catggtaaag	tagtggtgcc	agaggacat	gcaatacccg	tccaggactt	tcaagctctg	1860
agtgaagtg	ccaccattgt	gtacaacgaa	cgtgagttcg	taaacaggta	cctgcaccat	1920
attgccacac	atggaggagc	gctgaacact	gatgaagaat	attacaaaac	tgtcaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtacttg	ggctagggct	cacaggcgag	ctgggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccag	taccaaccat	aggggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatatgtg	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgacacac	agtcttccac	2520
aaaagcatct	ctcgccgttg	cactaaatct	gtgacttcgg	tcgtctcaac	cttgttttac	2580
gacaaaaaaa	tgagaacgac	gaatccgaaa	gagactaaga	ttgtgattga	cactaccggc	2640

agtaccaaac	ctaagcagga	cgatctcatt	ctcacttggt	tcagagggtg	ggtgaagcag	2700
ttgcaaatag	attacaaagg	caacgaaata	atgacggcag	ctgcctctca	agggtgacc	2760
cgtaaagggtg	tgtatgccgt	tcggtacaag	gtgaatgaaa	atcctctgta	cgcaccacc	2820
tcagaacatg	tgaacgtcct	actgacccgc	acggaggacc	gcatcgtgtg	gaaaacacta	2880
gccggcgacc	catggataaa	aacactgact	gccaaagtacc	ctgggaatth	cactgccacg	2940
atagaggagt	ggcaagcaga	gcatgatgcc	atcatgaggc	acatcttgga	gagaccggac	3000
cctaccgacg	tcttccagaa	taaggcaaac	gtgtgttggg	ccaaggctth	agtgccggtg	3060
ctgaagaccg	ctggcataga	catgaccact	gaacaatgga	acactgtgga	ttatthtgaa	3120
acggacaaaag	ctcactcagc	agagatagta	ttgaaccaac	tatgcgtgag	gttctthtgga	3180
ctcgatctgg	actccggtct	atthtctgca	cccactgttc	cgthtatccat	taggaataat	3240
cactgggata	actccccgtc	gcctaacatg	tacgggtgta	ataaagaagt	ggtccgtcag	3300
ctctctcgca	ggtaccacaca	actgcctcgg	gcagttgcca	ctggaagagt	ctatgacatg	3360
aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	ctttagtcct	ccaccataat	gaacaccac	agagtgactt	ttcttcattc	3480
gtcagcaaat	tgaagggcag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatggttg	actggttgtc	agaccggcct	gaggctacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gacccctac	3660
aagtaccacc	actaccagca	gtgcgaggac	cacgccatca	agctgagcat	gctgaccaag	3720
aaggcctgcc	tgcacttgaa	tccggcgga	acctgtgtca	gcataggtta	tggttacgt	3780
gacagggccca	gcgaaagcat	cattggtgct	atagcgcggc	agttcaagtt	ttcccggtta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	cggctacgac	3900
agaaaggccc	gtacgcacaa	tccttacaag	cttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
ggggtgtgcg	gagcgtgtga	taagaaattc	ccggaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaagggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatcttht	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaagggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgacag	aacctgatgc	agagctggtg	agggtgcac	cgaagagttc	tttggctgga	4560
aggaagggt	acagcacaag	cgatggcaaa	actthtctcat	atthtggaag	gaccaagtht	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgthtc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
ccgctcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgccttgctt	gtgcatccat	4800
gccatgactc	cagaaagagt	acagcgcta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgtcatcct	ttccattgcc	gaagtataga	atcactggtg	tgcagaagat	ccaatgtcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatth	atccaaggaa	gtatctcgtg	4980

gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgacccac	5160
caggtgctgc	aagtcgaggc	agacattcac	gggccgccct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgcctccg	gcacaagaac	accgtcactt	gcacccagca	gggcctgctc	gagaaccagc	5460
ctagttttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgctcac	gcactcctag	caggtcggtc	tcgagaacca	gcctggcttc	caacccgccca	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgcctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgtctacc	gaacctgtca	tcctgttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatggg	tgacggagct	6120
tcattgctgct	tagacactgc	cagtttttgc	cctgcaaagc	tgcgcagctt	tccaaagaaa	6180
cactcctatt	tggaaaccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240
cagaacgtcc	tggcagctgc	cacaaaaaga	aattgcaatg	tcacgcaa	gagagaattg	6300
cccgtatttg	attcggcggc	ctttaatgtg	gaatgcttca	agaagtacgc	ctgcaacaac	6360
gagtactggg	agacattcaa	agagaacccc	atccggctga	ccgaggaaaa	cgtggtcaac	6420
tacatcacca	agctgaaggg	ccccaaagcc	gccgctctgt	ttgccaagac	acacaacctg	6480
aacatgctgc	aggacatccc	catggacaga	ttcgtgatgg	acctgaagcg	ggacgtgaaa	6540
gtgacccctg	gcaccaagca	caccgaggaa	cggcctaagg	tgcaagtgat	ccaggccgct	6600
gatcctcttg	ccacagccta	tctgtgtggc	atccacagag	aactcgtgcg	gagactgaat	6660
gccgtgctgc	ttccgaacat	tcatacactg	tttgatatgt	cggctgaaga	ctttgacgct	6720
attatagccg	agcacttcca	gcctggggat	tgtgttctgg	aaactgacat	cgcgtcgttt	6780
gataaaaagt	aggacgacgc	catggctctg	accgcgctga	tgattctgga	agatctcgga	6840
gtggacgccg	agctgctgac	actgattgaa	gccgcctttg	gcgagatcag	cagcatccat	6900
ctgcctacca	agaccaagtt	caagtccggc	gccatgatga	aatctggaat	gttcctcaca	6960
ctgtttgtga	acacagtcac	taacattgta	atcgcaagca	gagtgttgag	agaacggcta	7020
acaggcagcc	cttgtgcagc	attcattgga	gatgacaata	tcgtgaaagg	agtcaaatec	7080
gacaaattaa	tggcagacag	gtgcgccacc	tggttgaata	tggaaagtaa	gattatagat	7140
gctgtgggtg	gcgagaaaag	gccttatttc	tgtggagggt	ttattttgtg	tgactccgtg	7200
accggcacag	cgtgccgtgt	ggcagacccc	ctaaaaaggc	tgtttaagct	tggcaaacct	7260
ctggcagcag	acgatgaaca	tgatgatgac	aggagaaggg	cattgcatga	agagtcaaca	7320

cgctggaacc	gagtgggtat	tctttcagag	ctgtgcaagg	cagtagaatc	aaggtatgaa	7380
accgtaggaa	cttccatcat	agttatggcc	atgactactc	tagctagcag	tgtaaataca	7440
ttcagctacc	tgagaggggc	ccctataact	ctctacggct	aa		7482
<210>	11					
<211>	7482					
<212>	DNA					
<213>	Artificial Sequence					
<220>						
<223>	Synthetic construct, C11 nsP coding sequence DNA					
<400>	11					
atggagaaaag	ttcacgttga	catcgaggaa	gacagcccat	tcctcagagc	tttgcagcgg	60
agcttccccgc	agtttgaggt	agaagccaag	caggtcactg	ataatgacca	tgctaattgcc	120
agagcgtttt	cgcactctggc	ttcaaaaactg	atcgaaacgg	aggtggaccc	atccgacacg	180
atccttgaca	ttggaagtgc	gcccggccgc	agaatgtatt	ctaagcacia	gtatcattgt	240
atctgtccga	tgagatgtgc	ggaagatccg	gacagattgt	ataagtatgc	aactaagctg	300
aagaaaaact	gtaaggaaat	aactgataag	gaattggaca	agaaaatgaa	ggagctcgcc	360
gccgtcatga	gcgaccctga	cctggaaact	gagactatgt	gcctccacga	cgacgagtcg	420
tgctgctacg	aagggcaagt	cgctgtttac	caggatgtat	atgccgtgga	tgccctaca	480
agcctgtacc	accaggccaa	caagggcgtc	agagtggcct	actggatcgg	cttcgacacc	540
acacctttca	tgttcaagaa	cctggctggc	gcttacccca	gctacagcac	aaactgggcc	600
gacgaaaccg	tgtaaacggc	cagaaatatc	ggcctgtgta	gcagcgacgt	gatggaaaga	660
tccagacggg	gcatgagcat	cctgcggaag	aagtacctga	agcctagcaa	caacgtgctg	720
ttcagcgtgg	gcagcaccat	ctaccacgag	aagagggacc	tgctgcggag	ctggcatctg	780
ccttccgtgt	ttcacctgag	aggcaagcag	aactacacct	gtagatgcga	gacaatcgtg	840
tcctgcgacg	gctacgtggt	caagcggatc	gccatttctc	ctggcctgta	cggcaagcct	900
tctggctatg	ccgccaccat	gcacagagaa	ggctttctgt	gttgcaaagt	gaccgacaca	960
ctgaacggcg	agcgggtgtc	ctttcctgtg	tgtacctatg	tgcccggcac	actgtgcgat	1020
cagatgacag	gcattctggc	caccgacgtg	tcagccgacg	atgccagaa	actgctcgtg	1080
ggcctgaacc	agagaatcgt	ggtcaacggc	agaaccacgc	ggaacaccaa	caccatgaag	1140
aactacctgc	tgccgtgtgt	ggcccaggcc	tttgccagat	ggcccaaaga	gtacaaagag	1200
gatcaagagg	acgagcggcc	cctgggcctg	agagatagac	aactggatcat	gggctgctgc	1260
tgggccttca	gaaggcacia	gatcaccagc	atctacaagc	ggcccagacac	acagaccatc	1320
atcaaagtga	acagcgactt	ccacagcttc	gtgctgcctc	ggatcggcag	caacacactg	1380
gaaatcggcc	tgccgacccg	gatcagaaaag	atgctggaag	aacacaaaga	gccctctcca	1440
ctgatcaccg	ccgaggatgt	gcaagaggcc	aatgtgccg	ccgacgaggc	taaagaagtg	1500
cgcgaagccg	aggaactgag	agccgcactt	cctcctctgg	ccgccgatgt	tgaagaaccc	1560
actctggaag	ccgacgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagc	gagaagctga	gctgcattca	ccctctggcc	1740

gagcaagtga	tcgtgatcac	acacagcggc	cggaaggcca	gatatgccgt	ggaaccttat	1800
cacggcaagg	tggtggtgcc	tgaggacac	gctattccag	tgcaggactt	tcaggccctg	1860
agcgagtctg	ccaccatcgt	gtacaacgag	cgcgagttcg	tgaacagata	cctgcaccac	1920
attgccacac	acggcggagc	cctgaacacc	gacgaagagt	actacaagac	cgtgaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtcaactg	ggctagggct	cacaggcgag	ctgggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccaag	taccaacat	agggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggat	agcgtgctgc	tgaacggctg	caagcacccc	2340
gtagagaccc	tgtatataga	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcctaagca	gtgcggcttc	2460
ttcaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgacaca	agtcttccac	2520
aagagcatca	gcagacgggtg	caccaagagc	gtgaccagcg	tggtgtctac	cctgtttctac	2580
gacaagaaga	tgcgagcgac	aaaccccaaa	gagacaaaga	tcgtcatcga	caccaccggc	2640
agcaccaagc	ctaagcagga	cgatctgac	ctgacctgct	tcagaggctg	ggtcaagcag	2700
ctgcagatcg	actacaaggg	caacgagatc	atgaccgccg	ctgcctctca	gggcctgaca	2760
agaaaaggcg	tgtacgccgt	gcggtacaaa	gtgaacgaga	accctctgta	cgcccctacc	2820
agcgagcatg	tgaatgtgct	gctgaccggg	accgaggacc	ggatcgtttg	gaaaacactg	2880
gccggcgatc	cctggatcaa	gaccctgaca	gccaagtatc	ccggcaactt	caccgccacc	2940
atcgaggaat	ggcaggccga	gcacgatgcc	atcatgcggc	acatcctgga	aagacccgat	3000
cctaccgacg	tgttccagaa	caaggccaac	gtgtgctggg	ccaaagctct	ggtgcctgtg	3060
ctgaaaaccg	ccggcatcga	tatgaccacc	gagcagtgga	acaccgtgga	ctacttcgag	3120
acagacaagg	cccacagcgc	cgagatcgtg	ctgaatcagc	tgtgcgtgcg	gttcttcggc	3180
ctggatctgg	atagcggcct	gttctctgct	cctaccgtgc	ctctgagcat	ccggaacaac	3240
cactgggaca	acagcccctc	tcctaatatg	tacggcctga	acaaagaagt	cgtgcggcag	3300
ctgagcagaa	gatacccaca	gctgcctaga	gccgtggcca	caggcagagt	gtacgacatg	3360
aataccggca	cactgcggaa	ctacgacccc	agaatcaatc	tggtgcccgt	gaacagaagg	3420
ctgccccacg	ctctggttct	gcaccacaat	gagcaccctc	agagcgactt	cagcagcttc	3480
gtgtccaagc	tgaaggcgag	aaccgtgctg	gttggtggcg	agaagctgtc	tgtgcctggc	3540
aagatggtgg	actggctgag	cgatagaccc	gaggccacct	ttagagccag	actggacctt	3600
ggaatccctg	gcgacgtgcc	caaatacgac	atcatcttcg	tgaacgtgcg	gacgccctac	3660
aagtaccacc	actaccagca	gtgcgaggac	cacgccatca	agctgagcat	gctgaccaag	3720
aaggcctgcc	tgcacctgaa	tcctggcggc	acctgtgtgt	ctatcggcta	cggctatgcc	3780
gacagagcca	gcgagtctat	catcggcgcc	attgccagac	agttcaagtt	cagcagagtg	3840
tgaagccca	agagcagcct	ggaagagaca	gaggtgctgt	tcgtgttcat	cggctatgac	3900
cggaaaggccc	gtacgcacaa	cccctacaag	ctgagcagca	ccctgaccaa	catctacacc	3960
ggcagcagac	tgcacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacagcca	cagaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080

ggggtgtgcg	gagcgctgta	taagaaattc	ccggaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgcacg	aacctgatgc	agagctgggt	agggtgcac	cgaagagttc	tttggctgga	4560
aggaagggt	acagcacaa	cgatggcaaa	actttctcat	atttggaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgccttgctt	gtgcatccat	4800
gccatgactc	cagaaaagagt	acagcgcta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgtcatcct	ttccattgcc	gaagtataga	atcactgggt	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatc	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgaccac	5160
caggtgctgc	aagtcgaggc	agacattcac	gggccgcct	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgtccgc	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagtttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgctac	gcactcctag	caggtcggtc	tcgagaacca	gcctggtctc	caaccgcga	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgcctacc	gaacctgca	tcctgttctt	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatggg	tgacggagct	6120
tcattgctgct	tagacactgc	cagtttttgc	cctgcaaagc	tgcgcagctt	tccaaagaaa	6180
cactcctatt	tggaaaccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240
cagaacgtcc	tggcagctgc	cacaaaaaga	aattgcaatg	tcacgcaaat	gagagaattg	6300
cccgtattgg	attcggcggc	ctttaatgtg	gaatgcttca	agaagtacgc	ctgcaacaac	6360
gagtactggg	agacattcaa	agagaacccc	atccggctga	ccgaggaaaa	cgtggtcaac	6420

```

tacatcacca agctgaaggg ccccaaagcc gccgctctgt ttgccaagac acacaacctg 6480
aacatgctgc aggacatccc catggacaga ttcgtgatgg acctgaagcg ggacgtgaaa 6540
gtgacccctg gcaccaagca caccgaggaa cggcctaagg tgcaagtgat ccaggccgct 6600
gatcctcttg ccacagccta tctgtgtggc atccacagag aactcgtgcg gagactgaat 6660
gccgtgctgc ttccgaacat tcatacactg tttgatatgt cggctgaaga ctttgacgct 6720
attatagccg agcacttcca gcctggggat tgtgttctgg aaactgacat cgcgtcgttt 6780
gataaaaagtg aggacgacgc catggctctg accgcgctga tgattctgga agatctcgga 6840
gtggacgccg agctgctgac actgattgaa gccgcctttg gcgagatcag cagcatccat 6900
ctgcctacca agaccaagtt caagttcggc gccatgatga aatctggaat gttcctcaca 6960
ctgtttgtga acacagtcac taacattgta atcgcaagca gagtgttgag agaacggcta 7020
acaggcagcc cttgtgcagc attcattgga gatgacaata tcgtgaaagg agtcaaatec 7080
gacaaattaa tggcagacag gtgcgccacc tggttgaata tggaagtcaa gattatagat 7140
gctgtggtgg gcgagaaaagc gccttatttc tgtggagggt ttattttgtg tgactccgtg 7200
accggcacag cgtgccgtgt ggacagacccc ctaaaaaggc tgtttaagct tggcaaacct 7260
ctggcagcag acgatgaaca tgatgatgac aggagaaggg cattgcatga agagtcaaca 7320
cgctggaacc gagtgggtat tctttcagag ctgtgcaagg cagtagaatc aaggtatgaa 7380
accgtaggaa cttccatcat agttatggcc atgactactc tagctagcag tgttaaatca 7440
ttcagctacc tgagaggggc ccctataact ctctacggt aa 7482

```

<210> 12

<211> 2513

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, C16 nsP-2A protein amino acid

<220>

<221> misc_feature

<222> (1880)..(1880)

<223> Xaa can be any naturally occurring amino acid

<400> 12

Met Glu Lys Val His Val Asp Ile Glu Glu Asp Ser Pro Phe Leu Arg

1 5 10 15

Ala Leu Gln Arg Ser Phe Pro Gln Phe Glu Val Glu Ala Lys Gln Val

20 25 30

Thr Asp Asn Asp His Ala Asn Ala Arg Ala Phe Ser His Leu Ala Ser

35 40 45

Lys Leu Ile Glu Thr Glu Val Asp Pro Ser Asp Thr Ile Leu Asp Ile

50 55 60

Gly Ser Ala Pro Ala Arg Arg Met Tyr Ser Lys His Lys Tyr His Cys

65 70 75 80

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

385	390	395	400
Asp Gln Glu Asp Glu Arg Pro Leu Gly Leu Arg Asp Arg Gln Leu Val			
	405	410	415
Met Gly Cys Cys Trp Ala Phe Arg Arg His Lys Ile Thr Ser Ile Tyr			
	420	425	430
Lys Arg Pro Asp Thr Gln Thr Ile Ile Lys Val Asn Ser Asp Phe His			
	435	440	445
Ser Phe Val Leu Pro Arg Ile Gly Ser Asn Thr Leu Glu Ile Gly Leu			
	450	455	460
Arg Thr Arg Ile Arg Lys Met Leu Glu Glu His Lys Glu Pro Ser Pro			
465	470	475	480
Leu Ile Thr Ala Glu Asp Val Gln Glu Ala Lys Cys Ala Ala Asp Glu			
	485	490	495
Ala Lys Glu Val Arg Glu Ala Glu Glu Leu Arg Ala Ala Leu Pro Pro			
	500	505	510
Leu Ala Ala Asp Val Glu Glu Pro Thr Leu Glu Ala Asp Val Asp Leu			
	515	520	525
Met Leu Gln Glu Ala Gly Ala Gly Ser Val Glu Thr Pro Arg Gly Leu			
	530	535	540
Ile Lys Val Thr Ser Tyr Ala Gly Glu Asp Lys Ile Gly Ser Tyr Ala			
545	550	555	560
Val Leu Ser Pro Gln Ala Val Leu Lys Ser Glu Lys Leu Ser Cys Ile			
	565	570	575
His Pro Leu Ala Glu Gln Val Ile Val Ile Thr His Ser Gly Arg Lys			
	580	585	590
Gly Arg Tyr Ala Val Glu Pro Tyr His Gly Lys Val Val Val Pro Glu			
	595	600	605
Gly His Ala Ile Pro Val Gln Asp Phe Gln Ala Leu Ser Glu Ser Ala			
	610	615	620
Thr Ile Val Tyr Asn Glu Arg Glu Phe Val Asn Arg Tyr Leu His His			
625	630	635	640
Ile Ala Thr His Gly Gly Ala Leu Asn Thr Asp Glu Glu Tyr Tyr Lys			
	645	650	655
Thr Val Lys Pro Ser Glu His Asp Gly Glu Tyr Leu Tyr Asp Ile Asp			
	660	665	670
Arg Lys Gln Cys Val Lys Lys Glu Leu Val Thr Gly Leu Gly Leu Thr			
	675	680	685
Gly Glu Leu Val Asp Pro Pro Phe His Glu Phe Ala Tyr Glu Ser Leu			
	690	695	700

Arg Thr Arg Pro Ala Ala Pro Tyr Gln Val Pro Thr Ile Gly Val Tyr			
705	710	715	720
Gly Val Pro Gly Ser Gly Lys Ser Gly Ile Ile Lys Ser Ala Val Thr			
	725	730	735
Lys Lys Asp Leu Val Val Ser Ala Lys Lys Glu Asn Cys Ala Glu Ile			
	740	745	750
Ile Arg Asp Val Lys Lys Met Lys Gly Leu Asp Val Asn Ala Arg Thr			
	755	760	765
Val Asp Ser Val Leu Leu Asn Gly Cys Lys His Pro Val Glu Thr Leu			
	770	775	780
Tyr Ile Asp Glu Ala Phe Ala Cys His Ala Gly Thr Leu Arg Ala Leu			
785	790	795	800
Ile Ala Ile Ile Arg Pro Lys Lys Ala Val Leu Cys Gly Asp Pro Lys			
	805	810	815
Gln Cys Gly Phe Phe Asn Met Met Cys Leu Lys Val His Phe Asn His			
	820	825	830
Glu Ile Cys Thr Gln Val Phe His Lys Ser Ile Ser Arg Arg Cys Thr			
	835	840	845
Lys Ser Val Thr Ser Val Val Ser Thr Leu Phe Tyr Asp Lys Lys Met			
	850	855	860
Arg Thr Thr Asn Pro Lys Glu Thr Lys Ile Val Ile Asp Thr Thr Gly			
865	870	875	880
Ser Thr Lys Pro Lys Gln Asp Asp Leu Ile Leu Thr Cys Phe Arg Gly			
	885	890	895
Trp Val Lys Gln Leu Gln Ile Asp Tyr Lys Gly Asn Glu Ile Met Thr			
	900	905	910
Ala Ala Ala Ser Gln Gly Leu Thr Arg Lys Gly Val Tyr Ala Val Arg			
	915	920	925
Tyr Lys Val Asn Glu Asn Pro Leu Tyr Ala Pro Thr Ser Glu His Val			
	930	935	940
Asn Val Leu Leu Thr Arg Thr Glu Asp Arg Ile Val Trp Lys Thr Leu			
945	950	955	960
Ala Gly Asp Pro Trp Ile Lys Thr Leu Thr Ala Lys Tyr Pro Gly Asn			
	965	970	975
Phe Thr Ala Thr Ile Glu Glu Trp Gln Ala Glu His Asp Ala Ile Met			
	980	985	990
Arg His Ile Leu Glu Arg Pro Asp Pro Thr Asp Val Phe Gln Asn Lys			
	995	1000	1005
Ala Asn Val Cys Trp Ala Lys Ala Leu Val Pro Val Leu Lys Thr			

1010	1015	1020
Ala Gly Ile Asp Met Thr Thr	Glu Gln Trp Asn Thr	Val Asp Tyr
1025	1030	1035
Phe Glu Thr Asp Lys Ala His	Ser Ala Glu Ile Val	Leu Asn Gln
1040	1045	1050
Leu Cys Val Arg Phe Phe Gly	Leu Asp Leu Asp Ser	Gly Leu Phe
1055	1060	1065
Ser Ala Pro Thr Val Pro Leu	Ser Ile Arg Asn Asn	His Trp Asp
1070	1075	1080
Asn Ser Pro Ser Pro Asn Met	Tyr Gly Leu Asn Lys	Glu Val Val
1085	1090	1095
Arg Gln Leu Ser Arg Arg Tyr	Pro Gln Leu Pro Arg	Ala Val Ala
1100	1105	1110
Thr Gly Arg Val Tyr Asp Met	Asn Thr Gly Thr Leu	Arg Asn Tyr
1115	1120	1125
Asp Pro Arg Ile Asn Leu Val	Pro Val Asn Arg Arg	Leu Pro His
1130	1135	1140
Ala Leu Val Leu His His Asn	Glu His Pro Gln Ser	Asp Phe Ser
1145	1150	1155
Ser Phe Val Ser Lys Leu Lys	Gly Arg Thr Val Leu	Val Val Gly
1160	1165	1170
Glu Lys Leu Ser Val Pro Gly	Lys Met Val Asp Trp	Leu Ser Asp
1175	1180	1185
Arg Pro Glu Ala Thr Phe Arg	Ala Arg Leu Asp Leu	Gly Ile Pro
1190	1195	1200
Gly Asp Val Pro Lys Tyr Asp	Ile Ile Phe Val Asn	Val Arg Thr
1205	1210	1215
Pro Tyr Lys Tyr His His Tyr	Gln Gln Cys Glu Asp	His Ala Ile
1220	1225	1230
Lys Leu Ser Met Leu Thr Lys	Lys Ala Cys Leu His	Leu Asn Pro
1235	1240	1245
Gly Gly Thr Cys Val Ser Ile	Gly Tyr Gly Tyr Ala	Asp Arg Ala
1250	1255	1260
Ser Glu Ser Ile Ile Gly Ala	Ile Ala Arg Gln Phe	Lys Phe Ser
1265	1270	1275
Arg Val Cys Lys Pro Lys Ser	Ser Leu Glu Glu Thr	Glu Val Leu
1280	1285	1290
Phe Val Phe Ile Gly Tyr Asp	Arg Lys Ala Arg Thr	His Asn Pro
1295	1300	1305

Tyr Lys	Leu Ser Ser Thr	Leu Thr Asn Ile Tyr Thr	Gly Ser Arg
1310		1315	1320
Leu His	Glu Ala Gly Cys Ala	Pro Ser Tyr His Val	Val Arg Gly
1325		1330	1335
Asp Ile	Ala Thr Ala Thr Glu	Gly Val Ile Ile Asn	Ala Ala Asn
1340		1345	1350
Ser Lys	Gly Gln Pro Gly Gly	Gly Val Cys Gly Ala	Leu Tyr Lys
1355		1360	1365
Lys Phe	Pro Glu Ser Phe Asp	Leu Gln Pro Ile Glu	Val Gly Lys
1370		1375	1380
Ala Arg	Leu Val Lys Gly Ala	Ala Lys His Ile Ile	His Ala Val
1385		1390	1395
Gly Pro	Asn Phe Asn Lys Val	Ser Glu Val Glu Gly	Asp Lys Gln
1400		1405	1410
Leu Ala	Glu Ala Tyr Glu Ser	Ile Ala Lys Ile Val	Asn Asp Asn
1415		1420	1425
Asn Tyr	Lys Ser Val Ala Ile	Pro Leu Leu Ser Thr	Gly Ile Phe
1430		1435	1440
Ser Gly	Asn Lys Asp Arg Leu	Thr Gln Ser Leu Asn	His Leu Leu
1445		1450	1455
Thr Ala	Leu Asp Thr Thr Asp	Ala Asp Val Ala Ile	Tyr Cys Arg
1460		1465	1470
Asp Lys	Lys Trp Glu Met Thr	Leu Lys Glu Ala Val	Ala Arg Arg
1475		1480	1485
Glu Ala	Val Glu Glu Ile Cys	Ile Ser Asp Asp Ser	Ser Val Thr
1490		1495	1500
Glu Pro	Asp Ala Glu Leu Val	Arg Val His Pro Lys	Ser Ser Leu
1505		1510	1515
Ala Gly	Arg Lys Gly Tyr Ser	Thr Ser Asp Gly Lys	Thr Phe Ser
1520		1525	1530
Tyr Leu	Glu Gly Thr Lys Phe	His Gln Ala Ala Lys	Asp Ile Ala
1535		1540	1545
Glu Ile	Asn Ala Met Trp Pro	Val Ala Thr Glu Ala	Asn Glu Gln
1550		1555	1560
Val Cys	Met Tyr Ile Leu Gly	Glu Ser Met Ser Ser	Ile Arg Ser
1565		1570	1575
Lys Cys	Pro Val Glu Glu Ser	Glu Ala Ser Thr Pro	Pro Ser Thr
1580		1585	1590
Leu Pro	Cys Leu Cys Ile His	Ala Met Thr Pro Glu	Arg Val Gln

1595	1600	1605
Arg Leu Lys Ala Ser Arg	Pro Glu Gln Ile Thr	Val Cys Ser Ser
1610	1615	1620
Phe Pro Leu Pro Lys Tyr Arg	Ile Thr Gly Val Gln	Lys Ile Gln
1625	1630	1635
Cys Ser Gln Pro Ile Leu Phe	Ser Pro Lys Val Pro	Ala Tyr Ile
1640	1645	1650
His Pro Arg Lys Tyr Leu Val	Glu Thr Pro Pro Val	Asp Glu Thr
1655	1660	1665
Pro Glu Pro Ser Ala Glu Asn	Gln Ser Thr Glu Gly	Thr Pro Glu
1670	1675	1680
Gln Pro Pro Leu Ile Thr Glu	Asp Glu Thr Arg Thr	Arg Thr Pro
1685	1690	1695
Glu Pro Ile Ile Ile Glu Glu	Glu Glu Glu Asp Ser	Ile Ser Leu
1700	1705	1710
Leu Ser Asp Gly Pro Thr His	Gln Val Leu Gln Val	Glu Ala Asp
1715	1720	1725
Ile His Gly Pro Pro Ser Val	Ser Ser Ser Ser Trp	Ser Ile Pro
1730	1735	1740
His Ala Ser Asp Phe Asp Val	Asp Ser Leu Ser Ile	Leu Asp Thr
1745	1750	1755
Leu Glu Gly Ala Ser Val Thr	Ser Gly Ala Thr Ser	Ala Glu Thr
1760	1765	1770
Asn Ser Tyr Phe Ala Lys Ser	Met Glu Phe Leu Ala	Arg Pro Val
1775	1780	1785
Pro Ala Pro Arg Thr Val Phe	Arg Asn Pro Pro His	Pro Ala Pro
1790	1795	1800
Arg Thr Arg Thr Pro Ser Leu	Ala Pro Ser Arg Ala	Cys Ser Arg
1805	1810	1815
Thr Ser Leu Val Ser Thr Pro	Pro Gly Val Asn Arg	Val Ile Thr
1820	1825	1830
Arg Glu Glu Leu Glu Ala Leu	Thr Pro Ser Arg Thr	Pro Ser Arg
1835	1840	1845
Ser Val Ser Arg Thr Ser Leu	Val Ser Asn Pro Pro	Gly Val Asn
1850	1855	1860
Arg Val Ile Thr Arg Glu Glu	Phe Glu Ala Phe Val	Ala Gln Gln
1865	1870	1875
Gln Xaa Arg Phe Asp Ala Gly	Ala Tyr Ile Phe Ser	Ser Asp Thr
1880	1885	1890

Gly Gln	Gly His	Leu Gln	Gln	Lys Ser	Val Arg	Gln	Thr Val	Leu
1895			1900			1905		
Ser Glu	Val Val	Leu Glu	Arg	Thr Glu	Leu Glu	Ile	Ser Tyr	Ala
1910			1915			1920		
Pro Arg	Leu Asp	Gln Glu	Lys	Glu Glu	Leu Leu	Arg	Lys Lys	Leu
1925			1930			1935		
Gln Leu	Asn Pro	Thr Pro	Ala	Asn Arg	Ser Arg	Tyr	Gln Ser	Arg
1940			1945			1950		
Lys Val	Glu Asn	Met Lys	Ala	Ile Thr	Ala Arg	Arg	Ile Leu	Gln
1955			1960			1965		
Gly Leu	Gly His	Tyr Leu	Lys	Ala Glu	Gly Lys	Val	Glu Cys	Tyr
1970			1975			1980		
Arg Thr	Leu His	Pro Val	Pro	Leu Tyr	Ser Ser	Ser	Val Asn	Arg
1985			1990			1995		
Ala Phe	Ser Ser	Pro Lys	Val	Ala Val	Glu Ala	Cys	Asn Ala	Met
2000			2005			2010		
Leu Lys	Glu Asn	Phe Pro	Thr	Val Ala	Ser Tyr	Cys	Ile Ile	Pro
2015			2020			2025		
Glu Tyr	Asp Ala	Tyr Leu	Asp	Met Val	Asp Gly	Ala	Ser Cys	Cys
2030			2035			2040		
Leu Asp	Thr Ala	Ser Phe	Cys	Pro Ala	Lys Leu	Arg	Ser Phe	Pro
2045			2050			2055		
Lys Lys	His Ser	Tyr Leu	Glu	Pro Thr	Ile Arg	Ser	Ala Val	Pro
2060			2065			2070		
Ser Ala	Ile Gln	Asn Thr	Leu	Gln Asn	Val Leu	Ala	Ala Ala	Thr
2075			2080			2085		
Lys Arg	Asn Cys	Asn Val	Thr	Gln Met	Arg Glu	Leu	Pro Val	Leu
2090			2095			2100		
Asp Ser	Ala Ala	Phe Asn	Val	Glu Cys	Phe Lys	Lys	Tyr Ala	Cys
2105			2110			2115		
Asn Asn	Glu Tyr	Trp Glu	Thr	Phe Lys	Glu Asn	Pro	Ile Arg	Leu
2120			2125			2130		
Thr Glu	Glu Asn	Val Val	Asn	Tyr Ile	Thr Lys	Leu	Lys Gly	Pro
2135			2140			2145		
Lys Ala	Ala Ala	Leu Phe	Ala	Lys Thr	His Asn	Leu	Asn Met	Leu
2150			2155			2160		
Gln Asp	Ile Pro	Met Asp	Arg	Phe Val	Met Asp	Leu	Lys Arg	Asp
2165			2170			2175		
Val Lys	Val Thr	Pro Gly	Thr	Lys His	Thr Glu	Glu	Arg Pro	Lys

2180	2185	2190
Val Gln Val Ile Gln Ala	Ala Asp Pro Leu Ala	Thr Ala Tyr Leu
2195	2200	2205
Cys Gly Ile His Arg Glu	Leu Val Arg Arg Leu	Asn Ala Val Leu
2210	2215	2220
Leu Pro Asn Ile His Thr	Leu Phe Asp Met Ser	Ala Glu Asp Phe
2225	2230	2235
Asp Ala Ile Ile Ala Glu	His Phe Gln Pro Gly	Asp Cys Val Leu
2240	2245	2250
Glu Thr Asp Ile Ala Ser	Phe Asp Lys Ser Glu	Asp Asp Ala Met
2255	2260	2265
Ala Leu Thr Ala Leu Met	Ile Leu Glu Asp Leu	Gly Val Asp Ala
2270	2275	2280
Glu Leu Leu Thr Leu Ile	Glu Ala Ala Phe Gly	Glu Ile Ser Ser
2285	2290	2295
Ile His Leu Pro Thr Lys	Thr Lys Phe Lys Phe	Gly Ala Met Met
2300	2305	2310
Lys Ser Gly Met Phe Leu	Thr Leu Phe Val Asn	Thr Val Ile Asn
2315	2320	2325
Ile Val Ile Ala Ser Arg	Val Leu Arg Glu Arg	Leu Thr Gly Ser
2330	2335	2340
Pro Cys Ala Ala Phe Ile	Gly Asp Asp Asn Ile	Val Lys Gly Val
2345	2350	2355
Lys Ser Asp Lys Leu Met	Ala Asp Arg Cys Ala	Thr Trp Leu Asn
2360	2365	2370
Met Glu Val Lys Ile Ile	Asp Ala Val Val Gly	Glu Lys Ala Pro
2375	2380	2385
Tyr Phe Cys Gly Gly Phe	Ile Leu Cys Asp Ser	Val Thr Gly Thr
2390	2395	2400
Ala Cys Arg Val Ala Asp	Pro Leu Lys Arg Leu	Phe Lys Leu Gly
2405	2410	2415
Lys Pro Leu Ala Ala Asp	Asp Glu His Asp Asp	Asp Arg Arg Arg
2420	2425	2430
Ala Leu His Glu Glu Ser	Thr Arg Trp Asn Arg	Val Gly Ile Leu
2435	2440	2445
Ser Glu Leu Cys Lys Ala	Val Glu Ser Arg Tyr	Glu Thr Val Gly
2450	2455	2460
Thr Ser Ile Ile Val Met	Ala Met Thr Thr Leu	Ala Ser Ser Val
2465	2470	2475

Lys Ser Phe Ser Tyr Leu Arg Gly Ala Pro Ile Thr Leu Tyr Gly
 2480 2485 2490
 Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu
 2495 2500 2505
 Glu Asn Pro Gly Pro
 2510

<210> 13

<211> 7539

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C16 nsP-2A coding sequence DNA

<400> 13

```

atggagaaag ttcacgttga catcgaggaa gacagcccat tcctcagagc tttgcagcgg      60
agcttcccgc agtttgaggt agaagccaag caggtcactg ataatgacca tgctaattgcc      120
agagcgTTTT cgcattctggc ttcaaaactg atcgaaacgg aggtggaccc atccgacacg      180
atccttgaca ttggaagtgc gcccggccgc agaattgtatt ctaagcacia gtatcattgt      240
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg      300
aagaaaaact gtaaggaaat aactgataag gaattggaca agaaaatgaa ggagctcgcc      360
gccgtcatga gcgaccctga cctggaaact gagactatgt gcctccacga cgacgagtcg      420
tgtcgctacg aagggaagt cgctgtttac caggatgtat acgcggttga cggaccgaca      480
agtctctatc accaagccaa taaggaggtt agagtcgcct actggatagg ctttgacacc      540
acccctttta tgtttaagaa cttggctgga gcatatccat catactctac caactgggcc      600
gacgaaaccg tgtaaacggc tcgtaacata ggcctatgca gctctgacgt tatggagcgg      660
tcacgtagag ggatgtccat tcttagaaag aagtatttga aaccatccaa caatgttcta      720
ttctctgttg gctcgaccat ctaccacgag aagagggact tactgaggag ctggcacctg      780
ccgtctgtat ttcaattacg tggcaagcaa aattacacat gtcggtgtga gactatagtt      840
agttgcgacg ggtacgtcgt taaaagaata gctatcagtc caggcctgta tgggaagcct      900
tcaggctatg ctgtacgat gcaccgcgag ggattcttgt gctgcaaagt gacagacaca      960
ttgaacgggg agagggtctc ttttccctg tgcacgtatg tgccagctac attgtgtgac     1020
caaatgactg gcatactggc aacagatgtc agtgcggacg acgcgcaaaa actgctggtt     1080
gggctcaacc agcgtatagt cgtcaacggc cgcaccaga gaaacaccaa taccatgaaa     1140
aattaccttt tgcccgtagt ggcccaggca tttgctaggt gggcaaagga atataaggaa     1200
gatcaagaag atgaaaggcc actaggacta cgagatagac agttagtcat ggggtgttgt     1260
tgggctttta gaaggcacia gataacatct atttataagc gcccggatac ccaaaccatc     1320
atcaaagtga acagcgattt ccaactcattc gtgctgcccc ggataggcag taacacattg     1380
gagatcgggc tgagaacaag aatcaggaaa atgttagagg agcacaagga gccgtcacct     1440
ctcattaccg ccgaggacgt acaagaagct aagtgcgcag ccgatgaggc taaggaggtg     1500
cgtgaagccg aggagttgcg cgcagctcta ccacctttgg cagctgatgt tgaggagccc     1560

```

actctggaag	ccgatgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagt	gaaaaattat	cttgcatcca	ccctctcgct	1740
gaacaagtca	tagtgataac	acactctggc	cgaaaagggc	gttatgccgt	ggaaccatac	1800
catggtaaag	tagtggtgcc	agagggacat	gcaatacccg	tccaggactt	tcaagctctg	1860
agtgaaagtg	ccaccattgt	gtacaacgaa	cgtgagttcg	taaacaggta	cctgcaccat	1920
attgccacac	atggaggagc	gctgaacact	gatgaagaat	attacaaaac	tgtcaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtacttg	ggctagggct	cacaggcgag	ctgggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccaag	taccaacat	aggggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatatgtg	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgcacaca	agtcttccac	2520
aaaagcatct	ctcgccgttg	cactaaatct	gtgacttcgg	tcgtctcaac	cttgttttac	2580
gacaaaaaaaa	tgagaacgac	gaatccgaaa	gagactaaga	ttgtgattga	cactaccggc	2640
agtaccaaac	ctaagcagga	cgatctcatt	ctcacttggt	tcagagggtg	ggtgaagcag	2700
ttgcaaatag	attacaaagg	caacgaaata	atgacggcag	ctgcctctca	agggtgacc	2760
cgtaaagggtg	tgtatgccgt	tcggtacaag	gtgaatgaaa	atcctctgta	cgcacccacc	2820
tcagaacatg	tgaacgtcct	actgacccgc	acggaggacc	gcatcgtgtg	gaaaacacta	2880
gccggcgacc	catggataaa	aacactgact	gccaaagtacc	ctgggaattt	cactgccacg	2940
atagaggagt	ggcaagcaga	gcatgatgcc	atcatgaggc	acatcttgga	gagaccggac	3000
cctaccgacg	tcttccagaa	taaggcaaac	gtgtgttggg	ccaaggcttt	agtgccggtg	3060
ctgaagaccg	ctggcataga	catgaccact	gaacaatgga	acactgtgga	ttattttgaa	3120
acggacaaaag	ctcactcagc	agagatagta	ttgaaccaac	tatgcgtgag	gttctttgga	3180
ctcgatctgg	actccggtct	attttctgca	cccactgttc	cgttatccat	taggaataat	3240
cactgggata	actccccgtc	gcctaacatg	tacgggctga	ataaagaagt	ggtccgtcag	3300
ctctctcgca	ggtaaccaca	actgcctcgg	gcagttgcc	ctggaagagt	ctatgacatg	3360
aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	cttttagtcct	ccaccataat	gaacacccac	agagtgactt	ttcttcattc	3480
gtcagcaaat	tgaagggcag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatggttg	actggttgtc	agaccggcct	gaggtacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gacccatat	3660
aaataaccatc	actatcagca	gtgtgaagac	catgccatta	agcttagcat	gttgaccaag	3720
aaagcttgtc	tgcacttgaa	tcccggcgga	acctgtgtca	gcataggtta	tggttacgct	3780
gacagggcca	gcgaaaagcat	cattggtgct	atagcgcggc	agttcaagtt	ttcccgggta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	tgggtacgat	3900

cgcaaggccc	gtacgcacaa	tccttacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
ggggtgtgcg	gagcgctgta	taagaaattc	ccggaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtcc	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtgacag	aacctgatgc	agagctggtg	agggtgcac	cgaagagttc	tttggctgga	4560
aggaagggt	acagcacaag	cgatggcaaa	actttctcat	atttggaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgccttgctt	gtgcatccat	4800
gccatgactc	cagaaaagagt	acagcgccta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgtcatcct	ttccattgcc	gaagtataga	atcactggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgctt	gcgtatatc	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgaccac	5160
caggtgctgc	aagtcgaggc	agacattcac	gggccgcctt	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaacag	tattcaggaa	ccctccacat	5400
cccgtccgc	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagttttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgtcac	gcactcctag	caggtcggtc	tcgagaacca	gcctggtctc	caaccgcga	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgccctga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820
aatcccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgetacc	gaaccctgca	tcctgttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggtcgcagt	ggaagcctgt	aacgccatgt	tgaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatggt	tgacggagct	6120
tcatgctgct	tagacactgc	cagtttttgc	cctgcaaagc	tgcgcagctt	tccaaagaaa	6180
cactcctatt	tggaaccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240

```

cagaacgtcc tggcagctgc cacaaaaaga aattgcaatg tcacgcaa at gagagaattg 6300
cccgatattgg attcggcggc ctttaatgtg gaatgcttca agaaatatgc gtgtaataat 6360
gaatatattggg aaacgtttta agaaaacccc atcaggctta ctgaagaaaa cgtggtaaat 6420
tacattacca aattaaaagg accaaaagct gctgctcttt ttgcgaagac acataatttg 6480
aatatgttgc aggacatacc aatggacagg tttgtaatgg acttaaagag agacgtgaaa 6540
gtgactccag gaacaaaaca tactgaagaa cggcccaagg tacaggtgat ccaggctgcc 6600
gatccgctag caacagcgta tctgtgcgga atccaccgag agctggtttag gagattaaat 6660
gcggtcctgc ttccgaacat tcatacactg tttgatattg cggctgaaga ctttgacgct 6720
attatagccg agcacttcca gcctggggat tgtgttctgg aaactgacat cgcgtcgttt 6780
gataaaaagtg aggacgacgc catggctctg accgcgttaa tgattctgga agacttaggt 6840
gtggacgcag agctgttgac gctgattgag gcggctttcg gcgaaatttc atcaatacat 6900
ttgcccacta aaactaaatt taagtctggc gccatgatga agtccggcat gtttctgacc 6960
ctgttcgtga acaccgtgat caacatcgtg atcgccagcc ggggtgctgag agagagactg 7020
acaggatctc cttgcgccgc cttcatcggc gacgacaata tcgtgaaggg cgtgaagtcc 7080
gacaagctga tggccgatag atgcgccacc tggctgaaca tggaagtga gatcatcgac 7140
gccgtcgtgg gcgagaaggc cccttatttt tgcggcggct tcatcctgtg cgacagcgtg 7200
acaggcacag cctgcagagt tgccgatact ctgaagcggc tgttcaagct gggaaaacct 7260
ctggccgccg acgacgagca cgacgacgat agacgtagag ccctgcacga ggaatccacc 7320
agatggaaca gaggggcat cctgagcgag ctgtgcaagg ccgtggaaag cagatacgag 7380
acagtgggca ccagcatcat tgtgatggca atgaccacac tggccagcag cgtgaaaagc 7440
ttcagctacc taaggggccc ccctatcaca ctgtacggct ctggcgaagg cagaggcagc 7500
cttctgacat gtggcgacgt ggaagagaac cccggacct 7539

```

<210> 14

<211> 128

<212> DNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C16 subgenomic promoter DNA

<400> 14

```

acttccatca tagttatggc catgactact ctagctagca gtgttaa atc attcagctac 60
ctgagagggg cccctataac tctctacggc taacctgaat ggactacgac atagtctagt 120
ccgccaag 128

```

<210> 15

<211> 2751

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, C16 nsP-2A-eGFP protein

<220>

<221> misc_feature
 <222> (1880)..(1880)
 <223> Xaa can be any naturally occurring amino acid
 <400> 15
 Met Glu Lys Val His Val Asp Ile Glu Glu Asp Ser Pro Phe Leu Arg
 1 5 10 15
 Ala Leu Gln Arg Ser Phe Pro Gln Phe Glu Val Glu Ala Lys Gln Val
 20 25 30
 Thr Asp Asn Asp His Ala Asn Ala Arg Ala Phe Ser His Leu Ala Ser
 35 40 45
 Lys Leu Ile Glu Thr Glu Val Asp Pro Ser Asp Thr Ile Leu Asp Ile
 50 55 60
 Gly Ser Ala Pro Ala Arg Arg Met Tyr Ser Lys His Lys Tyr His Cys
 65 70 75 80
 Ile Cys Pro Met Arg Cys Ala Glu Asp Pro Asp Arg Leu Tyr Lys Tyr
 85 90 95
 Ala Thr Lys Leu Lys Lys Asn Cys Lys Glu Ile Thr Asp Lys Glu Leu
 100 105 110
 Asp Lys Lys Met Lys Glu Leu Ala Ala Val Met Ser Asp Pro Asp Leu
 115 120 125
 Glu Thr Glu Thr Met Cys Leu His Asp Asp Glu Ser Cys Arg Tyr Glu
 130 135 140
 Gly Gln Val Ala Val Tyr Gln Asp Val Tyr Ala Val Asp Gly Pro Thr
 145 150 155 160
 Ser Leu Tyr His Gln Ala Asn Lys Gly Val Arg Val Ala Tyr Trp Ile
 165 170 175
 Gly Phe Asp Thr Thr Pro Phe Met Phe Lys Asn Leu Ala Gly Ala Tyr
 180 185 190
 Pro Ser Tyr Ser Thr Asn Trp Ala Asp Glu Thr Val Leu Thr Ala Arg
 195 200 205
 Asn Ile Gly Leu Cys Ser Ser Asp Val Met Glu Arg Ser Arg Arg Gly
 210 215 220
 Met Ser Ile Leu Arg Lys Lys Tyr Leu Lys Pro Ser Asn Asn Val Leu
 225 230 235 240
 Phe Ser Val Gly Ser Thr Ile Tyr His Glu Lys Arg Asp Leu Leu Arg
 245 250 255
 Ser Trp His Leu Pro Ser Val Phe His Leu Arg Gly Lys Gln Asn Tyr
 260 265 270
 Thr Cys Arg Cys Glu Thr Ile Val Ser Cys Asp Gly Tyr Val Val Lys

275	280	285
Arg Ile Ala Ile Ser Pro Gly Leu Tyr Gly Lys Pro Ser Gly Tyr Ala		
290	295	300
Ala Thr Met His Arg Glu Gly Phe Leu Cys Cys Lys Val Thr Asp Thr		
305	310	315
Leu Asn Gly Glu Arg Val Ser Phe Pro Val Cys Thr Tyr Val Pro Ala		
325	330	335
Thr Leu Cys Asp Gln Met Thr Gly Ile Leu Ala Thr Asp Val Ser Ala		
340	345	350
Asp Asp Ala Gln Lys Leu Leu Val Gly Leu Asn Gln Arg Ile Val Val		
355	360	365
Asn Gly Arg Thr Gln Arg Asn Thr Asn Thr Met Lys Asn Tyr Leu Leu		
370	375	380
Pro Val Val Ala Gln Ala Phe Ala Arg Trp Ala Lys Glu Tyr Lys Glu		
385	390	395
Asp Gln Glu Asp Glu Arg Pro Leu Gly Leu Arg Asp Arg Gln Leu Val		
405	410	415
Met Gly Cys Cys Trp Ala Phe Arg Arg His Lys Ile Thr Ser Ile Tyr		
420	425	430
Lys Arg Pro Asp Thr Gln Thr Ile Ile Lys Val Asn Ser Asp Phe His		
435	440	445
Ser Phe Val Leu Pro Arg Ile Gly Ser Asn Thr Leu Glu Ile Gly Leu		
450	455	460
Arg Thr Arg Ile Arg Lys Met Leu Glu Glu His Lys Glu Pro Ser Pro		
465	470	475
Leu Ile Thr Ala Glu Asp Val Gln Glu Ala Lys Cys Ala Ala Asp Glu		
485	490	495
Ala Lys Glu Val Arg Glu Ala Glu Glu Leu Arg Ala Ala Leu Pro Pro		
500	505	510
Leu Ala Ala Asp Val Glu Glu Pro Thr Leu Glu Ala Asp Val Asp Leu		
515	520	525
Met Leu Gln Glu Ala Gly Ala Gly Ser Val Glu Thr Pro Arg Gly Leu		
530	535	540
Ile Lys Val Thr Ser Tyr Ala Gly Glu Asp Lys Ile Gly Ser Tyr Ala		
545	550	555
Val Leu Ser Pro Gln Ala Val Leu Lys Ser Glu Lys Leu Ser Cys Ile		
565	570	575
His Pro Leu Ala Glu Gln Val Ile Val Ile Thr His Ser Gly Arg Lys		
580	585	590

Gly Arg Tyr Ala Val Glu Pro Tyr His Gly Lys Val Val Val Pro Glu		
595	600	605
Gly His Ala Ile Pro Val Gln Asp Phe Gln Ala Leu Ser Glu Ser Ala		
610	615	620
Thr Ile Val Tyr Asn Glu Arg Glu Phe Val Asn Arg Tyr Leu His His		
625	630	635
Ile Ala Thr His Gly Gly Ala Leu Asn Thr Asp Glu Glu Tyr Tyr Lys		
	645	650
Thr Val Lys Pro Ser Glu His Asp Gly Glu Tyr Leu Tyr Asp Ile Asp		
	660	665
Arg Lys Gln Cys Val Lys Lys Glu Leu Val Thr Gly Leu Gly Leu Thr		
	675	680
Gly Glu Leu Val Asp Pro Pro Phe His Glu Phe Ala Tyr Glu Ser Leu		
690	695	700
Arg Thr Arg Pro Ala Ala Pro Tyr Gln Val Pro Thr Ile Gly Val Tyr		
705	710	715
Gly Val Pro Gly Ser Gly Lys Ser Gly Ile Ile Lys Ser Ala Val Thr		
	725	730
Lys Lys Asp Leu Val Val Ser Ala Lys Lys Glu Asn Cys Ala Glu Ile		
	740	745
Ile Arg Asp Val Lys Lys Met Lys Gly Leu Asp Val Asn Ala Arg Thr		
	755	760
Val Asp Ser Val Leu Leu Asn Gly Cys Lys His Pro Val Glu Thr Leu		
770	775	780
Tyr Ile Asp Glu Ala Phe Ala Cys His Ala Gly Thr Leu Arg Ala Leu		
785	790	795
Ile Ala Ile Ile Arg Pro Lys Lys Ala Val Leu Cys Gly Asp Pro Lys		
	805	810
Gln Cys Gly Phe Phe Asn Met Met Cys Leu Lys Val His Phe Asn His		
	820	825
Glu Ile Cys Thr Gln Val Phe His Lys Ser Ile Ser Arg Arg Cys Thr		
	835	840
Lys Ser Val Thr Ser Val Val Ser Thr Leu Phe Tyr Asp Lys Lys Met		
850	855	860
Arg Thr Thr Asn Pro Lys Glu Thr Lys Ile Val Ile Asp Thr Thr Gly		
865	870	875
Ser Thr Lys Pro Lys Gln Asp Asp Leu Ile Leu Thr Cys Phe Arg Gly		
	885	890
Trp Val Lys Gln Leu Gln Ile Asp Tyr Lys Gly Asn Glu Ile Met Thr		

900	905	910
Ala Ala Ala Ser Gln Gly Leu Thr Arg Lys Gly Val Tyr Ala Val Arg		
915	920	925
Tyr Lys Val Asn Glu Asn Pro Leu Tyr Ala Pro Thr Ser Glu His Val		
930	935	940
Asn Val Leu Leu Thr Arg Thr Glu Asp Arg Ile Val Trp Lys Thr Leu		
945	950	955
Ala Gly Asp Pro Trp Ile Lys Thr Leu Thr Ala Lys Tyr Pro Gly Asn		
965	970	975
Phe Thr Ala Thr Ile Glu Glu Trp Gln Ala Glu His Asp Ala Ile Met		
980	985	990
Arg His Ile Leu Glu Arg Pro Asp Pro Thr Asp Val Phe Gln Asn Lys		
995	1000	1005
Ala Asn Val Cys Trp Ala Lys Ala Leu Val Pro Val Leu Lys Thr		
1010	1015	1020
Ala Gly Ile Asp Met Thr Thr Glu Gln Trp Asn Thr Val Asp Tyr		
1025	1030	1035
Phe Glu Thr Asp Lys Ala His Ser Ala Glu Ile Val Leu Asn Gln		
1040	1045	1050
Leu Cys Val Arg Phe Phe Gly Leu Asp Leu Asp Ser Gly Leu Phe		
1055	1060	1065
Ser Ala Pro Thr Val Pro Leu Ser Ile Arg Asn Asn His Trp Asp		
1070	1075	1080
Asn Ser Pro Ser Pro Asn Met Tyr Gly Leu Asn Lys Glu Val Val		
1085	1090	1095
Arg Gln Leu Ser Arg Arg Tyr Pro Gln Leu Pro Arg Ala Val Ala		
1100	1105	1110
Thr Gly Arg Val Tyr Asp Met Asn Thr Gly Thr Leu Arg Asn Tyr		
1115	1120	1125
Asp Pro Arg Ile Asn Leu Val Pro Val Asn Arg Arg Leu Pro His		
1130	1135	1140
Ala Leu Val Leu His His Asn Glu His Pro Gln Ser Asp Phe Ser		
1145	1150	1155
Ser Phe Val Ser Lys Leu Lys Gly Arg Thr Val Leu Val Val Gly		
1160	1165	1170
Glu Lys Leu Ser Val Pro Gly Lys Met Val Asp Trp Leu Ser Asp		
1175	1180	1185
Arg Pro Glu Ala Thr Phe Arg Ala Arg Leu Asp Leu Gly Ile Pro		
1190	1195	1200

Gly Asp	Val Pro Lys Tyr Asp	Ile Ile Phe Val Asn	Val Arg Thr
1205	1210	1215	
Pro Tyr	Lys Tyr His His Tyr	Gln Gln Cys Glu Asp	His Ala Ile
1220	1225	1230	
Lys Leu	Ser Met Leu Thr Lys	Lys Ala Cys Leu His	Leu Asn Pro
1235	1240	1245	
Gly Gly	Thr Cys Val Ser Ile	Gly Tyr Gly Tyr Ala	Asp Arg Ala
1250	1255	1260	
Ser Glu	Ser Ile Ile Gly Ala	Ile Ala Arg Gln Phe	Lys Phe Ser
1265	1270	1275	
Arg Val	Cys Lys Pro Lys Ser	Ser Leu Glu Glu Thr	Glu Val Leu
1280	1285	1290	
Phe Val	Phe Ile Gly Tyr Asp	Arg Lys Ala Arg Thr	His Asn Pro
1295	1300	1305	
Tyr Lys	Leu Ser Ser Thr Leu	Thr Asn Ile Tyr Thr	Gly Ser Arg
1310	1315	1320	
Leu His	Glu Ala Gly Cys Ala	Pro Ser Tyr His Val	Val Arg Gly
1325	1330	1335	
Asp Ile	Ala Thr Ala Thr Glu	Gly Val Ile Ile Asn	Ala Ala Asn
1340	1345	1350	
Ser Lys	Gly Gln Pro Gly Gly	Gly Val Cys Gly Ala	Leu Tyr Lys
1355	1360	1365	
Lys Phe	Pro Glu Ser Phe Asp	Leu Gln Pro Ile Glu	Val Gly Lys
1370	1375	1380	
Ala Arg	Leu Val Lys Gly Ala	Ala Lys His Ile Ile	His Ala Val
1385	1390	1395	
Gly Pro	Asn Phe Asn Lys Val	Ser Glu Val Glu Gly	Asp Lys Gln
1400	1405	1410	
Leu Ala	Glu Ala Tyr Glu Ser	Ile Ala Lys Ile Val	Asn Asp Asn
1415	1420	1425	
Asn Tyr	Lys Ser Val Ala Ile	Pro Leu Leu Ser Thr	Gly Ile Phe
1430	1435	1440	
Ser Gly	Asn Lys Asp Arg Leu	Thr Gln Ser Leu Asn	His Leu Leu
1445	1450	1455	
Thr Ala	Leu Asp Thr Thr Asp	Ala Asp Val Ala Ile	Tyr Cys Arg
1460	1465	1470	
Asp Lys	Lys Trp Glu Met Thr	Leu Lys Glu Ala Val	Ala Arg Arg
1475	1480	1485	
Glu Ala	Val Glu Glu Ile Cys	Ile Ser Asp Asp Ser	Ser Val Thr

1490	1495	1500
Glu Pro Asp Ala Glu Leu Val	Arg Val His Pro Lys	Ser Ser Leu
1505	1510	1515
Ala Gly Arg Lys Gly Tyr Ser	Thr Ser Asp Gly Lys	Thr Phe Ser
1520	1525	1530
Tyr Leu Glu Gly Thr Lys Phe	His Gln Ala Ala Lys	Asp Ile Ala
1535	1540	1545
Glu Ile Asn Ala Met Trp Pro	Val Ala Thr Glu Ala	Asn Glu Gln
1550	1555	1560
Val Cys Met Tyr Ile Leu Gly	Glu Ser Met Ser Ser	Ile Arg Ser
1565	1570	1575
Lys Cys Pro Val Glu Glu Ser	Glu Ala Ser Thr Pro	Pro Ser Thr
1580	1585	1590
Leu Pro Cys Leu Cys Ile His	Ala Met Thr Pro Glu	Arg Val Gln
1595	1600	1605
Arg Leu Lys Ala Ser Arg Pro	Glu Gln Ile Thr Val	Cys Ser Ser
1610	1615	1620
Phe Pro Leu Pro Lys Tyr Arg	Ile Thr Gly Val Gln	Lys Ile Gln
1625	1630	1635
Cys Ser Gln Pro Ile Leu Phe	Ser Pro Lys Val Pro	Ala Tyr Ile
1640	1645	1650
His Pro Arg Lys Tyr Leu Val	Glu Thr Pro Pro Val	Asp Glu Thr
1655	1660	1665
Pro Glu Pro Ser Ala Glu Asn	Gln Ser Thr Glu Gly	Thr Pro Glu
1670	1675	1680
Gln Pro Pro Leu Ile Thr Glu	Asp Glu Thr Arg Thr	Arg Thr Pro
1685	1690	1695
Glu Pro Ile Ile Ile Glu Glu	Glu Glu Glu Asp Ser	Ile Ser Leu
1700	1705	1710
Leu Ser Asp Gly Pro Thr His	Gln Val Leu Gln Val	Glu Ala Asp
1715	1720	1725
Ile His Gly Pro Pro Ser Val	Ser Ser Ser Ser Trp	Ser Ile Pro
1730	1735	1740
His Ala Ser Asp Phe Asp Val	Asp Ser Leu Ser Ile	Leu Asp Thr
1745	1750	1755
Leu Glu Gly Ala Ser Val Thr	Ser Gly Ala Thr Ser	Ala Glu Thr
1760	1765	1770
Asn Ser Tyr Phe Ala Lys Ser	Met Glu Phe Leu Ala	Arg Pro Val
1775	1780	1785

Pro Ala	Pro Arg	Thr Val	Phe	Arg Asn	Pro Pro	His	Pro Ala	Pro											
1790				1795		1800													
Arg Thr	Arg Thr	Pro Ser	Leu	Ala Pro	Ser Arg	Ala	Cys Ser	Arg											
1805				1810		1815													
Thr Ser	Leu Val	Ser Thr	Pro	Pro Gly	Val Asn	Arg	Val Ile	Thr											
1820				1825		1830													
Arg Glu	Glu Leu	Glu Ala	Leu	Thr Pro	Ser Arg	Thr	Pro Ser	Arg											
1835				1840		1845													
Ser Val	Ser Arg	Thr Ser	Leu	Val Ser	Asn Pro	Pro	Gly Val	Asn											
1850				1855		1860													
Arg Val	Ile Thr	Arg Glu	Glu	Phe Glu	Ala Phe	Val	Ala Gln	Gln											
1865				1870		1875													
Gln Xaa	Arg Phe	Asp Ala	Gly	Ala Tyr	Ile Phe	Ser	Ser Asp	Thr											
1880				1885		1890													
Gly Gln	Gly His	Leu Gln	Gln	Lys Ser	Val Arg	Gln	Thr Val	Leu											
1895				1900		1905													
Ser Glu	Val Val	Leu Glu	Arg	Thr Glu	Leu Glu	Ile	Ser Tyr	Ala											
1910				1915		1920													
Pro Arg	Leu Asp	Gln Glu	Lys	Glu Glu	Leu Leu	Arg	Lys Lys	Leu											
1925				1930		1935													
Gln Leu	Asn Pro	Thr Pro	Ala	Asn Arg	Ser Arg	Tyr	Gln Ser	Arg											
1940				1945		1950													
Lys Val	Glu Asn	Met Lys	Ala	Ile Thr	Ala Arg	Arg	Ile Leu	Gln											
1955				1960		1965													
Gly Leu	Gly His	Tyr Leu	Lys	Ala Glu	Gly Lys	Val	Glu Cys	Tyr											
1970				1975		1980													
Arg Thr	Leu His	Pro Val	Pro	Leu Tyr	Ser Ser	Ser	Val Asn	Arg											
1985				1990		1995													
Ala Phe	Ser Ser	Pro Lys	Val	Ala Val	Glu Ala	Cys	Asn Ala	Met											
2000				2005		2010													
Leu Lys	Glu Asn	Phe Pro	Thr	Val Ala	Ser Tyr	Cys	Ile Ile	Pro											
2015				2020		2025													
Glu Tyr	Asp Ala	Tyr Leu	Asp	Met Val	Asp Gly	Ala	Ser Cys	Cys											
2030				2035		2040													
Leu Asp	Thr Ala	Ser Phe	Cys	Pro Ala	Lys Leu	Arg	Ser Phe	Pro											
2045				2050		2055													
Lys Lys	His Ser	Tyr Leu	Glu	Pro Thr	Ile Arg	Ser	Ala Val	Pro											
2060				2065		2070													
Ser Ala	Ile Gln	Asn Thr	Leu	Gln Asn	Val Leu	Ala	Ala Ala	Thr											

2075	2080	2085
Lys Arg Asn Cys Asn Val Thr	Gln Met Arg Glu Leu	Pro Val Leu
2090	2095	2100
Asp Ser Ala Ala Phe Asn Val	Glu Cys Phe Lys Lys	Tyr Ala Cys
2105	2110	2115
Asn Asn Glu Tyr Trp Glu Thr	Phe Lys Glu Asn Pro	Ile Arg Leu
2120	2125	2130
Thr Glu Glu Asn Val Val Asn	Tyr Ile Thr Lys Leu	Lys Gly Pro
2135	2140	2145
Lys Ala Ala Ala Leu Phe Ala	Lys Thr His Asn Leu	Asn Met Leu
2150	2155	2160
Gln Asp Ile Pro Met Asp Arg	Phe Val Met Asp Leu	Lys Arg Asp
2165	2170	2175
Val Lys Val Thr Pro Gly Thr	Lys His Thr Glu Glu	Arg Pro Lys
2180	2185	2190
Val Gln Val Ile Gln Ala Ala	Asp Pro Leu Ala Thr	Ala Tyr Leu
2195	2200	2205
Cys Gly Ile His Arg Glu Leu	Val Arg Arg Leu Asn	Ala Val Leu
2210	2215	2220
Leu Pro Asn Ile His Thr Leu	Phe Asp Met Ser Ala	Glu Asp Phe
2225	2230	2235
Asp Ala Ile Ile Ala Glu His	Phe Gln Pro Gly Asp	Cys Val Leu
2240	2245	2250
Glu Thr Asp Ile Ala Ser Phe	Asp Lys Ser Glu Asp	Asp Ala Met
2255	2260	2265
Ala Leu Thr Ala Leu Met Ile	Leu Glu Asp Leu Gly	Val Asp Ala
2270	2275	2280
Glu Leu Leu Thr Leu Ile Glu	Ala Ala Phe Gly Glu	Ile Ser Ser
2285	2290	2295
Ile His Leu Pro Thr Lys Thr	Lys Phe Lys Phe Gly	Ala Met Met
2300	2305	2310
Lys Ser Gly Met Phe Leu Thr	Leu Phe Val Asn Thr	Val Ile Asn
2315	2320	2325
Ile Val Ile Ala Ser Arg Val	Leu Arg Glu Arg Leu	Thr Gly Ser
2330	2335	2340
Pro Cys Ala Ala Phe Ile Gly	Asp Asp Asn Ile Val	Lys Gly Val
2345	2350	2355
Lys Ser Asp Lys Leu Met Ala	Asp Arg Cys Ala Thr	Trp Leu Asn
2360	2365	2370

Met Glu Val Lys Ile Ile Asp Ala Val Val Gly Glu Lys Ala Pro		
2375	2380	2385
Tyr Phe Cys Gly Gly Phe Ile Leu Cys Asp Ser Val Thr Gly Thr		
2390	2395	2400
Ala Cys Arg Val Ala Asp Pro Leu Lys Arg Leu Phe Lys Leu Gly		
2405	2410	2415
Lys Pro Leu Ala Ala Asp Asp Glu His Asp Asp Asp Arg Arg Arg		
2420	2425	2430
Ala Leu His Glu Glu Ser Thr Arg Trp Asn Arg Val Gly Ile Leu		
2435	2440	2445
Ser Glu Leu Cys Lys Ala Val Glu Ser Arg Tyr Glu Thr Val Gly		
2450	2455	2460
Thr Ser Ile Ile Val Met Ala Met Thr Thr Leu Ala Ser Ser Val		
2465	2470	2475
Lys Ser Phe Ser Tyr Leu Arg Gly Ala Pro Ile Thr Leu Tyr Gly		
2480	2485	2490
Ser Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu		
2495	2500	2505
Glu Asn Pro Gly Pro Val Ser Lys Gly Glu Glu Leu Phe Thr Gly		
2510	2515	2520
Val Val Pro Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His		
2525	2530	2535
Lys Phe Ser Val Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly		
2540	2545	2550
Lys Leu Thr Leu Lys Phe Ile Cys Thr Thr Gly Lys Leu Pro Val		
2555	2560	2565
Pro Trp Pro Thr Leu Val Thr Thr Leu Thr Tyr Gly Val Gln Cys		
2570	2575	2580
Phe Ser Arg Tyr Pro Asp His Met Lys Gln His Asp Phe Phe Lys		
2585	2590	2595
Ser Ala Met Pro Glu Gly Tyr Val Gln Glu Arg Thr Ile Phe Phe		
2600	2605	2610
Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu Val Lys Phe Glu		
2615	2620	2625
Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly Ile Asp Phe		
2630	2635	2640
Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr Asn Tyr		
2645	2650	2655
Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn Gly		

2660	2665	2670
Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser		
2675	2680	2685
Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly Asp		
2690	2695	2700
Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser		
2705	2710	2715
Lys Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu		
2720	2725	2730
Leu Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu		
2735	2740	2745
Leu Tyr Lys		
2750		
<210> 16		
<211> 8256		
<212> DNA		
<213> Artificial Sequence		
<220>		
<223> Synthetic construct, C16 nsP-2A-eGFP coding sequence DNA		
<400> 16		
atggagaaag ttcacgttga catcgaggaa gacagcccat tcctcagagc tttgcagcgg	60	
agcttccccgc agtttgaggt agaagccaag caggtcactg ataatgacca tgctaattgcc	120	
agagcgtttt cgcacttggc ttcaaaactg atcgaaacgg aggtggaccc atccgacacg	180	
atccttgaca ttggaagtgc gcccggccgc agaattgtatt ctaagcacia gtatcattgt	240	
atctgtccga tgagatgtgc ggaagatccg gacagattgt ataagtatgc aactaagctg	300	
aagaaaaact gtaaggaaat aactgataag gaattggaca agaaaatgaa ggagctcgcc	360	
gccgtcatga gcgaccctga cctggaaact gagactatgt gcctccacga cgacgagtcg	420	
tgctgctacg aagggaagt cgctgtttac caggatgtat acgcggttga cggaccgaca	480	
agtctctatc accaagccaa taaggaggt agagtcgcct actggatagg ctttgacacc	540	
acccctttta tgtttaagaa cttggctgga gcatatccat catactctac caactgggcc	600	
gacgaaaccg tgtaaacggc tcgtaacata ggcctatgca gctctgacgt tatggagcgg	660	
tcacgtagag ggatgtccat tcttagaaag aagtatttga aaccatccaa caatgttcta	720	
ttctctgttg gctcgaccat ctaccacgag aagagggact tactgaggag ctggcacctg	780	
ccgtctgtat ttacttacg tggcaagcaa aattacacat gtcggtgtga gactatagtt	840	
agttgcgacg ggtacgtcgt taaaagaata gctatcagtc caggcctgta tgggaagcct	900	
tcaggctatg ctgctacgat gcaccgcgag ggattcttgt gctgcaaagt gacagacaca	960	
ttgaacgggg agagggtctc ttttcccggtg tgcacgtatg tgccagctac attgtgtgac	1020	
caaatgactg gcatactggc aacagatgtc agtgcggacg acgcgcaaaa actgctggtt	1080	
gggctcaacc agcgtatagt cgtcaacggc cgcaccacga gaaacaccaa taccatgaaa	1140	

aattaccttt	tgcccgtagt	ggcccaggca	tttgctaggt	gggcaaagga	atataaggaa	1200
gatcaagaag	atgaaaggcc	actaggacta	cgagatagac	agttagtcat	ggggtgttgt	1260
tgggctttta	gaaggcacia	gataacatct	atttataagc	gcccggatac	ccaaaccatc	1320
atcaaagtga	acagcgattt	ccactcattc	gtgctgccc	ggataggcag	taacacattg	1380
gagatcgggc	tgagaacaag	aatcaggaaa	atgttagagg	agcacaagga	gccgtcacct	1440
ctcattaccg	ccgaggacgt	acaagaagct	aagtgcgcag	ccgatgaggc	taaggaggtg	1500
cgtgaagccg	aggagttgcg	cgcagctcta	ccacctttgg	cagctgatgt	tgaggagccc	1560
actctggaag	ccgatgtcga	cttgatgtta	caagaggctg	gggccggctc	agtggagaca	1620
cctcgtggct	tgataaaggt	taccagctac	gctggcgagg	acaagatcgg	ctcttacgct	1680
gtgctttctc	cgcaggctgt	actcaagagt	gaaaaattat	cttgcattca	ccctctcgct	1740
gaacaagtca	tagtgataac	acactctggc	cgaaaagggc	gttatgccgt	ggaaccatac	1800
catggtaaag	tagtggtgcc	agagggacat	gcaatacccg	tccaggactt	tcaagctctg	1860
agtgaagtg	ccaccattgt	gtacaacgaa	cgtgagttcg	taaacaggta	cctgcaccat	1920
attgccacac	atggaggagc	gctgaacact	gatgaagaat	attacaaaac	tgtcaagccc	1980
agcgagcacg	acggcgaata	cctgtacgac	atcgacagga	aacagtgcgt	caagaaagaa	2040
ctagtcaactg	ggctagggct	cacaggcgag	ctgggtggatc	ctcccttcca	tgaattcgcc	2100
tacgagagtc	tgagaacacg	accagccgct	ccttaccag	taccaaccat	aggggtgtat	2160
ggcgtgccag	gatcaggcaa	gtctggcatc	attaaaagcg	cagtcaccaa	aaaagatcta	2220
gtggtgagcg	ccaagaaaaga	aaactgtgca	gaaattataa	gggacgtcaa	gaaaatgaaa	2280
gggctggacg	tcaatgccag	aactgtggac	tcagtgtctt	tgaatggatg	caaacacccc	2340
gtagagaccc	tgtatataga	cgaagctttt	gcttgtcatg	caggtactct	cagagcgctc	2400
atagccatta	taagacctaa	aaaggcagtg	ctctgcgggg	atcccaaaca	gtgcggtttt	2460
tttaacatga	tgtgcctgaa	agtgcatttt	aaccacgaga	tttgacaca	agtcttccac	2520
aaaagcatct	ctcgccgttg	cactaaatct	gtgacttcgg	tcgtctcaac	cttgttttac	2580
gacaaaaaaaa	tgagaacgac	gaatccgaaa	gagactaaga	ttgtgattga	cactaccggc	2640
agtaccaaac	ctaagcagga	cgatctcatt	ctcacttggt	tcagagggtg	ggtgaagcag	2700
ttgcaaatag	attacaaaag	caacgaaata	atgacggcag	ctgcctctca	agggctgacc	2760
cgtaaagggtg	tgtatgccgt	tcggtacaag	gtgaatgaaa	atcctctgta	cgcaccacc	2820
tcagaacatg	tgaacgtcct	actgaccgcg	acggaggacc	gcacgtgtg	gaaaacacta	2880
gccggcgacc	catggataaa	aacactgact	gccaaagtacc	ctgggaattt	cactgccacg	2940
atagaggagt	ggcaagcaga	gcatgatgcc	atcatgaggc	acatcttgga	gagaccggac	3000
cctaccgacg	tcttccagaa	taaggcaaac	gtgtgttggg	ccaaggcttt	agtgccggtg	3060
ctgaagaccg	ctggcataga	catgaccact	gaacaatgga	acactgtgga	ttattttgaa	3120
acggacaaaag	ctcactcagc	agagatagta	ttgaaccaac	tatgcgtgag	gttctttgga	3180
ctcgatctgg	actccggtct	atcttctgca	cccactgttc	cgttatccat	taggaataat	3240
cactgggata	actccccgtc	gcctaacatg	tacgggctga	ataaagaagt	ggtccgtcag	3300
ctctctcgca	ggtaccacac	actgcctcgg	gcagttgcca	ctggaagagt	ctatgacatg	3360
aacactggta	cactgcgcaa	ttatgatccg	cgcataaacc	tagtacctgt	aaacagaaga	3420
ctgcctcatg	ctttagtcct	ccaccataat	gaacaccac	agagtgcatt	ttcttcattc	3480

gtcagcaaat	tgaagggcag	aactgtcctg	gtggtcgggg	aaaagttgtc	cgtcccaggc	3540
aaaatggttg	actggttgtc	agaccggcct	gaggctacct	tcagagctcg	gctggattta	3600
ggcatcccag	gtgatgtgcc	caaatatgac	ataatatttg	ttaatgtgag	gaccccatat	3660
aaataccatc	actatcagca	gtgtgaagac	catgccatta	agcttagcat	gttgaccaag	3720
aaagcttgtc	tgcacttgaa	tcccggcgga	acctgtgtca	gcataggtta	tggttacgct	3780
gacagggcca	gcgaaagcat	cattgggtgct	atagcgcggc	agttcaagtt	ttcccgggta	3840
tgcaaaccga	aatcctcact	tgaagagacg	gaagttctgt	ttgtattcat	tggttacgat	3900
cgcaaggccc	gtacgcacaa	tccttacaag	ctttcatcaa	ccttgaccaa	catttataca	3960
ggttccagac	tccacgaagc	cggatgtgca	ccctcatatc	atgtgggtgcg	aggggatatt	4020
gccacggcca	ccgaaggagt	gattataaat	gctgctaaca	gcaaaggaca	acctggcgga	4080
ggggtgtgcg	gagcgtgta	taagaaattc	ccggaaagct	tcgatttaca	gccgatcgaa	4140
gtaggaaaag	cgcgactggt	caaaggtgca	gctaaacata	tcattcatgc	cgtaggacca	4200
aacttcaaca	aagtttcgga	ggttgaaggt	gacaaacagt	tggcagaggc	ttatgagtec	4260
atcgctaaga	ttgtcaacga	taacaattac	aagtcagtag	cgattccact	gttgtccacc	4320
ggcatctttt	ccgggaacaa	agatcgacta	acccaatcat	tgaaccattt	gctgacagct	4380
ttagacacca	ctgatgcaga	tgtagccata	tactgcaggg	acaagaaatg	ggaaatgact	4440
ctcaaggaag	cagtggctag	gagagaagca	gtggaggaga	tatgcatatc	cgacgactct	4500
tcagtacag	aacctgatgc	agagctggtg	agggtgcatc	cgaagagttc	tttggctgga	4560
aggaagggct	acagcacaag	cgatggcaaa	actttctcat	atttggaagg	gaccaagttt	4620
caccaggcgg	ccaaggatat	agcagaaatt	aatgccatgt	ggcccgttgc	aacggaggcc	4680
aatgagcagg	tatgcatgta	tatcctcgga	gaaagcatga	gcagtattag	gtcgaaatgc	4740
cccgtcgaag	agtcggaagc	ctccacacca	cctagcacgc	tgcccttgctt	gtgcatccat	4800
gccatgactc	cagaaagagt	acagcgccta	aaagcctcac	gtccagaaca	aattactgtg	4860
tgctcatcct	ttccattgcc	gaagtataga	atcactgggtg	tgcagaagat	ccaatgctcc	4920
cagcctatat	tgttctcacc	gaaagtgcct	gcgtatatte	atccaaggaa	gtatctcgtg	4980
gaaacaccac	cggtagacga	gactccggag	ccatcggcag	agaaccaatc	cacagagggg	5040
acacctgaac	aaccaccact	tataaccgag	gatgagacca	ggactagaac	gcctgagccg	5100
atcatcatcg	aagaggaaga	agaggatagc	ataagtttgc	tgtcagatgg	cccgaccac	5160
cagggtgctgc	aagtcgaggc	agacattcac	gggccgcctt	ctgtatctag	ctcatcctgg	5220
tccattcctc	atgcatccga	ctttgatgtg	gacagtttat	ccatacttga	caccctggag	5280
ggagctagcg	tgaccagcgg	ggcaacgtca	gccgagacta	actcttactt	cgcaaagagt	5340
atggagtttc	tggcgcgacc	ggtgcctgcg	cctcgaaacg	tattcaggaa	ccctccacat	5400
cccgtcctgc	gcacaagaac	accgtcactt	gcaccagca	gggcctgctc	gagaaccagc	5460
ctagttttcca	ccccgccagg	cgtgaatagg	gtgatcacta	gagaggagct	cgaggcgctt	5520
accccgtcac	gcactcctag	caggtcggtc	tcgagaacca	gcctggtctc	caaccgccea	5580
ggcgtaaata	gggtgattac	aagagaggag	tttgaggcgt	tcgtagcaca	acaacaatga	5640
cggtttgatg	cgggtgcata	catcttttcc	tccgacaccg	gtcaagggca	tttacaacaa	5700
aaatcagtaa	ggcaaacggt	gctatccgaa	gtggtgttgg	agaggaccga	attggagatt	5760
tcgtatgccc	cgcgcctcga	ccaagaaaaa	gaagaattac	tacgcaagaa	attacagtta	5820

aatccacac	ctgctaacag	aagcagatac	cagtccagga	aggtggagaa	catgaaagcc	5880
ataacagcta	gacgtattct	gcaaggccta	gggcattatt	tgaaggcaga	aggaaaagtg	5940
gagtgtacc	gaaccctgca	tcctgttcct	ttgtattcat	ctagtgtgaa	ccgtgccttt	6000
tcaagcccca	aggctgcagt	ggaagcctgt	aacgccatgt	tgaaagagaa	ctttccgact	6060
gtggcttctt	actgtattat	tccagagtac	gatgcctatt	tggacatggg	tgacggagct	6120
tcatgtctgt	tagacactgc	cagtttttgc	cctgcaaagc	tgcgagctt	tccaaagaaa	6180
cactcctatt	tggaacccac	aatacgatcg	gcagtgcctt	cagcgatcca	gaacacgctc	6240
cagaacgtcc	tggcagctgc	cacaaaaaga	aattgcaatg	tcacgcaa	gagagaattg	6300
cccgtattgg	attcggcggc	ctttaatgtg	gaatgcttca	agaaatatgc	gtgtaataat	6360
gaatattggg	aaacgtttaa	agaaaacccc	atcaggctta	ctgaagaaaa	cgtggtaaat	6420
tacattacca	aattaaaagg	accaaaagct	gctgctcttt	ttgcgaagac	acataatttg	6480
aatatgttgc	aggacatacc	aatggacagg	tttgtaatgg	acttaaagag	agacgtgaaa	6540
gtgactccag	gaacaaaaca	tactgaagaa	cggcccaagg	tacaggtgat	ccaggctgcc	6600
gatccgctag	caacagcgta	tctgtgcgga	atccaccgag	agctggttag	gagattaaat	6660
gcggctctgc	ttccgaacat	tcatacactg	tttgatatgt	cggctgaaga	ctttgacgct	6720
attatagccg	agcacttcca	gcctggggat	tgtgttctgg	aaactgacat	cgcgtcgttt	6780
gataaaagtg	aggacgacgc	catggctctg	accgcgttaa	tgattctgga	agacttaggt	6840
gtggacgcag	agctgttgac	gctgattgag	gcggctttcg	gcgaaatttc	atcaatacat	6900
ttgcccacta	aaactaaatt	taagttcggc	gccatgatga	agtccggcat	gtttctgacc	6960
ctgttcgtga	acaccgtgat	caacatcgtg	atcgccagcc	gggtgctgag	agagagactg	7020
acaggatctc	cttgcgccgc	cttcacggc	gacgacaata	tcgtgaaggg	cgtgaagtcc	7080
gacaagctga	tgccgatag	atgcgccacc	tggtgaaca	tggaagtga	gatcatcgac	7140
gccgtcgtgg	gcgagaaggc	cccttatttt	tgcggcggct	tcctcctgtg	cgacagcgtg	7200
acaggcacag	cctgcagagt	tgccgatcct	ctgaagcggc	tgttcaagct	gggaaaacct	7260
ctggccgccg	acgacgagca	cgacgacgat	agacgtagag	ccctgcacga	ggaatccacc	7320
agatggaaca	gagtgggcat	cctgagcgag	ctgtgcaagg	ccgtggaaag	cagatacgag	7380
acagtgggca	ccagcatcat	tgtgatggca	atgaccacac	tgccagcag	cgtgaaaagc	7440
ttcagctacc	taaggggcgc	ccctatcaca	ctgtacggct	ctggcgaagg	cagaggcagc	7500
cttctgacat	gtggcgacgt	ggaagagaac	cccggacctg	tgtctaaggg	cgaagaactg	7560
tttaccggcg	tggtgccc	cctggtggaa	ctggatgggg	atgtgaacgg	ccacaagttc	7620
agcgttagcg	gagaaggcga	aggcgacgcc	acatacgga	agctgaccct	gaagttcatc	7680
tgcaccaccg	gcaagctgcc	tgtgccatgg	cctacactgg	tcaccacact	gacatacggc	7740
gtgcagtgt	tcagcagata	ccccgacct	atgaagcagc	acgacttctt	caagagcgcc	7800
atgcctgagg	gctacgtgca	agagcggacc	atcttcttta	aggacgacgg	caactacaag	7860
accagggccg	aagtgaagtt	cgagggcgac	accctggtca	accggatcga	gctgaagggc	7920
atgcacttca	aagaggacgg	caatatcctg	ggccacaagc	tcgagtacaa	ctacaacagc	7980
cacaacgtgt	acatcatggc	cgacaagcag	aaaaacggca	tcaaagtga	cttcaagatc	8040
cggcacaaca	tcgaggacgg	ctctgtgcag	ctggccgatc	actaccagca	gaacacaccc	8100
atcgagatg	gccctgtgct	gctgcccgat	aaccactacc	tgagcaccca	gagcaagctg	8160

agcaaggacc ccaacgagaa gcgggaccac atgggtgctgc tggaatttgt gacagccgcc	8220
ggaatcaccc tcggcatgga tgagctgtac aagtga	8256
<210> 17	
<211> 11447	
<212> DNA	
<213> Venezuelan equine encephalitis virus	
<220>	
<221> misc_feature	
<222> (1) .. (11447)	
<223> VEEV genome L01442.2	
<400> 17	
atgggcggcg catgagagaa gcccagacca attacctacc caaaatggag aaagttcacg	60
ttgacatcga ggaagacagc ccattcctca gagctttgca gcggagcttc ccgcagtttg	120
aggtagaagc caagcaggtc actgataatg accatgctaa tgccagagcg ttttcgcatac	180
tggttcaaaa actgatcgaa acggaggtgg acccatccga cacgatacctt gacattggaa	240
gtgcgccccgc ccgcagaatg tattctaagc acaagtatca ttgtatctgt ccgatgagat	300
gtgcggaaga tccggacaga ttgtataagt atgcaactaa gctgaagaaa aactgtaagg	360
aaataactga taagggaattg gacaagaaaa tgaaggagct cgccgccgtc atgagcgacc	420
ctgacctgga aactgagact atgtgcctcc acgacgacga gtcgtgtcgc tacgaagggc	480
aagtcgctgt ttaccaggat gtatacgcgg ttgacggacc gacaagtctc tatcaccaag	540
ccaataaggg agttagagtc gcctactgga taggctttga caccaccctt tttatgttta	600
agaacttggc tggagcatat ccatcact ctaccaactg ggccgacgaa accgtgttaa	660
cggctcgtaa cataggccta tgcagctctg acgttatgga gcggtcacgt agagggatgt	720
ccattcttag aaagaagtat ttgaaacat ccaacaatgt tctattctct gttggctcga	780
ccatctacca cgagaagagg gacttactga ggagctggca cctgccgtct gtatttact	840
tacgtggcaa gcaaaattac acatgtcggg gtgagactat agttagttgc gacgggtacg	900
tcgttaaaaag aatagctatc agtccaggcc tgtatgggaa gccttcaggc tatgctgcta	960
cgatgcaccg cgagggattc ttgtgctgca aagtacaga cacattgaac ggggagaggg	1020
tctcttttcc cgtgtgcacg tatgtgccag ctacattgtg tgaccaaag actggcatac	1080
tggaacaga tgtcagtgcg gacgacgcg aaaaactgct ggttgggctc aaccagcgta	1140
tagtcgtcaa cggtcgcacc cagagaaaca ccaataccat gaaaaattac cttttgccc	1200
tagtggccca ggcatthtct aggtgggcaa aggaatataa ggaagatcaa gaagatgaaa	1260
ggccactagg actacgagat agacagttag tcatgggggtg ttgttgggct tttagaaggc	1320
acaagataac atctatttat aagcgcccgg ataccaaac catcatcaaa gtgaacagcg	1380
atttccactc attcgtgctg cccaggatag gcagtaacac attggagatc gggctgagaa	1440
caagaatcag gaaaatgtta gaggagcaca aggagccgtc acctctcatt accgccgagg	1500
acgtacaaga agctaagtgc gcagccgatg aggctaagga ggtgcgtgaa gccgaggagt	1560
tgcgcgcagc tctaccacct ttggcagctg atgttgagga gccactctg gaagccgatg	1620
tcgacttgat gttacaagag gctggggccg gctcagtga gacacctcgt ggcttgataa	1680

aggttaccag	ctacgctggc	gaggacaaga	tcggctctta	cgctgtgctt	tctccgcagg	1740
ctgtactcaa	gagtgaaaaa	ttatcttgca	tccacctct	cgctgaacaa	gtcatagtga	1800
taacacactc	tggccgaaaa	gggcgttatg	ccgtggaacc	ataccatggg	aaagtagtgg	1860
tgccagaggg	acatgcaata	cccgtccagg	actttcaagc	tctgagtga	agtgccacca	1920
ttgtgtacaa	cgaacgtgag	ttcgtaaaca	ggtacctgca	ccatattgcc	acacatggag	1980
gagcgctgaa	cactgatgaa	gaatattaca	aaactgtcaa	gcccagcgag	cacgacggcg	2040
aatacctgta	cgacatcgac	aggaaacagt	gcgtcaagaa	agaactagtc	actgggctag	2100
ggctcacagg	cgagctgggt	gacctccct	tccatgaatt	cgctacgag	agtctgagaa	2160
cacgaccagc	cgctccttac	caagtaccaa	ccataggggt	gtatggcgtg	ccaggatcag	2220
gcaagtctgg	catcattaaa	agcgcagtca	ccaaaaaga	tctagtgggt	agcgccaaga	2280
aagaaaactg	tgcagaaatt	ataagggacg	tcaagaaaat	gaaagggtg	gacgtcaatg	2340
ccagaactgt	ggactcagt	ctcttgaatg	gatgcaaaca	ccccgtagag	accctgtata	2400
ttgacgaagc	ttttgcttgt	catgcaggta	ctctcagagc	gctcatagcc	attataagac	2460
ctaaaaaggc	agtgtctctg	ggggatccca	aacagtgcgg	tttttttaac	atgatgtgcc	2520
tgaaagtgca	ttttaaccac	gagatttgca	cacaagtctt	ccacaaaagc	atctctcgcc	2580
gttgcaactaa	atctgtgact	tcggtcgtct	caaccttggt	ttacgacaaa	aaaatgagaa	2640
cgacgaatcc	gaaagagact	aagattgtga	ttgacactac	cggcagtacc	aaacctaagc	2700
aggacgatct	cattctcact	tgtttcagag	ggtgggtgaa	gcagttgcaa	atagattaca	2760
aaggcaacga	aataatgacg	gcagctgcct	ctcaagggt	gacccgtaaa	ggtgtgtatg	2820
ccgttcggta	caaggtgaat	gaaaatcctc	tgtacgcacc	cacctcagaa	catgtgaacg	2880
tcctactgac	ccgcacggag	gaccgcatcg	tgtggaaaac	actagccggc	gacccatgga	2940
taaaaacact	gactgccaag	taccctggga	atttactgc	cacgatagag	gagtggcaag	3000
cagagcatga	tgccatcatg	aggcacatct	tggagagacc	ggaccctacc	gacgtcttcc	3060
agaataaggc	aaacgtgtgt	tgggccaagg	ctttagtgcc	ggtgctgaag	accgctggca	3120
tagacatgac	cactgaacaa	tggaaactg	tggattat	tgaaacggac	aaagctcact	3180
cagcagagat	agtattgaac	caactatgcg	tgaggttctt	tggactcgat	ctggactccg	3240
gtctattttc	tgcacccact	gttccgttat	ccattagga	taatcactgg	gataactccc	3300
cgtcgcctaa	catgtacggg	ctgaataaag	aagtgggtccg	tcagctctct	cgcaggtacc	3360
cacaactgcc	tcgggcagtt	gccactggaa	gagtctatga	catgaacact	ggtacactgc	3420
gcaattatga	tccgcgcata	aacctagtac	ctgtaaacag	aagactgcct	catgctttag	3480
tcctccacca	taatgaacac	ccacagagt	acttttcttc	attcgtcagc	aaattgaagg	3540
gcagaactgt	cctgggtggc	ggggaaaagt	tgtccgtccc	aggcaaaatg	gttgactggt	3600
tgtcagaccg	gcctgaggt	accttcagag	ctcggttgga	tttaggcata	ccaggtgatg	3660
tgccccaaata	tgacataata	tttgttaatg	tgaggacccc	atataaatac	catcactatc	3720
agcagtgtga	agaccatgcc	attaagctta	gcatgttgac	caagaaagct	tgtctgcata	3780
tgaatcccgg	cggaaacctgt	gtcagcatag	gttatggtta	cgctgacagg	gccagcgaaa	3840
gcatcattgg	tgctatagcg	cggcagttca	agttttcccg	ggtatgcaaa	ccgaaatcct	3900
cacttgaaga	gacggaagtt	ctgtttgtat	tcattgggta	cgatcgcaag	gcccgtacgc	3960
acaatcctta	caagctttca	tcaaccttga	ccaacattta	tacaggttcc	agactccacg	4020

aagccg gatg	tgcaccc tca	tatcatgtgg	tgcgagggga	tattgccacg	gccaccgaag	4080
gagt gattat	aaatgctgct	aacagcaaag	gacaacctgg	cggaggggtg	tgcggagcgc	4140
tgtataagaa	attccccgaa	agcttcgatt	tacagccgat	cgaagtagga	aaagcgcgac	4200
tgg tcaaagg	tgcagctaaa	catatcattc	atgccgtagg	accaaacttc	aacaaagttt	4260
cggaggttga	aggtgacaaa	cagttggcag	aggcttatga	gtccatcgct	aagattgtca	4320
acgataacaa	ttacaagtca	gtagcgattc	cactgttgct	caccggcatc	ttttccggga	4380
acaaagatcg	actaacccaa	tcattgaacc	atttgc t gac	agcttttagac	accactgatg	4440
cagatgtagc	catatactgc	agggacaaga	aatgggaaat	gactctcaag	gaagcagtgg	4500
ctaggagaga	agcagtggag	gagatatgca	tatccgacga	ctcttcagt	acagaacctg	4560
atgcagagct	ggtgaggggtg	catccgaaga	gttctttggc	tggaaggaag	ggctacagca	4620
caagcgatgg	caaaactttc	tcataatttg	aagggaacca	gtttcaccag	gcggccaagg	4680
atatagcaga	aattaatgcc	atgtggcccc	ttgcaacgga	ggccaatgag	caggtatgca	4740
tgtatatcct	cggagaaaagc	atgagcagta	ttaggtcgaa	atgccccgtc	gaagagtcgg	4800
aagcctccac	accacctagc	acgtgcctt	gcttgtgcat	ccatgccatg	actccagaaa	4860
gagtacagcg	cctaaaagcc	tcacgtccag	aacaaattac	tgtgtgctca	tcctttccat	4920
tgccgaagta	tagaatcact	ggtgtgcaga	agatccaatg	ctcccagcct	atattgttct	4980
caccgaaagt	gcctgcgtat	attcatccaa	ggaagtatct	cgtggaaaca	ccaccgtag	5040
acgagactcc	ggagccatcg	gcagagaacc	aatccacaga	ggggacacct	gaacaaccac	5100
cacttataac	cgaggatgag	accaggacta	gaacgcctga	gccgatcatc	atcgaagagg	5160
aagaagagga	tagcataagt	ttgctgtcag	atggccccgac	ccaccaggtg	ctgcaagtcg	5220
aggcagacat	tcacgggccg	ccctctgtat	ctagctcatc	ctggtccatt	cctcatgcat	5280
ccgactttga	tgtggacagt	ttatccatac	ttgacaccct	ggaggagct	agcgtgacca	5340
gcggggcaac	gtcagccgag	actaactctt	acttcgcaaa	gagtatggag	tttctggcgc	5400
gaccggtgcc	tgcgcctcga	acagtattca	ggaaccctcc	acatcccgtc	ccgcgcacaa	5460
gaacaccgtc	acttgcaccc	agcagggcct	gctcgagaac	cagcctagtt	tccaccccgc	5520
caggcgtgaa	tagggtgata	actagagagg	agctcgaggc	gcttaccctg	tcacgcactc	5580
ctagcaggtc	ggtctcgaga	accagcctgg	tctccaacct	gccaggcgta	aatagggtga	5640
ttacaagaga	ggagtttgag	gcgttcgtag	cacaacaaca	atgacggttt	gatgcgggtg	5700
catacatctt	ttcctccgac	accggtcaag	ggcatttaca	acaaaaatca	gtaaggcaaa	5760
cggtgctatc	cgaagtgggt	ttggagagga	ccgaattgga	gatttcgtat	gccccgcgcc	5820
tcgaccaaga	aaaagaagaa	ttactacgca	agaaattaca	gttaaattccc	acacctgcta	5880
acagaagcag	ataccagtcc	aggaaggtgg	agaacatgaa	agccataaca	gctagacgta	5940
ttctgcaagg	cctagggcat	tatttgaagg	cagaaggaaa	agtggagtgc	taccgaacct	6000
tgcac tctgt	tcctttgtat	tcactctagt	tgaaccgtgc	cttttcaagc	cccaaggtcg	6060
cagtggaaagc	ctgtaacgcc	atgttgaaag	agaactttcc	gactgtggct	tcttactgta	6120
ttattccaga	gtacgatgcc	tatttggaca	tggttgacgg	agcttcatgc	tgttagaca	6180
ctgccagttt	ttgccctgca	aagctgcgca	gctttccaaa	gaaacactcc	tatttggaac	6240
ccacaatacg	atcggcagtg	ccttcagcga	tccagaacac	gctccagaac	gtcctggcag	6300
ctgccacaaa	aagaaattgc	aatgtcacgc	aatgagaga	attgcccgtg	ttggattcgg	6360

cggccttttaa	tgtggaatgc	ttcaagaaat	atgcgtgtaa	taatgaatat	tgggaaacgt	6420
ttaaagaaaa	ccccatcagg	cttactgaag	aaaacgtggt	aaattacatt	accaaattaa	6480
aaggaccaaaa	agctgctgct	ctttttgcga	agacacataa	tttgaatatg	ttgcaggaca	6540
taccaatgga	caggtttgta	atggacttaa	agagagacgt	gaaagtgact	ccaggaacaa	6600
aacatactga	agaacggccc	aaggtacagg	tgatccaggc	tgccgatccg	ctagcaacag	6660
cgtatctgtg	cggaatccac	cgagagctgg	ttaggagatt	aatgcggtc	ctgcttccga	6720
acattcatac	actgtttgat	atgtcggctg	aagactttga	cgctattata	gccgagcact	6780
tccagcctgg	ggatttgttt	ctggaaactg	acatcgcgtc	gtttgataaa	agtgaggacg	6840
acgccatggc	tctgaccgcg	ttaatgattc	tggaagactt	aggtgtggac	gcagagctgt	6900
tgacgctgat	tgaggcggct	ttcggcgaaa	tttcatcaat	acatttgccc	actaaaacta	6960
aattttaaat	cggagccatg	atgaaatctg	gaatgttctt	cacactgttt	gtgaacacag	7020
tcattaacat	tgtaatcgca	agcagagtgt	tgagagaacg	gctaaccgga	tcaccatgtg	7080
cagcattcat	tggagatgac	aatatcgtga	aaggagtcaa	atcggacaaa	ttaatggcag	7140
acaggtgcgc	cacctggttg	aatatggaag	tcaagattat	agatgctgtg	gtgggcgaga	7200
aagcgcctta	tttctgtgga	gggtttat	tgtgtgactc	cgtgaccggc	acagcgtgcc	7260
gtgtggcaga	ccccctaaaa	aggctgttta	agcttggcaa	acctctggca	gcagacgatg	7320
aacatgatga	tgacaggaga	agggcattgc	atgaagagtc	aacacgctgg	aaccgagtgg	7380
gtattctttc	agagctgtgc	aaggcagtag	aatcaaggta	tgaaaccgta	ggaacttcca	7440
tcatagtatt	ggccatgact	actctagcta	gcagtgttaa	atcattcagc	tacctgagag	7500
gggcccctat	aactctctac	ggctaacctg	aatggactac	gacatagtct	agtccgccaa	7560
gatgttcccc	ttccagccaa	tgtatccgat	gcagccaatg	ccctatcgca	acccgttcgc	7620
ggccccgcgc	aggccctggt	tccccagaac	cgaccctttt	ctggcgatgc	aggtgcagga	7680
attaacccgc	tcgatggcta	acctgacgtt	caagcaacgc	cgggacgcgc	cacctgaggg	7740
gccatccgct	aagaaaccga	agaaggaggc	ctcgcaaaaa	cagaaagggg	gaggccaagg	7800
gaagaagaag	aagaaccaag	ggaagaagaa	ggctaagaca	gggccgccta	atccgaaggc	7860
acagaatgga	aacaagaaga	agaccaacaa	gaaaccaggc	aagagacagc	gcatggtcat	7920
gaaattggaa	tctgacaaga	cgttcccaat	catgttggaa	gggaagataa	acggctacgc	7980
ttgtgtggtc	ggagggaagt	tattcaggcc	gatgcatgtg	gaaggcaaga	tcgacaacga	8040
cgttctggcc	gcgcttaaga	cgaagaaagc	atccaaatac	gatcttgagt	atgcagatgt	8100
gccacagaac	atgcgggccg	atacattcaa	atacaccat	gagaaacccc	aaggctatta	8160
cagctggcat	catggagcag	tccaatatga	aaatgggcgt	ttcacgggtgc	cgaaaggagt	8220
tggggccaag	ggagacagcg	gacgacccat	tctggataac	cagggacggg	tggtcgctat	8280
tgtgctggga	ggtgtgaatg	aaggatctag	gacagccctt	tcagtcgtca	tgtggaacga	8340
gaagggagtt	accgtgaagt	atactccgga	gaactgcgag	caatggtcac	tagtgaccac	8400
catgtgtctg	ctcgccaatg	tgacgttccc	atgtgctcaa	ccaccaat	gctacgacag	8460
aaaaccagca	gagactttgg	ccatgctcag	cgttaacgtt	gacaacccgg	gctacgatga	8520
gctgctggaa	gcagctgtta	agtccccgg	aaggaaaagg	agatccaccg	aggagctgtt	8580
taaggagtat	aagctaacgc	gcccttacat	ggccagatgc	atcagatgtg	cagttgggag	8640
ctgccatagt	ccaatagcaa	tcgaggcagt	aaagagcgac	gggcacgacg	gttatgttag	8700

acttcagact	tcctcgcagt	atggcctgga	ttcctccggc	aacttaaagg	gcaggaccat	8760
gcggtatgac	atgcacggga	ccattaaaga	gataccacta	catcaagtgt	cactccatac	8820
atctcgcctg	tgtcacattg	tggatgggca	cggttatttc	ctgcttgcca	ggtgcccggc	8880
aggggactcc	atcaccatgg	aatttaagaa	agattccgtc	acacactcct	gctcgggtgcc	8940
gtatgaagtg	aaattttaatc	ctgtaggcag	agaactctat	actcatcccc	cagaacacgg	9000
agtagagcaa	gcgtgccaag	tctacgcaca	tgatgcacag	aacagaggag	cttatgtcga	9060
gatgcacctc	ccgggctcag	aagtggacag	cagtttggtt	tccttgagcg	gcagttcagt	9120
caccgtgaca	cctcctgttg	ggactagcgc	cctgggtggaa	tgcgagtgtg	gcggcacaaa	9180
gatctccgag	accatcaaca	agacaaaaca	gttcagccag	tgcacaaaga	aggagcagtg	9240
cagagcatat	cggctgcaga	acgataagtg	ggtgtataat	tctgacaaac	tgcccaaagc	9300
agcgggagcc	accttaaaaag	gaaaactgca	tgtcccatte	ttgctggcag	acggcaaattg	9360
caccgtgcct	ctagcaccag	aacctatgat	aacctttggt	ttcagatcag	tgtcactgaa	9420
actgcaccct	aagaatccca	catatctaac	caccgccea	cttgctgatg	agcctcacta	9480
cacgcacgag	ctcatatctg	aaccagctgt	taggaatttt	accgtcaccg	aaaaagggtg	9540
ggagtttgta	tggggaaacc	acccgccgaa	aaggttttgg	gcacaggaaa	cagcaccceg	9600
aaatccacat	gggctaccgc	acgaggtgat	aactcattat	taccacagat	accctatgtc	9660
caccatcctg	ggtttgtcaa	tttgtgccgc	cattgcaacc	gtttccgttg	cagcgtctac	9720
ctggctgttt	tgcagatcta	gagttgcgtg	cctaactcct	taccggctaa	cacctaacgc	9780
taggatacca	ttttgtctgg	ctgtgctttg	ctgcgcccgc	actgcccggg	ccgagaccac	9840
ctgggagtcc	ttggatcacc	tatggaacaa	taaccaacag	atgttctgga	ttcaattgct	9900
gatccctctg	gccgccttga	tcgtagtac	tcgcctgctc	aggtgcgtgt	gctgtgtcgt	9960
gcctttttta	gtcatggccg	gcgccgcagg	cgccggcgcc	tacgagcacg	cgaccacgat	10020
gccgagccaa	gcgggaatct	cgtataacac	tatagtcaac	agagcaggct	acgcaccact	10080
ccctatcagc	ataacaccaa	caaagatcaa	gctgatacct	acagtgaact	tggagtacgt	10140
cacctgccac	tacaaaacag	gaatggattc	accagccatc	aatgctgcg	gatctcagga	10200
atgcactcca	acttacaggc	ctgatgaaca	gtgcaaagtc	ttcacagggg	tttaccggtt	10260
catgtggggg	ggtgcatatt	gcttttgcca	cactgagaac	acccaagtca	gcaaggccta	10320
cgtaatgaaa	tctgacgact	gccttgcgga	tcatgctgaa	gcatataaag	cgcacacagc	10380
ctcagtgcag	gcgttcctca	acatcacagt	gggagaacac	tctattgtga	ctaccgtgta	10440
tgtgaatgga	gaaactcctg	tgaatttcaa	tgggggtcaaa	ttaactgcag	gtccgctttc	10500
cacagcttgg	acaccctttg	atcgcaaaat	cgtgcagtat	gccggggaga	tctataatta	10560
tgatttttcct	gagtatgggg	caggacaacc	aggagcattt	ggagatatac	aatccagaac	10620
agtctcaagc	tcagatctgt	atgccaatac	caacctagt	ctgcagagac	ccaaagcagg	10680
agcgatccac	gtgccataca	ctcaggcacc	ttcgggtttt	gagcaatgga	agaaagataa	10740
agctccatca	ttgaaattta	ccgccccctt	cggatgcgaa	atatatacaa	accccatteg	10800
cgccgaaaac	tgtgctgtag	ggtcaattcc	attagccttt	gacattcccg	acgccttggt	10860
caccagggtg	tcagaaacac	cgacactttc	agcgcccgaa	tgcactctta	acgagtgcgt	10920
gtattcttcc	gactttgggt	ggatcgccac	ggtcaagtac	tcggccagca	agtcaggcaa	10980
gtgcgcagtc	catgtgccat	cagggactgc	taccctaaaa	gaagcagcag	tcgagctaac	11040

cgagcaaggg tcggcgacta tccatttctc gaccgcaa atccacccgg agttcaggct 11100
 ccaaatatgc acatcatatg ttactgtcaa aggtgattgt cccccccga aagaccatat 11160
 tgtgacacac cctcagtatc acgcccacac atttacagcc gcggtgtcaa aaaccgcgtg 11220
 gacgtgggta acatccctgc tgggaggatc agccgtaatt attataattg gcttggtgct 11280
 ggctactatt gtggccatgt acgtgctgac caaccagaaa cataattgaa tacagcagca 11340
 attggcaagc tgcttacata gaactcgcgg cgattggcat gccgccttaa aattttttatt 11400
 ttatttttttc ttttcttttc cgaatcggat tttgttttta atatttc 11447

<210> 18

<211> 558

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, VEEV rep RNA nt 6966-7523 with silent mutations

<400> 18

aaguucggcg ccaugaugaa guccggcaug uuucugaccc uguucgugaa caccgugauc 60
 aacaucguga ucgccagccg ggugcugaga gagagacuga caggauccucc uugcgccgcc 120
 uucaucggcg acgacaauau cgugaaggcg gugaaguccg acaagcugau ggccgauaga 180
 ugcgccaccu ggcugaacau ggaagugaag aucaucgacg ccgucguggg cgagaaggcc 240
 ccuuauuuuu gcggcgggcu cauccugugc gacagcguga caggcacagc cugcagaguu 300
 gccgauccuc ugaagcgggc guucaagcug ggaaaaccuc uggccgccga cgacgagcac 360
 gacgacgaua gacguagagc ccugcacgag gaauccacca gauggaacag agugggcauc 420
 cugagcgagc ugugcaaggc cguggaaagc agauacgaga cagugggcac cagcaucauu 480
 gugauggcaa ugaccacacu ggccagcagc gugaaaagcu ucagcuaccu aaggggcgcc 540
 ccuaucacac uguacggc 558

<210> 19

<211> 1821

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, RNA nsP4 engineered RNA +silent mutations

<400> 19

uacaucuuuu ccuccgacac cggucaaggg cauuuacaac aaaaucagu aaggcaaagc 60
 gugcuauccg aagugguguu ggagaggacc gaauuggaga uuucguaugc cccgcgccuc 120
 gaccaagaaa aagaagaauu acuacgcaag aaauuacagu uaaaucccac accugcuaac 180
 agaagcagau accaguccag gaagguggag aacaugaaag ccuaaacagc uagacguauu 240
 cugcaaggcc uagggcuuu uuugaaggca gaaggaaaag uggagugcua ccgaaccug 300
 cauccuguuc cuuuguauuc aucuagugug aaccgugccu uuucaagccc caaggucgca 360
 guggaagccu guaagccau guugaaagag aacuuuccga cuguggcuuc uuacuguauu 420

auuccagagu	acgaugccua	uuuggacaug	guugacggag	cuucaugcug	cuuagacacu	480
gccaguuuuu	gcccugcaaa	gcugcgcagc	uuuccaaaga	aacacuccua	uuuggaaccc	540
acaauacgau	cggcagugcc	uucagcgauc	cagaacacgc	uccagaacgu	ccuggcagcu	600
gccacaaaaa	gaaaauugcaa	ugucacgcaa	augagagaa	ugcccguauu	ggauucggcg	660
gccuuuaa	uggaauvcuu	caagaaauau	gcguguaaua	augaauauug	ggaaacguuu	720
aaagaaaacc	ccaucaggcu	uacugaagaa	aacgugguaa	auuacauuac	caaaauaaaa	780
ggacccaaaag	cugcugcucu	uuuugcgaag	acacauaaau	ugaauauguu	gcaggacaua	840
ccaauvgaca	gguuuuguaa	ggacuuaaaag	agagacguga	aagugacucc	aggaacaaaa	900
cauacugaag	aacggcccaa	gguacaggug	auccaggcug	ccgauccgcu	agcaacagcg	960
uauvcugvcg	gaauccaccg	agagcugguu	aggagauuaa	augcgguccu	gcuuccgaac	1020
auucauacac	uguuugauau	gucggcugaa	gacuuugacg	cuauuauagc	cgagcacuuc	1080
cagccugggg	auuguguucu	ggaaacugac	aucgcgucgu	uugauaaaag	ugaggacgac	1140
gccauvgcuc	ugaccgcguu	aaugauucug	gaagacuuag	guguggacgc	agagcuguug	1200
acgcugauug	aggcggcuuu	cggcgaaaau	ucaucaauac	auuugcccac	uaaaacuaaa	1260
uuuaaguucg	gcgccaugau	gaaguccggc	auguuucuga	cccuguucgu	gaacaccgug	1320
aucaacaucg	ugaucgccag	ccgggugcug	agagagagac	ugacaggau	uccuugcgcc	1380
gccuucaucg	gcgacgacaa	uauvcugaag	ggcgugaagu	ccgacaagcu	gauggccgau	1440
agaugcgcca	ccuggcugaa	cauggaagug	aagaucaucg	acgccgucgu	ggcgagaaag	1500
gccccuuauu	uuugcggcgg	cuucauccug	ugcgacagcg	ugacaggcac	agccugcaga	1560
guugccgauc	cucugaagcg	gcuguucaag	cugggaaaac	cucuggccgc	cgacgacgag	1620
cacgacgacg	auagacguag	agcccugcac	gaggaaucua	ccagauggaa	cagagugggc	1680
auccugagcg	agcugugcaa	ggccguggaa	agcagauacg	agacaguggg	caccagcauc	1740
auugugaugg	caaugaccac	acuggccagc	agcgugaaaa	gcuucagcua	ccuaaggggc	1800
gccccuauca	cacuguacgg	c				1821

<210> 20

<211> 7561

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C01 nsP1-4 ORF and SGP RNA

<400> 20

augggcggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccaauccuca	gagcuuugca	gcggagcuuc	ccgcaguuug	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgcauc	180
uggcuucaaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaaguauc	uuguaucugu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuua	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480

aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaaggg	aguuaagaguc	gccuacugga	uaggcuuuga	caccaccccu	uuuauguuaa	600
agaacuuggc	uggagcauau	ccaucauacu	cuaccaacug	ggccgacgaa	accguguuaa	660
cggcucguua	cauaggccua	ugcagcucug	acguuaugga	gcggucacgu	agagggaugu	720
ccauucuuag	aaagaaguau	uugaaaccau	ccaacaaugu	ucuauucucu	guuggcucga	780
ccaucuaacca	cgagaagagg	gacuuacuga	ggagcuggca	ccugccgucu	guauuucacu	840
uacguggcaa	gcaaaaauuac	acaugucggg	gugagacuau	aguuauguuc	gacggguacg	900
ucguuaaaaag	aauagcuau	aguccaggcc	uguauuggga	gccuucaggc	uauugcucua	960
cgaugcaccg	cgagggaauu	uugugcugca	aagugacaga	cacauugaac	ggggagaggg	1020
ucucuuuuucc	cgugugcacg	uauugccag	cuacauugug	ugaccaaau	acuggcauac	1080
uggcaacaga	ugucagugcg	gacgacgcgc	aaaaacugcu	gguugggcuc	aaccagcgua	1140
uagucgucaa	cggucgcacc	cagagaaaca	ccaauaccu	gaaaaauuac	uuuuugcccc	1200
uaguggccca	ggcauuugcu	aggugggcaa	aggaauuaa	ggaagaucaa	gaagaugaaa	1260
ggccacuagg	acuacgagau	agacaguua	ucauggggug	uuuugggcu	uuuagaaggc	1320
acaagauaac	aucuauuuau	aagcgcccgg	auacccaaac	caucaucaa	gugaacagcg	1380
auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagau	gggcugagaa	1440
caagaauacg	gaaaauguua	gaggagcaca	aggagccguc	accucucau	accgccgagg	1500
acguacaaga	agcuagugc	gcagccgaug	aggcuagga	ggugcgugaa	gccgaggagu	1560
ugcgcgcagc	ucuaccaccu	uuggcagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucua	cgcugugcu	ucuccgcagg	1740
cuguacucaa	gagugaaaaa	uuaucuugca	uccaccucuc	cgcugaacaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuaug	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuaaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgcugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcua	2100
ggcucacagg	cgagcuggug	gauccuccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuuaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaaggggag	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaaca	ccccguagag	accuguaaua	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggaucucca	aacagucggg	uuuuuuuac	augaugugcc	2520
ugaaagugca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagaaa	2640
cgaagaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaagacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguaug	2820

ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaacacu	gacugccaag	uaccucggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggaccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuugugcc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaucg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacu	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaag	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuaua	caugaacacu	gguacacugc	3420
gcauuuaua	uccgcgcua	aaccuaguac	cuguaaacag	aagacugccu	caugcuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaua	ugacauaaua	uuuguuaaag	ugaggacccc	auauaaaauac	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaauccggg	cggaaaccugu	gucagcauag	guuaugguua	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuua	caagcuuua	ucaaccuuga	ccaacauua	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucu	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuau	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgauug	caaaacuuc	ucauauuugg	aagggaccaa	guuucaccag	gcggccaagg	4680
auauagcaga	aauuauugcc	auguggcccg	uugcaacgga	ggccaaugag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaauacacu	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgua	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggauag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160

aagaagagga	uagcauaagu	uugcugucag	auggcccgcac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggaggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uagggugauc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaauucc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uuucgcaagg	ccuagggcua	uauuuagaag	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuaug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguuagaa	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaagu	acgccugcaa	caacagaguac	ugggagacau	6420
ucaaagagaa	ccccauccgg	cugaccgagg	aaaacguggu	caacuacauc	accaagcuga	6480
agggccccaa	agccgccgcu	cuguuugcca	agacacacaa	ccugaacaug	cugcaggaca	6540
uccccaugga	cagauucgug	auggaccuga	agcgggacgu	gaaagugacc	ccuggcacca	6600
agcacaccga	ggaacggccu	aaggugcaag	ugauccaggc	cgcugauccu	cuggccacag	6660
ccuauucugug	uggcauccac	agagaacucg	ugcggagacu	gaaugccgug	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	cugaugauuc	uggaagauc	cggaguggac	gccgagcugc	6900
ugacacugau	ugaagccgcc	uuuggcgaga	ucagcagcau	ccaucugccu	accaagacca	6960
aguucaaguu	cggcgccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguuaucgca	agcagagugu	ugagagaacg	gcuaacaggc	agcccuugug	7080
cagcauucau	uggagaugac	aaauucguga	aaggagucua	aucggacaaa	uuauuggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgcuuua	uuucugugga	ggguuuuuuu	ugugugacuc	cugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aaauaugaga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucuaaggua	ugaaaccgua	ggaacuucca	7440
ucauaguuau	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500

gggccccuau aacucucucac ggcuaaccug aauggacuac gacauagucu aguccgccaa	7560
g	7561
<210> 21	
<211> 7561	
<212> RNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic construct, C02 nsP1-4 ORF and SGP RNA	
<400> 21	
augggcgggcg caugagagaa gcccagacca auuaccuacc caaaauggag aaaguucacg	60
uugacaucga ggaagacagc ccauuccuca gagcuuugca gcgagacuuc ccgcaguuug	120
agguagaagc caagcagguc acugauaaug accaugcuua ugccagagcg uuuucgcauc	180
uggcuucaaa acugaucgaa acggaggugg acccauccga cacgauccuu gacauuggaa	240
gugcgcccg cgcagaaug uauucuaagc acaaguauca uuguauugcu ccgaugagau	300
gugcggaaga uccggacaga uuguauaagu augcaacuaa gcugaagaaa aacuguaagg	360
aaauaacuga uaaggaauug gacaagaaaa ugaaggagcu cgccgccguc augagcgacc	420
cugaccugga aacugagacu augugccucc acgacgacga gucgugucgc uacgaagggc	480
aagucgcugu uuaccaggau guauaugccg uggauggccc uacaagccug uaccaccagg	540
ccaacaaggg cgucagagug gccuacugga ucggcuucga caccacaccu uucauguuca	600
agaaccuggc uggcgcuuac cccagcuaca gcacaaacug ggccgacgaa accguguuaa	660
cggcucguua cauaggccua ugcagcucug acguuaugga gcgguacgu agagggaugu	720
ccaauucuuag aaagaaguau uugaaaccau ccaacaaugu ucuauucucu guuggcucga	780
ccaucuaacca cgagaagagg gacuuacuga ggagcuggca ccugccgucu guauuucacu	840
uacguggcaa gcaaaaauac acaugucggu gugagacuau aguuauguu gacggguacg	900
ucguuaaaaag aaauagcuac aguccaggcc uguauuggga gccuucaggc uaugcugcu	960
cgaugcaccg cgagggaauu uugugcugca aagugacaga cacauugaac ggggagaggg	1020
ucucuuuuucc cgugugcacg uaugugccag cuacauugug ugacaaaug acuggcauac	1080
uggcaacaga ugucagugcg gacgacgcg aaaaacugcu gguugggcuc aaccagcgua	1140
uagucgucaa cggucgcacc cagagaaaca ccauaccuau gaaaaauuac cuuuugcccg	1200
uaguggccca ggcauuugcu agguugggcaa aggaauauaa ggaagaucaa gaagaugaaa	1260
ggccacuagg acuacgagau agacaguuaug ucauggggug uuguugggcu uuuagaaggc	1320
acaagauaac aucuauuuau aagcgcccgg auacccaaac caucaucaa gugaacagcg	1380
auuuccacuc auucgugcug cccaggauag gcaguaacac auuggagau gggcugagaa	1440
caagaauacg gaaaauguua gaggagcaca aggagccguc accucucauu accgccgagg	1500
acguacaaga agcuagugc gcagccgaug aggcuagga ggugcgugaa gccgaggagu	1560
ugcgcgagc ucuaccaccu uuggcagcug auguugagga gcccacucug gaagccgaug	1620
ucgacuugau guuacaagag gcuggggccg gcucagugga gacaccucgu ggcuugauaa	1680
agguuaccag cuacgcuggc gaggacaaga ucggcucuua cgcugugcuu ucuccgcagg	1740
cuguacucaa gagugaaaaa uuaucuugca uccaccucuc cgcugaacaa gucauaguga	1800

uaacacacuc	uggccgaaaa	gggcguuau	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagagggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgcugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauag	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaagggacg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaca	ccccguagag	accuguaau	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagugcgg	uuuuuuuac	augaugugcc	2520
ugaaauguca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucugccc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaugagaa	2640
cgacgaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaagacg	gcagcugccu	cuaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaacacu	gacugccaag	uacccuggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggacccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuagugcc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaucg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacac	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaag	aagugguucc	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuaua	caugaacacu	gguacacugc	3420
gcauuuaua	uccgcgcaua	aaccuaguac	cuguuaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaauagaacac	ccacagagug	acuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaau	ugacauaaau	uuuuguuaug	ugaggacccc	auauaaaau	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaaucccgg	cggaaaccugu	gucagcauag	guuaugguuu	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacgggaugu	cuguuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuuu	caagcuuuca	ucaaccuuga	ccaacauuuu	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucu	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuuu	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140

uguauaagaa	auucccgga	agcuucgau	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgaau	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aaggggacaa	guuucaccag	gcggccaagg	4680
auauagcaga	aauuauugcc	auguggcccc	uugcaacgga	ggccaauag	caggaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaaucacu	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcguau	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggaugag	accaggacua	gaacgccuga	gccgaucac	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggagggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcaccc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uaggguaguc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaauucc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcgu	uauuugaagg	cagaaggaaa	aguggagugc	uaccgaaccc	6000
ugcauccugu	uccuuugua	ucaucuaug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaaau	augcguguua	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480

aaggacccaaa	agcugcugcu	uuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaauugga	cagguuugua	auggacuuaa	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugaucaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuaaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuuuuuu	cggagccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguuauucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucau	uggagaugac	aaauaucguga	aaggagucua	aucggacaaa	uuauuggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aacaugauga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaaggua	ugaaaccgua	ggaacuucca	7440
ucauaguua	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucuac	ggcuuaccug	aauggacuac	gacauagucu	aguccgccaa	7560
g						7561

<210> 22

<211> 7561

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C03 nsP1-4 ORF and SGP RNA

<400> 22

augggcggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccauuccuca	gagcuuugca	gcggagcuuc	ccgcaguuuu	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgauc	180
uggcuuacaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaaguauca	uuguauucgu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuua	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaaggg	aguuaagagc	gccuacugga	uaggcuuuga	caccaccccu	uuuauuuua	600
agaacuuggc	uggagcauau	ccaucuuacu	cuaccaacug	ggccgacgaa	accguguuaa	660
cggcucguua	cauaggccua	ugcagcucug	acguuaugga	gcggucacgu	agagggaugu	720
ccaauucuuag	aaagaaguau	uugaaaccau	ccaacaauug	ucuauucucu	guuggcucga	780

ccaucuacca	cgagaagagg	gacuuacuga	ggagcuggca	ccugccgucu	guauuucacu	840
uacguggcaa	gcaaaaauac	acaugucggg	gugagacuau	aguuaugu	gacggguacg	900
ucguuaaaag	aaauagcuau	aguccaggcc	uguaugggaa	gccuucaggc	uaugcugcua	960
cgaugcaccg	cgagggauuc	uugugcugca	aagugacaga	cacauugaac	ggggagaggg	1020
ucucuuuucc	cgugugcacg	uaugugccag	cuacauugug	ugaccaaau	acuggcauac	1080
uggcaacaga	ugucagugcg	gacgacgcgc	aaaaacugcu	gguggggcuc	aaccagcgua	1140
uagucgucaa	cggucgcacc	cagagaaaca	ccaauaccu	gaaaaauuac	uuuugccccg	1200
uaguggccca	ggcauuugcu	aggugggcaa	aggauauaa	ggaagaucaa	gaagaugaaa	1260
ggccacuagg	acuacgagau	agacaguua	ucauggggug	uuuugggcu	uuuagaaggc	1320
acaagauaac	aucuauuuau	aagcgcccg	auacccaaac	caucaucaa	gugaacagcg	1380
auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagauc	gggcugagaa	1440
caagaaucag	gaaaaugua	gaggagcaca	aggagccguc	accucucau	accgccgagg	1500
acguacaaga	agcuaaugc	gcagccgaug	aggcuagga	ggugcgugaa	gccgaggagu	1560
ugcgcgcagc	ucuaccaccu	uuggcagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuua	cgcuugcucu	ucuccgcagg	1740
cuguacucaa	gagugaaaaa	uuauuuugca	uccaccucuc	cgcuagaaca	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuaug	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuuguuacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgcugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaagggacg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaca	ccccguagag	accuguaaua	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccea	aacagugcgg	uuuuuuuaac	augaugugcc	2520
ugaaauguca	uuuuuaccac	gagauuuugca	cacaagucuu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagagaa	2640
cgacgaaucc	gaaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaagacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uacccuggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggacccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuuaugucc	ggugcugaag	accgcuggca	3120

uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaucg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacac	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaa	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuauga	caugaacacu	gguacacugc	3420
gcauuuauga	uccgcgcuaa	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaau	ugacauaaau	uuuguuaaau	ugaggacccc	cuacaaguac	caccacuacc	3720
agcagugcga	ggaccacgcc	aucaagcuga	gcaugcugac	caagaaggcc	ugccugcauc	3780
ugaaucccgg	cggaaccugu	gucagcauag	guuaugguua	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuuguau	ucaucggcuu	cgacagaaag	gcccguacgc	3960
acaauccuua	caagcuuuca	ucaaccuuga	ccaacauuuu	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucua	uaucaugugg	ugcgagggga	uaaugccacg	gccaccgaag	4080
gagugauuuu	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aagggaacca	guuucaccag	gcggccaagg	4680
auauagcaga	aauuauugcc	auguggcccc	uugcaacgga	ggccaauag	cagguaugca	4740
uguauauccu	cggagaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaauacac	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgua	auucauccaa	ggaaguauuc	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuaauaac	cgaggauag	accaggacua	gaacgccuga	gccgaucac	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	auggcccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggaggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguaauca	ggaaccucc	acaucggcu	ccgcgcacaa	5460

gaacaccguc	acuugcaccc	agcagggccu	gcucgagaac	cagccuaguu	uccaccccgc	5520
caggcgugaa	uaggguaguc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuugag	gcuuucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaaauc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcau	uauuugaagg	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuaug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaaau	augcguguaa	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggaccaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaauugga	cagguuugua	auaggacuua	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuga	cgcuaauua	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuauugauuc	uggaagacu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aaauuaaaau	cggagccaug	augaaaucug	gaauuuuccu	cacacuguuu	gugaacacag	7020
ucauuacau	uguauucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucan	uggagaugac	aaauucguga	aaggagucan	aucggacaaa	uuauuggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuua	agcuuggcaa	accucuggca	gcagacgaug	7320
aacauauga	ugacaggaga	agggcauugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaggua	ugaaaccgua	ggaacuucca	7440
ucauaguau	ggccaugacu	acucuaagcu	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucua	ggcuuaccug	aauggacuac	gacauagucu	aguccgcaa	7560
g						7561
<210>	23					
<211>	7561					
<212>	RNA					

<213> Artificial Sequence

<220>

<223> Synthetic construct, C04 nsP1-4 ORF and SGP RNA

<400> 23

augggcgggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccauuccuca	gagcuuugca	gcgagagcuuc	ccgcaguuug	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgcauc	180
uggcuucaaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgccccg	ccgcagaaug	uauucuaagc	acaaguauca	uuguauucugu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuua	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaaggg	aguuaagaguc	gccuacugga	uaggcuuuga	caccaccccu	uuuauuuua	600
agaacuuggc	uggagcauau	ccaucauacu	cuaccaacug	ggccgacgaa	accguguuaa	660
cggcucguua	cauaggccua	ugcagcucug	acguuaugga	gcggucacgu	agagggaugu	720
ccaauucuuag	aaagaaguau	uugaaaccu	ccaacaangu	ucuauucucu	guuggcucga	780
ccaucuaacca	cgagaagagg	gacuuacuga	ggagcuggca	ccugccgucu	guauuucacu	840
uacguggcaa	gcaaaaauac	acaugucggu	gugagacuau	aguuauguuc	gacggguacg	900
ucguuaaaag	aauagcuau	aguccaggcc	uguaugggaa	gccuucaggc	uauugcucua	960
cgaugcaccg	cgagggaauuc	uugugcugca	aagugacaga	cacauugaac	ggggagaggg	1020
ucucuuuuucc	cgugugcacg	uauugccag	cuacauugug	ugaccaaau	acuggcauac	1080
uggcaacaga	ugucagugcg	gacgacgcgc	aaaaacugcu	gguugggcuc	aaccagcgua	1140
uagucgucaa	cggucgcacc	cagagaaaca	ccaauaccau	gaaaaauac	cuuugccccg	1200
uaguggccca	ggcauuugcu	agguggggcaa	aggaauauaa	ggaagaucaa	gaagaugaaa	1260
ggccacuagg	acuacgagau	agacaguua	ucauggggug	uuguugggcu	uuuagaaggc	1320
acaagauaac	aucuauuuau	aagcgcccgg	auacccaaac	caucaucaa	gugaacagcg	1380
auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagau	gggcugagaa	1440
caagaauucag	gaaaauguua	gaggagcaca	aggagccguc	accucucauu	accgccgagg	1500
acguacaaga	agcuuaaguc	gcagccgaug	aggcuuaagga	ggugcgugaa	gccgaggagu	1560
ugcgcgagc	ucuaccaccu	uuggcagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuua	cgcugugcuu	ucuccgcagg	1740
cuguacucaa	gagugaaaaa	uuaucuugca	uccaccucuc	cgcugaacaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuaug	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100

ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauag	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaagggacg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaca	ccccguagag	accuguaau	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagugcgg	uuuuuuuaac	augaugugcc	2520
ugaaaugugc	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucugcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagaa	2640
cgacgaaucc	gaaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaagacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguau	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uacccuggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggaccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuugugcc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaucg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacacu	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaa	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuauga	caugaacacu	gguacacugc	3420
gcauuuauga	uccgcgcuaa	aaccuaguac	cuguaaacag	aagacugccu	caugcuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccgucce	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaau	ugacauaaau	uuuguuaaau	ugaggacccc	auauaaaau	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaaucccg	cggaaaccugu	gucagcauag	guuaugguu	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaaccccu	caagcugagc	agcaccucga	ccaacaucua	caccggcagc	agacugcacg	4020
aagccggau	ugcaccucua	uaucaugugg	ugcgagggga	uauugccaca	gccacagaag	4080
gagugauuu	aaauugcugc	aacagcaaag	gacaaccugg	cggaggggug	ugcgagcg	4140
uguauaagaa	auucccgga	agcuucgau	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuagac	accacugaug	4440

cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aagggaccaa	guuucaccag	gcgccaagg	4680
auauagcaga	aauuauugcc	auguggcccc	uugcaacgga	ggccaauagag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaauacacu	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgua	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggaugag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucugua	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacaccu	ggaggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucgcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uagggugauc	acuagagagg	agcucgaggc	gcuuacccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaauucc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcau	uauuugaagg	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuagug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaaau	augcguguua	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggacaaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaauugga	cagguuuugua	auaggacuuua	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cagagcuggg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780

uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuaaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuaaaau	cggagccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuaacau	uguaaucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucau	uggagaugac	aauaucguga	aaggagucua	aucggacaaa	uuaauggcag	7140
acaggugcgc	caccugguug	aauauggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuauuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	cccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aacaugauga	ugacaggaga	agggcauugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaaggua	ugaaaccgua	ggaacuucca	7440
ucauaguua	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucucac	ggcuaaccug	aauggacuac	gacauagucu	aguccgcca	7560
g						7561

<210> 24

<211> 7561

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C05 nsP1-4 ORF and SGP RNA

<400> 24

augggcggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccaauccuca	gagcuuugca	gcggagcuuc	ccgcaguuug	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgauc	180
uggcuucaaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgccccg	ccgcagaaug	uauucuaagc	acaaguauca	uuguauucugu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuua	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaaggg	aguuaagaguc	gccuacugga	uaggcuuuga	caccaccccu	uuuauguuuu	600
agaacuuggc	uggagcauau	ccaucacuac	cuaccaacug	ggccgacgaa	accguguuaa	660
cggcucguua	cauaggccua	ugcagcucug	acguuaugga	gcggucacgu	agaggggaugu	720
ccaauucuuag	aaagaaguau	uugaaaccu	ccaacaugu	ucuauucucu	guuggcucga	780
ccaucuaacca	cgagaagagg	gacuuacuga	ggagcuggca	ccugccgucu	guauuucacu	840
uacguggcaa	gcaaaaauac	acaugucggu	gugagacuau	aguuauguuc	gacggguacg	900
ucguuaaaag	aaauagcuau	aguccaggcc	uguaugggaa	gccuucaggc	uauugcucua	960
cgaugcaccg	cgagggaauuc	uugugcugca	aagugacaga	cacauugaac	ggggagaggg	1020
ucucuuuucc	cgugugcacg	uauugccag	cuacauugug	ugaccaaau	acuggcauac	1080

uggcaacaga	ugucagugcg	gacgacgcgc	aaaaacugcu	gguugggcuc	aaccagcgua	1140
uagucgucaa	cggucgcacc	cagagaaaca	ccaauaccu	gaaaaauuac	uuuugcccc	1200
uaguggccca	ggcauuugcu	aggugggcaa	aggaauuaa	ggaagaucaa	gaagaugaaa	1260
ggccacuagg	acuacgagau	agacaguua	ucauggggug	uuuugggcu	uuuagaaggc	1320
acaagauaac	aucuauuuau	aagcgcccgg	auacccaaac	caucaucaa	gugaacagcg	1380
auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagauc	gggcugagaa	1440
caagaauacg	gaaaauguua	gaggagcaca	aggagccguc	accucucauu	accgccgagg	1500
acguacaaga	agcuagugc	gcagccgaug	aggcuagga	ggugcgugaa	gccgaggagu	1560
ugcgcgcgag	ucuaccaccu	uuggcgagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuua	cgcugugcuu	ucuccgcagg	1740
cuguacucaa	gagcgagaag	cugagcugca	uucaccucuc	ggccgagcaa	gugaucguga	1800
ucacacacag	cggccggaag	ggcagauaug	ccguggaacc	uuauacggc	aagguggugg	1860
ugccugaggg	acacgcuauu	ccagugcagg	acuuucaggc	ccugagcgag	ucugccacca	1920
ucguguacaa	cgagcgcgag	uucgugaaca	gauaccugca	ccacauugcc	acacacggcg	1980
gagcccugaa	caccgacgaa	gaguacuaca	agaccgugaa	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuaaa	agcgaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaaggagcg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggauagcgug	cugcugaacg	gcugcaagca	ccccguagag	accuguaaua	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccua	agcagugcgg	cuucuuaac	augaugugcc	2520
ugaaggugca	cuucaaccac	gagaucugca	cccagguguu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagaaa	2640
cgacgaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaugacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uaccucggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggaccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuugagucc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaucg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacac	guuccguuuu	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaa	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuauga	caugaacacu	gguacacugc	3420

gcaauuauga	uccgcgcgcaua	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuuucuc	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccgucce	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaaua	ugacauaaua	uuuguuauug	ugaggacccc	auauaaaauac	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaauccccg	cggaaaccugu	gucagcauag	guuaugguuu	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuuguau	ucauuggguu	cgaucgcaag	gcccguacgc	3960
acaauccuua	caagcuuuca	ucaaccuuga	ccaacauuuu	uacagguucc	agacuccacg	4020
aagccggau	ugcaccuuc	uaucaugugg	ugcgaggggga	uauugccacg	gccaccgaag	4080
gagugauuuu	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaagauucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aagggaacca	guuucaccag	gcggccaagg	4680
auauagcaga	aaauaaugcc	auguggcccc	uugcaacgga	ggccaauag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaaucau	ggugugcaga	agaucuaug	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgau	auucauccaa	ggaaguauuc	cugggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	auuccacaga	ggggacaccu	gaacaaccac	5100
cacuuaauaac	cgaggauag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccc	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggaggggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucuc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccaccccg	5520
caggcgugaa	uaggguaguc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauagguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaaauca	guaaggcaaa	5760

cggugcuauc	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaaaucce	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcgu	uauuugaagg	cagaaggaaa	aguggagugc	uaccgaaccc	6000
ugcauccugu	uccuuuguau	ucaucuagug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaaauugc	aaugucacgc	aaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaauugc	uucaagaaau	augcguguaa	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggacaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaaugga	cagguuugua	auggacuuaa	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgaagcuggg	uuaggagauu	aaugcgguc	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuuaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aaauuaaaau	cggagccaug	augaaaucug	gaauuguuccu	cacacuguuu	gugaacacag	7020
ucauuacau	uguaaucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucan	uggagaugac	aaauucguga	aaggagucan	aucggacaaa	uuauuggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgcuuu	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcugucc	7260
guguggcaga	cccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aacaugauga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucagguua	ugaaaccgua	ggaacuucca	7440
ucauaguua	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucua	ggcuuaccug	aauggacuac	gacauagucu	aguccgcaa	7560
g						7561
<210>	25					
<211>	7561					
<212>	RNA					
<213>	Artificial Sequence					
<220>						
<223>	Synthetic construct, C06 nsP1-4 ORF and SGP RNA					
<400>	25					
augggcggcg	caugagagaa	gccagacca	auuaccuacc	caaauggag	aaaguucacg	60

uugacaucga	ggaagacagc	ccaauccuca	gagcuuugca	gcggagcuuc	ccgcaguuug	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgauc	180
uggcuucaaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaaguauca	uuguaucugu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuaa	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaagg	aguuagaguc	gccuacugga	uaggcuuuga	caccaccccu	uuuauuuua	600
agaacuuggc	uggagcauau	ccaucacuac	cuaccaacug	ggccgacgaa	accguguuaa	660
cggccagaaa	uaucggccug	uguagcagcg	acgugaugga	aagauccaga	cggggcauga	720
gcauccugcg	gaagaaguac	cugaagccua	gcaacaacgu	gcuguucagc	gugggcagca	780
ccaucuaacca	cgagaagagg	gaccugcugc	ggagcuggca	ucugccuucc	guguuucacc	840
ugagaggcaa	gcagaacuac	accuguagau	gcgagacaau	cguguccugc	gacggcuacg	900
uggucaagcg	gaucgccauu	ucuccuggcc	uguacggcaa	gccuucuggc	uaucccgcca	960
ccaugcacag	agaaggcuuu	cuguguugca	aagugaccga	cacacugaac	ggcgagcggg	1020
uguccuuucc	uguguguacc	uauugccccg	ccacacugug	cgaucagaug	acaggcauuc	1080
uggccaccga	cgugucagcc	gacgaugccc	agaaacugcu	cgugggccug	aaccagagaa	1140
ucguggucaa	cggcagaacc	cagcgggaaca	ccaacaccau	gaagaacuac	cugcugccug	1200
ugguggccca	ggccuuugcc	agaugggcca	aagaguacaa	agaggaucaa	gaggacgagc	1260
ggccccuggg	ccugagagau	agacaacugg	ucaugggcug	cugcugggcc	uucagaaggc	1320
acaagaucac	cagcaucuac	aagcggcccc	acacacagac	caucaucaa	gugaacagcg	1380
acuuccacag	cuucgugcug	ccucggaucg	gcagcaacac	acuggaaauc	ggccugcgga	1440
cccggauca	aaagaugcug	gaagaacaca	aagagcccuc	uccacugauc	accgccgagg	1500
augugcaaga	ggccaaaugu	gccgccgacg	aggcuaaaaga	agugcgcgaa	gccgaggaac	1560
ugagagccgc	acuuccuccu	cuggccgccg	auguugaaga	accacucug	gaagccgacg	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuua	cgcugugcuu	ucuccgcagg	1740
cuguacucaa	gagugaaaaa	uuaucuugca	uccaccucuc	cgcugaacaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuau	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagagg	acaugcaaua	cccguccagg	acuuaaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaucugca	ccaauuugcc	acacauggag	1980
gagcgugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaauu	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccaauaggggu	guauggcgug	ccaggaucag	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaaggagcg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaca	ccccguagag	accuguaaua	2400

uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagugcgg	uuuuuuuaac	augaugugcc	2520
ugaaagugca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagaaa	2640
cgacgaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaagacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uaccucggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggacccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaaag	cuuuagugcc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaucg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacac	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaa	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuauga	caugaacacu	gguacacugc	3420
gcaauuauga	uccgcgcuaa	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaauagaacac	ccacagagug	acuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccgucgc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaaua	ugacauaaua	uuuguuaaag	ugaggacccc	auauaaaauac	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaaucccg	cggaaccugu	gucagcauag	guuauugguu	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuua	caagcuuua	ucaaccuuga	ccaacauuu	uacagguucc	agacuccacg	4020
aagccggau	ugcaccucca	uaucaugugg	ugcgaggggga	uauugccacg	gccaccgaag	4080
gagugauuu	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aagggaacaa	guuucaccag	gcggccaagg	4680
auauagcaga	aauuuaugcc	auguggcccc	uugcaacgga	ggccaauag	cagguaugca	4740

uguauauccu	cggagaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaaauac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaaucacu	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcguau	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggauag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggcccac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcuauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggagggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcaccc	agcagggccu	gcucgagaac	cagccuaguu	uccaccccg	5520
caggcgugaa	uaggguaguc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuugag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaaauc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcau	uauuuagaag	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuagug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaaauugc	aaugucacgc	aaauagagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaauugc	uucaagaaa	augcguguua	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggacaaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaaugga	cagguuugua	auggacuuaa	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuuaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuaauuu	cggagccaug	augaaaucug	gaauuguuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguaaucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080

cagcauuc	au	uggagaug	ac	aaauauc	guga	aaggagu	caa	aucggaca	aaa	uuaauggc	cag	7140	
acaggugc	gc	caccuggu	ug	aaauaug	gaag	ucaagau	uau	agaugcu	gug	gugggc	gaga	7200	
aagcgccu	ua	uuucugug	ga	ggguuu	uuu	ugugug	acuc	cgugacc	ggc	acagcgug	cc	7260	
guguggc	aga	ccccua	aaaa	aggcu	guuu	aa	agcuugg	caa	accucug	gca	gcagacgaug	7320	
aacauga	ga	ugacagg	aga	agggc	auug	c	augaag	aguc	aacacgc	ugg	aaccgagugg	7380	
guauucu	uu	agagcu	gugc	aaggc	agu	ag	aauc	aaggua	ugaaacc	gua	ggaacuucca	7440	
ucauagu	uu	au	ggccaug	acu	cuag	cua	gcagug	uu	aa	aucau	ucagc	uaccugagag	7500
gggcccc	uau	aacucuc	uac	ggcu	aacc	ug	aaugga	cuac	gacau	aguc	aguccgcca	7560	
g												7561	
<210>	26												
<211>	7561												
<212>	RNA												
<213>	Artificial Sequence												
<220>													
<223>	Synthetic construct, C07 nsP1-4 ORF and SGP RNA												
<400>	26												
augggc	ggcg	caugaga	gaa	gcccaga	cca	auuaccu	acc	caaaaugg	ag	aaaguuc	acg	60	
uugaca	ucga	ggaagac	agc	ccaaucc	uca	gagcuu	ugca	gcggagc	uuc	ccgcagu	uu	120	
agguaga	aagc	caagcagg	uc	acugau	aaug	accaugc	uaa	ugccagag	cgc	uuuucga	uc	180	
uggcuu	caaa	acugauc	gaa	acggagg	ugg	acccauc	cga	cacgauc	cuu	gacauug	gaa	240	
gugcgcc	cg	ccgcaga	aug	uauuc	uaagc	acaagua	uca	uuguauc	ugu	ccgaug	agau	300	
gugcgga	aga	uccggac	aga	uuguau	aagu	augcaac	uaa	gcugaag	aaa	aacugua	aagg	360	
aaaua	acuga	uaagga	auug	gacaaga	aaaa	ugaagg	agcu	cgccg	ccguc	augagc	gacc	420	
cugaccu	gga	aacugag	acu	augugc	cucc	acgacg	acga	gucgug	ucgc	uacga	agggc	480	
aagucgc	ugu	uuaccag	gau	guauac	gcgg	uugacg	gacc	gacaag	ucuc	uauacca	aag	540	
ccaaua	aggg	aguuag	aguc	gccuac	ugga	uaggcu	uu	gaa	ccccu	uuuau	guu	600	
agaacu	uggc	uggagc	auau	ccauc	auacu	cuacca	acug	ggccg	acgaa	accgu	guu	660	
cggcuc	gu	cauagg	ccua	ugcagc	ucug	acguu	augga	gcgguc	acgu	agaggga	ugu	720	
ccaucu	uuag	aaaga	aguau	uugaa	accau	ccaaca	augu	ucuau	ucuc	guuggc	ucga	780	
ccaucu	acca	cgaga	aagagg	gacuu	acuga	ggagc	uggca	ccugc	cguc	guauu	ucacu	840	
uacgug	gcaa	gcaaaa	uuac	acaug	ucggu	gugag	acuau	aguua	guugc	gacggg	uacg	900	
ucguu	aaaaag	aauag	cuauc	agucc	aggcc	uguau	gggaa	gccu	ucaggc	uau	gcugcua	960	
cgaugc	caccg	cgaggga	uu	uugug	cugca	aagug	acaga	cacau	ugaac	gggg	agaggg	1020	
ucucu	uuuucc	cgugug	cacg	uaugug	ccag	cuacau	ugug	ugacca	aaug	acuggc	auac	1080	
uggca	acaga	ugucag	ugcg	gacg	acgcgc	aaaaa	cugcu	gguug	ggcuc	aacc	agcgua	1140	
uaguc	gucaa	cgguc	gcacc	cagaga	aaaca	ccaau	accu	gaaaa	uuac	cuuu	ugccccg	1200	
uagugg	ccca	ggcau	uuugcu	aggugg	ggcaa	agga	auauaa	ggaag	aucaa	gaaga	ugaaa	1260	
ggccac	uagg	acuac	gagau	agac	aguuag	ucaugg	ggug	uuu	ugggcu	uuu	agaaggc	1320	
acaagau	aac	aucu	auuuau	aagcg	ccccg	auaccc	aaac	cauca	caaaa	guga	acagcg	1380	

auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagauc	gggcugagaa	1440
caagaaucag	gaaaaugua	gaggagcaca	aggagccguc	accucucauu	accgccgagg	1500
acguacaaga	agcuaagugc	gcagccgaug	aggcuaagga	ggugcgugaa	gccgaggagu	1560
ugcgcgcagc	ucuaccaccu	uuggcagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucgggcucuua	cgcugugcuu	ucuccgcagg	1740
cuguacucua	gagugaaaaa	uuaucuugca	uccacccucu	cgcugaacaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuaug	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauaggag	1980
gagcgcugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaui	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaui	auaaggagcg	ucaagaaaui	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaauu	gaugcaaaca	ccccguagag	accuguaui	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagugcgg	uuuuuuuac	augaugugcc	2520
ugaaagugca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaagagc	aucagcagac	2580
ggugcaccac	gagcugagac	agcguuggugu	cuaccucguu	cuacgacaag	aagaugcgga	2640
cgacaaaccc	caaagagaca	aagaucguca	ucgacaccac	cggcagcacc	aagccuaagc	2700
aggacgaucu	gauccugacc	ugcuucagag	gcugggucaa	gcagcugcag	aucgacuaca	2760
agggcaacga	gaucaugacc	gccgcugccu	cucagggccu	gacaagaaaa	ggcguguacg	2820
ccgugcggua	caaagugaac	gagaacccuc	uguacgcccc	uaccagcgag	caugugaauu	2880
ugcugcugac	ccggaccgag	gaccggauug	uuuggaaaac	acuggccggc	gaucccugga	2940
ucaagacccu	gacagccaag	uaucccggca	acuucaccgc	caccaucgag	gaauggcagg	3000
ccgagcacga	ugccaucaug	cggcacaucc	uggaaagacc	cgauccuacc	gacguguucc	3060
agaacaaggc	caacgugugc	ugggccaaaag	cucuggugcc	ugugcugaaa	accgccggca	3120
ucgauaugac	caccgagcag	uggaacaccg	uggacuacuu	cgagacagac	aaggcccaca	3180
gcgccgagau	cgugcugaau	cagcugugcg	ugcgguuuuu	cggccuggau	cuggauagcg	3240
gccuguucuc	ugcuccuacc	gugccucuga	gcauccggaa	caaccacugg	gacaacagcc	3300
ccucuccuaa	uauguacggc	cugaacaaaag	aagucgugcg	gcagcugagc	agaagauacc	3360
cacagcugcc	uagagccgug	gccacaggca	gaguguacga	caugaauacc	ggcacacugc	3420
ggaacuacga	ccccagaauc	aaucuggugc	ccgugaacag	aaggcugccc	cacgcucugg	3480
uucugcacca	caaugagcac	ccucagagcg	acuucagcag	cuucgugucc	aagcugaagg	3540
gcagaaccgu	gcugguugug	ggcgagaagc	ugucugugcc	uggcaagaug	guggacuggc	3600
ugagcgauag	acccgaggcc	accuuuagag	ccagacugga	ccuuggaauc	ccuggcgacg	3660
ugcccaaaui	cgacaucauc	uucgugaacg	ugcggacgcc	cuacaaguac	caccacuacc	3720

agcagugcga	ggaccacgcc	aucaagcuga	gcaugcugac	caagaaggcc	ugccugcacc	3780
ugaauccugg	cggcaccugu	gugucuaucg	gcuacggcua	ugccgacaga	gccagcgagu	3840
cuaucaucgg	cgccauugcc	agacaguuca	aguucagcag	aguugcaag	cccaagagca	3900
gccuggaaga	gacagaggug	cuguucgugu	ucaucggcua	ugaccggaag	gcccguacgc	3960
acaauccuua	caagcuuua	ucaaccuuga	ccaacauua	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucca	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuau	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aagggaacca	guuucaccag	gcggccaagg	4680
auauagcaga	aauuauugcc	auguggcccc	uugcaacgga	ggccaugag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaauacau	ggugugcaga	agauccaaug	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgua	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggauag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacaccuu	ggagggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uaggguaga	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaaauc	acaccugcu	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcgu	uauuuagaag	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuagug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060

caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaauugc	uucaagaaau	augcguguua	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggaccaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaauugga	cagguuugua	auggacuuaa	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauuacauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuaaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuaaaau	cggagccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguaaucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucan	uggagaugac	aaauucguga	aaggagucan	aucggacaaa	uuauuggcag	7140
acaggucgcg	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuauuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aacaugauga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaaggua	ugaaaccgua	ggaacuucca	7440
ucauaguua	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucuc	ggcuuaccug	aauggacuac	gacauagucu	aguccgcaa	7560
g						7561

<210> 27

<211> 7561

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C08 nsP1-4 ORF and SGP RNA

<400> 27

augggcggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggaa	aaggugcacg	60
uggacaucga	agaggacagc	ccauuccuga	gagcccugca	gagaagcuuc	ccucaguucg	120
agguggaagc	caaacaagug	accgacaacg	aucacgcca	cgcgagggcc	uucucucacc	180
ucgcaagcaa	gcugaucgag	acagaggugg	accccagcga	caccauccug	gauauuggau	240
cugccccugc	cagaagaauug	uauucuaagc	acaaguauca	uuguaucugu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuua	gcugaagaaa	aacuguaagg	360

aaauaacuga	uaaggaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaaggg	aguuagaguc	gccuacugga	uaggcuuuga	caccaccccu	uuuauuuua	600
agaacuuggc	uggagcauau	ccaucacuac	cuaccaacug	ggccgacgaa	accguguuaa	660
cggcucguaa	cauaggccua	ugcagcucug	acguuaugga	gcggucacgu	agagggaugu	720
ccauucuuag	aaagaaguau	uugaaaccu	ccaacaauu	ucuauucucu	guuggcucga	780
ccaucuaacca	cgagaagagg	gacuuacuga	ggagcuggca	ccugccgucu	guauuucacu	840
uacguggcaa	gcaaaaauac	acaugucgg	gugagacuau	aguuaguugc	gacggguacg	900
ucguaaaaag	aauagcuau	aguccaggcc	uguaugggaa	gccuucaggc	uauugcucua	960
cgaugcaccg	cgagggaau	uugugcugca	aagugacaga	cacauugaac	ggggagagg	1020
ucucuuuuucc	cgugugcacg	uaugugccag	cuacauugug	ugaccaaau	acuggcauac	1080
uggcaacaga	ugucagugcg	gacgacgcgc	aaaaacugcu	gguugggcuc	aaccagcgua	1140
uagucgucaa	cggucgcacc	cagagaaaca	ccaauaccu	gaaaaauac	cuuuugccc	1200
uaguggccca	ggcauuugcu	aggugggcaa	aggaauuaa	ggaagaucaa	gaagaugaaa	1260
ggccacuagg	acuacgagau	agacaguua	ucauggggug	uuuugggcu	uuuagaagg	1320
acaagauaac	aucuauuuau	aagcgcccgg	auacccaaac	caucaucaa	gugaacagcg	1380
auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagau	gggcugagaa	1440
caagaauacg	gaaaaugua	gaggagcaca	aggagccguc	accucucau	accgccgagg	1500
acguacaaga	agcuagugc	gcagccgaug	aggcuagga	ggugcgugaa	gccgaggagu	1560
ugcgcgcagc	ucuaccaccu	uuggcagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuua	cgcugugcuu	ucuccgcagg	1740
cuguacucaa	gagugaaaa	uuaucuugca	uccaccucuc	cgcugaacaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuaug	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgcugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuuaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaaggggacg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaca	ccccguagag	accuguaaua	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagucgg	uuuuuuuac	augaugugcc	2520
ugaaagugca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagaaa	2640
cgacgaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700

aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aauaaugacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uacccuggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggacccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuuagugcc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaucg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacu	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaag	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuaua	caugaacacu	gguacacugc	3420
gcauuuauga	uccgcgcaua	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccgucce	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaua	ugacauaaua	uuuguuauug	ugaggacccc	auauaaaauac	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaaucccg	cggaaucugu	gucagcauag	guuaugguua	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuugua	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuua	caagcuuua	ucaaccuuga	ccaacauua	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucu	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuau	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aagggaacca	guuucaccag	gcggccaagg	4680
auauagcaga	aauuuaugcc	auguggcccc	uugcaacgga	ggccaaugag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaauacacu	ggugugcaga	agauccaaug	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgua	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040

acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggaugag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccc	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggaggaggcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uaggguaguc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuca	guuaauccc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcau	uauuuagaag	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuaug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaaa	augcguguaa	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggaccaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaaugga	cagguuugua	auggacuuua	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuuga	cgcuauuuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuuaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuaaaau	cggagccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguaaucgca	agcagagugu	ugagagaacg	gcuuaccgga	ucaccaugug	7080
cagcauucau	uggagaugac	aaauucguga	aaggagucua	aucggacaaa	uuauuggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgcuuua	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aacaugauga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380

guauucuuuc	agagcugugc	aaggcaguag	aaucaaggua	ugaaaccgua	ggaacuucca	7440
ucauaguuaa	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucuac	ggcuaaccug	aauggacuac	gacauagucu	aguccgccaa	7560
g						7561
<210>	28					
<211>	7561					
<212>	RNA					
<213>	Artificial Sequence					
<220>						
<223>	Synthetic construct, C09 nsP1-4 ORF and SGP RNA					
<400>	28					
augggcggcg	caugagagaa	gcccagacca	auuaccuacc	caaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccauuccuca	gagcuuugca	gcggagcuuc	ccgcaguuug	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgcauc	180
uggcuucaaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaaguauca	uuguaucugu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuaa	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaaggg	aguuagaguc	gccuacugga	uaggcuuuga	caccaccccu	uuuauuuua	600
agaacuuggc	uggagcauau	ccaucauacu	cuaccaacug	ggccgacgaa	accguguuaa	660
cggccagaaa	uauccggccug	uguagcagcg	acgugaugga	aagauccaga	cggggcauga	720
gcauccugcg	gaagaaguac	cugaagccua	gcaacaacgu	gcuguucagc	gugggcagca	780
ccaucuacca	cgagaagagg	gaccugcugc	ggagcuggca	ucugccuucc	guguuucacc	840
ugagaggcaa	gcagaacuac	accuguagau	gcgagacaau	cguguccugc	gacggcuacg	900
uggucaagcg	gaucgccauu	ucuccuggcc	uguacggcaa	gccuucuggc	uaugccgcca	960
ccaugcacag	agaaggcuuu	cuguguugca	aagugaccga	cacacugaac	ggcgagcggg	1020
uguccuuucc	uguguguacc	uaugugcccc	ccacacugug	cgaucaugau	acaggcauuc	1080
uggccaccga	cgugucagcc	gacgaugccc	agaaacugcu	cgugggccug	aaccagagaa	1140
ucguggucaa	cggcagaacc	cagcgggaaca	ccaacaccau	gaagaacuac	cugcugccug	1200
ugguggccca	ggccuuugcc	agaugggcca	aagaguacaa	agaggaucaa	gaggacgagc	1260
ggccccuggg	ccugagagau	agacaacugg	ucaugggcug	cugcugggcc	uucagaaggc	1320
acaagaucac	cagcaucuac	aagcggcccc	acacacagac	caucaucaa	gugaacagcg	1380
acuuccacag	cuucgugcug	ccucggaucc	gcagcaacac	acuggaaauc	ggccugcgga	1440
cccggauacg	aaagaugcug	gaagaacaca	aagagcccuc	uccacugauc	accgccgagg	1500
augugcaaga	ggccaaaugu	gccgccgacg	aggcuuaaga	agugcgcgaa	gccgaggaac	1560
ugagagccgc	acuuccuccu	cuggccgccg	auguugaaga	accacucug	gaagccgacg	1620
ucgacuugau	guuacaagag	gcuggggccc	gcucagugga	gacaccucgu	ggcuugauaa	1680

agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuaa	cgcugugcuu	ucuccgcagg	1740
cuguacucua	gagugaaaaa	uuaucuugca	uccaccucuc	cgcugaacaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuaug	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgcugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucucc	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaagggacg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaca	ccccguagag	accuguaaua	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagugcgg	uuuuuuuac	augaugugcc	2520
ugaaagugca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaagagc	aucagcagac	2580
ggugcaccaa	gagcgugacc	agcguggugu	cuaccuguu	cuacgacaag	aagaugcgga	2640
cgacaaaacc	caaagagaca	aagaucguca	ucgacaccac	cggcagcacc	aagccuaagc	2700
aggacgaucu	gauccugacc	ugcuucagag	gcugggucaa	gcagcugcag	aucgacuaca	2760
agggcaacga	gaucaugacc	gccgcugccu	cucagggccu	gacaagaaaa	ggcguguacg	2820
ccgugcggua	caaagugaac	gagaaccuc	uguacgcccc	uaccagcgag	caugugaau	2880
ugcugcugac	ccggaccgag	gaccggau	uuuggaaaac	acuggccggc	gauccugga	2940
ucaagacccu	gacagccaag	uaucgggca	acuucaccgc	caccaucgag	gaauggcagg	3000
ccgagcacga	ugccaucaug	cggcacaucc	uggaaagacc	cgauccuacc	gacguguucc	3060
agaacaaggc	caacgugugc	ugggccaaag	cucuggugcc	ugugcugaaa	accgccggca	3120
ucgauaugac	caccgagcag	uggaacaccg	uggacuacuu	cgagacagac	aaggcccaca	3180
gcgccgagau	cgugcugaau	cagcugugcg	ugcgguuuuu	cggccuggau	cuggauagcg	3240
gccuguucuc	ugcuccuacc	gugccucuga	gcauccggaa	caaccacugg	gacaacagcc	3300
ccucuccuaa	uauguacggc	cugaacaaag	aagucgugcg	gcagcugagc	agaagauacc	3360
cacagcugcc	uagagccgug	gccacaggca	gaguguacga	caugaauacc	ggcacacugc	3420
ggaacuacga	ccccagaauc	aaucuggugc	ccgugaacag	aaggcugccc	cacgcucugg	3480
uucugcacca	caaugagcac	ccucagagcg	acuucagcag	cuucgugucc	aagcugaagg	3540
gcagaaccgu	gcugguugug	ggcgagaagc	ugucugugcc	uggcaaug	guggacuggc	3600
ugagcgauag	acccgaggcc	accuuuagag	ccagacugga	ccuuggaau	ccuggcgacg	3660
ugcccaaua	cgacaucauc	uucgugaacg	ugcggacgcc	cuacaaguac	caccacuacc	3720
agcagugcga	ggaccacgcc	aucaagcuga	gcaugcugac	caagaaggcc	ugccugcacc	3780
ugaauccugg	cggcaccugu	gugucuaucg	gcuacggcuu	ugccgacaga	gccagcgagu	3840
cuaucaucgg	cgccauugcc	agacaguuca	aguucagcag	agugugcaag	cccaagagca	3900
gccuggaaga	gacagaggug	cuguucgugu	ucaucggcuu	ugaccggaag	gcccguacgc	3960
acaauccuua	caagcuuua	ucaaccuuga	ccaacauuuu	uacagguucc	agacuccacg	4020

aagccggaug	ugcaccucca	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuau	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aaggggacca	guuucaccag	gcggccaagg	4680
auauagcaga	aauuauugcc	auguggcccc	uugcaacgga	ggccaauag	caggaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacacgc	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaaucacu	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcguau	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggaugag	accaggacua	gaacgccuga	gccgaucac	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucac	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggaggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucgccu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uagggugauc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guaaaaucc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcua	uauuuagaag	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuaug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360

cggccuuuaa	uguggaaugc	uucaagaaa	augcguguaa	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggacccaaa	agcugcugcu	uuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaauugga	cagguuugua	auggacuuaa	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauuacauac	acuguuugau	augucggcug	aagacuuuga	cgcuauuuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuaaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuaaaau	cggagccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuacau	uguaaucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucan	uggagaugac	aaauaucguga	aaggagucan	aucggacaaa	uuaauggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aacauauga	ugacaggaga	agggcauugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaaggua	ugaaaccgua	ggaacuucca	7440
ucauaguau	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucuac	ggcuuaccug	aauggacuac	gacauagucu	aguccgccaa	7560
g						7561
<210>	29					
<211>	7561					
<212>	RNA					
<213>	Artificial Sequence					
<220>						
<223>	Synthetic construct, C10 nsP1-4 ORF and SGP RNA					
<400>	29					
augggcgggc	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccauuccuca	gagcuuugca	gcggagcuuc	ccgcaguuug	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgcauc	180
uggcuucaaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaaguauca	uuguauucgu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuua	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaaauu	gacaagaaaa	ugaaggagcu	cggcgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaaggg	aguuaagagc	gccuacugga	uaggcuuuga	caccaccccu	uuuauguuuu	600
agaacuuggc	uggagcauau	ccaucauacu	cuaccaacug	ggccgacgaa	accguguuaa	660

cggcucguaa	cauaggccua	ugcagcucug	acguuaugga	gcggucacgu	agagggaugu	720
ccaauucuuag	aaagaaguau	uugaaaccau	ccaacaauug	ucuauucucu	guuggcucga	780
ccaucuuacca	cgagaagagg	gacuuacuga	ggagcuggca	ccugccgucu	guauuucacu	840
uacguggcaa	gcaaaaauac	acaugucggg	gugagacuau	aguuauguuc	gacggguacg	900
ucguuaaaag	aauagcuau	aguccaggcc	uguauuggaa	gccuucaggc	uaugcugcua	960
cgaugcaccg	cgagggauuc	uugugcugca	aagugacaga	cacauugaac	ggggagaggg	1020
ucucuuuuucc	cgugugcacg	uaugugccag	cuacauugug	ugaccaaau	acuggcauac	1080
uggcaacaga	ugucagugcg	gacgacgcgc	aaaaacugcu	ggauugggcuc	aaccagcgua	1140
uagucgucaa	cggucgcacc	cagagaaaca	ccaauaccu	gaaaaauuac	uuuuugcccc	1200
uaguggccca	ggcauuugcu	aggugggcaa	aggaauauaa	ggaagaucaa	gaagaugaaa	1260
ggccacuagg	acuacgagau	agacaguua	ucauggggug	uuuuggggcu	uuuagaaggc	1320
acaagauaac	aucuauuuau	aagcgccccg	auacccaaac	caucaucaa	gugaacagcg	1380
auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagau	gggcugagaa	1440
caagaaucag	gaaaaugua	gaggagcaca	aggagccguc	accucucau	accgccgagg	1500
acguacaaga	agcuuagug	gcagccgaug	aggcuuagga	ggugcgugaa	gccgaggagu	1560
ugcgcgcagc	ucuaccaccu	uuggcagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuua	cgcuugcucu	ucuccgcagg	1740
cuguacucaa	gagugaaaaa	uuaucuugca	uccaccucuc	cgcuuagaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuau	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaaag	ucugagugaa	agugccacca	1920
uuuguuacaa	cgaacgugag	uucguuuaa	gguaccugca	ccauauugcc	acacauagg	1980
gagcgugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aauaccugua	cgacaucgac	aggaaacagu	gcguuagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccauuauu	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuuaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaaggggag	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaa	ccccguagag	accuguaau	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaaggc	agugcucugc	ggggauccca	aacagugcgg	uuuuuuuac	augaugugcc	2520
ugaaauguca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagaaa	2640
cgcagaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaagacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uacccuggga	auuucacugc	cacgauagag	gaguggcaag	3000

cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggacccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaaagg	cuuuagugcc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaugcg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacacu	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaaag	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuaua	caugaacacu	gguacacugc	3420
gcaauuauga	uccgcgcuaa	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaauagaacac	ccacagagug	acuuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaaua	ugacauaaaua	uuuguuaaag	ugaggacccc	cuacaaguac	caccacuacc	3720
agcagugcga	ggaccacgcc	aucaagcuga	gcaugcugac	caagaaggcc	ugccugcauc	3780
ugaaucccgg	cggaaaccugu	gucagcauag	guuaugguua	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuuguau	ucaucggcuu	cgacagaaag	gcccguacgc	3960
acaauccuua	caagcuuuca	ucaaccuuga	ccaacauuuu	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucu	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuuu	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcgagcgcc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgauug	caaaacuuc	ucauauuugg	aagggaacca	guuucaccag	gcggccaagg	4680
auauagcaga	auuuauugcc	auguggcccc	uugcaacgga	ggccaauag	caggaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaaucaacu	ggugugcaga	agauccaaug	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcguau	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	auccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	caggaugag	accaggacua	gaacgccuga	gccgaucac	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucac	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggagggagcu	agcgugacca	5340

gcggggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguaauca	ggaaccucc	acaucgccu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uagggugauc	acuagagagg	agcucgaggc	gcuuacccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuugag	gcuuucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaaauc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcau	uauuugaagg	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuaug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaauugc	aaugucacgc	aaauagagaa	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaagu	acgccugcaa	caacgaguac	ugggagacau	6420
ucaaagagaa	ccccauccgg	cugaccgagg	aaaacguggu	caacuacauc	accaagcuga	6480
agggccccaa	agccgccgcu	cuguuugcca	agacacacaa	ccugaacaug	cugcaggaca	6540
uccccaugga	cagauucgug	auggaccuga	agcgggacgu	gaaagugacc	ccuggcacca	6600
agcacaccga	ggaacggccu	aaggugcaag	ugauccaggc	cgcugauccu	cuggccacag	6660
ccuauucugug	uggcauccac	agagaacucg	ugcggagacu	gaaugccgug	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	cugaugauuc	uggaagauc	cggaguggac	gccgagcugc	6900
ugacacugau	ugaagccgcc	uuuggcgaga	ucagcagcau	ccaucugccu	accaagacca	6960
aguucaaguu	cggcgccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuacau	uguaaucgca	agcagagugu	ugagagaacg	gcuaacaggc	agcccuugug	7080
cagcauucan	uggagaugac	aaauaucguga	aaggagucan	aucggacaaa	uuauaggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuauuu	ugugugacuc	cugagaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aaauaugaga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaggua	ugaaaccgua	ggaacuucca	7440
ucauaguuau	ggccaugacu	acucuaagcu	gcaguguuua	aucauucagc	uaccugagag	7500
gggccccuau	aacucucua	ggcuuaccug	aauggacuac	gacauagucu	aguccgcaa	7560
g						7561

<210> 30

<211> 7561

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C11 nsP1-4 ORF and SGP RNA

<400> 30

augggcgggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccauuccuca	gagcuuugca	gcggagcuuc	ccgcaguuug	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgcauc	180
uggcuucaaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaaguauc	uuguaucugu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuaa	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauaugccg	uggaugggcc	uacaagccug	uaccaccagg	540
ccaacaaggg	cgucagagug	gccuacugga	ucggcuucga	caccacaccu	uucauguuca	600
agaaccuggc	uggcgcuuac	cccagcuaca	gcacaaacug	ggccgacgaa	accguguuaa	660
cggccagaaa	uaucggccug	uguagcagcg	acgugaugga	aagauccaga	cggggcauga	720
gcauccugcg	gaagaaguac	cugaagccua	gcaacaacgu	gcuguucagc	gugggcagca	780
ccaucuacca	cgagaagagg	gaccugcugc	ggagcuggca	ucugccuucc	guguuucacc	840
ugagaggcaa	gcagaacuac	accuguagau	gcgagacaau	cguguccugc	gacggcuacg	900
uggucaagcg	gaucgccauu	ucuccuggcc	uguacggcaa	gccuucuggc	uaucccgcca	960
ccaugcacag	agaaggcuuu	cuguguugca	aagugaccga	cacacugaac	ggcgagcggg	1020
uguccuuucc	uguguguacc	uauugcccg	ccacacugug	cgaucaaug	acaggcauuc	1080
uggccaccga	cgugucagcc	gacgaugccc	agaaacugcu	cgugggccug	aaccagagaa	1140
ucguggucaa	cggcagaacc	cagcggaaac	ccaacaccau	gaagaacuac	cugcugccug	1200
ugguggccca	ggccuuugcc	agaugggcca	aagaguacaa	agaggaucaa	gaggacgagc	1260
ggccccuggg	ccugagagau	agacaacugg	ucaugggcug	cugcugggcc	uucagaaggc	1320
acaagaucac	cagcaucuac	aagcggcccc	acacacagac	caucaucaa	gugaacagcg	1380
acuuccacag	cuucgugcug	ccucggauug	gcagcaaac	acuggaaauc	ggccugcgga	1440
cccggauca	aaagaugcug	gaagaacaca	aagagcccuc	uccacugauc	accgccgagg	1500
augugcaaga	ggccaaaugu	gccgccgacg	aggcuaaaag	agugcgcgaa	gccgaggaac	1560
ugagagccgc	acuuccuccu	cuggccgccc	auguugaaga	acccacucug	gaagccgacg	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuuu	cgcugugcuu	ucuccgcagg	1740
cuguacucaa	gagcgagaag	cugagcugca	uucaccucuu	ggccgagcaa	gugaucguga	1800
ucacacacag	cggccggaag	ggcagauaug	ccguggaacc	uuauacggc	aagguggugg	1860
ugccugaggg	acacgcuaau	ccagugcagg	acuuucaggc	ccugagcgag	ucugccacca	1920
ucguguacaa	cgagcgcgag	uucgugaaca	gauaccugca	ccacauugcc	acacacggcg	1980

gagcccugaa	caccgacgaa	gaguacuaca	agaccgugaa	gcccagcgag	cacgacggcg	2040
aaauaccgua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaagggacg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggauagcgug	cugcugaacg	gcugcaagca	ccccguagag	accuguaaua	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaagg	agugcucugc	ggggauccua	agcagugcgg	cuucuucaac	augaugugcc	2520
ugaaauggca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaagagc	aucagcagac	2580
ggugcaccaa	gagcgugacc	agcguuggugu	cuaccuguu	cuacgacaag	aagaugcgga	2640
cgacaaaacc	caaagagaca	aagaucguca	ucgacaccac	cggcagcacc	aagccuaagc	2700
aggacgaucu	gauccugacc	ugcuucagag	gcugggucaa	gcagcugcag	aucgacuaca	2760
agggcaacga	gaucaugacc	gccgcugccu	cucagggccu	gacaagaaaa	ggcguguacg	2820
ccgugcggua	caaagugaac	gagaaccuc	uguacgcccc	uaccagcgag	caugugaaug	2880
ugcugcugac	ccggaccgag	gaccggaucc	uuuggaaaac	acuggccggc	gaucccugga	2940
ucaagacccu	gacagccaag	uaucgggca	acuucaccgc	caccaucgag	gaauggcagg	3000
ccgagcacga	ugccaucaug	cggcacaucc	uggaaagacc	cgauccuacc	gacguguucc	3060
agaacaaggc	caacgugugc	ugggccaaaag	cucuggugcc	ugugcugaaa	accgccggca	3120
ucgauaugac	caccgagcag	uggaacaccg	uggacuacuu	cgagacagac	aaggcccaca	3180
gcgccgagau	cgugcugaau	cagcugugcg	ugcgguuuuu	cggccuggau	cuggauagcg	3240
gccuguuuc	ugcuccuacc	gugccucuga	gcauccggaa	caaccacugg	gacaacagcc	3300
ccucuccuaa	uauguacggc	cugaacaaaag	aagucgugcg	gcagcugagc	agaagauacc	3360
cacagcugcc	uagagccgug	gccacaggca	gaguguacga	caugaauacc	ggcacacugc	3420
ggaacuacga	ccccagaauc	aaucuggugc	ccgugaacag	aaggcugccc	cacgcucugg	3480
uucugcacca	caaugagcac	ccucagagcg	acuucagcag	cuucgugucc	aagcugaagg	3540
gcagaaccgu	gcugguugug	ggcgagaagc	ugucugugcc	uggcaaugug	guggacuggc	3600
ugagcgauag	acccgaggcc	accuuuagag	ccagacugga	ccuuggaauc	ccuggcgacg	3660
ugcccaaaua	cgacaucauc	uucgugaacg	ugcggacgcc	cuacaaguac	caccacuacc	3720
agcagugcga	ggaccacgcc	aucaagcuga	gcaugcugac	caagaaggcc	ugccugcacc	3780
ugaauccugg	cggcaccugu	gugucuaucg	gcuacggcuu	ugccgacaga	gccagcgagu	3840
cuaucaucgg	cgccauugcc	agacaguuca	aguucagcag	aguugcaag	cccaagagca	3900
gccuggaaga	gacagaggug	cuguucgugu	ucaucggcuu	ugaccggaag	gcccguacgc	3960
acaaccccuu	caagcugagc	agcaccucga	ccaacaucua	caccggcagc	agacugcacg	4020
aagccggau	ugcaccucua	uaucaugugg	ugcgagggga	uaaugccaca	gccacagaag	4080
gagugauuau	aaauugcugc	aacagcaaag	gacaaccugg	cggaggggug	ugcgagcgcg	4140
uguauaagaa	auuuccggaa	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaaagg	ugcagcuaaa	cauaucuuuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320

acgauaaca	uuacaaguc	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuuc	ucauauuugg	aagggaccaa	guuucaccag	gcgccaagg	4680
auauagcaga	aauuauugcc	auguggcccc	uugcaacgga	ggccaugag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaauacac	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgauu	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggauag	accaggacua	gaacgccuga	gccgaucac	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucac	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggagggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucgccu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uagggugauc	acuagagagg	agcucgaggc	gcuuacccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauagguga	5640
uuacaagaga	ggaguuugag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaaaucc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcgu	uauuugaagg	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuaug	ugaaccguc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaagu	acgccugcaa	caacgaguac	ugggagacau	6420
ucaagagaa	ccccauccgg	cugaccgagg	aaaacguggu	caacuacauc	accaagcuga	6480
agggccccaa	agccgccgcu	cuguuugcca	agacacacaa	ccugaacaug	cugcaggaca	6540
uccccaugga	cagauucgug	auggaccuga	agcgggacgu	gaaagugacc	ccuggcacca	6600
agcacaccga	ggaacggccu	aaggugcaag	ugauccaggc	cgcugauccu	cuggccacag	6660

ccuauucugug	uggcauccac	agagaacucg	ugcggagacu	gaaugccgug	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuuga	cgcuauuaua	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	cugaugauuc	uggaagauuc	cggaguggac	gccgagcugc	6900
ugacacugau	ugaagccgcc	uuuggcgaga	ucagcagcau	ccaucugccu	accaagacca	6960
aguucaaguu	cggcgccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguauucgca	agcagagugu	ugagagaacg	gcuaacaggc	agcccuugug	7080
cagcauucau	uggagaugac	aaauaucguga	aaggagucua	aucggacaaa	uuauuggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	cccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aacaugauga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaggua	ugaaaccgua	ggaacuucca	7440
ucauaguau	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucucac	ggcuuaccug	aauggacuac	gacauagucu	aguccgcaa	7560
g						7561

<210> 31

<211> 8428

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C13 nsP1-4_GFP ORF and SGP RNA

<400> 31

augggcggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccauuccuca	gagcuuugca	gcggagcuuc	ccgcaguuu	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgauc	180
uggcuucaaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaaguauca	uuguauucgu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuua	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaaau	gacaagaaaa	ugaaggagcu	cgccgcccug	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaagg	aguuaagag	gccuacugga	uaggcuuuga	caccaccccu	uuuauuuua	600
agaacuuggc	uggagcauau	ccaucuuacu	cuaccaacug	ggccgacgaa	accguguuaa	660
cggcucguua	cauaggccua	ugcagcucug	acguuauugga	gcggucacgu	agagggaugu	720
ccaauucuu	aaagaaguau	uugaaaccuu	ccaacaauuu	ucuauucucu	guuggcucga	780
ccaucuuacc	cgagaagagg	gacuuacuga	ggagcuggca	ccugccgucu	guuuuucacu	840
uacguggcaa	gcaaaaauac	acaugucgg	gugagacuau	aguuaugu	gacggguacg	900
ucguuuuuag	aaauagcuau	aguccaggcc	uguauuggaa	gccuucaggc	uauugcucua	960

cgaugcaccg	cgagggauuc	uugugcugca	aagugacaga	cacauugaac	ggggagaggg	1020
ucucuuuucc	cgugugcacg	uaugugccag	cuacauugug	ugaccaaauug	acuggc auac	1080
uggcaacaga	ugucagugcg	gacgacgcgc	aaaaacugcu	gguugggcuc	aaccagcgua	1140
uagucgucaa	cggucgcacc	cagagaaaca	ccaauacc au	gaaaaauuac	cuuugccccg	1200
uaguggccca	ggcauuugcu	agguggggcaa	aggaauauaa	ggaagaucaa	gaagaugaaa	1260
ggccacuagg	acuacgagau	agacaguua g	ucauggggug	uuguugggcu	uuuagaaggc	1320
acaagauaac	aucuauuuau	aagcgccccg	auacccaaac	caucaucaa	gugaacagcg	1380
auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagau c	gggcugagaa	1440
caagaaucag	gaaaauguu a	gaggagcaca	aggagccguc	accucucauu	accgccgagg	1500
acguacaaga	agcuaauguc	gcagccgaug	aggcuaa gga	ggugcgugaa	gccgaggagu	1560
ugcgcgcgagc	ucuaccaccu	uuggcgagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuua	cgcugugcuu	ucuccgcagg	1740
cuguacucaa	gagugaaaaa	uuaucuugca	uccaccucuc	cgcugaacaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuaug	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgugaa	cacugaugaa	gaauauuaca	aaacuguc aa	gcccagcgag	cacgacggcg	2040
aauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaa uu	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauca g	2220
gcaagucugg	caucauuaaa	agcgcgauca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaa uu	auaaggggacg	ucaagaaa au	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau g	gaugcaaaca	ccccguagag	accugua u a	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagugcgg	uuuuuuu aac	augaugugcc	2520
ugaaauguca	uuuu aaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagaa	2640
cgacgaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aauaaugacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uacccuggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggacc cuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaaagg	cuuuaugucc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuau c g	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacac	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300

cgucgccuaa	cauguacggg	cugaauaaag	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuaua	caugaacacu	gguacacugc	3420
gcaauuauga	uccgcgcuaa	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuuucu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaaua	ugacauaaau	uuuguuaaau	ugaggacccc	auauaaaauac	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaaucccgg	cggaaaccugu	gucagcauag	guuaugguua	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuua	caagcuuuca	ucaaccuuga	ccaacauua	uacagguucc	agacuccacg	4020
aagccggau	ugcaccuua	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuau	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaaca	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaagauucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgau	caaaacuuc	ucauauuugg	aagggaacaa	guuucaccag	gcggccaagg	4680
auauagcaga	aaauaaugcc	auguggcccc	uugcaacgga	ggccaauag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaauacau	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgau	auucauccaa	ggaaguauu	cguaggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	caggaugag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcuauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacaccu	ggagggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucggcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uaggguaguc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauagguga	5640

uuacaagaga	ggaguuugag	gcuuucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacauuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaaaucucc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uucugcaagg	ccuagggcga	uauuugaagg	cagaaggaaa	aguggagugc	uaccgaaccc	6000
ugcauccugu	uccuuugua	ucaucuagug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaag	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gucccagaac	guccuggcag	6300
cugccacaaa	aagaaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaaau	augcguguaa	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggaccaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaaugga	cagguuugua	auggacuuua	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuaaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuaaguu	cggcgccaug	augaaguccg	gcauguuucu	gaccuguuuc	gugaacaccg	7020
ugaucaacau	cgugaucgcc	agccgggugc	ugagagagag	acugacagga	ucuccuugcg	7080
ccgccuucan	cggcgacgac	aaauucguga	agggcgugaa	guccgacaag	cugauggccg	7140
auagaugcgc	caccuggcug	aacauugaag	ugaagaucan	cgacgccguc	gugggcgaga	7200
aggcccccua	uuuuugcggc	ggcuucaucc	ugugcgacag	cgugacaggc	acagccugca	7260
gaguugccga	uccucugaag	cggcuguuca	agcugggaaa	accucuggcc	gccgacgacg	7320
agcacgacga	cgauagacgu	agagcccugc	acgaggaauc	caccagaugg	aacagagugg	7380
gcauccugag	cgagcugugc	aaggccgugg	aaagcagaua	cgagacagug	ggcaccagca	7440
ucauugugau	ggcaaugacc	acacuggcca	gcagcgugaa	aagcuucagc	uaccuaaggg	7500
gcgccccuau	cacacuguac	ggcuucggcg	aaggcagagg	cagccuucug	acauguggcg	7560
acuggaaga	gaaccccgga	ccugugucua	agggcgaaag	acuguuuacc	ggcguggugc	7620
ccauccuggu	ggaacuggau	ggggauguga	acggccacaa	guucagcguu	agcggagaag	7680
gcgaaggcga	cgccacauac	ggaaagcuga	cccugaaguu	caucugcacc	accggcaagc	7740
ugccugugcc	auggccuaca	cuggucacca	cacugacaua	cggcgugcag	ugcuucagca	7800
gauaccccga	ccauaugaag	cagcacgacu	ucuucaagag	cgccaugccu	gagggcuacg	7860
ugcaagagcg	gaccaucuu	uuuaaggacg	acggcaacua	caagaccagg	gccgaaguga	7920
aguucgaggg	cgacaccucg	gucaaccgga	ucgagcugaa	gggcaucgac	uucaaagagg	7980

acggcaauau ccugggccac aagcucgagu acaacuacaa cagccacaac guguacauca	8040
uggccgacaa gcagaaaaac ggcaucaaag ugaacuucac gaucggcac acaucgagg	8100
acggcucugu gcagcuggcc gaucacuacc agcagaacac acccaucgga gauggccug	8160
ugcugcugcc cgauaaccac uaccugagca cccagagcaa gcugagcaag gacccaacg	8220
agaagcggga ccacauggug cugcuggaau uugugacagc cgccggauc acccucggca	8280
uggaugagcu guacaaguga acuucauca uaguuauggc caugacuacu cuagcuagca	8340
guguuaaauc auucagcuac cugagagggg ccccuauaac ucucuacggc uaaccugaau	8400
ggacuacgac auagucuagu ccgccaag	8428
<210> 32	
<211> 8353	
<212> RNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic construct, nsP1-4_3Cpro and SGP_VEEVrep-nsP4[3C]	
<400> 32	
augggcggcg caugagagaa gcccagacca auuaccuacc caaauggag aaaguucacg	60
uugacaucga ggaagacagc ccauuccuca gagcuuugca gcggagcuuc ccgaguuug	120
agguagaagc caagcagguc acugauaaug accaugcuua ugccagagcg uuuucgcauc	180
uggcuuacaa acugaucgaa acggaggugg acccauccga cacgauccuu gacauuggaa	240
gugcgcccg cgcagaaug uauucuaagc acaaguauca uuguaucugu ccgaugagau	300
gugcggaaga uccggacaga uuguauaagu augcaacuaa gcugaagaaa aacuguaagg	360
aaauaacuga uaaggaaug gacaagaaaa ugaaggagcu cgccgccguc augagcgacc	420
cugaccugga aacugagacu augugccucc acgacgacga gucgugucgc uacgaagggc	480
aagucgcugu uuaccaggau guauacgcgg uugacggacc gacaagucuc uaucaccaag	540
ccaauaaggg aguuagaguc gccuacugga uaggcuuuga caccacccu uuuauuuua	600
agaacuuggc uggagcauau ccaucauacu cuaccaacug ggccgacgaa accguguuaa	660
cggcucguua cauaggccua ugcagcucug acguuaugga gcggucacgu agagggaugu	720
ccaauucuuag aaagaaguau uugaaaccu ccaacaangu ucuauucucu guuggcucga	780
ccaucuaacca cgagaagagg gacuuacuga ggagcuggca ccugccgucu guauuucacu	840
uacguggcaa gcaaaaauac acaugucggu gugagacuau aguuaguugc gacggguacg	900
ucguuaaaaag aauagcuau aguccaggcc uguauaggga gccuucaggc uaugcugcua	960
cgaugcaccg cgagggauc uugugcugca aagugacaga cacauugaac ggggagaggg	1020
ucucuuuuucc cgugugcacg uaugugccag cuacauugug ugaccaaau acuggcauac	1080
uggcaacaga ugucagugcg gacgacgcgc aaaaacugcu gguugggcuc aaccagcgua	1140
uagucgucaa cggucgcacc cagagaaaca ccauaccu gaaaaauac cuuuugcccg	1200
uaguggccca ggcauuugcu agguaggcaa aggaauuaa ggaagauca gaagaugaa	1260
ggccacuagg acuacgagau agacaguua ucauggggug uuguugggc uuuagaaggc	1320
acaagauaac aucuauuuau aagcgcccgg auacccaaac caucauaaa gugaacagcg	1380
auuuccacuc auucgugcug cccaggauag gcaguaacac auuggagauc gggcugagaa	1440

caagaaucag	gaaaauguua	gaggagcaca	aggagccguc	accucucuu	accgccgagg	1500
acguacaaga	agcuaagugc	gcagccgaug	aggcuaagga	ggugcgugaa	gccgaggagu	1560
ugcgcgcagc	ucuaccaccu	uuggcagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuu	cgcugugcuu	ucuccgcagg	1740
cuguacucua	gagugaaaaa	uuaucuugca	uccaccucuu	cgcugaacaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuaug	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccgagagg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgcugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaagggacg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaca	ccccguagag	accuguaua	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagucgg	uuuuuuuac	augaugugcc	2520
ugaaagugca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagaaa	2640
cgacgaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaugacg	gcagcugccu	cuaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uaccucggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggaccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuuagugcc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuauugc	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacacu	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaa	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuauga	caugaacacu	gguacacugc	3420
gcauuuaua	uccgcgcuaa	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaauagaacac	ccacagagug	acuuuucuu	auucgucagc	aaaugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaua	ugacauaaua	uuuguuuau	ugaggacccc	auauaaaauac	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780

ugaaucccg	cggaaccugu	gucagcauag	guuaugguua	cgugacag	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuua	caagcuuua	ucaaccuuga	ccaacauua	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucu	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuau	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaaca	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aaugggacaa	guuucaccag	gchgccaagg	4680
auauagcaga	aauuauugcc	auguggcccc	uugcaacgga	ggccaauag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaaucacu	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgauu	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggaugag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuga	uguggacagu	uuauccauac	uugacacccu	ggagggagcu	agcgugacca	5340
gchgggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uagggugauc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaauucc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uuucgcaagg	ccuagggcua	uauuuagaag	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuuguau	ucaucuagug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaa	agaacuucc	gacuguggcu	ucuuacugua	6120

uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaaaugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucagaaaau	augcguguaa	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggacccaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaauugga	cagguuugua	auggacuuaa	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauuacauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuuaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuaaguu	cggcgccaug	augaaguccg	gcauguuucu	gaccucguuc	gugaacaccg	7020
ugaucaacau	cgugaucgcc	agccgggugc	ugagagagag	acugacagga	ucuccuugcg	7080
ccgccuucau	cggcgacgac	aaauucguga	agggcgugaa	guccgacaag	cugauggccg	7140
auagaugcgc	caccuggcug	aacauuggaag	ugaagaucan	cgacgccguc	gugggcgaga	7200
aggcccccua	uuuuugcggc	ggcuucaucc	ugugcgacag	cgugacaggc	acagccugca	7260
gaguugccga	uccucugaag	cggcuguuca	agcugggaaa	accucuggcc	gccgacgacg	7320
agcacgacga	cgauagacgu	agagcccugc	acgaggaau	caccagaugg	aacagagugg	7380
gcauccugag	cgagcugugc	aaggccgugg	aaagcagaua	cgagacagug	ggcaccagca	7440
ucauugugau	ggcaaugacc	acacuggcca	gcagcgugaa	aagcuucagc	uaccuaaggg	7500
gcgccccuau	cacacuguac	ggcucuggcg	aaggcagagg	cagccuucug	acaugggcg	7560
acguggaaga	gaaccccgga	ccuaguggug	ccccaccgac	cgacuugcaa	aagaugguca	7620
ugggcaacac	aaagccuguu	gagcucaucc	uugacgggaa	gacaguagcc	aucuguugug	7680
cuacuggagu	guuuggcacu	gcuuaccucg	ugccucguca	ucuuuucgca	gagaaguaug	7740
acaagaucan	gcuggauggc	agagccauga	cagacaguga	cuacagagug	uuugaguuu	7800
agauuaaagu	aaaaggacag	gacaugcucu	cagacgcugc	gcucauggug	cuccaccgug	7860
ggaaccgcgu	gagagauauc	acgaaacacu	uucgugauac	agcaagaau	aagaaaggca	7920
cccccgucgu	cggugugguc	aacaacgccg	acguugggag	acugauuuuc	ucuggugagg	7980
cccucaccua	caaggauauu	guagugugca	uggacggaga	caccaugccu	agccucuug	8040
ccuacaaaagc	cgccaccaag	gcaggcuacu	guggaggagc	cguucucgcc	aaggacgggg	8100
ccgacacuuu	caucgucggc	acucacuccg	caggaggcaa	uggaguugga	uacugcucau	8160
gcguuuccaa	guccaugcuu	cucagaauga	aggcacacgu	ugaccucgaa	ccacaacacg	8220
aguagacuu	caucanaguu	auggccauga	cuacucuagc	uagcaguguu	aaaucauua	8280
gcuaccugag	aggggccccu	auaacucucu	acggcuuacc	ugaauggacu	acgacauagu	8340
cuaguccgcc	aag					8353

<210> 33

- <211> 2322
 <212> RNA
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, 01 Manisa P1 pr _ RNA coding sequence (SG ORF)
 <400> 33

```

auggggagcug gacagagcuc cccugcuacc ggcagccaga aucaguccgg caacaccggc      60
uccaucauca acaacuacua caugcagcag uaccagaaca gcauggacac ccagcugggc      120
gacaacgcua ccagcggcgg cuccaacgag ggaagcaccg acaccaccuc caccacacacc      180
accaacacccc agaacaacga cugguucucc aagcuggcca gcagcgccuu cagcggccug      240
uucggagccc ugcuggccga caagaagacc gaggagacaa cccugcugga ggaccggauc      300
cugaccacca gaaacggaca cacuaccucc accaccaga gcuccguggg agugaccuac      360
ggauacgcua ccgcugagga cuucgugucc ggaccaaaca ccuccggccu ggagacaagg      420
guggcucagg cugagagauu cuucaagacc caccuguucg acugggugac cagcgaccca      480
uucggaagau gccaccugcu ggagcugccc accgaccaca agggcgugua cggcagccug      540
accgacuccu acgccuacau gagaaacgga ugggacgugg aggugaccgc ugugggaaac      600
caguucaacg gcggaugccu gcugguggc uaggugcccg agcugugcuc cauccagaag      660
agggagcugu accagcugac ccuguucccc caccaguuca ucaaccccag aaccaacaug      720
accgcucaca ucaccgugcc cuucguggga gugaaccggu acgaccagua caaggugcac      780
aagcccugga cccugguggu caugguggug gccccacuga ccgugaacag cgagggagcc      840
ccacagauca agguhuacgc caacaucgcu ccaaccaacg ugcacguggc uggagaguuc      900
ccuccaaag agggcaucuu cccuguggcu ugcagcgacg gauacggcgg ccuggugacc      960
accgacccaa agaccgcuga cccagccuac ggcaaggugu ucaaccacc ccggaacaug      1020
cugccuggac gcuuaccaa cuuccuggac guggcugagg ccugcccaac cuuccugcac      1080
uucgagggcg acgugcccua cgugaccacc aagaccgacu ccgacagagu gcuggcccag      1140
uucgaccugu cccuggccgc caagcacaug agcaacaccu uccuggccgg ccuggcccag      1200
uacuacaccc aguacagcgg caccaucaac cugcacuua uguuacccgg accaaccgac      1260
gcuaaggcua gguacaugau cgcuuacgcc ccaccggaa uggagcccc aaagaccccu      1320
gaggcugcug cucacugcau ccacgccgag ugggacaccg gccugaacag caaguucacc      1380
uucuccaucc ccuaccugag cgcugcugac uacgcuuaca ccgccuccga caccgcugag      1440
acaaccaacg ugcagggcug ggugugccug uuccagauc cccacggcaa ggcugacggc      1500
gacgcucugg uggugcuggc cagcgccggc aaggacuucg agcugaggcu gcccguggac      1560
gcuagaaccc agaccaccag cgcuggagag uccgcugacc cagugaccgc caccguggag      1620
aacuacggcg gcgagacaca ggugcagcgg aggcagcaca ccgacguguc cuucauccug      1680
gacagauucg ugaaggugac ccccaaggac cagaucacg ugcuggaccu gaugcagacc      1740
ccugcucaca cccugguggg agcccugcug cggaccgcca ccuacuacuu cgccgaccug      1800
gagguggccg ugaagcacga gggcaaccug accugggugc ccaacggagc uccugaggcu      1860
gcccuggaca acaccaccaa ccccaccgcu uaccauaagg ccccacugac ccggcuggcc      1920

```

cugcccuaca	ccgccccaca	ccgcgugcug	gccaccgugu	acaacggcaa	cugcaaguac	1980
ggcgacggca	ccguggccaa	cgugcggggc	gaccugcagg	ugcuggccca	gaaggcugcu	2040
agagcccugc	ccaccagcuu	caacuacggc	gccaucaagg	ccaccagagu	gaccgagcug	2100
cuguacagga	ugaagagagc	cgagacauac	ugccccaggc	cccugcuggc	cauccacca	2160
gaccaggcca	gacacaagca	gaagaucgug	gccccguga	agcagcugcu	gaacuucgac	2220
cugcugaagc	uggccggcga	cguggagucc	aaccucggac	ccuucuucuu	cagcgacgug	2280
cgcagcaacu	ucuccaagcu	gguggagaca	aucaaccagu	ga		2322
<210>	34					
<211>	8848					
<212>	RNA					
<213>	Artificial Sequence					
<220>						
<223>	Synthetic construct, nsP1-4-STING pr ORF, and SGP_VEEVrep-nsP4[STING]					
<400>	34					
augggcgggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccaauccuca	gagcuuugca	gcggagcuuc	ccgcaguuu	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgcauc	180
uggcuucaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaagauca	uuguaucugu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguaaaugu	augcaacuua	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaaau	gacaagaaaa	ugaaggagcu	cgcccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaagg	aguuagaguc	gccuacugga	uaggcuuuga	caccaccccu	uuuauuuua	600
agaacuuggc	uggagcauau	ccaucuuacu	cuaccaacug	ggccgacgaa	accguguuaa	660
cggcucguua	cauaggccua	ugcagcucug	acguuaugga	gcggucacgu	agagggau	720
ccaauucuuag	aaagaaguau	uugaaaccu	ccaacaugu	ucuauucucu	guuggcucga	780
ccaucuaacca	cgagaagagg	gacuuacuga	ggagcuggca	ccugccgucu	guauuucacu	840
uacguggcaa	gcaaaaauac	acaugucggg	gugagacuau	aguuaugu	gacggguacg	900
ucguuuuuuag	aaauagcuac	aguccaggcc	uguaugggaa	gccuucaggc	uauugcucua	960
cgaugcaccg	cgagggaau	uugugcugca	aagugacaga	cacauugaac	ggggagaggg	1020
ucucuuuuucc	cgugugcacg	uaugugccag	cuacauugug	ugaccaaau	acuggcauac	1080
uggcaacaga	ugucagugcg	gacgacgcgc	aaaaacugcu	ggauugggcuc	aaccagcgua	1140
uagucgucaa	cggucgcacc	cagagaaaca	ccaauaccu	gaaaaauac	uuuuugccc	1200
uaguggccca	ggcauuugcu	aggugggcaa	aggaauuaa	ggaagaucaa	gaagaugaaa	1260
ggccacuagg	acuacgagau	agacaguua	ucauggggug	uuuugggcu	uuuagaaggc	1320
acaagauaac	aucuauuuau	aagcgcccgg	auacccaaac	caucaucaa	gugaacagcg	1380
auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagauc	gggcugagaa	1440

caagaaucag	gaaaauguua	gaggagcaca	aggagccguc	accucucauu	accgccgagg	1500
acguacaaga	agcuaagugc	gcagccgaug	aggcuaagga	ggugcgugaa	gccgaggagu	1560
ugcgcgcagc	ucuaccaccu	uuggcagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuuu	cgcugugcuu	ucuccgcagg	1740
cuguacucua	gagugaaaaa	uuaucuugca	uccaccucuc	cgcugaacaa	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuaug	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgcugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaagggacg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaca	ccccguagag	accuguaua	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagugcgg	uuuuuuuac	augaugugcc	2520
ugaaagugca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagaaa	2640
cgacgaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaugacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uaccucggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggaccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuuagugcc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuauugc	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuaauuu	ugcaccacac	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaa	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuauga	caugaacacu	gguacacugc	3420
gcaauuauga	uccgcgcava	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuucuu	auucgucagc	aaaugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaauua	ugacauaaua	uuuguuuau	ugaggacccc	auauaaaau	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780

ugaaucccg	cggaaccugu	gucagcauag	guuaugguua	cgugacag	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuua	caagcuuua	ucaaccuuga	ccaacauua	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucu	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuau	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaaca	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aaugggacaa	guuucaccag	gcggccaagg	4680
auauagcaga	aauuauugcc	auguggcccc	uugcaacgga	ggccaauag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaaucacu	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgauu	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggaugag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuga	uguggacagu	uuauccauac	uugacacccu	ggagggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uagggugauc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaauucc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uuucgcaagg	ccuagggcua	uauuuagaag	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuuguau	ucaucuagug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaa	agaacuucc	gacuguggcu	ucuuacugua	6120

uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaauugc	uucaagaaa	augcguguaa	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggaccaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaauugga	cagguuugua	auggacuuaa	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauuacauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuuaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuaaguu	cggcgccaug	augaaguccg	gcauguuucu	gaccucguuc	gugaacaccg	7020
ugaucaacau	cgugaucgcc	agccgggugc	ugagagagag	acugacagga	ucuccuugcg	7080
ccgccuucau	cggcgacgac	aaauaucguga	agggcgugaa	guccgacaag	cugauggccg	7140
auagaugcgc	caccuggcug	aacauuggaag	ugaagaucan	cgacgccguc	gugggcgaga	7200
aggcccccua	uuuuugcggc	ggcuucaucc	ugugcgacag	cgugacaggc	acagccugca	7260
gaguugccga	uccucugaag	cggcuguuca	agcugggaaa	accucuggcc	gccgacgacg	7320
agcacgacga	cgauagacgu	agagcccugc	acgaggaau	caccagaugg	aacagagugg	7380
gcauccugag	cgagcugugc	aaggccgugg	aaagcagaua	cgagacagug	ggcaccagca	7440
ucauugugau	ggcaaugacc	acacuggcca	gcagcgugaa	aagcuucagc	uaccuaaggg	7500
gcgccccuau	cacacuguac	ggcucuggcg	aaggcagagg	cagccuucug	acauguggcg	7560
acguggaaga	gaaccccgga	ccuccucaca	gcucucugca	cccuagcauc	ccuuguccua	7620
gaggacacgg	cgcccagaaa	gcugcucugg	uucugcuguc	ugccugccug	guuacacugu	7680
ggggacuggg	agagccuccu	gagcacacac	ugagauaccu	ggugcugcac	cuggcuucuc	7740
ugcagcuggg	acugcugcug	aacggcgugu	guucucuggc	cgaggaacug	agacacaucc	7800
acagcagaua	caggggcagc	uacuggcgga	cagucagagc	uugucugggc	ugcccucuaa	7860
gaagaggcgc	ucugcugcug	cuguccaucu	acuucuacua	cagccugccu	aacgcugugg	7920
gcccuccuuu	cacauuggaug	cuggcucugc	uggggccuguc	ucaggcucug	aaauuccugc	7980
uggggcgugaa	aggccuggc	ccugcugaaa	ucucugccgu	gugcgagaag	ggcaacuucu	8040
cuguggcuca	cggccuggcc	ugguccuacu	acaucggaua	ccugagacug	auccugccug	8100
agcugcaggc	cagaaucaga	accuacaacc	agcacuacaa	caaccugcug	agaggcgug	8160
ugucccagag	acuguacau	cugcugccuc	uggacugugg	cgugccagac	aaccugucua	8220
uggccgaucc	uaacauccgg	uuucuggaca	agcugccuca	gcagacaggc	gaucacgccg	8280
gcaucaagga	cagaguguac	agcaacagca	ucuacgagcu	gcuggaaaac	ggccagcgag	8340
ccggaacaug	cgugcuggaa	uaugccacac	cucugcagac	ccuguucgcc	augagccagu	8400
acucucaggc	cggcuucagc	cgcgaggaua	ugcuggaaca	ggccaagcug	uucugcagga	8460

cccucgagga cauccuggcu gaugcuccag agagccagaa caacugcaga cugaucgccu	8520
accaagagcc ugccgacgac agcagcuua gccugagcca agaggugcug agacaccuga	8580
gacaagaaga gaaagaggaa gucaccgucg gcagccugaa aaccagcgcc gugccuagca	8640
ccagcaccuau gucucaagaa cccgaacugc ugaucuccgg cauggaaaag cccugccuc	8700
ugagaaccga cuucagcuga acuucauca uaguuauggc caugacuacu cuagcuagca	8760
guguuaaauc auucagcuac cugagagggg ccccuauaac ucucuacggc uaaccugaau	8820
ggacuacgac auagucuagu ccgccaag	8848
<210> 35	
<211> 7561	
<212> RNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic construct, C12_nsP1-4 ORF, and SGP	
<400> 35	
augggcggcg caugagagaa gcccagacca auuaccuacc caaaauggag aaaguucacg	60
uugacaucga ggaagacagc ccauuccuca gagcuuugca gcggagcuuc ccgcaguug	120
agguagaagc caagcagguc acugauaaug accaugcuua ugccagagcg uuuucgcauc	180
uggcuucaaa acugaucgaa acggaggugg acccauccga cacgauccuu gacauuggaa	240
gugcgcccg ccgcagaaug uauucuaagc acaaguauca uuguaucugu ccgaugagau	300
gugcggaaga uccggacaga uuguauaagu augcaacuaa gcugaagaaa aacuguaagg	360
aaauaacuga uaaggaauug gacaagaaaa ugaaggagcu cgccgcccugc augagcgacc	420
cugaccugga aacugagacu augugccucc acgacgacga gucgugucgc uacgaagggc	480
aagucgcugu uuaccaggau guauacgcgg uugacggacc gacaagucuc uaucaccaag	540
ccaauaaggg aguuagaguc gccuacugga uaggcuuuga caccaccccu uuuauuuua	600
agaacuuggc uggagcauau ccaucauacu cuaccaacug ggccgacgaa accguguuaa	660
cggccagaaa caucggccug ugcagcagcg acgugaugga acggucuaga cggggcauga	720
gcauccugcg gaagaaguac cugaagccua gcaacaacgu gcuguucagc gugggcagca	780
ccaucuacca cgagaagagg gaccugcugc ggagcuggca ucugccuucc guguuuacc	840
ugagaggcaa gcagaacuac accuguagau gcgagacaau cguguccugc gacggcuacg	900
uggucaagcg gaucgccauu ucuccuggcc uguacggcaa gccuucuggc uaugccgcca	960
ccaugcacag agaaggcuuu cuguguugca aagugaccga cacacugaac ggcgagcggg	1020
uguccuuucc uguguguacc uaugugcccg ccacacugug cgaucagaug acaggcauuc	1080
uggccaccga cgugucagcc gacgaugccc agaaacugcu cgugggccug aaccagagaa	1140
ucguggucaa cggcagaacc cagcggaca ccaacaccau gaagaacuac cugcugccug	1200
ugguggccca ggccuuugcc agaugggcca aagaguacaa agaggaucaa gaggacgagc	1260
ggccccuggg ccugagagau agacaacugg ucaugggcug cugcugggcc uucagaaggc	1320
acaagaucac cagcaucuac aagcggcccc acacacagac caucauaaa gugaacagcg	1380
acuuccacag cuucgugcug ccucggaucg gcagcaacac acuggaaauc ggccugcgga	1440
cccggaucag aaagaugcug gaagaacaca aagagcccuc uccacugauc accgccgagg	1500

augugcaaga	ggccaaaugu	gccgccgacg	aggcuaaaga	agugcgcgaa	gccgaggaac	1560
ugagagccgc	acuuccuccu	cuggccgccg	auguugaaga	accacucug	gaagccgacg	1620
uggaccugau	gcuucaagaa	gccggcgug	gcagcgugga	aacaccuaga	ggacugauca	1680
aagucaccag	cuacgccggc	gaggacaaga	ucggauuuu	ugccgugcug	agcccucagg	1740
cugugcugaa	gucugagaag	cugagcugca	uucaccacu	ggccgagcaa	gugaucguga	1800
ucacacacag	cggccggaag	ggcagauaug	ccguggaacc	uuauacggc	aagguggugg	1860
ugccugaggg	acacgcuaau	ccagugcagg	acuucaggc	ccugagcgag	ucugccacca	1920
ucguguacaa	cgagcgcgag	uucgugaaca	gauaccugca	ccacauugcc	acacacggcg	1980
gagcccugaa	uaccgacgaa	gaguacuaca	agaccgugaa	gcccagcgag	cacgacggcg	2040
aguaccugua	cgacaucgac	agaaagcagu	gcgugaagaa	agagcugguc	accggcuugg	2100
gacugacagg	cgaacuggug	gauccuccau	uccacgaguu	ugccuacgag	agccugagaa	2160
ccagaccugc	cgcuccuuac	caggugccaa	caaucggagu	guauggcgug	ccaggcucug	2220
gcaagagcgg	caucauuuag	agcgccguga	ccaaaaagga	ccugguggug	uccgccaaga	2280
aagagaacug	cgccgagau	auccgggacg	ugaagaagau	gaagggccuc	gacgugaacg	2340
ccagaaccgu	ggauagcgug	cugcugaacg	gaugcaagca	ccccguggaa	accuguaca	2400
ucgaugaggc	cuucgccugc	caugccggaa	cacugagagc	acugaucgcc	aucaucagac	2460
ccaagaaagc	cgugcugugc	ggcgaccua	agcagugugg	cuucuuaau	augaugugcc	2520
ugaaggugca	cuucaaccac	gagaucugca	cccagguguu	ccacaagagc	aucagcagac	2580
ggugcaccaa	gagcgugacc	agcguggugu	cuaccuguu	cuacgacaag	aagaugcgga	2640
cgacaaaccc	caaagagacu	aagaucguca	ucgacaccac	cggcuccacc	aagccuaagc	2700
aggacgaucu	gauccugacc	ugcuucagag	gcugggucaa	gcagcugcag	aucgacuaca	2760
agggcaacga	gaucaugacc	gcugccgcuu	cucagggacu	gaccagaaaa	ggcguguacg	2820
ccgugcggua	caaagugaau	gagaaccuc	uguacgccc	uaccuccgag	caugugaau	2880
uccugcugac	caggaccgag	gaccggau	uguggaaaac	acuggcuggc	gacccugga	2940
ucaagacccu	gacagccaag	uauccggca	acuucaccgc	cacaucgag	gaauggcagg	3000
ccgagcacga	ugccauuau	cggcacaucc	uggaacggcc	cgauccuacc	gauguguucc	3060
agaacaaggc	caacgugugc	ugggcuaaag	cccuggugcc	agugcugaaa	accgccggca	3120
ucgauaugac	caccgagcag	uggaauaccg	uggacuacuu	cgagacagac	aaggcccacu	3180
cugccgagau	cgugcugaau	cagcugugcg	ugcgguuuu	cggccuggau	cuggauagcg	3240
gccuguuuag	cgcuccuacc	gugccucuga	gcauccggaa	caaccacugg	gacaacagcc	3300
ccucuccuaa	uauguacggc	cugaacaaag	aaguugugcg	gcagcugagc	agaagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuaua	caugaacacu	gguacacugc	3420
gcaauuauga	uccgcgcua	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaua	ugacauaaua	uuuguuaau	ugaggacccc	auauaaaau	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuu	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaaucccg	cggaaaccugu	gucagcauag	guuaugguu	cgcugacagg	gccagcgaaa	3840

gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuugua	ucauugggua	cgauccgaag	gcccguacgc	3960
acaauccuua	caagcuuuc	ucaaccuuga	ccaacauua	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccuc	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuau	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgau	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgaugg	caaaacuuc	ucauauuugg	aagggaacca	guuucaccag	gcggccaagg	4680
auauagcaga	auuuauugcc	auguggcccc	uugcaacgga	ggccaugag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaauacau	ggugugcaga	agauccaau	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgua	auucauccaa	ggaaguauc	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggauag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160
aagaagagga	uagcauaagu	uugcugucag	augggccgac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacaccu	ggagggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccaccccg	5520
caggcgugaa	uaggguag	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacauuu	uuucccgac	accggucaag	ggcauuuaca	acaaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaaaucc	acaccugcu	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uuucgcaagg	ccuagggc	uauuuagaag	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuaug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguugaaa	agaacuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180

cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaaa	augcguguaa	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggaccaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaaugga	cagguuugua	auggacuuaa	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cgagagcugg	uuaggagauu	aaaugcgguc	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuuga	cgcuaauuaa	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuaaugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggcu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aauuuaaaau	cggagccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguaaucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucau	uggagaugac	aaauaucguga	aaggagucua	aucggacaaa	uuaauggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuua	agcuuggcaa	accucuggca	gcagacgaug	7320
aacaugauga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaaggua	ugaaaccgua	ggaacuucca	7440
ucauaguuau	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucuac	ggcuaaccug	aauggacuac	gacauagucu	aguccgccaa	7560
g						7561

<210> 36

<211> 7561

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C13_nsP1-4 ORF, and SGP

<400> 36

augggcggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccaauccuca	gagcuuugca	gcggagcuuc	ccgcaguuug	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgcauc	180
uggcuucaaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaaguauca	uuguaucugu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuua	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480

aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaaggg	aguuaagaguc	gccuacugga	uaggcuuuga	caccaccccu	uuuauguuaa	600
agaacuuggc	uggagcauau	ccaucauacu	cuaccaacug	ggccgacgaa	accguguuaa	660
cggccagaaa	caucggccug	ugcagcagcg	acgugaugga	acggagcaga	cggggcauga	720
gcauccugcg	gaagaaguac	cugaagccca	gcaacaacgu	gcuguucagc	gugggcagca	780
ccaucuaacca	cgagaagagg	gaccugcugc	ggagcuggca	ccugcccagc	guguuccacc	840
ugagaggcaa	gcagaacuac	accugcagau	gcgagacaau	cgugagcugc	gacggcuacg	900
uggucaagcg	gaucgccaua	agccccggcc	uguacggcaa	gcccagcggc	uacgccgcca	960
ccaugcacag	agaaggcuuc	cugugcugca	aagugaccga	cacacugaac	ggcgagcggg	1020
ugagcuuccc	cgugugcacc	uacgugcccc	ccacacugug	cgaccagaug	acaggcauac	1080
uggccaccga	cgugagcgcc	gacgacgccc	agaaacugcu	cguggggccug	aaccagagaa	1140
ucguggucaa	cggcagaacc	cagcgggaaca	ccaacaccu	gaagaacuac	cugcugccag	1200
ugguggccca	ggccuucgcc	agaugggcca	aagaguacaa	agaggaccaa	gaggacgagc	1260
ggccccuggg	ccugagagac	agacaacugg	ucaugggcug	cugcuggggc	uucagaaggc	1320
acaagaucac	cagcaucuac	aagcggcccc	acacacagac	caucaucaa	gugaacagcg	1380
acuuccacag	cuucgugcug	ccacggaucg	gcagcaacac	acuggaaauc	ggccugcgga	1440
cccggaucag	aaagaugcug	gaagaacaca	aagagccag	cccacugauc	accgccgagg	1500
acgugcaaga	ggccaaaugc	gccgccgacg	aggcaaaaga	agugcgcgaa	gccgaggaac	1560
ugagagccgc	acucccccca	cuggccgccc	acguagaaga	acccaccug	gaagccgacg	1620
uggaccugau	gcuccaagaa	gccggcgag	gcagcgugga	aacaccgaga	ggacugauca	1680
aagucaccag	cuacgccggc	gaggacaaga	ucggaagcu	cgccgugcug	agcccccagg	1740
ccgugcugaa	aagcgagaag	cugagcugca	uccaccacu	ggccgagcaa	gugaucguga	1800
ucacacacag	cggccggaag	ggcagauacg	ccguggaacc	cuaccacggc	aagguggugg	1860
ugccagaggg	acacgcaauc	ccagugcagg	acuuccaggc	ccugagcgag	agcgccacca	1920
ucguguacaa	cgagcgcgag	uucgugaaca	gauaccugca	ccacauagcc	acacacggcg	1980
gagcccugaa	caccgacgaa	gaguacuaca	agaccgugaa	gcccagcgag	cacgacggcg	2040
aguaccugua	cgacaucgac	agaaagcagu	gcgugaagaa	agagcugguc	accggccuag	2100
gacugacagg	cgaacuggug	gacccacccu	uccacgagu	cgccuacgag	agccugagaa	2160
ccagaccagc	cgcacccuac	caggugccaa	caaucggagu	guacggcgug	ccaggcagcg	2220
gcaagagcgg	caucauaaaag	agcggcguga	ccaaaaagga	ccugguggug	agcgccaaga	2280
aagagaacug	cggcgagauc	auccgggacg	ugaagaagau	gaagggccuc	gacgugaacg	2340
ccagaaccgu	ggacagcgug	cugcugaacg	gaugcaagca	ccccguggaa	accuguaca	2400
ucgacgaggc	cuucgccugc	cacgccggaa	cacugagagc	acugaucgcc	aucaucagac	2460
ccaagaaaagc	cgugcugugc	ggcgacccca	agcagugcgg	cuucuucaac	augaugugcc	2520
ugaaggugca	cuucaaccac	gagaucugca	cccagguguu	ccacaagagc	aucagcagac	2580
ggugcaccaa	gagcgugacc	agcgugguga	gcaccuguu	cuacgacaag	aagaugcgga	2640
cgacaaaacc	caaagagacc	aagaucguca	ucgacaccac	cggcagcacc	aagccaaagc	2700
aggacgaccu	gauccugacc	ugcuucagag	gcugggucaa	gcagcugcag	aucgacuaca	2760
agggcaacga	gaucaugacc	gccgccgcaa	gccagggacu	gaccagaaaa	ggcguguacg	2820

ccgugcggua	caaagugaac	gagaaccac	uguacgcccc	caccagcgag	cacgugaacg	2880
uccugcugac	caggaccgag	gaccggaucg	uguggaaaac	acuggccggc	gacccugga	2940
ucaagacccu	gacagccaag	uaccccgga	acuucaccgc	cacaucgag	gaauggcagg	3000
ccgagcacga	cgccauaaug	cggcacaucc	uggaacggcc	cgaccaacc	gacguguucc	3060
agaacaaggc	caacgugugc	ugggcaaaag	cccuggugcc	agugcugaaa	accgccggca	3120
ucgacaugac	caccgagcag	uggaacaccg	uggacuacuu	cgagacagac	aaggcccaca	3180
gcgccgagau	cgugcugaac	cagcugugcg	ugcgguuuu	cggccuggac	cuggacagcg	3240
gccuguucag	cgcacccacc	gugccccuga	gcauccggaa	caaccacugg	gacaacagcc	3300
ccagcccaaa	cauguacggc	cugaacaaag	aagucgugcg	gcagcugagc	agaagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuaua	caugaacacu	gguacacugc	3420
gcauuuaua	uccgcgcuaa	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuuuuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaau	ugacauaaau	uuuguuaaag	ugaggacccc	auauaaaauac	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaauccggg	cggaaaccugu	gucagcauag	guuaugguua	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuua	caagcuuuca	ucaaccuuga	ccaacauuuu	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucu	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuau	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagcgc	4140
uguauaagaa	auucccgga	agcuucgauu	uacagccgau	cgaaguagga	aaagcgcgac	4200
uggucaaaagg	ugcagcuaaa	cauaucauuc	augccguagg	accaaacuuc	aacaaaguuu	4260
cggagguuga	aggugacaaa	caguuggcag	aggcuuauga	guccaucgcu	aagauuguca	4320
acgauaacia	uuacaaguca	guagcgauuc	cacuguuguc	caccggcauc	uuuuccggga	4380
acaaagaucg	acuaacccaa	ucauugaacc	auuugcugac	agcuuuagac	accacugaug	4440
cagauguagc	cauauacugc	agggacaaga	aaugggaaau	gacucucaag	gaagcagugg	4500
cuaggagaga	agcaguggag	gagauaugca	uauccgacga	cucuucagug	acagaaccug	4560
augcagagcu	ggugagggug	cauccgaaga	guucuuuggc	uggaaggaag	ggcuacagca	4620
caagcgauug	caaaacuuc	ucauauuugg	aagggaccaa	guuucaccag	gcggccaagg	4680
auauagcaga	aauuuaugcc	auguggcccc	uugcaacgga	ggccaaugag	cagguaugca	4740
uguauauccu	cggagaaaagc	augagcagua	uuaggucgaa	augccccguc	gaagagucgg	4800
aagccuccac	accaccuagc	acgcugccuu	gcuugugcau	ccaugccaug	acuccagaaa	4860
gaguacagcg	ccuaaaagcc	ucacguccag	aacaaauuac	ugugugcuca	uccuuuccau	4920
ugccgaagua	uagaauacau	ggugugcaga	agauccaaug	cucccagccu	auauuguucu	4980
caccgaaagu	gccugcgua	auucauccaa	ggaaguaucu	cguggaaaca	ccaccgguag	5040
acgagacucc	ggagccaucg	gcagagaacc	aauccacaga	ggggacaccu	gaacaaccac	5100
cacuuauaac	cgaggauag	accaggacua	gaacgccuga	gccgaucauc	aucgaagagg	5160

aagaagagga	uagcauaagu	uugcugucag	auggcccgcac	ccaccaggug	cugcaagucg	5220
aggcagacau	ucacgggccg	cccucuguau	cuagcucauc	cugguccauu	ccucaugcau	5280
ccgacuuuga	uguggacagu	uuauccauac	uugacacccu	ggaggagcu	agcgugacca	5340
gcggggcaac	gucagccgag	acuaacucuu	acuucgcaa	gaguauggag	uuucuggcgc	5400
gaccggugcc	ugcgccucga	acaguauuca	ggaaccucc	acaucccgcu	ccgcgcacaa	5460
gaacaccguc	acuugcacc	agcagggccu	gcucgagaac	cagccuaguu	uccacccgc	5520
caggcgugaa	uagggugauc	acuagagagg	agcucgaggc	gcuuaccccg	ucacgcacuc	5580
cuagcagguc	ggucucgaga	accagccugg	ucuccaacc	gccaggcgua	aaauaggguga	5640
uuacaagaga	ggaguuuag	gcguucguag	cacaacaaca	augacgguuu	gaugcgggug	5700
cauacaucuu	uuccuccgac	accggucaag	ggcauuuaca	acaaaaauca	guaaggcaaa	5760
cggugcuau	cgaaguggug	uuggagagga	ccgaauugga	gauuucguau	gccccgcgcc	5820
ucgaccaaga	aaaagaagaa	uuacuacgca	agaaauuaca	guuaauucc	acaccugcua	5880
acagaagcag	auaccagucc	aggaaggugg	agaacaugaa	agccauaaca	gcuagacgua	5940
uuucgcaagg	ccuagggcua	uauuuagaag	cagaaggaaa	aguggagugc	uaccgaacc	6000
ugcauccugu	uccuuugua	ucaucuaug	ugaaccgugc	cuuuucaagc	cccaaggucg	6060
caguggaagc	cuguaacgcc	auguuuagaa	agaacuuucc	gacuguggcu	ucuuacugua	6120
uuauuccaga	guacgaugcc	uauuuaggaca	ugguugacgg	agcuucaugc	ugcuuagaca	6180
cugccaguuu	uugcccugca	aagcugcgca	gcuuuccaaa	gaaacacucc	uauuuggaac	6240
ccacaauacg	aucggcagug	ccuucagcga	uccagaacac	gcuccagaac	guccuggcag	6300
cugccacaaa	aagaaauugc	aaugucacgc	aaaugagaga	auugcccgua	uuggauucgg	6360
cggccuuuaa	uguggaaugc	uucaagaaa	augcguguua	uaaugaauau	ugggaaacgu	6420
uuaaagaaaa	ccccaucagg	cuuacugaag	aaaacguggu	aaauuacauu	accaaauuaa	6480
aaggaccaaa	agcugcugcu	cuuuuugcga	agacacauaa	uuugaauaug	uugcaggaca	6540
uaccaaugga	cagguuugua	auggacuuaa	agagagacgu	gaaagugacu	ccaggaacaa	6600
aacauacuga	agaacggccc	aagguacagg	ugauccaggc	ugccgauccg	cuagcaacag	6660
cguaucugug	cggaauccac	cagagcugg	uuaggagauu	aaaugcggu	cugcuuccga	6720
acauucauac	acuguuugau	augucggcug	aagacuuga	cgcuauuaua	gccgagcacu	6780
uccagccugg	ggauuguguu	cuggaaacug	acaucgcguc	guuugauaaa	agugaggacg	6840
acgccauggc	ucugaccgcg	uuauugauuc	uggaagacuu	agguguggac	gcagagcugu	6900
ugacgcugau	ugaggcggu	uucggcgaaa	uuucaucaau	acauuugccc	acuaaaacua	6960
aaauuuuuuu	cggagccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguaaucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucau	uggagaugac	aaauaucguga	aaggagucua	aucggacaaa	uuauuggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgcuuua	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aaauugauga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaaggua	ugaaaccgua	ggaacuucca	7440
ucauaguua	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500

gggccccuau aacucucucac ggcuaaccug aauggacuac gacauagucu aguccgccaa	7560
g	7561
<210> 37	
<211> 7561	
<212> RNA	
<213> Artificial Sequence	
<220>	
<223> Synthetic construct, C14_nsP1-4 ORF, and SGP	
<400> 37	
augggcgggcg caugagagaa gcccagacca auuaccuacc caaaauggag aaaguucacg	60
uugacaucga ggaagacagc ccauuccuca gagcuuugca gcggagcuuc ccgcaguuug	120
agguagaagc caagcagguc acugauaaug accaugcuua ugccagagcg uuuucgcauc	180
uggcuucaaa acugaucgaa acggaggugg acccauccga cacgauccuu gacauuggaa	240
gugcgcccg cgcagaaug uauucuaagc acaaguauca uuguauugcu ccgaugagau	300
gugcggaaga uccggacaga uuguauaagu augcaacuaa gcugaagaaa aacuguaagg	360
aaauaacuga uaaggaaauug gacaagaaaa ugaaggagcu cgccgcccugc augagcgacc	420
cugaccugga aacugagacu augugccucc acgacgacga gucgugucgc uacgaagggc	480
aagucgcugu uuaccaggau guauacgcgg uugacggacc gacaagucuc uaucaccaag	540
ccaauaaggg aguuagaguc gccuacugga uaggcuuuga caccaccccu uuuauuuua	600
agaacuuggc uggagcauau ccaucauacu cuaccaacug ggccgacgaa accguguuaa	660
cggcucguua cauaggccua ugcagcucug acguuaugga gcggucacgu agagggaugu	720
ccaauucuuag aaagaaguau uugaaaccau ccaacaaugu ucuauucucu guuggcucga	780
ccaucuaacca cgagaagagg gacuuacuga ggagcuggca ccugccgucu guauuucacu	840
uacguggcaa gcaaaaauac acaugucggg uugagacuau aguuaguugc gacggguacg	900
ucguuaaaaag aauagcuauac aguccaggcc uguauuggga gccuucaggc uaugcugcua	960
cgaugcaccg cgagggaauu uugugcugca aagugacaga cacauugaac ggggagaggg	1020
ucucuuuuucc cgugugcacg uaugugccag cuacauugug ugacaaaug acuggcauac	1080
uggcaacaga ugucagugcg gacgacgcgc aaaaacugcu gguugggcuc aaccagcgua	1140
uagucgucaa cggucgcacc cagagaaaca ccauaccuau gaaaaauuac cuuuugcccg	1200
uaguggccca ggcauuugcu agguugggcaa aggaauauaa ggaagaucaa gaagaugaaa	1260
ggccacuagg acuacgagau agacaguuaug ucauggggug uuguugggcu uuuagaaggc	1320
acaagauaac aucuauuuau aagcgcccgg auacccaaac caucauaaa gugaacagcg	1380
auuuccacuc auucgugcug cccaggauag gcaguaacac auuggagauac gggcugagaa	1440
caagaauacg gaaaauguua gaggagcaca aggagccguc accucucauu accgccgagg	1500
acguacaaga agcuagugc gcagccgaug aggcuagga ggugcgugaa gccgaggagu	1560
ugcgcgcgagc ucuaccaccu uuggcagcug auguugagga gcccacucug gaagccgaug	1620
ucgacuugau guuacaagag gcuggggccg gcucagugga gacaccucgu ggcuugauaa	1680
agguuaccag cuacgcuggc gaggacaaga ucggcucuaa cgcugugcuu ucuccgcagg	1740
cuguacucaa gagugaaaaa uuaucuugca uccaccucuc cgcugaacaa gucauaguga	1800

uaacacacuc	uggccgaaaa	gggcguuau	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuguguacaa	cgaacgugag	uucguaaaca	gguaccugca	ccauauugcc	acacauggag	1980
gagcgcugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauag	2220
gcaagucugg	caucauuaaa	agcgcaguca	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaagggacg	ucaagaaaau	gaaagggcug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaca	ccccguagag	accuguaau	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccca	aacagugcgg	uuuuuuuac	augaugugcc	2520
ugaaagugca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucugcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaugagaa	2640
cgacgaaucc	gaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaagacg	gcagcugccu	cuaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaacacu	gacugccaag	uacccuggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggaccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuagugcc	ggugcugaag	accgcuggca	3120
uagacaugac	cacugaacaa	uggaacacug	uggauuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaucg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacac	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaag	aagugguucc	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuaua	caugaacacu	gguacacugc	3420
gcauuuaua	uccgcgcaua	aaccuaguac	cuguuaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaauagaacac	ccacagagug	acuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaau	ugacauaaau	uuuguuaau	ugaggacccc	auauaaaau	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaaucccgg	cggaaaccugu	gucagcauag	guuauggua	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuuu	caagcuuuca	ucaaccuuga	ccaacauuuu	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucua	uaucaugugg	ugcgagggga	uauugccacg	gccaccgaag	4080
gagugauuuu	aaaugcugcu	aacagcaaag	gacaaccugg	cggagggguu	ugcggagccc	4140

uguacaagaa	guuccccgag	agcuucgacc	ugcagccuau	cgaagugggc	aaagccagac	4200
ugguuaaggg	cgcugccaag	cacaucaucc	augccguggg	acccaacuuc	aacaaggugu	4260
ccgaggugga	aggcgacaag	cagcuggccg	aggccuaua	gucuaucgcc	aagaucguga	4320
acgacaacaa	cuacaagagc	guggccauuc	cucugcugag	caccggcauc	uucagcggca	4380
acaaggacag	acugacccag	agccugaacc	aucugcugac	agcccuggau	accaccgaug	4440
ccgauguggc	caucuacugc	cgggacaaga	aauggggagau	gaccugaaa	gaagccgugg	4500
ccagacgcga	ggccguggaa	gagaucugua	ucagcgacga	cagcagcgug	accgagccug	4560
augccgaacu	cguuagagug	caccccaagu	cuagccuggc	cggcagaaa	ggcuacagca	4620
ccucugaugg	caagaccuuc	agcuaccugg	aaggcaccaa	guuccaccag	gccgccaagg	4680
auaucgccga	gaucaacgcu	auguggcccc	uggccaccga	ggccaaugaa	caagugugca	4740
uguacauccu	gggcgagagc	augagcagca	uccgcagcaa	guguccugug	gaagaguccg	4800
aggccagcac	accuccuagc	acacugccuu	gccugugcau	ccacgccaug	acaccugaga	4860
gagugcagcg	gcugaaggcc	ucuagaccug	agcagaucac	cgugugcagc	agcuucccac	4920
ugccuaagua	cagaauccac	ggcgugcaga	aaauccagug	cagccagccu	auccuguuca	4980
gcccuagggu	gcccgccuac	auucacccca	gaaaguaucu	gguggaaacc	ccaccugugg	5040
acgagacacc	ugaaccuagc	gccgagaau	agagcaccga	gggaacacca	gaacagccuc	5100
cacugaucac	cgaggacgaa	accaggacca	gaacaccgga	gccuaucauc	aucgaggaag	5160
aagaagagga	cagcaucagc	cugcugagcg	acggaccuac	acaccaggug	cugcaagugg	5220
aagccgacau	ccauggaccu	ccaagcgugu	ccagcagcag	cuggucuauu	ccucacgcca	5280
gcgacuucga	cguggacagc	cugagcaucc	uggacacacu	ggaaggcgcc	ucugugacau	5340
cuggcgccac	auccgccgag	acaaacagcu	acuucgcaa	gagcauggaa	uuucuggcca	5400
ggccugugcc	ugcuccucgg	accguguuuu	gaaaccucc	ucauccugcu	ccuagaacca	5460
ggacaccuuc	ucuggccccc	agcagagccu	guagcagaac	cagccuggug	ucuacaccuc	5520
cuggcgugaa	cagagugauc	accagagagg	aacuggaagc	ccugacaccu	agcaggaccc	5580
caagcagauc	cgugucuaga	acaucccugg	uguccaauc	uccaggcguc	aaccggguca	5640
ucacccggga	agaguuuag	gccuuugugg	cucagcagca	gugaagauuc	gacgcaggcg	5700
ccuacaucuu	cuccagcgau	acaggccagg	gccaucugca	gcagaaaagc	gucagacaga	5760
ccgugcuguc	cgaaguggug	cuggaaagaa	ccgagcugga	aaucagcuac	gccccucggc	5820
uggaccaaga	gaaaagaaga	cugcugcgga	agaaacugca	gcugaacccc	acaccagcca	5880
acagaagcag	auaccagagc	cgggaaggugg	aaaacaugaa	ggccaucacc	gccagacgga	5940
uccugcaagg	ccugggccau	uaucugaagg	ccgagggcaa	aguggaauuc	uacagaacac	6000
ugcaccgccu	gccucugua	agcagcuccg	ugaauagggc	cuuuagcagc	ccaaaggugg	6060
ccgucgaagc	cugcaacgcc	augcugaaa	aaaacuucc	uaccguggcc	uccuacugca	6120
ucauccccga	guacgacgcc	uaccuggaca	ugguggaugg	cgcuagcugu	ugucuggaua	6180
ccgccagcuu	cugccccgcc	aagcugagaa	gcuuccccaa	gaagcacagc	uaucuggaac	6240
ccaccaucag	auccgccgug	ccuuccgcca	uucagaauac	ccugcagaac	gugcuggccg	6300
cugccaccaa	gaggaauugc	aacgugaccc	agaugcgcg	gcugcccgua	cuggauagcg	6360
ccgccuucua	cguggaaugu	uucaagaagu	acgccugcaa	caacgaguac	ugggagacau	6420
ucaaagagaa	ccccauccgg	cugaccgagg	aaaacguggu	caacuacauc	accaagcuga	6480

agggcccca	agccgccgc	cuguuugcca	agacacacaa	ccugaacaug	cugcaggaca	6540
uccccaugga	cagauucgug	auggaccuga	agcgggacgu	gaaagugacc	ccuggcacca	6600
agcacaccga	ggaacggccu	aaagugcaag	ugauccaggc	cgcugauccu	cuggccacag	6660
ccuauucugug	uggcauccac	agagaacucg	ugcggagacu	gaaugccgug	cugcugccca	6720
acauccacac	acuguucgac	augagcgccg	aggacuucga	ugccaucauu	gccgagcacu	6780
uccagccugg	cgaauugcgug	cucgagacag	auaucgccuc	cuucgacaag	agcgaggacg	6840
acgccauggc	ucugacugcc	cugaugaucc	uggaagaucu	gggcguugac	gccgagcugc	6900
ugacacugau	ugaagccgcc	uuuggcgaga	ucagcuccau	ccaucugccu	accaagacca	6960
aaauuuuuuu	cggagccaug	augaaaucug	gaauguuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguuauucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucau	uggagaugac	aaauaucguga	aaggagucua	aucggacaaa	uuauuggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aaauugauga	ugacaggaga	agggcuuugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaggua	ugaaaccgua	ggaacuucca	7440
ucauaguauu	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucuac	ggcuuaccug	aauggacuac	gacauagucu	aguccgcca	7560
g						7561

<210> 38

<211> 7561

<212> RNA

<213> Artificial Sequence

<220>

<223> Synthetic construct, C15_nsP1-4 ORF, and SGP

<400> 38

augggcggcg	caugagagaa	gcccagacca	auuaccuacc	caaaauggag	aaaguucacg	60
uugacaucga	ggaagacagc	ccauuccuca	gagcuuugca	gcggagcuuc	ccgcaguuuu	120
agguagaagc	caagcagguc	acugauaaug	accaugcuua	ugccagagcg	uuuucgauc	180
uggcuuacaa	acugaucgaa	acggaggugg	acccauccga	cacgauccuu	gacauuggaa	240
gugcgcccg	ccgcagaaug	uauucuaagc	acaaguauca	uuguauucgu	ccgaugagau	300
gugcggaaga	uccggacaga	uuguauaagu	augcaacuua	gcugaagaaa	aacuguaagg	360
aaauaacuga	uaaggaauug	gacaagaaaa	ugaaggagcu	cgccgccguc	augagcgacc	420
cugaccugga	aacugagacu	augugccucc	acgacgacga	gucgugucgc	uacgaagggc	480
aagucgcugu	uuaccaggau	guauacgcgg	uugacggacc	gacaagucuc	uauaccaag	540
ccaauaaggg	aguuaagagc	gccuacugga	uaggcuuuga	caccaccccu	uuuauuuuu	600
agaacuuggc	uggagcauuu	ccaucuuacu	cuaccaacug	ggccgacgaa	accguguuaa	660
cggcucguua	cauaggccua	ugcagcucug	acguuaugga	gcggucacgu	agagggaugu	720
ccaauucuuag	aaagaaguau	uugaaaccau	ccaacaauug	ucuauucucu	guuggcucga	780

ccaucuacca	cgagaagagg	gacuuacuga	ggagcuggca	ccugccgucu	guauuucacu	840
uacguggcaa	gcaaaaauac	acaugucggg	gugagacuau	aguuaugu	gacggguacg	900
ucguuaaaag	aaauagcuau	aguccaggcc	uguaugggaa	gccuucaggc	uauugcugua	960
cgaugcaccg	cgagggauuc	uugugcugca	aagugacaga	cacauugaac	ggggagaggg	1020
ucucuuuucc	cgugugcacg	uauugccag	cuacauugug	ugaccaaau	acuggcauac	1080
uggcaacaga	ugucagugcg	gacgacgcgc	aaaaacugcu	ggugggguc	aaccagcgua	1140
uagucgucaa	cggucgcacc	cagagaaaca	ccaauaccu	gaaaaauac	uuuugccc	1200
uaguggccca	ggcauuugcu	aggugggcaa	aggaauauaa	ggaagaucaa	gaagaugaaa	1260
ggccacuagg	acuacgagau	agacaguua	ucauggggug	uuuugggcu	uuuagaaggc	1320
acaagauaac	aucuauuuau	aagcgcccgg	auacccaaac	caucaucaa	gugaacagcg	1380
auuuccacuc	auucgugcug	cccaggauag	gcaguaacac	auuggagauc	gggcugagaa	1440
caagaaucag	gaaaaugua	gaggagcaca	aggagccguc	accucucau	accgccgagg	1500
acguacaaga	agcuuagugc	gcagccgaug	aggcuagga	ggugcgugaa	gccgaggagu	1560
ugcgcgagc	ucuaccaccu	uuggcagcug	auguugagga	gcccacucug	gaagccgaug	1620
ucgacuugau	guuacaagag	gcuggggccg	gcucagugga	gacaccucgu	ggcuugauaa	1680
agguuaccag	cuacgcuggc	gaggacaaga	ucggcucuua	cgcuugcugu	ucuccgcagg	1740
cuguacucaa	gagugaaaaa	uuaucuugca	uccaccucuc	cgcuagaaca	gucauaguga	1800
uaacacacuc	uggccgaaaa	gggcguuau	ccguggaacc	auaccauggu	aaaguagugg	1860
ugccagaggg	acaugcaaua	cccguccagg	acuuucaagc	ucugagugaa	agugccacca	1920
uuuguuacaa	cgaacgugag	uucguaaaaca	gguaccugca	ccauauugcc	acacauaggag	1980
gagcgugaa	cacugaugaa	gaauauuaca	aaacugucua	gcccagcgag	cacgacggcg	2040
aaauaccugua	cgacaucgac	aggaaacagu	gcgucaagaa	agaacuaguc	acugggcuag	2100
ggcucacagg	cgagcuggug	gauccucccu	uccaugaaau	cgccuacgag	agucugagaa	2160
cacgaccagc	cgcuccuuac	caaguaccaa	ccauaggggu	guauggcgug	ccaggauacg	2220
gcaagucugg	caucauuaaa	agcgagucua	ccaaaaaaga	ucuaguggug	agcgccaaga	2280
aagaaaacug	ugcagaaaau	auaaggagc	ucaagaaaau	gaaaggcgug	gacgucaaug	2340
ccagaacugu	ggacucagug	cucuugaau	gaugcaaaaca	ccccguagag	accuguaaua	2400
uugacgaagc	uuuugcuugu	caugcaggua	cucucagagc	gcucauagcc	auuauaagac	2460
cuaaaaaggc	agugcucugc	ggggauccea	aacagugcgg	uuuuuuuac	augaugugcc	2520
ugaaauguca	uuuuuaccac	gagauuugca	cacaagucuu	ccacaaaagc	aucucucgcc	2580
guugcacuaa	aucugugacu	ucggucgucu	caaccuuguu	uuacgacaaa	aaaauagagaa	2640
cgacgaaucc	gaaaagagacu	aagauuguga	uugacacuac	cggcaguacc	aaaccuaagc	2700
aggacgaucu	cauucucacu	uguuucagag	ggugggugaa	gcaguugcaa	auagauuaca	2760
aaggcaacga	aaauaagacg	gcagcugccu	cucaagggcu	gacccguaaa	gguguguaug	2820
ccguucggua	caaggugaau	gaaaauccuc	uguacgcacc	caccucagaa	caugugaacg	2880
uccuacugac	ccgcacggag	gaccgcaucg	uguggaaaac	acuagccggc	gacccaugga	2940
uaaaaaacacu	gacugccaag	uacccuggga	auuucacugc	cacgauagag	gaguggcaag	3000
cagagcauga	ugccaucaug	aggcacaucu	uggagagacc	ggaccuacc	gacgucuucc	3060
agaauaaggc	aaacgugugu	ugggccaagg	cuuauugcc	ggugcugaag	accgcuggca	3120

uagacaugac	cacugaacaa	uggaacacug	uggauuuuuu	ugaaacggac	aaagcucacu	3180
cagcagagau	aguauugaac	caacuaugcg	ugagguucuu	uggacucgau	cuggacuccg	3240
gucuauuuuc	ugcaccacac	guuccguuau	ccauuaggaa	uaaucacugg	gauaacuccc	3300
cgucgccuaa	cauguacggg	cugaauaaaag	aagugguccg	ucagcucucu	cgcagguacc	3360
cacaacugcc	ucgggcaguu	gccacuggaa	gagucuauga	caugaacacu	gguacacugc	3420
gcauuuauga	uccgcgcuaa	aaccuaguac	cuguaaacag	aagacugccu	caugcuuuag	3480
uccuccacca	uaaugaacac	ccacagagug	acuuuuucuu	auucgucagc	aaauugaagg	3540
gcagaacugu	ccuggugguc	ggggaaaagu	uguccguccc	aggcaaaaug	guugacuggu	3600
ugucagaccg	gccugaggcu	accuucagag	cucggcugga	uuuaggcauc	ccaggugaug	3660
ugcccaaaaua	ugacauaaaua	uuuguuaaau	ugaggacccc	auauaaaauac	caucacuauc	3720
agcaguguga	agaccaugcc	auuaagcuua	gcauguugac	caagaaagcu	ugucugcauc	3780
ugaaucccg	cggaaccugu	gucagcauag	guuaugguua	cgcugacagg	gccagcgaaa	3840
gcaucauugg	ugcuauagcg	cggcaguuca	aguuuucccg	gguaugcaaa	ccgaaauccu	3900
cacuugaaga	gacggaaguu	cuguuuuguau	ucauugggua	cgaucgcaag	gcccguacgc	3960
acaauccuua	caagcuuuca	ucaaccuuga	ccaacauuuu	uacagguucc	agacuccacg	4020
aagccggaug	ugcaccucu	uaucaugugg	ugcgagggga	uaaugccacg	gccaccgaag	4080
gagugauuuu	aaaugcugcu	aacagcaaag	gacaaccugg	cggaggggug	ugcggagccc	4140
uguacaagaa	guuccccgag	agcuucgacc	ugcagccaau	cgaagugggc	aaagccagac	4200
uggucaagg	cgcagccaag	cacaucaucc	acgccguggg	acccaacuuc	aacaagguga	4260
gcgaggugga	aggcgacaag	cagcuggccg	aggccuacga	gagcaucgcc	aagaucguga	4320
acgacaacaa	cuacaagagc	guggccauac	cccugcugag	caccggcauc	uucagcggca	4380
acaaggacag	acugaccag	agccugaacc	accugcugac	agcccuggac	accaccgacg	4440
ccgacguggc	caucuacugc	cgggacaaga	aaugggagau	gaccugaaa	gaagccgugg	4500
ccagacgcga	ggccguggaa	gagaucugca	ucagcgacga	cagcagcgug	accgagccag	4560
acgccgaacu	cgucagagug	cacccaaga	gcagccuggc	cggcagaaaag	ggcuacagca	4620
ccagcgacgg	caagaccuuc	agcuaccugg	aaggcaccaa	guuccaccag	gccgccaagg	4680
acaucgccga	gaucaacgca	auguggcccc	uggccaccga	ggccaacgaa	caagugugca	4740
uguacauccu	gggcgagagc	augagcagca	uccgcagcaa	gugccccgug	gaagagagcg	4800
aggccagcac	accacccagc	acacugcccu	gccugugcau	ccacgccaug	acaccagaga	4860
gagugcagcg	gcugaaggcc	agcagacccg	agcagaucac	cgugugcagc	agcuucccac	4920
ugcccaagua	cagaauacac	ggcgugcaga	aaauccagug	cagccagcca	auccuguuca	4980
gccccaaaggu	gcccgccuac	auacacccca	gaaaguaccu	gguggaaacc	ccaccgugg	5040
acgagacacc	cgaaccaagc	gccgagaacc	agagcaccga	gggaacacca	gaacagcccc	5100
cacugaucac	cgaggacgaa	accaggacca	gaacacccga	gccaaucauc	aucgaggaag	5160
aagaagagga	cagcaucagc	cugcugagcg	acggaccac	acaccaggug	cugcaagugg	5220
aagccgacau	ccacggaccc	ccaagcguga	gcagcagcag	cuggagcauc	ccccacgcca	5280
gcgacuucga	cguggacagc	cugagcaucc	uggacacacu	ggaaggcgcc	agcgugacaa	5340
gcggcgccac	aagcgccgag	acaaaacagcu	acuuucgcaa	gagcauggaa	uuccuggcca	5400
ggcccugucc	agcaccgccg	accguguuca	gaaaccccc	ucaccagcc	cccagaacca	5460

ggacacccag	ccuggcccca	agcagagccu	gcagcagaac	cagccuggug	agcacacccc	5520
ccggcgugaa	cagagugauc	accagagagg	aacuggaagc	ccugacaccc	agcaggaccc	5580
caagcagaag	cgugagcaga	acaagccugg	ugagcaaccc	accaggcguc	aaccggguca	5640
ucacccggga	agaguucgag	gccuucgugg	cacagcagca	gugaagauuc	gacgcaggcg	5700
ccuacaucuu	cagcagcgac	acaggccagg	gccaccugca	gcagaaaagc	gucagacaga	5760
ccgugcugag	cgaaguggug	cuggaaagaa	ccgagcugga	aaucagcuac	gccccccggc	5820
uggaccaaga	gaaagaagaa	cugcugcgga	agaaacugca	gcugaacccc	acaccagcca	5880
acagaagcag	auaccagagc	cggaggugg	aaaacaugaa	ggccaucacc	gccagacgga	5940
uccugcaagg	ccugggccac	uaccugaagg	ccgagggcaa	aguggaaugc	uacagaacac	6000
ugcaccccg	gccacugua	agcagcagcg	ugaacagggc	cuucagcagc	ccaaaggugg	6060
ccgucgaagc	cugcaacgcc	augcugaaag	aaaacuuccc	caccguggcc	agcuacugca	6120
ucauucccga	guacgacgcc	uaccuggaca	ugguggacgg	cgcaagcugc	ugccuggaca	6180
ccgccagcuu	cugccccgcc	aagcugagaa	gcuuccccaa	gaagcacagc	uaccuggaac	6240
ccaccaucag	aagcgccgug	ccaagcgcca	uucagaacac	ccugcagaac	gugcuggccg	6300
cagccaccaa	gaggaacugc	aacgugaccc	agaugcgca	gcugcccguc	cuggacagcg	6360
ccgccuucaa	cguggaaugc	uucaagaagu	acgccugcaa	caacgaguac	ugggagacau	6420
ucaaaagagaa	ccccauccgg	cugaccgagg	aaaacguggu	caacuacauc	accaagcuga	6480
agggccccaa	agccgcagcc	cuguucgcca	agacacacaa	ccugaacaug	cugcaggaca	6540
uccccaugga	cagauucgug	auggaccuga	agcgggacgu	gaaagugacc	ccaggcacca	6600
agcacaccga	ggaacggccc	aaagugcaag	ugauccaggc	cgcagacccc	cuggccacag	6660
ccuaccugug	cggcauccac	agagaacucg	ugcggagacu	gaacgccgug	cugcugccca	6720
acauccacac	acuguucgac	augagcgccg	aggacuucga	cgccaucaua	gccgagcacu	6780
uccagccagg	cgacugcgug	cucgagacag	acaucgccag	cuucgacaag	agcgaggacg	6840
acgccauggc	acugaccgcc	cugaugaucc	uggaagaccu	gggcguggac	gccgagcugc	6900
ugacacugau	agaagccgcc	uucggcgaga	ucagcagcau	ccaccugccc	accaagacca	6960
aauuuaaaau	cggagccaug	augaaaucug	gaauuuuccu	cacacuguuu	gugaacacag	7020
ucauuuacau	uguuaucgca	agcagagugu	ugagagaacg	gcuaaccgga	ucaccaugug	7080
cagcauucau	uggagaugac	aaauaucguga	aaggagucua	aucggacaaa	uuauuggcag	7140
acaggugcgc	caccugguug	aaauuggaag	ucaagauuau	agaugcugug	gugggcgaga	7200
aagcgccuua	uuucugugga	ggguuuuuuu	ugugugacuc	cgugaccggc	acagcgugcc	7260
guguggcaga	ccccuaaaa	aggcuguuuu	agcuuggcaa	accucuggca	gcagacgaug	7320
aacauauga	ugacaggaga	agggcauugc	augaagaguc	aacacgcugg	aaccgagugg	7380
guauucuuuc	agagcugugc	aaggcaguag	aaucaaggua	ugaaaccgua	ggaacuucca	7440
ucauaguua	ggccaugacu	acucuagcua	gcaguguuaa	aucauucagc	uaccugagag	7500
gggccccuau	aacucucua	ggcuuaccug	aauggacuac	gacauagucu	aguccgcca	7560
g						7561
<210>	39					
<211>	627					
<212>	PRT					

<213> Artificial Sequence

<220>

<223> Synthetic construct, VEEV nsP4-2A polyprotein

<400> 39

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

275	280	285
Leu Lys Arg Asp Val Lys Val Thr Pro Gly Thr Lys His Thr Glu Glu		
290	295	300
Arg Pro Lys Val Gln Val Ile Gln Ala Ala Asp Pro Leu Ala Thr Ala		
305	310	315
Tyr Leu Cys Gly Ile His Arg Glu Leu Val Arg Arg Leu Asn Ala Val		
325	330	335
Leu Leu Pro Asn Ile His Thr Leu Phe Asp Met Ser Ala Glu Asp Phe		
340	345	350
Asp Ala Ile Ile Ala Glu His Phe Gln Pro Gly Asp Cys Val Leu Glu		
355	360	365
Thr Asp Ile Ala Ser Phe Asp Lys Ser Glu Asp Asp Ala Met Ala Leu		
370	375	380
Thr Ala Leu Met Ile Leu Glu Asp Leu Gly Val Asp Ala Glu Leu Leu		
385	390	395
Thr Leu Ile Glu Ala Ala Phe Gly Glu Ile Ser Ser Ile His Leu Pro		
405	410	415
Thr Lys Thr Lys Phe Lys Phe Gly Ala Met Met Lys Ser Gly Met Phe		
420	425	430
Leu Thr Leu Phe Val Asn Thr Val Ile Asn Ile Val Ile Ala Ser Arg		
435	440	445
Val Leu Arg Glu Arg Leu Thr Gly Ser Pro Cys Ala Ala Phe Ile Gly		
450	455	460
Asp Asp Asn Ile Val Lys Gly Val Lys Ser Asp Lys Leu Met Ala Asp		
465	470	475
Arg Cys Ala Thr Trp Leu Asn Met Glu Val Lys Ile Ile Asp Ala Val		
485	490	495
Val Gly Glu Lys Ala Pro Tyr Phe Cys Gly Gly Phe Ile Leu Cys Asp		
500	505	510
Ser Val Thr Gly Thr Ala Cys Arg Val Ala Asp Pro Leu Lys Arg Leu		
515	520	525
Phe Lys Leu Gly Lys Pro Leu Ala Ala Asp Asp Glu His Asp Asp Asp		
530	535	540
Arg Arg Arg Ala Leu His Glu Glu Ser Thr Arg Trp Asn Arg Val Gly		
545	550	555
Ile Leu Ser Glu Leu Cys Lys Ala Val Glu Ser Arg Tyr Glu Thr Val		
565	570	575
Gly Thr Ser Ile Ile Val Met Ala Met Thr Thr Leu Ala Ser Ser Val		
580	585	590

Lys Ser Phe Ser Tyr Leu Arg Gly Ala Pro Ile Thr Leu Tyr Gly Ser
 595 600 605
 Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn
 610 615 620
 Pro Gly Pro
 625
 <210> 40
 <211> 840
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, VEEV nsP4-2A-3Cpro
 <400> 40
 Tyr Ile Phe Ser Ser Asp Thr Gly Gln Gly His Leu Gln Gln Lys Ser
 1 5 10 15
 Val Arg Gln Thr Val Leu Ser Glu Val Val Leu Glu Arg Thr Glu Leu
 20 25 30
 Glu Ile Ser Tyr Ala Pro Arg Leu Asp Gln Glu Lys Glu Glu Leu Leu
 35 40 45
 Arg Lys Lys Leu Gln Leu Asn Pro Thr Pro Ala Asn Arg Ser Arg Tyr
 50 55 60
 Gln Ser Arg Lys Val Glu Asn Met Lys Ala Ile Thr Ala Arg Arg Ile
 65 70 75 80
 Leu Gln Gly Leu Gly His Tyr Leu Lys Ala Glu Gly Lys Val Glu Cys
 85 90 95
 Tyr Arg Thr Leu His Pro Val Pro Leu Tyr Ser Ser Ser Val Asn Arg
 100 105 110
 Ala Phe Ser Ser Pro Lys Val Ala Val Glu Ala Cys Asn Ala Met Leu
 115 120 125
 Lys Glu Asn Phe Pro Thr Val Ala Ser Tyr Cys Ile Ile Pro Glu Tyr
 130 135 140
 Asp Ala Tyr Leu Asp Met Val Asp Gly Ala Ser Cys Cys Leu Asp Thr
 145 150 155 160
 Ala Ser Phe Cys Pro Ala Lys Leu Arg Ser Phe Pro Lys Lys His Ser
 165 170 175
 Tyr Leu Glu Pro Thr Ile Arg Ser Ala Val Pro Ser Ala Ile Gln Asn
 180 185 190
 Thr Leu Gln Asn Val Leu Ala Ala Ala Thr Lys Arg Asn Cys Asn Val
 195 200 205

Thr	Gln	Met	Arg	Glu	Leu	Pro	Val	Leu	Asp	Ser	Ala	Ala	Phe	Asn	Val
210						215					220				
Glu	Cys	Phe	Lys	Lys	Tyr	Ala	Cys	Asn	Asn	Glu	Tyr	Trp	Glu	Thr	Phe
225					230					235					240
Lys	Glu	Asn	Pro	Ile	Arg	Leu	Thr	Glu	Glu	Asn	Val	Val	Asn	Tyr	Ile
				245						250					255
Thr	Lys	Leu	Lys	Gly	Pro	Lys	Ala	Ala	Ala	Leu	Phe	Ala	Lys	Thr	His
				260						265					270
Asn	Leu	Asn	Met	Leu	Gln	Asp	Ile	Pro	Met	Asp	Arg	Phe	Val	Met	Asp
				275						280					285
Leu	Lys	Arg	Asp	Val	Lys	Val	Thr	Pro	Gly	Thr	Lys	His	Thr	Glu	Glu
				290						295					300
Arg	Pro	Lys	Val	Gln	Val	Ile	Gln	Ala	Ala	Asp	Pro	Leu	Ala	Thr	Ala
305						310					315				320
Tyr	Leu	Cys	Gly	Ile	His	Arg	Glu	Leu	Val	Arg	Arg	Leu	Asn	Ala	Val
				325						330					335
Leu	Leu	Pro	Asn	Ile	His	Thr	Leu	Phe	Asp	Met	Ser	Ala	Glu	Asp	Phe
				340						345					350
Asp	Ala	Ile	Ile	Ala	Glu	His	Phe	Gln	Pro	Gly	Asp	Cys	Val	Leu	Glu
				355						360					365
Thr	Asp	Ile	Ala	Ser	Phe	Asp	Lys	Ser	Glu	Asp	Asp	Ala	Met	Ala	Leu
				370						375					380
Thr	Ala	Leu	Met	Ile	Leu	Glu	Asp	Leu	Gly	Val	Asp	Ala	Glu	Leu	Leu
385						390					395				400
Thr	Leu	Ile	Glu	Ala	Ala	Phe	Gly	Glu	Ile	Ser	Ser	Ile	His	Leu	Pro
				405						410					415
Thr	Lys	Thr	Lys	Phe	Lys	Phe	Gly	Ala	Met	Met	Lys	Ser	Gly	Met	Phe
				420						425					430
Leu	Thr	Leu	Phe	Val	Asn	Thr	Val	Ile	Asn	Ile	Val	Ile	Ala	Ser	Arg
				435						440					445
Val	Leu	Arg	Glu	Arg	Leu	Thr	Gly	Ser	Pro	Cys	Ala	Ala	Phe	Ile	Gly
				450						455					460
Asp	Asp	Asn	Ile	Val	Lys	Gly	Val	Lys	Ser	Asp	Lys	Leu	Met	Ala	Asp
465						470					475				480
Arg	Cys	Ala	Thr	Trp	Leu	Asn	Met	Glu	Val	Lys	Ile	Ile	Asp	Ala	Val
				485						490					495
Val	Gly	Glu	Lys	Ala	Pro	Tyr	Phe	Cys	Gly	Gly	Phe	Ile	Leu	Cys	Asp
				500						505					510
Ser	Val	Thr	Gly	Thr	Ala	Cys	Arg	Val	Ala	Asp	Pro	Leu	Lys	Arg	Leu

515	520	525
Phe Lys Leu Gly Lys Pro Leu Ala Ala Asp Asp Glu His Asp Asp Asp		
530	535	540
Arg Arg Arg Ala Leu His Glu Glu Ser Thr Arg Trp Asn Arg Val Gly		
545	550	555
Ile Leu Ser Glu Leu Cys Lys Ala Val Glu Ser Arg Tyr Glu Thr Val		
565	570	575
Gly Thr Ser Ile Ile Val Met Ala Met Thr Thr Leu Ala Ser Ser Val		
580	585	590
Lys Ser Phe Ser Tyr Leu Arg Gly Ala Pro Ile Thr Leu Tyr Gly Ser		
595	600	605
Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn		
610	615	620
Pro Gly Pro Ser Gly Ala Pro Pro Thr Asp Leu Gln Lys Met Val Met		
625	630	635
Gly Asn Thr Lys Pro Val Glu Leu Ile Leu Asp Gly Lys Thr Val Ala		
645	650	655
Ile Cys Cys Ala Thr Gly Val Phe Gly Thr Ala Tyr Leu Val Pro Arg		
660	665	670
His Leu Phe Ala Glu Lys Tyr Asp Lys Ile Met Leu Asp Gly Arg Ala		
675	680	685
Met Thr Asp Ser Asp Tyr Arg Val Phe Glu Phe Glu Ile Lys Val Lys		
690	695	700
Gly Gln Asp Met Leu Ser Asp Ala Ala Leu Met Val Leu His Arg Gly		
705	710	715
Asn Arg Val Arg Asp Ile Thr Lys His Phe Arg Asp Thr Ala Arg Met		
725	730	735
Lys Lys Gly Thr Pro Val Val Gly Val Val Asn Asn Ala Asp Val Gly		
740	745	750
Arg Leu Ile Phe Ser Gly Glu Ala Leu Thr Tyr Lys Asp Ile Val Val		
755	760	765
Cys Met Asp Gly Asp Thr Met Pro Ser Leu Phe Ala Tyr Lys Ala Ala		
770	775	780
Thr Lys Ala Gly Tyr Cys Gly Gly Ala Val Leu Ala Lys Asp Gly Ala		
785	790	795
Asp Thr Phe Ile Val Gly Thr His Ser Ala Gly Gly Asn Gly Val Gly		
805	810	815
Tyr Cys Ser Cys Val Ser Lys Ser Met Leu Leu Arg Met Lys Ala His		
820	825	830

Val Asp Pro Glu Pro Gln His Glu
 835 840
 <210> 41
 <211> 865
 <212> PRT
 <213> Artificial Sequence
 <220>
 <223> Synthetic construct, VEEV nsP4-2A-GFP
 <400> 41
 Tyr Ile Phe Ser Ser Asp Thr Gly Gln Gly His Leu Gln Gln Lys Ser
 1 5 10 15
 Val Arg Gln Thr Val Leu Ser Glu Val Val Leu Glu Arg Thr Glu Leu
 20 25 30
 Glu Ile Ser Tyr Ala Pro Arg Leu Asp Gln Glu Lys Glu Glu Leu Leu
 35 40 45
 Arg Lys Lys Leu Gln Leu Asn Pro Thr Pro Ala Asn Arg Ser Arg Tyr
 50 55 60
 Gln Ser Arg Lys Val Glu Asn Met Lys Ala Ile Thr Ala Arg Arg Ile
 65 70 75 80
 Leu Gln Gly Leu Gly His Tyr Leu Lys Ala Glu Gly Lys Val Glu Cys
 85 90 95
 Tyr Arg Thr Leu His Pro Val Pro Leu Tyr Ser Ser Ser Val Asn Arg
 100 105 110
 Ala Phe Ser Ser Pro Lys Val Ala Val Glu Ala Cys Asn Ala Met Leu
 115 120 125
 Lys Glu Asn Phe Pro Thr Val Ala Ser Tyr Cys Ile Ile Pro Glu Tyr
 130 135 140
 Asp Ala Tyr Leu Asp Met Val Asp Gly Ala Ser Cys Cys Leu Asp Thr
 145 150 155 160
 Ala Ser Phe Cys Pro Ala Lys Leu Arg Ser Phe Pro Lys Lys His Ser
 165 170 175
 Tyr Leu Glu Pro Thr Ile Arg Ser Ala Val Pro Ser Ala Ile Gln Asn
 180 185 190
 Thr Leu Gln Asn Val Leu Ala Ala Ala Thr Lys Arg Asn Cys Asn Val
 195 200 205
 Thr Gln Met Arg Glu Leu Pro Val Leu Asp Ser Ala Ala Phe Asn Val
 210 215 220
 Glu Cys Phe Lys Lys Tyr Ala Cys Asn Asn Glu Tyr Trp Glu Thr Phe
 225 230 235 240

Lys	Glu	Asn	Pro	Ile	Arg	Leu	Thr	Glu	Glu	Asn	Val	Val	Asn	Tyr	Ile	245	250	255
Thr	Lys	Leu	Lys	Gly	Pro	Lys	Ala	Ala	Ala	Leu	Phe	Ala	Lys	Thr	His	260	265	270
Asn	Leu	Asn	Met	Leu	Gln	Asp	Ile	Pro	Met	Asp	Arg	Phe	Val	Met	Asp	275	280	285
Leu	Lys	Arg	Asp	Val	Lys	Val	Thr	Pro	Gly	Thr	Lys	His	Thr	Glu	Glu	290	295	300
Arg	Pro	Lys	Val	Gln	Val	Ile	Gln	Ala	Ala	Asp	Pro	Leu	Ala	Thr	Ala	305	310	315
Tyr	Leu	Cys	Gly	Ile	His	Arg	Glu	Leu	Val	Arg	Arg	Leu	Asn	Ala	Val	325	330	335
Leu	Leu	Pro	Asn	Ile	His	Thr	Leu	Phe	Asp	Met	Ser	Ala	Glu	Asp	Phe	340	345	350
Asp	Ala	Ile	Ile	Ala	Glu	His	Phe	Gln	Pro	Gly	Asp	Cys	Val	Leu	Glu	355	360	365
Thr	Asp	Ile	Ala	Ser	Phe	Asp	Lys	Ser	Glu	Asp	Asp	Ala	Met	Ala	Leu	370	375	380
Thr	Ala	Leu	Met	Ile	Leu	Glu	Asp	Leu	Gly	Val	Asp	Ala	Glu	Leu	Leu	385	390	395
Thr	Leu	Ile	Glu	Ala	Ala	Phe	Gly	Glu	Ile	Ser	Ser	Ile	His	Leu	Pro	405	410	415
Thr	Lys	Thr	Lys	Phe	Lys	Phe	Gly	Ala	Met	Met	Lys	Ser	Gly	Met	Phe	420	425	430
Leu	Thr	Leu	Phe	Val	Asn	Thr	Val	Ile	Asn	Ile	Val	Ile	Ala	Ser	Arg	435	440	445
Val	Leu	Arg	Glu	Arg	Leu	Thr	Gly	Ser	Pro	Cys	Ala	Ala	Phe	Ile	Gly	450	455	460
Asp	Asp	Asn	Ile	Val	Lys	Gly	Val	Lys	Ser	Asp	Lys	Leu	Met	Ala	Asp	465	470	475
Arg	Cys	Ala	Thr	Trp	Leu	Asn	Met	Glu	Val	Lys	Ile	Ile	Asp	Ala	Val	485	490	495
Val	Gly	Glu	Lys	Ala	Pro	Tyr	Phe	Cys	Gly	Gly	Phe	Ile	Leu	Cys	Asp	500	505	510
Ser	Val	Thr	Gly	Thr	Ala	Cys	Arg	Val	Ala	Asp	Pro	Leu	Lys	Arg	Leu	515	520	525
Phe	Lys	Leu	Gly	Lys	Pro	Leu	Ala	Ala	Asp	Asp	Glu	His	Asp	Asp	Asp	530	535	540
Arg	Arg	Arg	Ala	Leu	His	Glu	Glu	Ser	Thr	Arg	Trp	Asn	Arg	Val	Gly			

545	550	555	560
Ile Leu Ser Glu Leu Cys Lys Ala Val Glu Ser Arg Tyr Glu Thr Val			
	565	570	575
Gly Thr Ser Ile Ile Val Met Ala Met Thr Thr Leu Ala Ser Ser Val			
	580	585	590
Lys Ser Phe Ser Tyr Leu Arg Gly Ala Pro Ile Thr Leu Tyr Gly Ser			
	595	600	605
Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn			
	610	615	620
Pro Gly Pro Val Ser Lys Gly Glu Glu Leu Phe Thr Gly Val Val Pro			
625	630	635	640
Ile Leu Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val			
	645	650	655
Ser Gly Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys			
	660	665	670
Phe Ile Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val			
	675	680	685
Thr Thr Leu Thr Tyr Gly Val Gln Cys Phe Ser Arg Tyr Pro Asp His			
	690	695	700
Met Lys Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val			
705	710	715	720
Gln Glu Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg			
	725	730	735
Ala Glu Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu			
	740	745	750
Lys Gly Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu			
	755	760	765
Glu Tyr Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln			
	770	775	780
Lys Asn Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp			
785	790	795	800
Gly Ser Val Gln Leu Ala Asp His Tyr Gln Gln Asn Thr Pro Ile Gly			
	805	810	815
Asp Gly Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Thr Gln Ser			
	820	825	830
Lys Leu Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu			
	835	840	845
Glu Phe Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr			
850	855	860	

Lys

865

<210> 42

<211> 1005

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, VEEV nsP4-2A-STING

<400> 42

Tyr Ile Phe Ser Ser Asp Thr Gly Gln Gly His Leu Gln Gln Lys Ser

1 5 10 15

Val Arg Gln Thr Val Leu Ser Glu Val Val Leu Glu Arg Thr Glu Leu

20 25 30

Glu Ile Ser Tyr Ala Pro Arg Leu Asp Gln Glu Lys Glu Glu Leu Leu

35 40 45

Arg Lys Lys Leu Gln Leu Asn Pro Thr Pro Ala Asn Arg Ser Arg Tyr

50 55 60

Gln Ser Arg Lys Val Glu Asn Met Lys Ala Ile Thr Ala Arg Arg Ile

65 70 75 80

Leu Gln Gly Leu Gly His Tyr Leu Lys Ala Glu Gly Lys Val Glu Cys

85 90 95

Tyr Arg Thr Leu His Pro Val Pro Leu Tyr Ser Ser Ser Val Asn Arg

100 105 110

Ala Phe Ser Ser Pro Lys Val Ala Val Glu Ala Cys Asn Ala Met Leu

115 120 125

Lys Glu Asn Phe Pro Thr Val Ala Ser Tyr Cys Ile Ile Pro Glu Tyr

130 135 140

Asp Ala Tyr Leu Asp Met Val Asp Gly Ala Ser Cys Cys Leu Asp Thr

145 150 155 160

Ala Ser Phe Cys Pro Ala Lys Leu Arg Ser Phe Pro Lys Lys His Ser

165 170 175

Tyr Leu Glu Pro Thr Ile Arg Ser Ala Val Pro Ser Ala Ile Gln Asn

180 185 190

Thr Leu Gln Asn Val Leu Ala Ala Ala Thr Lys Arg Asn Cys Asn Val

195 200 205

Thr Gln Met Arg Glu Leu Pro Val Leu Asp Ser Ala Ala Phe Asn Val

210 215 220

Glu Cys Phe Lys Lys Tyr Ala Cys Asn Asn Glu Tyr Trp Glu Thr Phe

225 230 235 240

Lys	Glu	Asn	Pro	Ile	Arg	Leu	Thr	Glu	Glu	Asn	Val	Val	Asn	Tyr	Ile	245	250	255
Thr	Lys	Leu	Lys	Gly	Pro	Lys	Ala	Ala	Ala	Leu	Phe	Ala	Lys	Thr	His	260	265	270
Asn	Leu	Asn	Met	Leu	Gln	Asp	Ile	Pro	Met	Asp	Arg	Phe	Val	Met	Asp	275	280	285
Leu	Lys	Arg	Asp	Val	Lys	Val	Thr	Pro	Gly	Thr	Lys	His	Thr	Glu	Glu	290	295	300
Arg	Pro	Lys	Val	Gln	Val	Ile	Gln	Ala	Ala	Asp	Pro	Leu	Ala	Thr	Ala	305	310	315
Tyr	Leu	Cys	Gly	Ile	His	Arg	Glu	Leu	Val	Arg	Arg	Leu	Asn	Ala	Val	325	330	335
Leu	Leu	Pro	Asn	Ile	His	Thr	Leu	Phe	Asp	Met	Ser	Ala	Glu	Asp	Phe	340	345	350
Asp	Ala	Ile	Ile	Ala	Glu	His	Phe	Gln	Pro	Gly	Asp	Cys	Val	Leu	Glu	355	360	365
Thr	Asp	Ile	Ala	Ser	Phe	Asp	Lys	Ser	Glu	Asp	Asp	Ala	Met	Ala	Leu	370	375	380
Thr	Ala	Leu	Met	Ile	Leu	Glu	Asp	Leu	Gly	Val	Asp	Ala	Glu	Leu	Leu	385	390	395
Thr	Leu	Ile	Glu	Ala	Ala	Phe	Gly	Glu	Ile	Ser	Ser	Ile	His	Leu	Pro	405	410	415
Thr	Lys	Thr	Lys	Phe	Lys	Phe	Gly	Ala	Met	Met	Lys	Ser	Gly	Met	Phe	420	425	430
Leu	Thr	Leu	Phe	Val	Asn	Thr	Val	Ile	Asn	Ile	Val	Ile	Ala	Ser	Arg	435	440	445
Val	Leu	Arg	Glu	Arg	Leu	Thr	Gly	Ser	Pro	Cys	Ala	Ala	Phe	Ile	Gly	450	455	460
Asp	Asp	Asn	Ile	Val	Lys	Gly	Val	Lys	Ser	Asp	Lys	Leu	Met	Ala	Asp	465	470	475
Arg	Cys	Ala	Thr	Trp	Leu	Asn	Met	Glu	Val	Lys	Ile	Ile	Asp	Ala	Val	485	490	495
Val	Gly	Glu	Lys	Ala	Pro	Tyr	Phe	Cys	Gly	Gly	Phe	Ile	Leu	Cys	Asp	500	505	510
Ser	Val	Thr	Gly	Thr	Ala	Cys	Arg	Val	Ala	Asp	Pro	Leu	Lys	Arg	Leu	515	520	525
Phe	Lys	Leu	Gly	Lys	Pro	Leu	Ala	Ala	Asp	Asp	Glu	His	Asp	Asp	Asp	530	535	540
Arg	Arg	Arg	Ala	Leu	His	Glu	Glu	Ser	Thr	Arg	Trp	Asn	Arg	Val	Gly			

545	550	555	560
Ile Leu Ser Glu Leu Cys Lys Ala Val Glu Ser Arg Tyr Glu Thr Val			
	565	570	575
Gly Thr Ser Ile Ile Val Met Ala Met Thr Thr Leu Ala Ser Ser Val			
	580	585	590
Lys Ser Phe Ser Tyr Leu Arg Gly Ala Pro Ile Thr Leu Tyr Gly Ser			
	595	600	605
Gly Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn			
	610	615	620
Pro Gly Pro Pro His Ser Ser Leu His Pro Ser Ile Pro Cys Pro Arg			
625	630	635	640
Gly His Gly Ala Gln Lys Ala Ala Leu Val Leu Leu Ser Ala Cys Leu			
	645	650	655
Val Thr Leu Trp Gly Leu Gly Glu Pro Pro Glu His Thr Leu Arg Tyr			
	660	665	670
Leu Val Leu His Leu Ala Ser Leu Gln Leu Gly Leu Leu Leu Asn Gly			
	675	680	685
Val Cys Ser Leu Ala Glu Glu Leu Arg His Ile His Ser Arg Tyr Arg			
	690	695	700
Gly Ser Tyr Trp Arg Thr Val Arg Ala Cys Leu Gly Cys Pro Leu Arg			
705	710	715	720
Arg Gly Ala Leu Leu Leu Ser Ile Tyr Phe Tyr Tyr Ser Leu Pro			
	725	730	735
Asn Ala Val Gly Pro Pro Phe Thr Trp Met Leu Ala Leu Leu Gly Leu			
	740	745	750
Ser Gln Ala Leu Asn Ile Leu Leu Gly Leu Lys Gly Leu Ala Pro Ala			
	755	760	765
Glu Ile Ser Ala Val Cys Glu Lys Gly Asn Phe Ser Val Ala His Gly			
	770	775	780
Leu Ala Trp Ser Tyr Tyr Ile Gly Tyr Leu Arg Leu Ile Leu Pro Glu			
785	790	795	800
Leu Gln Ala Arg Ile Arg Thr Tyr Asn Gln His Tyr Asn Asn Leu Leu			
	805	810	815
Arg Gly Ala Val Ser Gln Arg Leu Tyr Ile Leu Leu Pro Leu Asp Cys			
	820	825	830
Gly Val Pro Asp Asn Leu Ser Met Ala Asp Pro Asn Ile Arg Phe Leu			
	835	840	845
Asp Lys Leu Pro Gln Gln Thr Gly Asp His Ala Gly Ile Lys Asp Arg			
	850	855	860

Val Tyr Ser Asn Ser Ile Tyr Glu Leu Leu Glu Asn Gly Gln Arg Ala
 865 870 875 880
 Gly Thr Cys Val Leu Glu Tyr Ala Thr Pro Leu Gln Thr Leu Phe Ala
 885 890 895
 Met Ser Gln Tyr Ser Gln Ala Gly Phe Ser Arg Glu Asp Met Leu Glu
 900 905 910
 Gln Ala Lys Leu Phe Cys Arg Thr Leu Glu Asp Ile Leu Ala Asp Ala
 915 920 925
 Pro Glu Ser Gln Asn Asn Cys Arg Leu Ile Ala Tyr Gln Glu Pro Ala
 930 935 940
 Asp Asp Ser Ser Phe Ser Leu Ser Gln Glu Val Leu Arg His Leu Arg
 945 950 955 960
 Gln Glu Glu Lys Glu Glu Val Thr Val Gly Ser Leu Lys Thr Ser Ala
 965 970 975
 Val Pro Ser Thr Ser Thr Met Ser Gln Glu Pro Glu Leu Leu Ile Ser
 980 985 990
 Gly Met Glu Lys Pro Leu Pro Leu Arg Thr Asp Phe Ser
 995 1000 1005

<210> 43

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, thosea asigna virus 2A protein

<400> 43

Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu Asn Pro
 1 5 10 15

Gly Pro

<210> 44

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> Synthetic construct, porcine teschovirus -1_2A protein

<400> 44

Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn
 1 5 10 15

Pro Gly Pro

<210> 45

<211> 22
<212> PRT
<213> Foot-and-mouth disease virus
<220>
<221> MISC_FEATURE
<222> (1) .. (22)
<223> FMDV 2A protein
<400> 45
Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val
1 5 10 15
Glu Ser Asn Pro Gly Pro
20
<210> 46
<211> 20
<212> PRT
<213> Artificial Sequence
<220>
<223> Syntehtic construct, equine rhinitis A virus 2A protein
<400> 46
Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp Val Glu Ser
1 5 10 15
Asn Pro Gly Pro
20

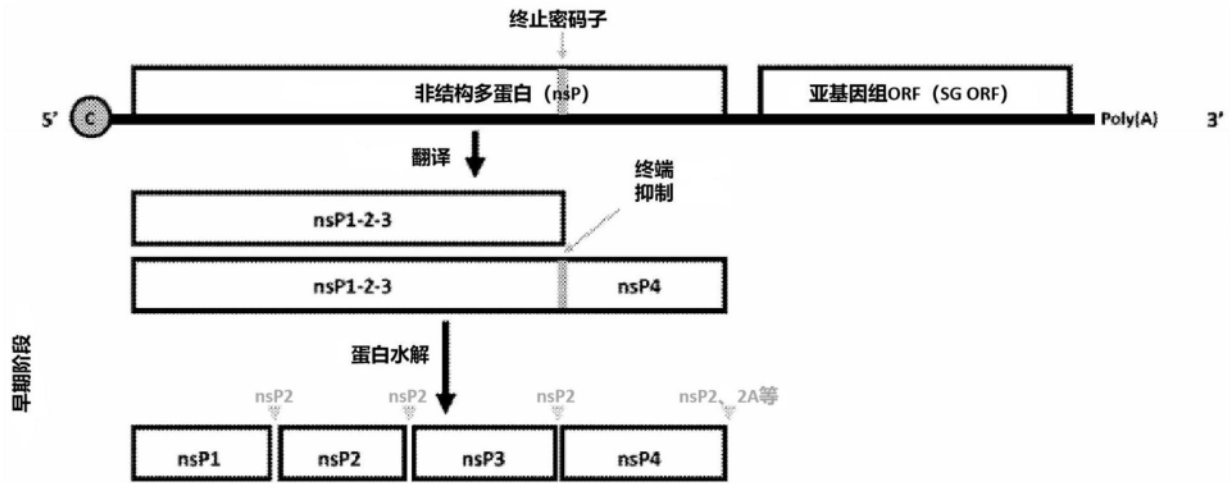


图1

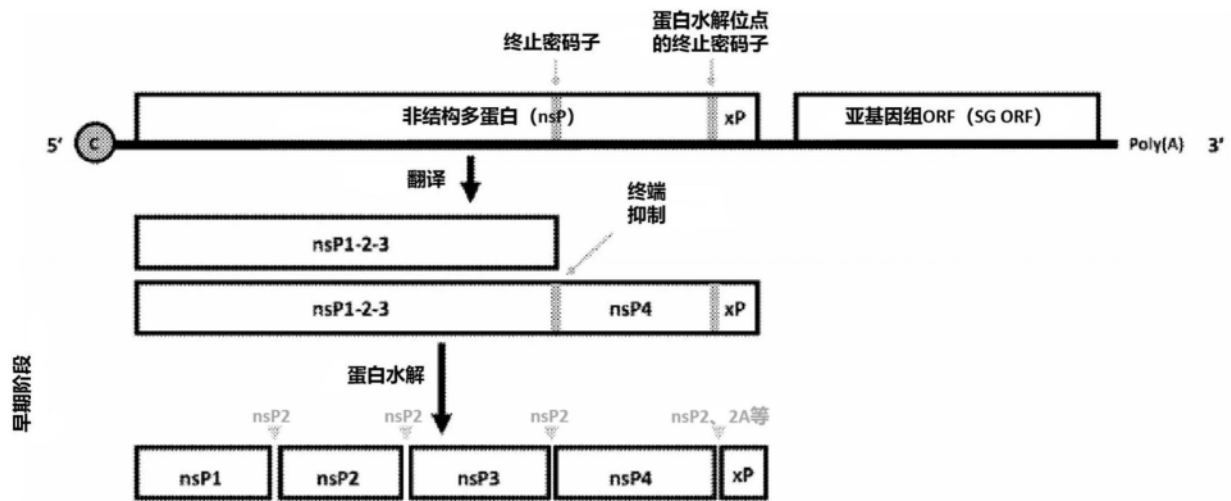


图2

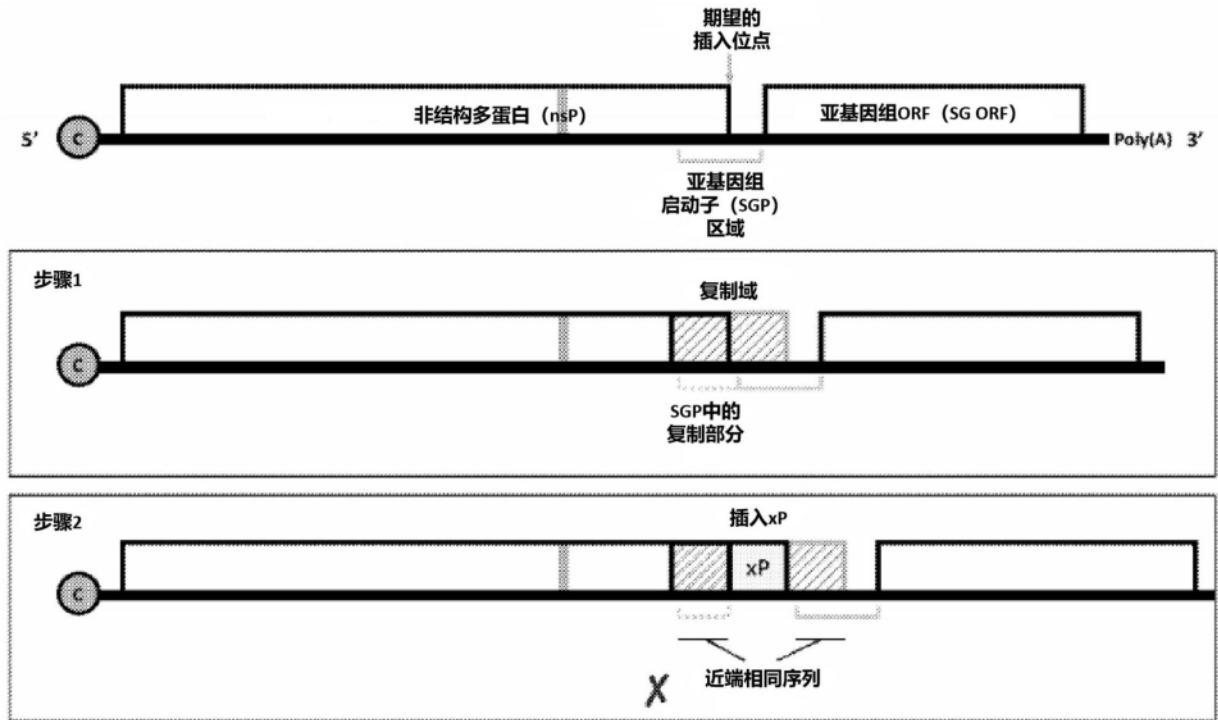


图3

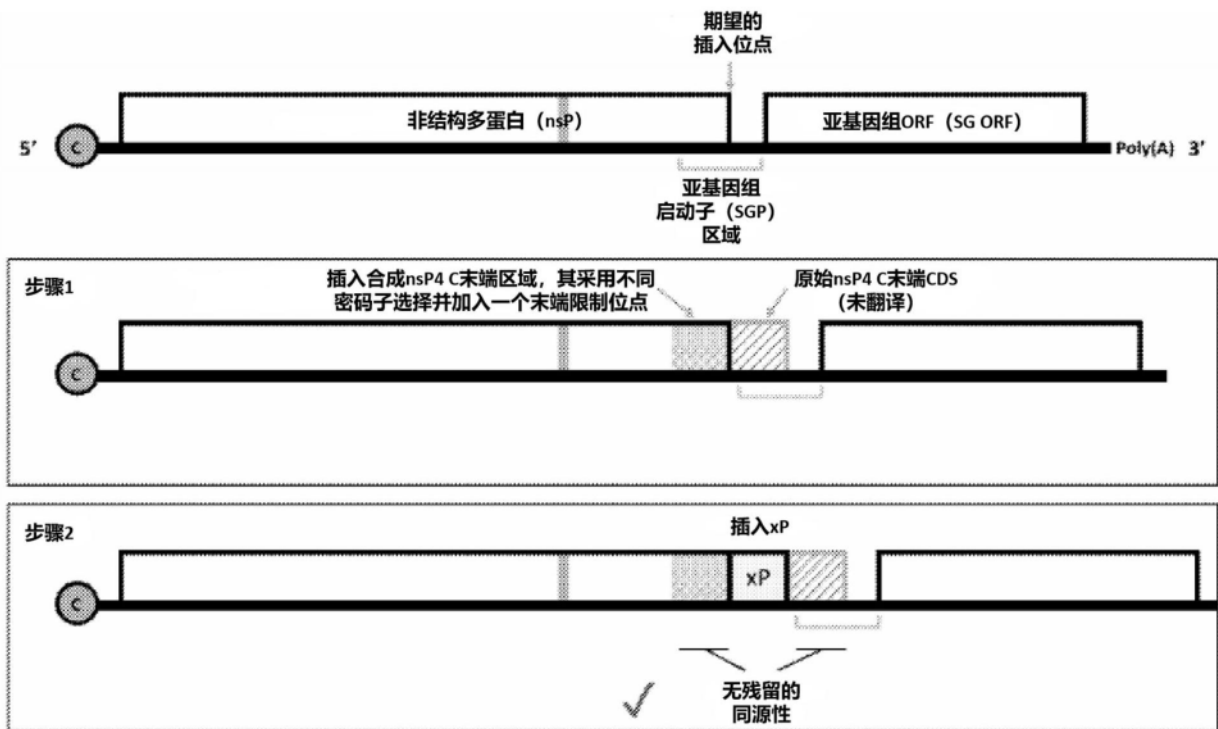


图4



图5

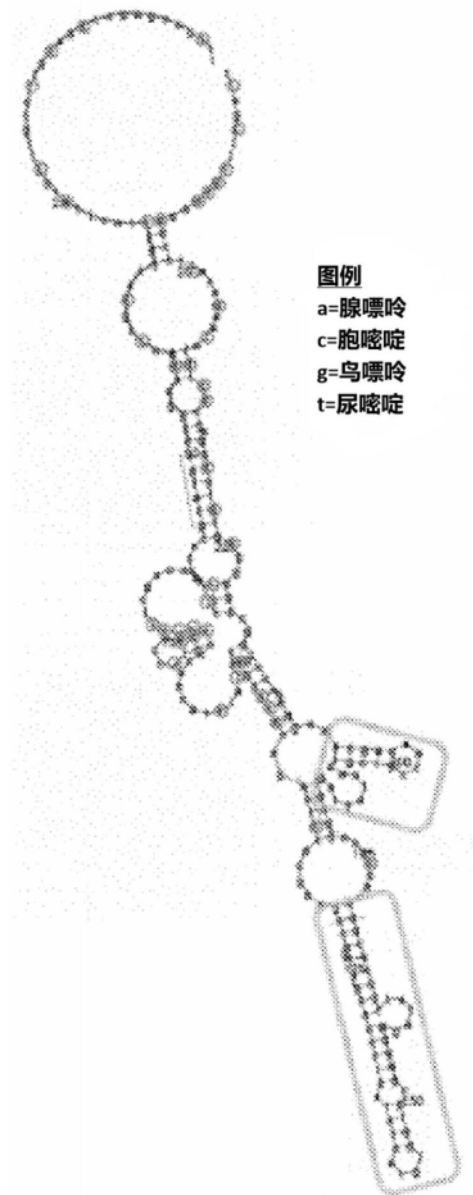


图6

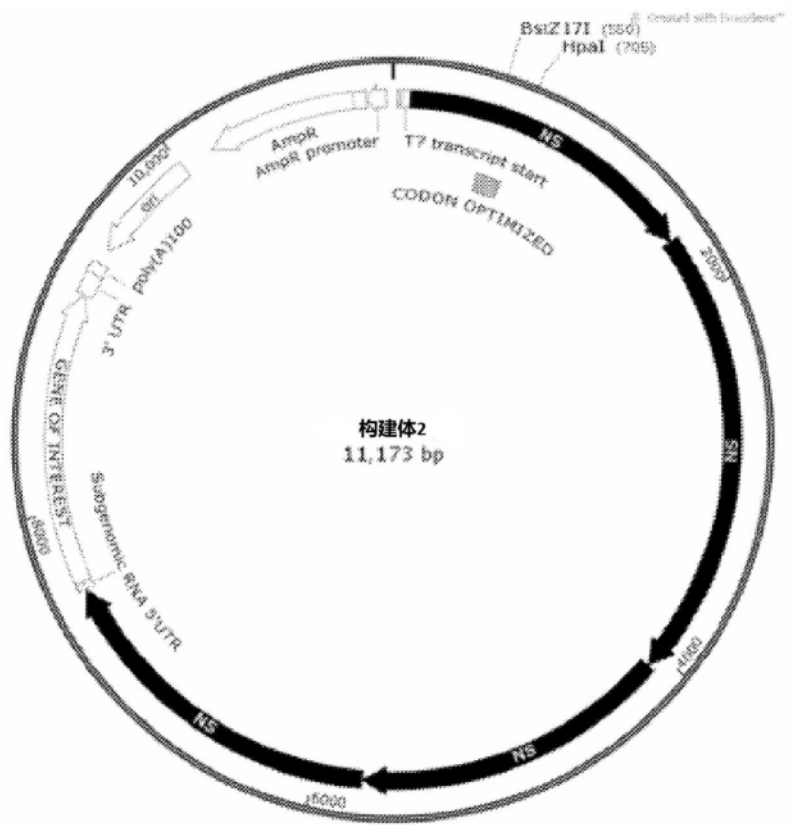


图7

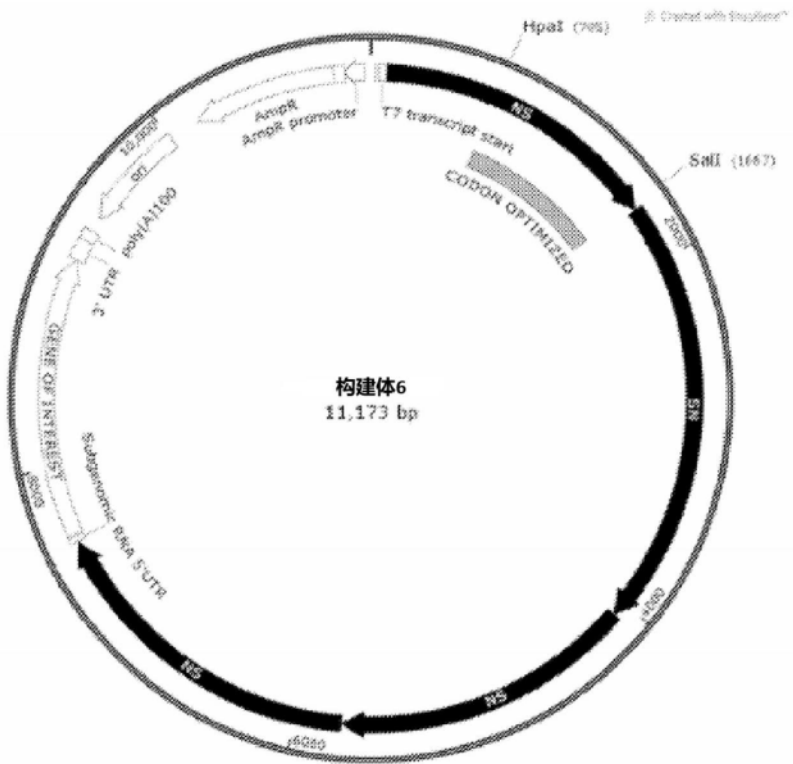


图8

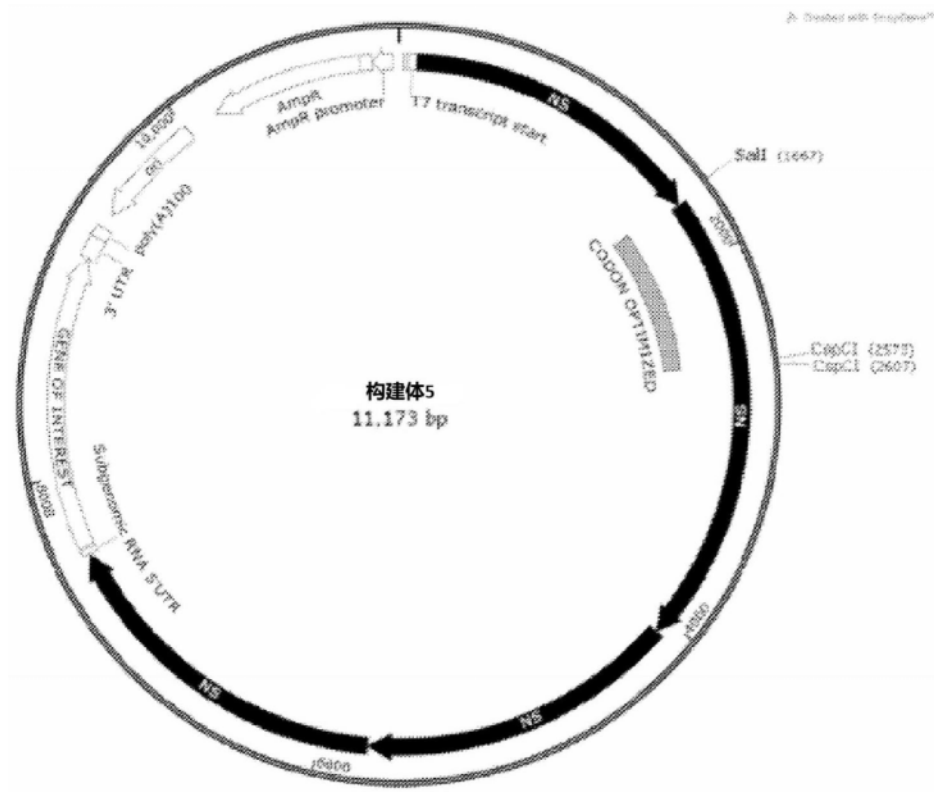


图9

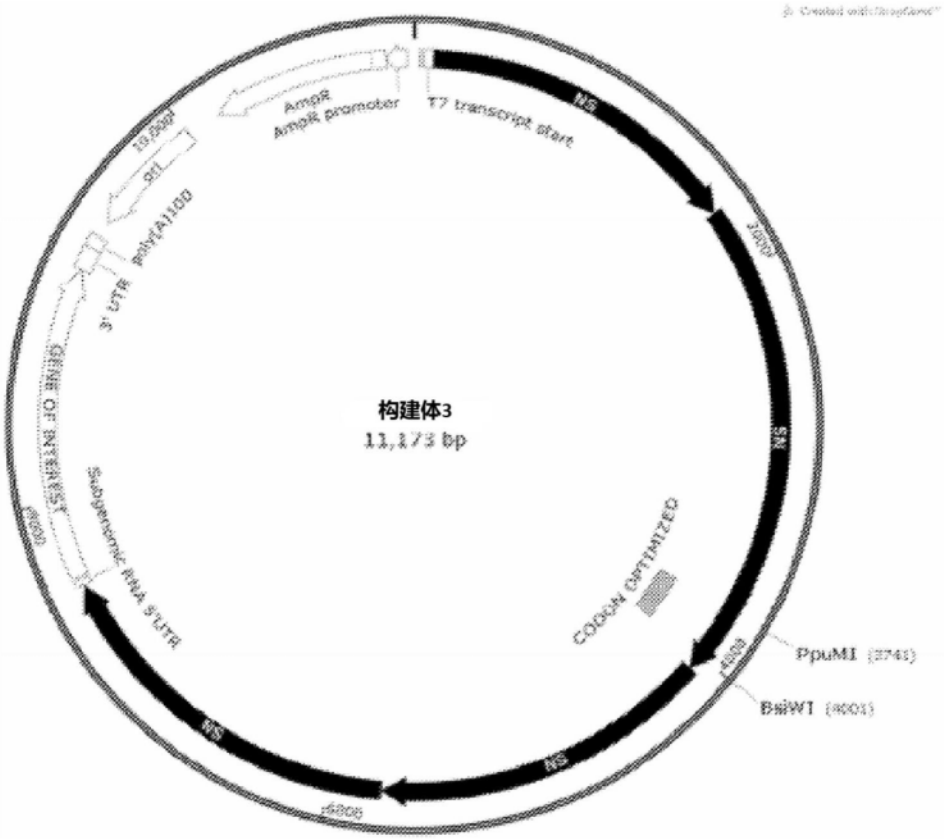


图10

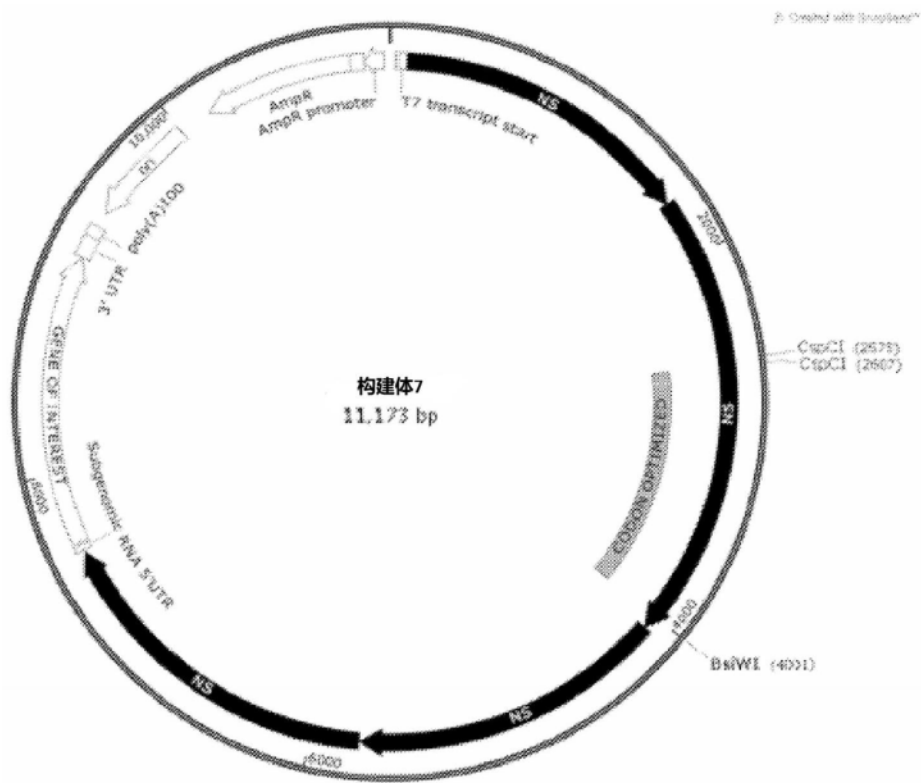


图11

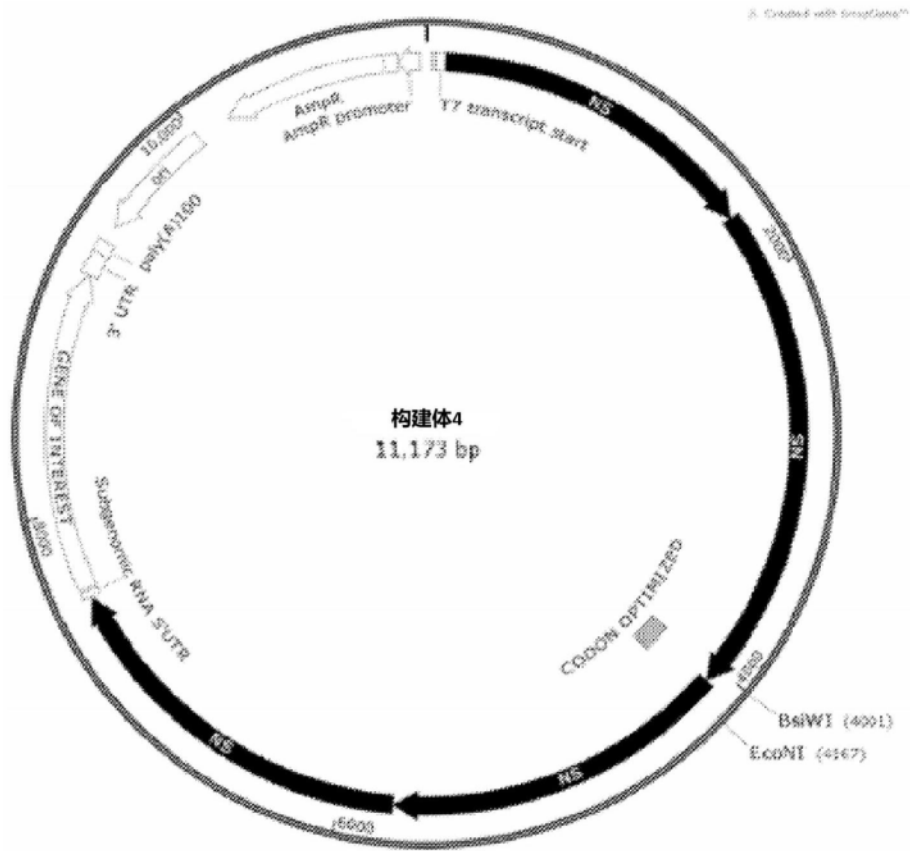


图12

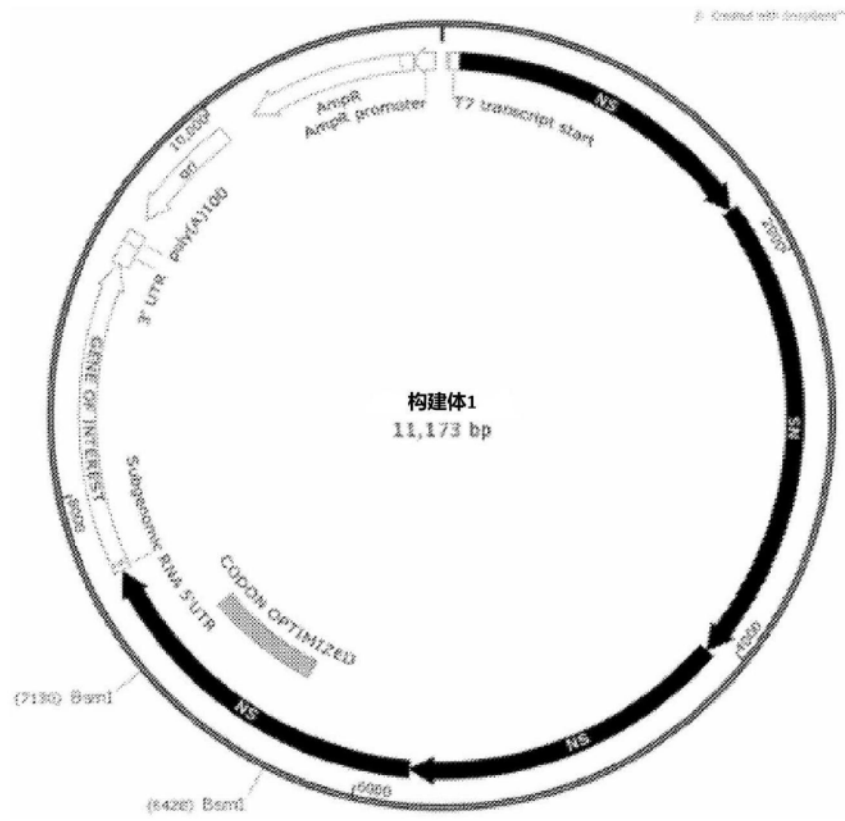


图13

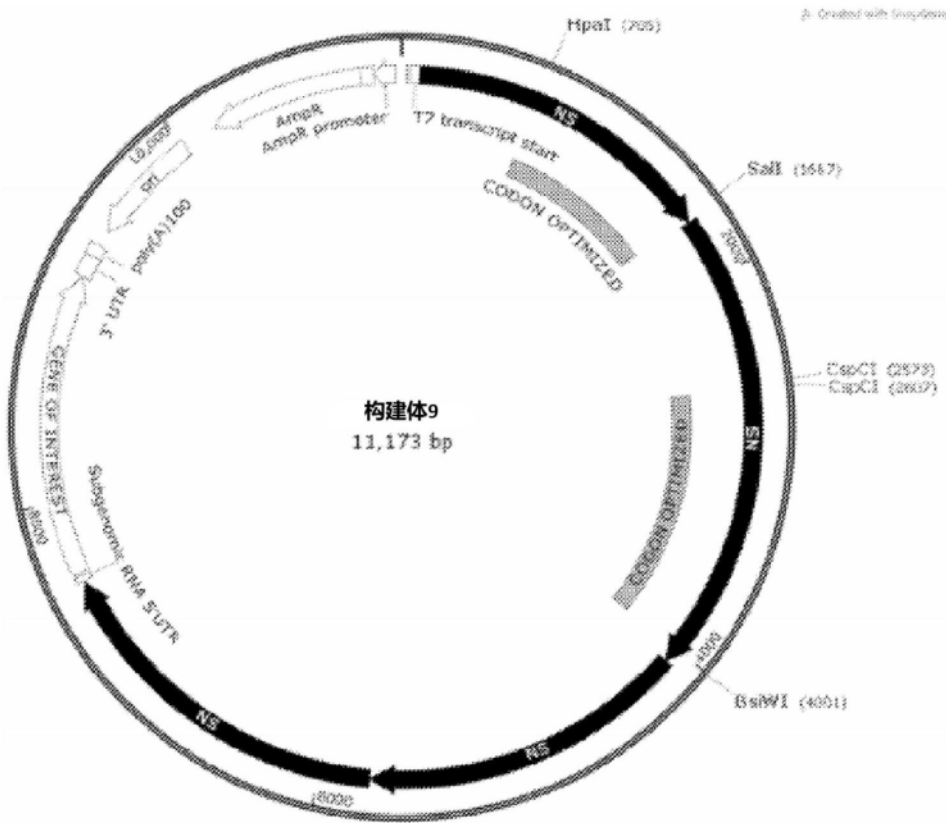


图14



图15

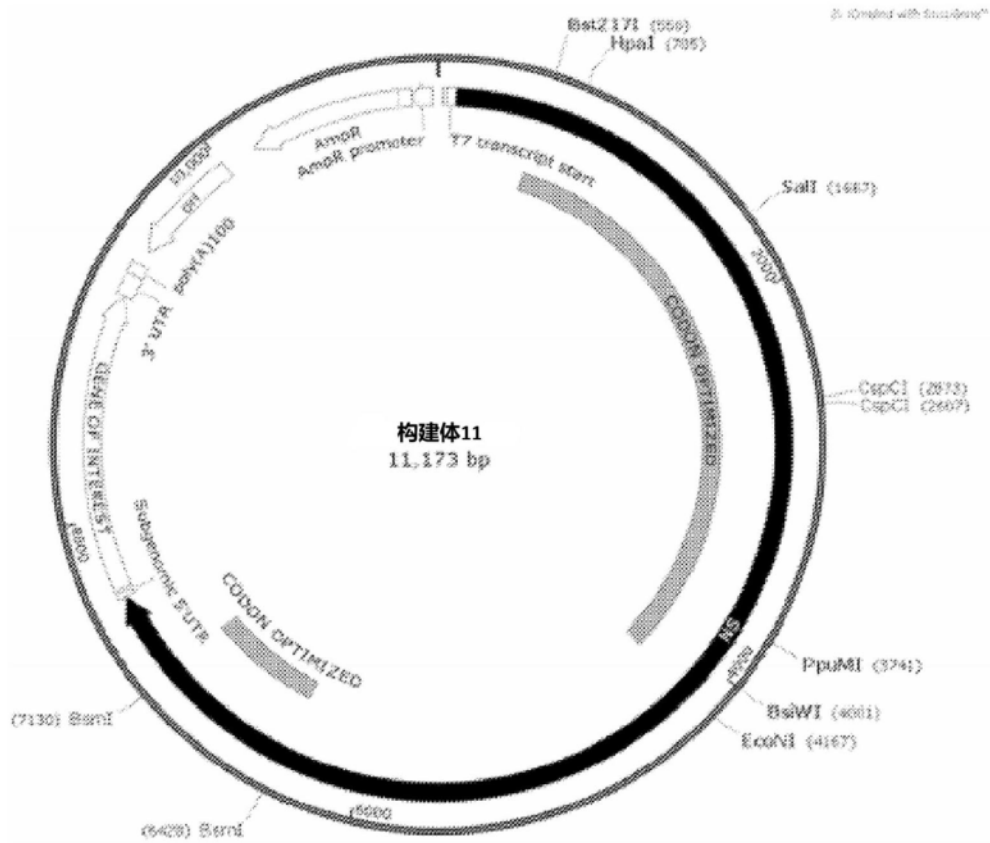


图16

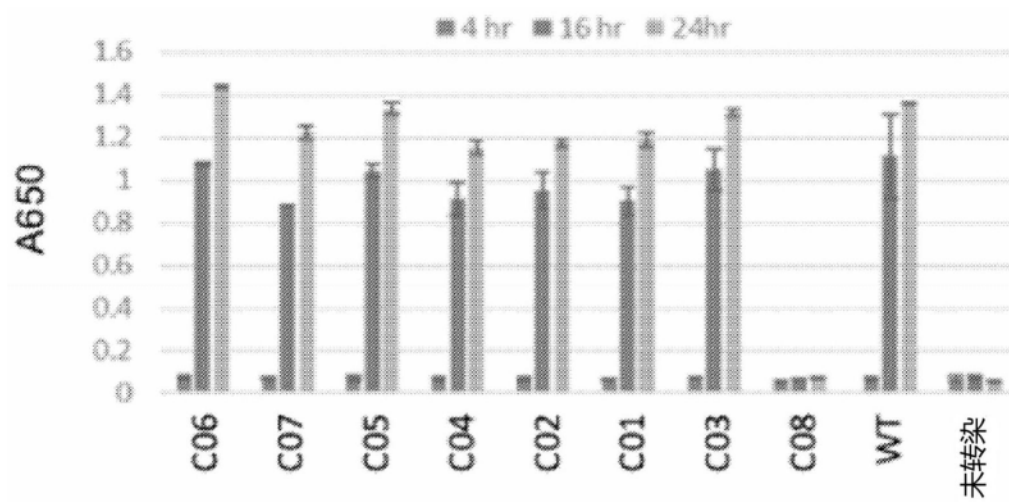


图17A

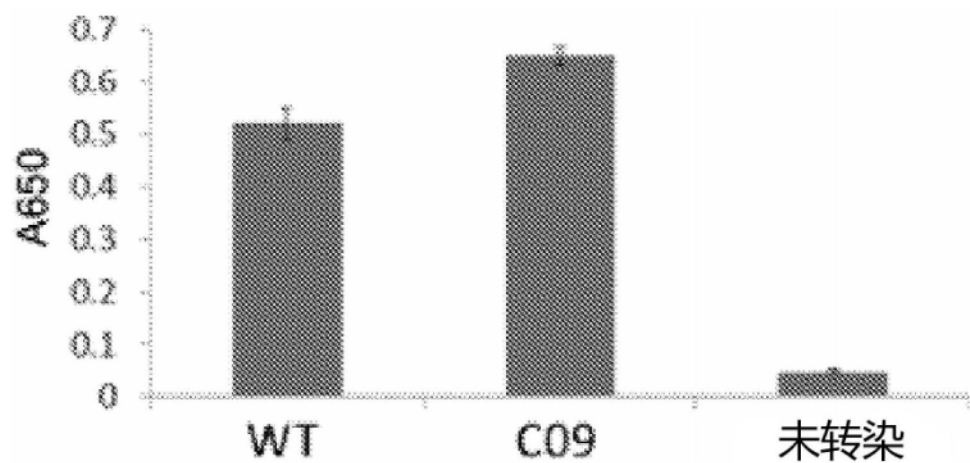


图17B

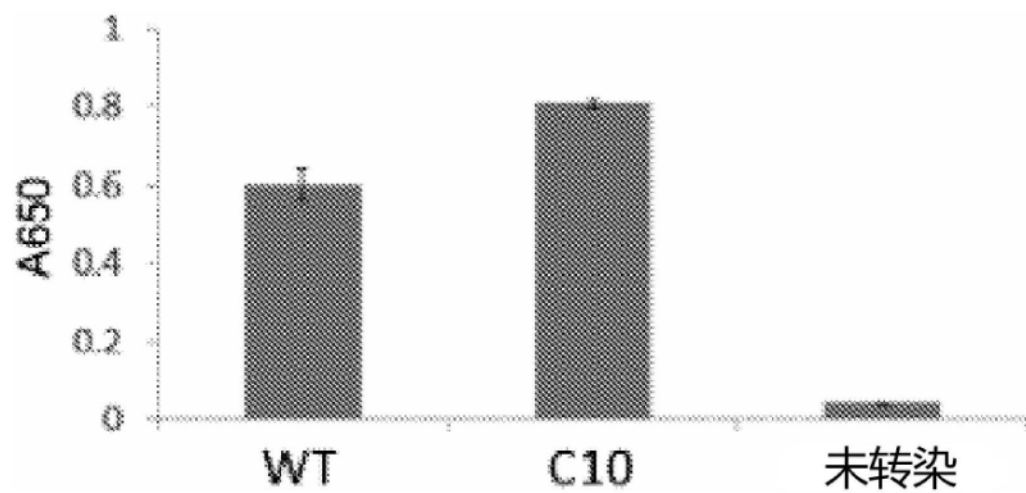


图17C

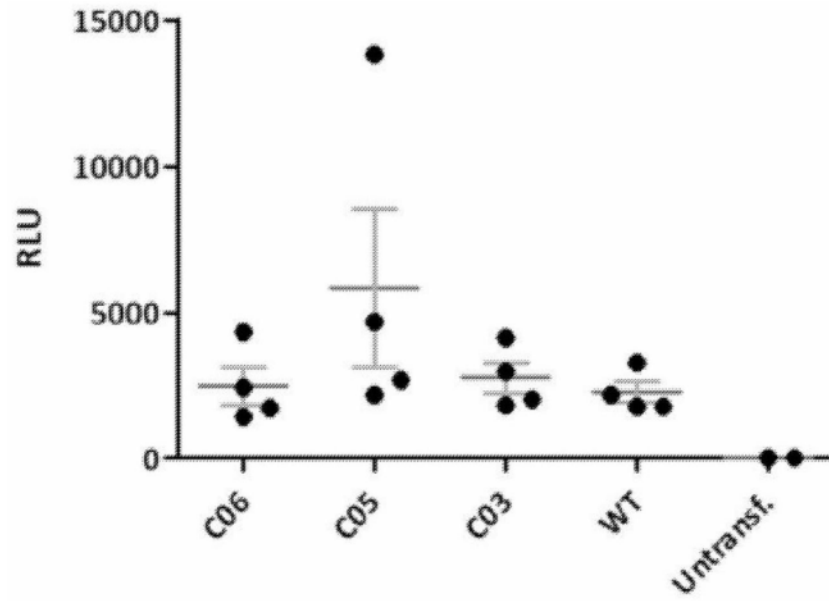


图17D

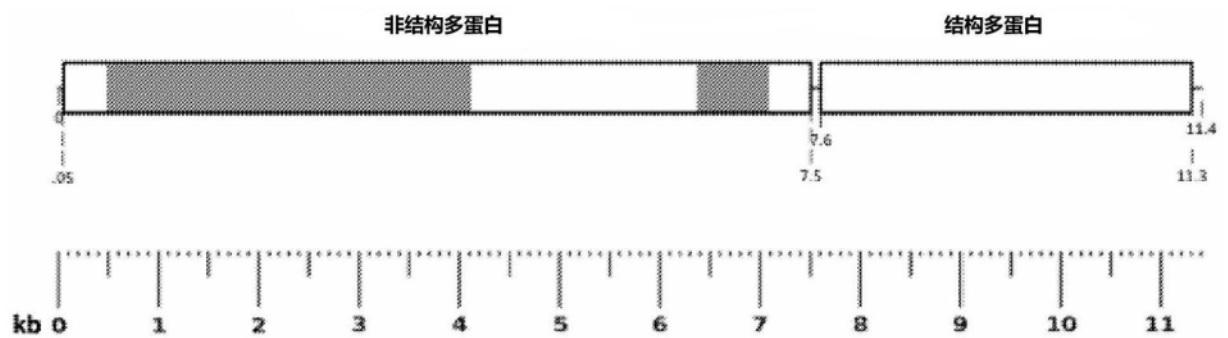


图18

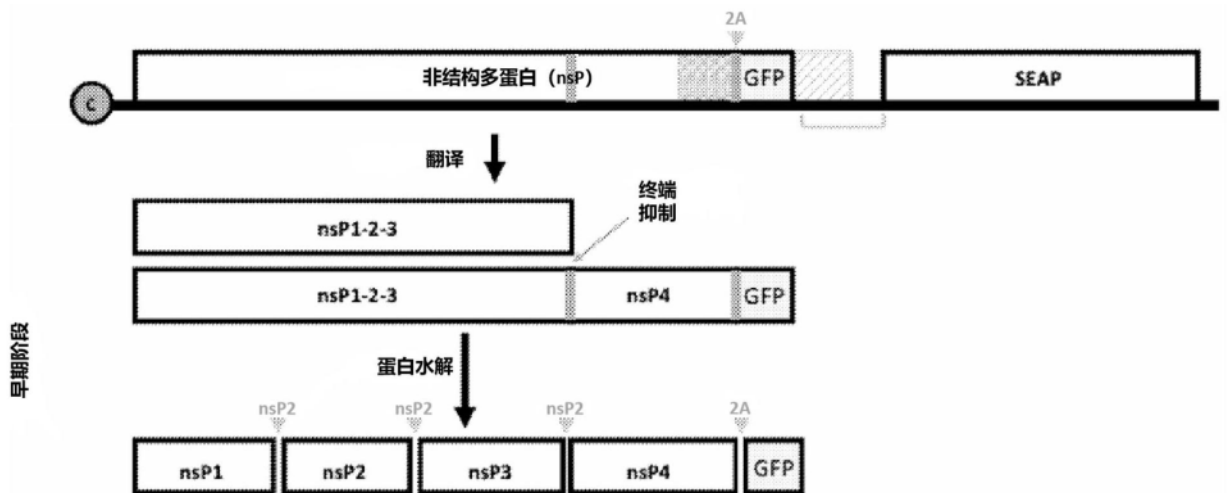


图19

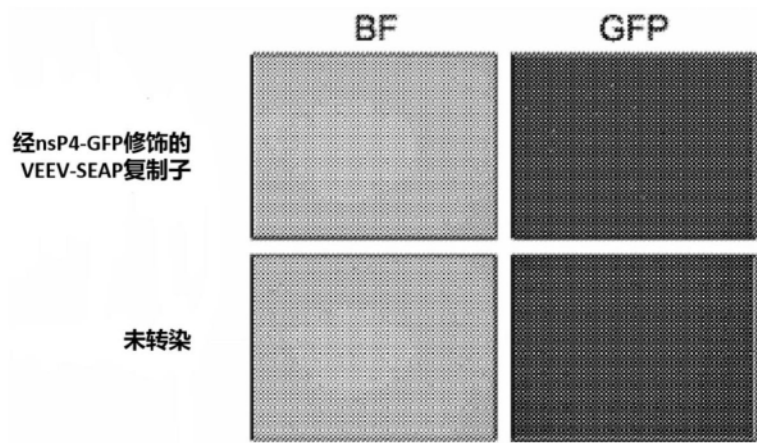


图20A

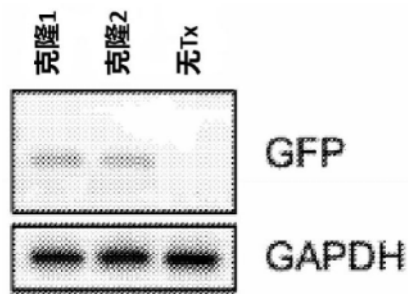


图20B

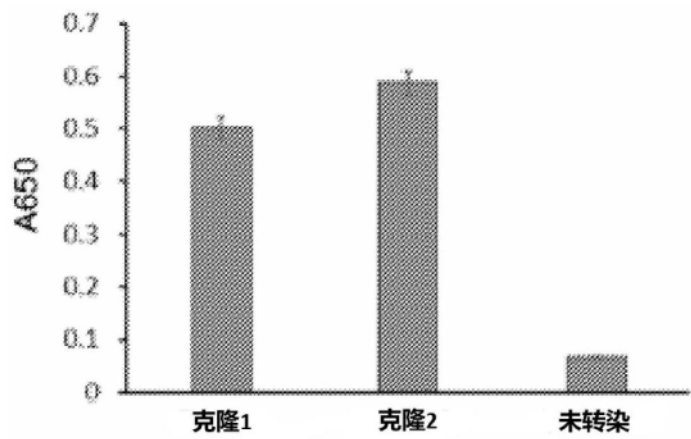


图20C

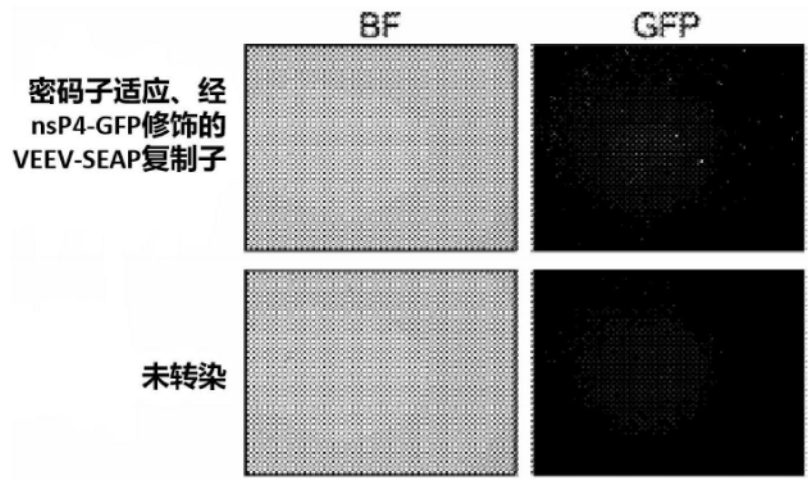


图21A

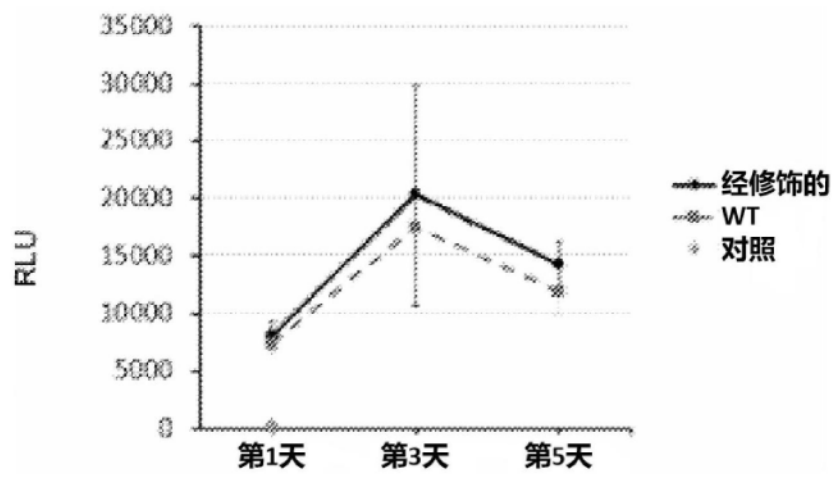


图21B

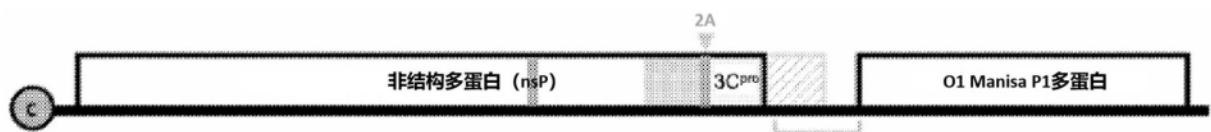


图22A

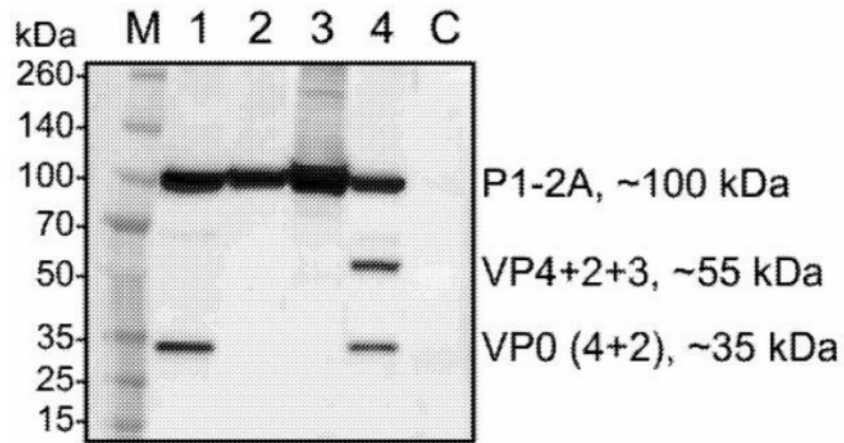


图22B

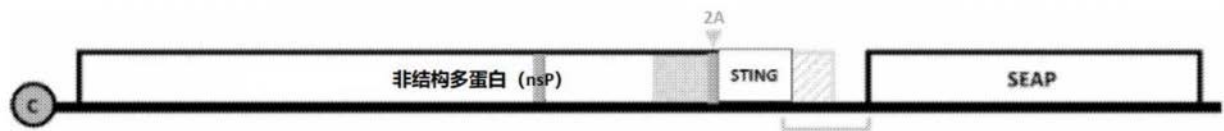


图23A

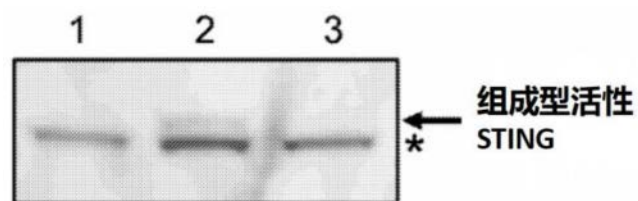


图23B

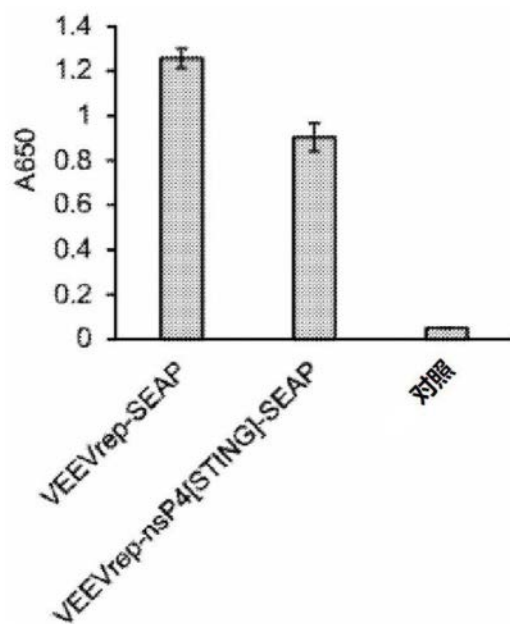


图23C

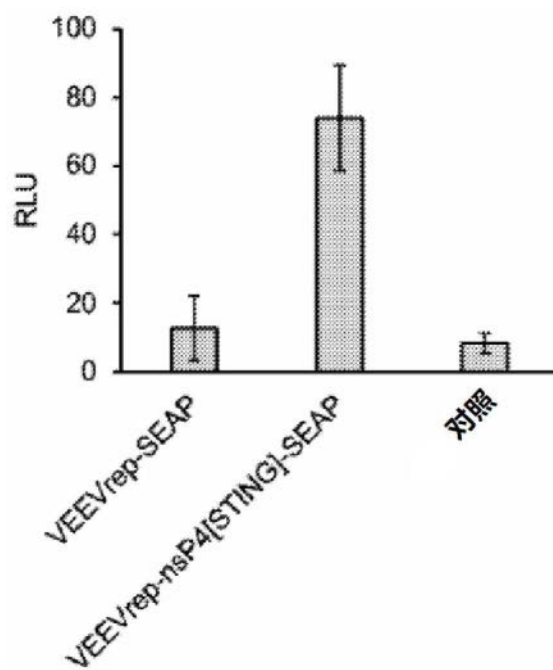


图23D

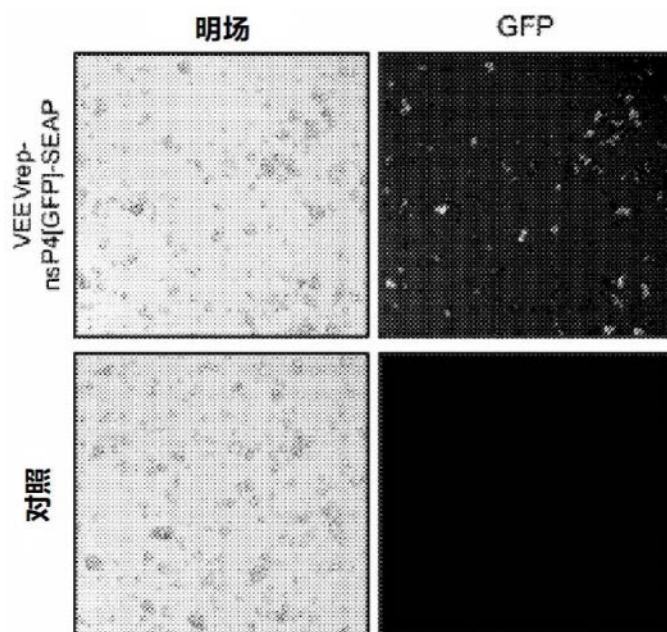


图24A

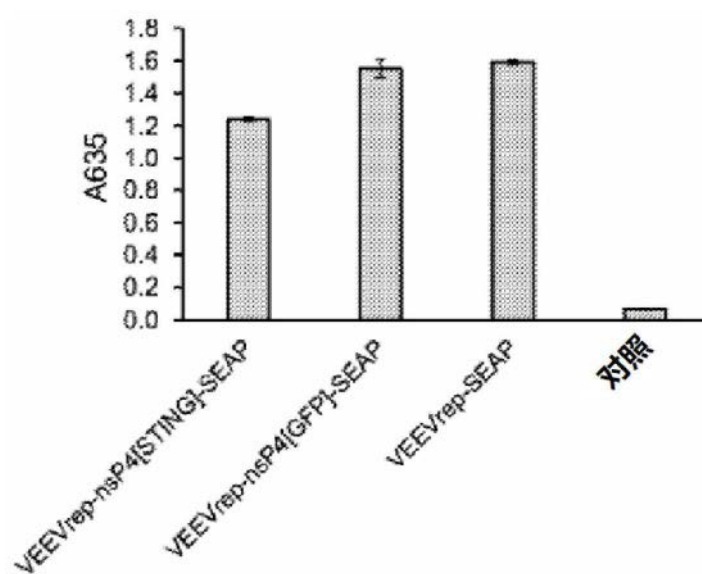


图24B

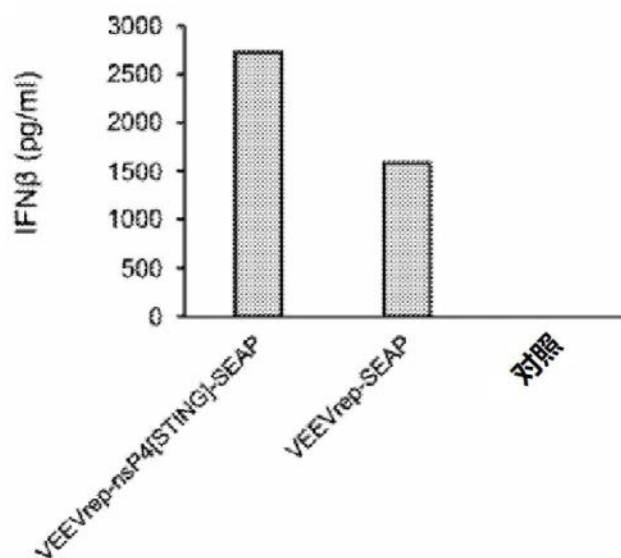


图24C

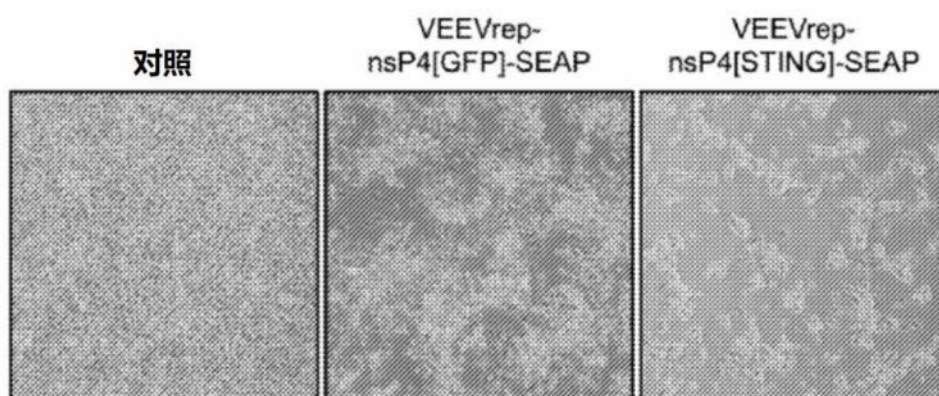


图24D

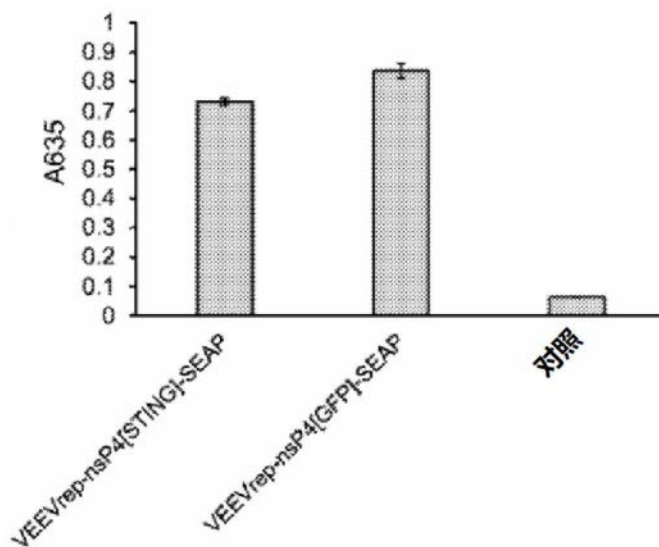


图25A

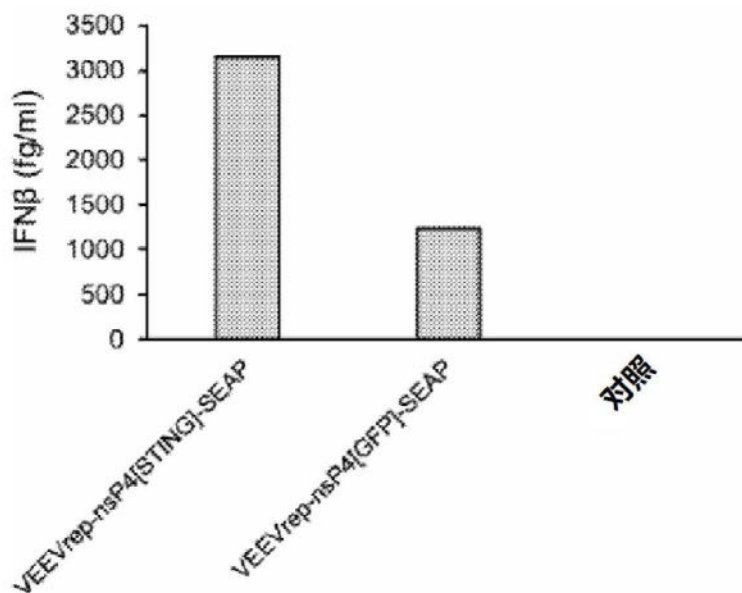


图25B

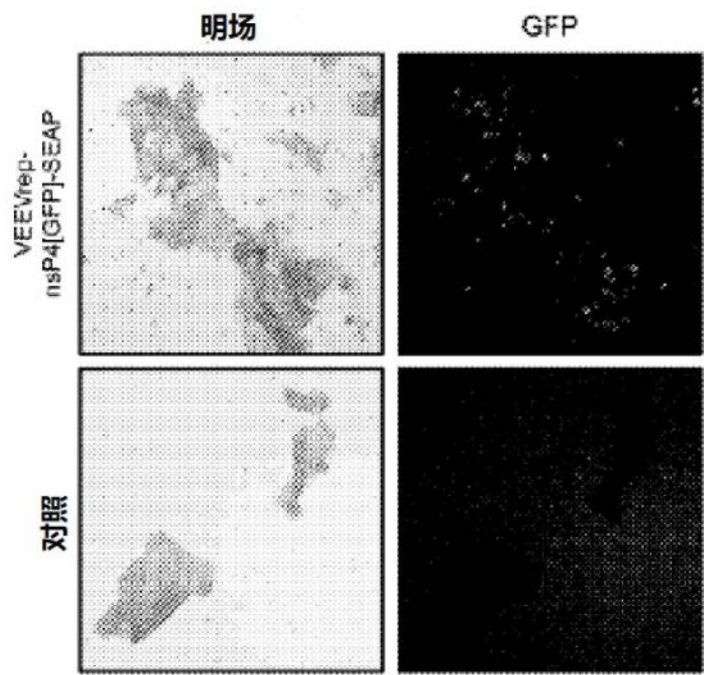


图25C

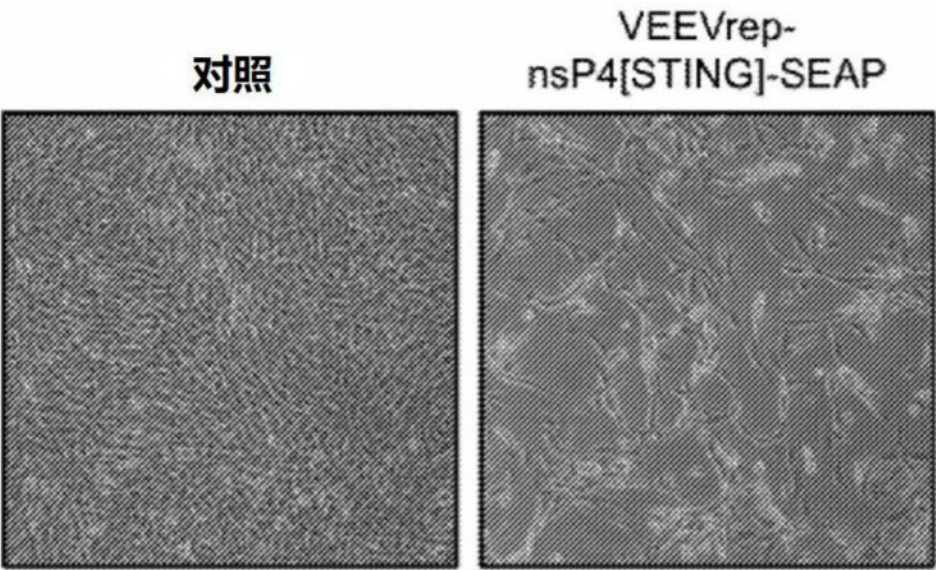


图25D

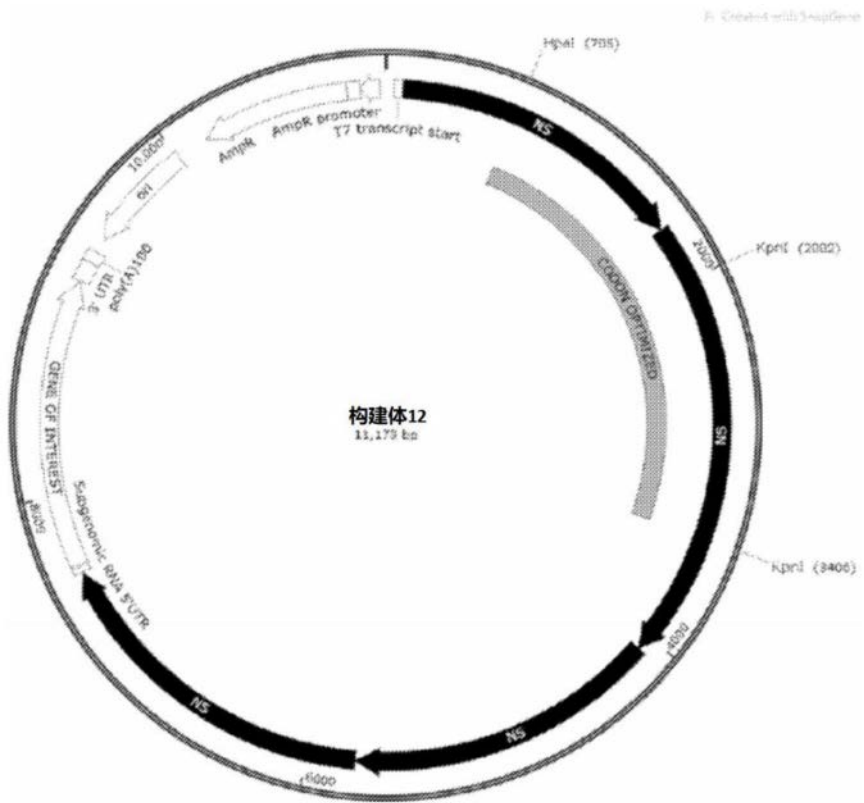


图26A

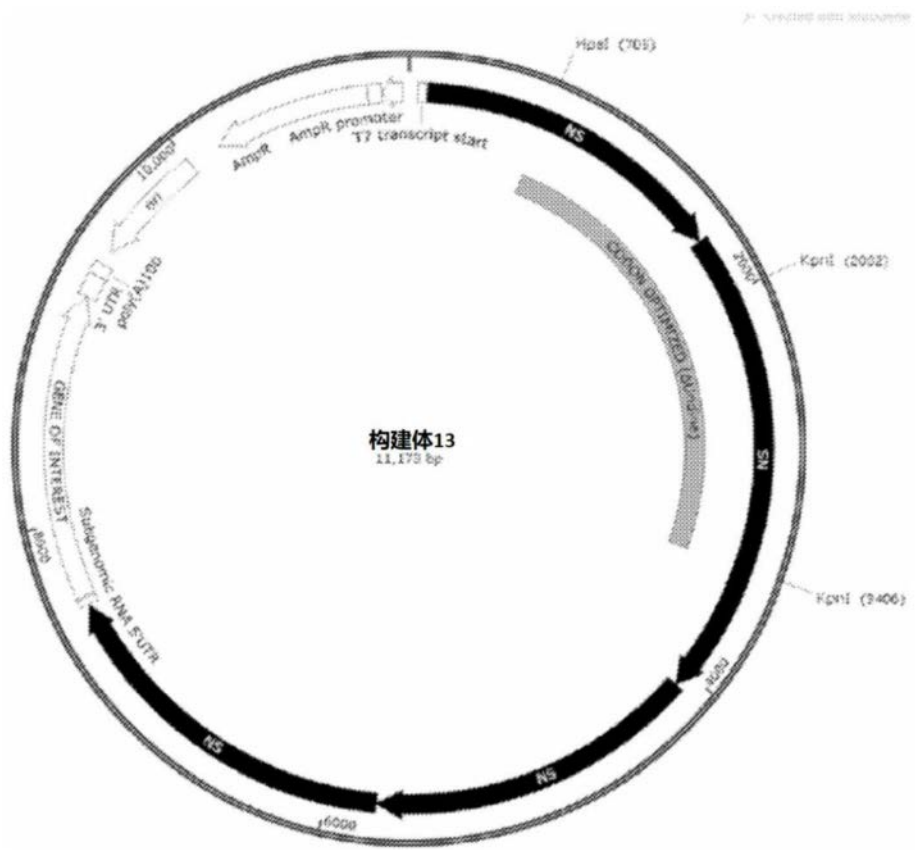


图26B

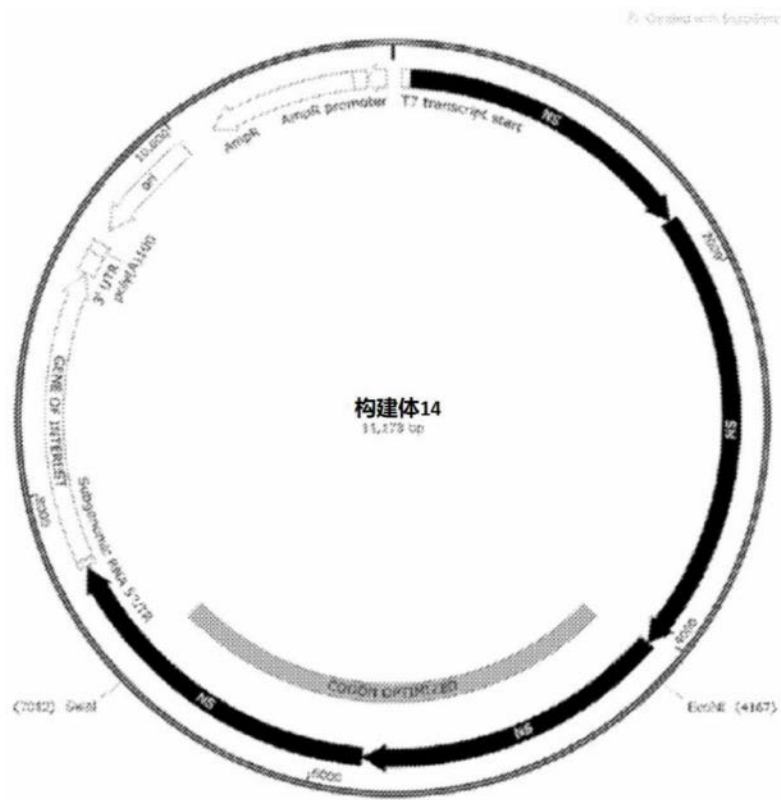


图26C



图26D

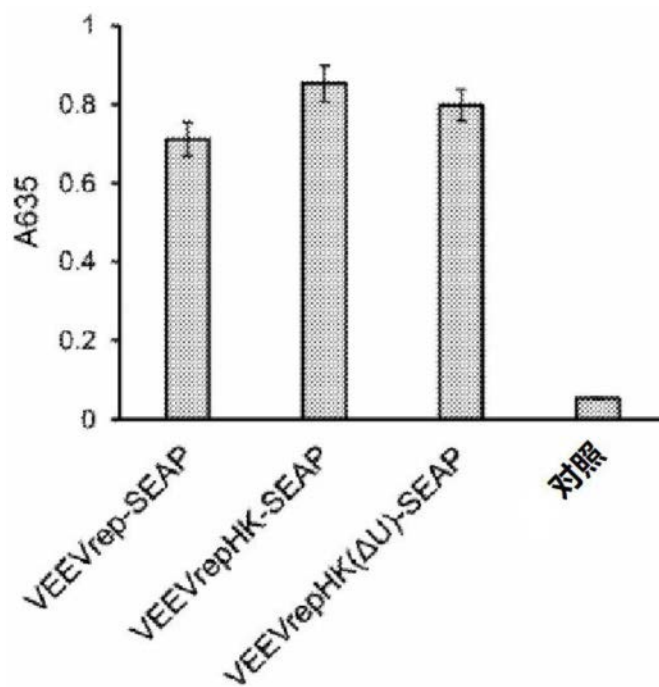


图26E

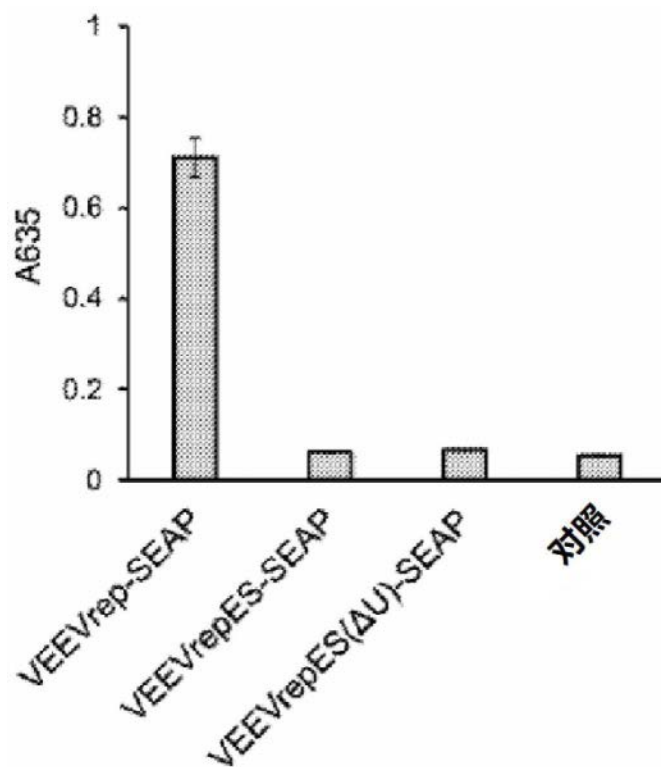


图26F

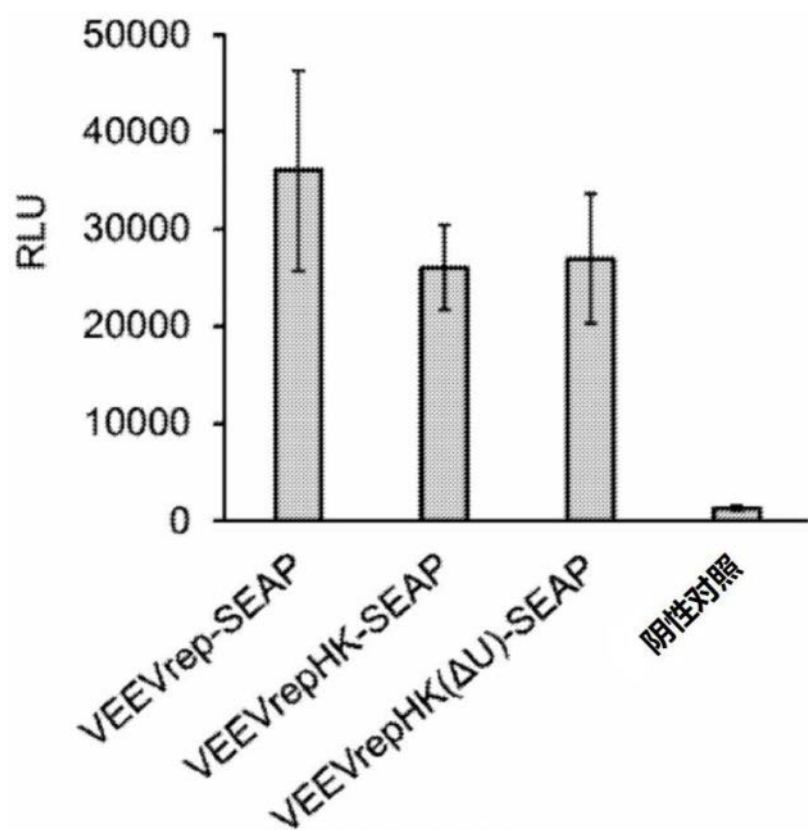


图26G

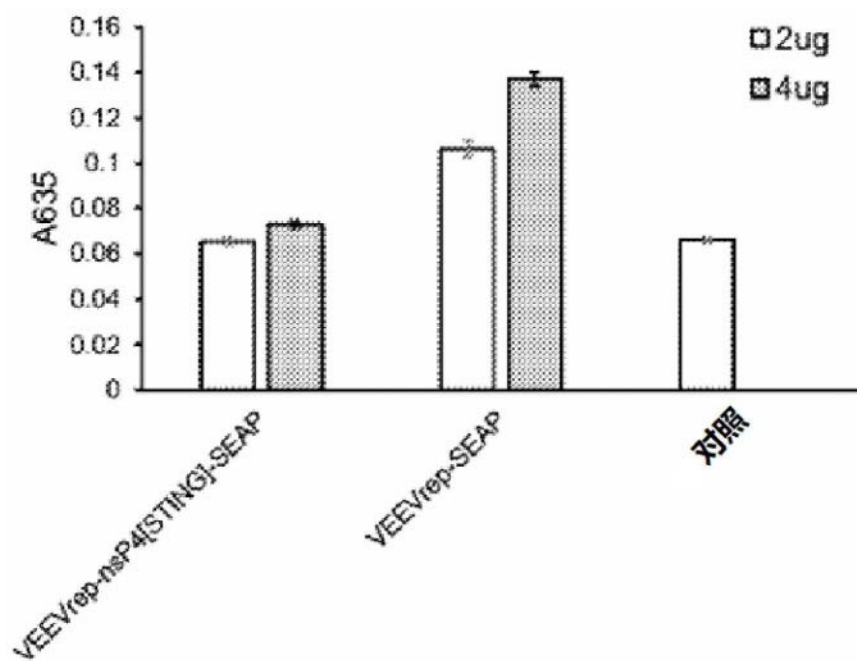


图27A

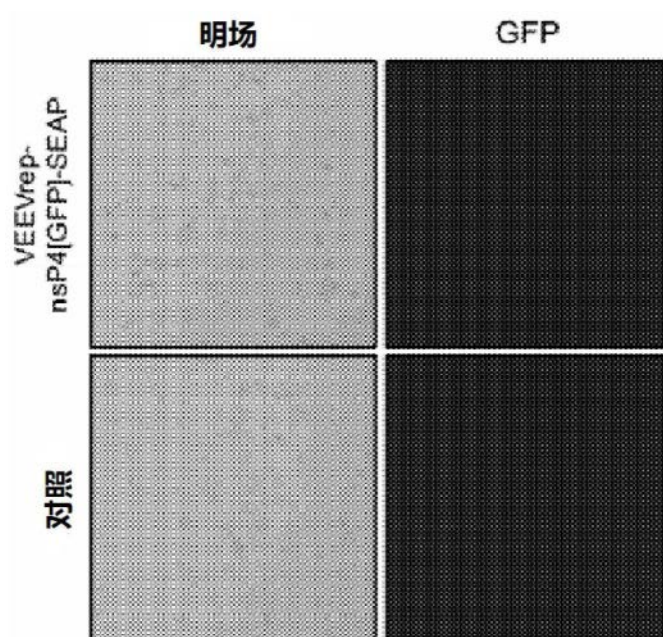


图27B

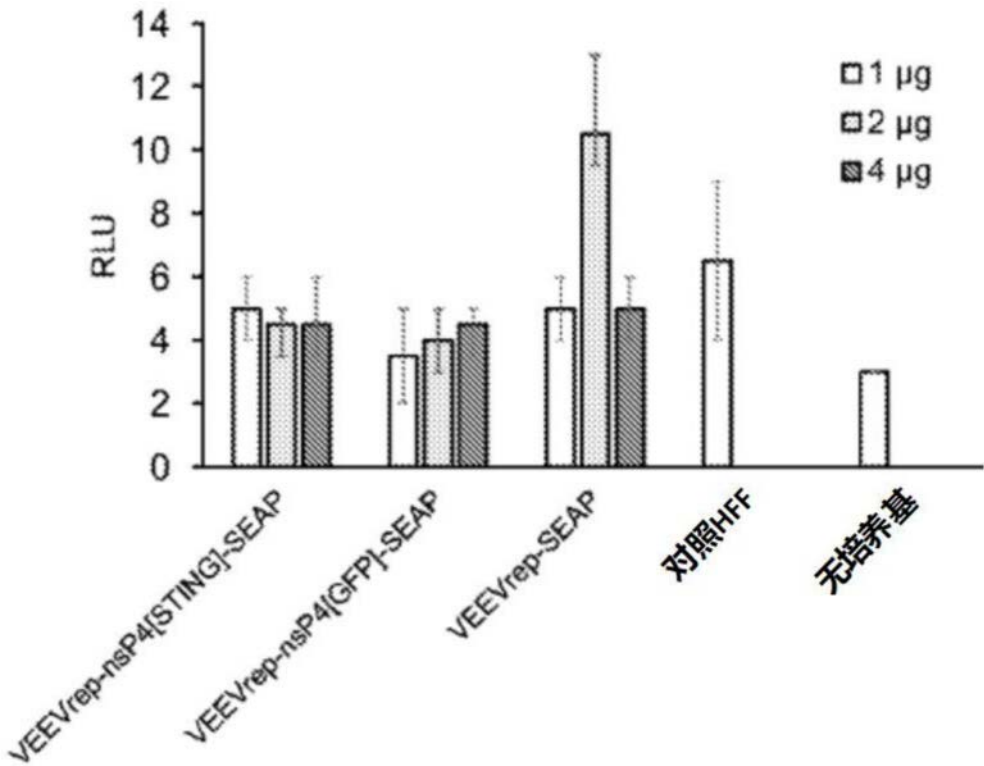


图27C

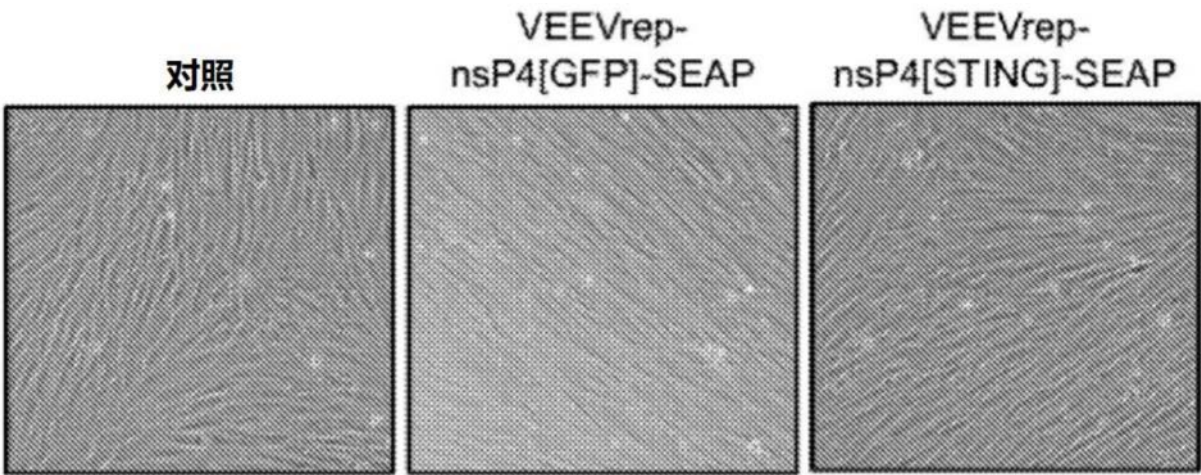


图27D